NUI MAYNOOTH

Ollscoil na hÉireann Má Naud

The Modelling of Scattering in Corrugated Waveguides

Presented by

Patrick Tully Peacocke

Thesis submitted in fulfilment of the requirements for the degree of Ph.D.

Department of Experimental Physics,
Faculty of Science,
National University of Ireland Maynooth

26th May 2012

Head of Department and Thesis Supervisor

Professor J.A. Murphy, M.Sc. Ph.D.

Acknowledgement

I would like thank to Professor Anthony Murphy, thesis supervisor and colleague in the Experimental Physics department, NUI Maynooth, to whom I am indebted for my time in the department and with whom I had many interesting discussions over the years that I spent in the department. Most of the work in this thesis was undertaken on contracts funded by PRODEX to work either for the European Space Agency Technical division, ESTEC, or for the Planck High Frequency Instrument consortium. Without the funding from PRODEX none of the work would have been possible. I also thank the staff of the Experimental Physics department, Brendan Crill of CalTech, Antoine Chamballu of Imperial College, London, Dominic Doyle and Jan Tauber of ESTEC, and Giorgio Savini and the members of Peter Ade's group at Cardiff University who provided some of the experimental data that is presented in support of the numerical models that constitute much of the work to be presented in the following chapters.

The designs for the horns studied in this work were provided by the Planck HFI team and by Peter Ade of Cardiff University, and the telescope reflector metrology data and telescope design was provided by ESTEC and are the work of the various European industrial groups cited in the text.

Contents

List of Figures

Chapter 1

Introduction

An optical instrument can be thought of as determining an operator that acts upon the direct sum of the Hilbert spaces of all input and output electromagnetic fields. The operator describes what happens to the fields: the reflection, scattering and transformation of the fields that are variously transmitted or scattered back to the function spaces from which they originated. The operator, the Hilbert spaces, and the vectors in those spaces that describe the input, constitute the model of the system. To analyse the system for performance prediction it is essential to have a good model – a good approximation to that operator – and a way of calculating the transformations that the fields undergo (the action of the operator on its domain) reliably and in a reasonable time, and with appropriate indicators of the accuracy of the result. The systems of particular concern here are horns and waveguides used in the millimetre and sub-millimetre regions of the spectrum, principally for cosmic microwave background studies. The work was driven by the need to understand the properties of the Planck multi-mode horns and the Planck telescope, and to provide the Planck High Frequency Instrument (HFI) team with accurate predictions of the broad band beam pattern on the celestial sphere formed by the four 545 GHz and four 857 GHz pixels. The High Frequency Instrument on the Planck telescope has been designed to measure the anisotropies in the Cosmic Microwave Background using single mode horns at 100 GHz, 143 GHz, 217 GHz and 353 GHz while simultaneously observing the astronomical foreground using multi-mode horns at 545 GHz and 857 GHz. The ability to simultaneously observe both the CMB and the foregrounds is unique to Planck among CMB experiments and was implemented to give improved foreground subtraction, and hence better CMB maps.

To derive the relative positions of the patterns on the sky required an accurate system model that included a model of the telescope with reflector shapes and positions as they are at L2, a correctly positioned focal plane assembly, and broad band predictions of the beam patterns of all thirty six pixels in the HFI focal plane assembly. The development of the software for the modelling of the horns, the derivation of the telescope construction parameters from the metrology and calibration data, and the findings of the study are the subject matter of this work.

To understand and to describe the instrument requires that the operators of the component parts be found. This process of constructing the operators is the building of the system model – an abstract mathematical model. The model describes the action of the instrument on the electromagnetic fields, and the process of abstraction throws into relief the assumptions made about the physical instrument and processes. Abstraction also clarifies the extent of applicability of the model process: if assumptions about the physical process are required to build the model (to render tractable the mathematics) then it is evident that the model is not applicable to any situation in which those physical assumptions do not hold at least very closely.

The concern of this thesis is applied engineering, primarily the analytical methods used to model the scattering of electromagnetic fields in corrugated waveguides and the transmission of the fields through optical systems. The background theory has been known for a long time; most of the theoretical development of the waveguide scattering analysis was worked out between the 1930's and the 1950's (see the extensive biography in [38]). The early work was analytic and the aim was to describe the optical properties of the guide and the scattering in terms of delay lines, notably in the work of Schwinger (see [42] and the references therein). When computers became accessible to university and industrial research groups, numerical methods of analysing scattering and propagation began to be developed – numerical solutions to differential and integral equations, and moment methods and mode matching. Within this work only mode matching and integral methods will be used. These methods have a long pedigree and the emphasis here is upon mathematical rigour and the reformulation of the equations into forms that can be computed quickly and accurately with reliable run-time estimates of the accuracy and validity of the results. Speed and precision have proven to be essential for the detailed analysis of the Planck multi-mode horns and to the study of the sensitivity of radiated field structure to manufacturing tolerances in any corrugated waveguide.

The second subject is the building of the Planck telescope models required for deriving accurate beam pattern predictions on the sky. These are not models of idealised telescopes, but attempts to build as precise a model of the telescope, as it is in operation at L2, as is possible given the data available at the time that the work was undertaken. The scientific context of the instruments is the concern of the cosmologists, not of the engineer, and has not been discussed.

Chapter 2 presents a short introduction to mode matching and the modelling of horns and waveguides. It then proceeds to look at the results that the numerical models gave for the far field beam patterns of the telescope for the multi-mode channels. Extensive numerical modelling, supported by measurements of the power transmission of the Planck pixels by Cardiff University, show that the design of the horn assemblies renders them very sensitive to manufacturing tolerances, and that matter is examined.

Chapter 3 presents the theoretical background for the formalism of mode matching and prepares the approach taken to the coding of the scattering equations for numerical simulation of waveguides, then goes on to discuss some of the consequences of the formalism.

Chapter 4 develops the scattering formalism for circular and rectangular waveguides and ends with an examination of the problem of predicting performance when assemblies of circular waveguides are not perfectly aligned. The misaligned case is analytically complex and leads to numerically intractable equations, so approximations and an approach to their solution are sought.

Chapter 5 gives an account of the scattering software developed for modelling the Planck multi-mode horns, software that was used for all of the simulations of waveguides and horns presented in this thesis and delivered to ESTEC and the HFI consortium.

Chapter 6 takes the process of abstraction of the field models further. The primary aim was to produce the most concise description of the fields in the horn aperture that is possible: a set of vectors of minimal size consistent with accurate modelling. There follows a further abstraction to look at model perturbations by infinitesimal group actions on the output of the models. This gives a way of simulating manufacturing errors in the horns.

In chapter 7 the thermo-elastic deformations of the Planck reflectors is taken into account in the preparation of the engineering model of the telescope, work that was done under contract for ESTEC.

Chapter 8 gives an account of the work done on building a ‘reverse engineered’ model of the telescope as it is at L2 in the light of the (then very limited) data from the preliminary in-flight calibration of the HFI. This work was also undertaken for ESTEC under contract.

There follows an appendix containing some basic mathematical definitions and equations, some comments on Fourier spectra, and the connection between the eigenfunctions of the Fourier transforms on $L^2(\mathbb{R})$ and Gaussian beam modes.

Chapter 2

Models of horn assemblies

Corrugated horn assemblies are the most common structures used to feed bolometer cavities for power detection in cosmic microwave background experiments. They have been used in ClOVER [2], WMAP and Planck [40, 44, 65], and they were used to couple the image of the sky formed by the telescope onto the detectors for the Herschel Heterodyne Interferometer for the Far-Infrared spectrometer at frequencies in the range 480 GHz to 1120 GHz, [49]. Traditionally horns have been single-mode (meaning the radiated field in the horn aperture is described by a single, frequency dependent, function) whether polarised or unpolarised, with a main beam profile that is approximately a Gaussian power distribution. single-mode corrugated horns have long been in common use for millimetre wavelength applications, are now routinely manufactured for sub-millimetre wavelengths, and are being developed by groups such as the Rutherford Appleton Laboratory for applications at a few Terahertz. When there is no requirement to know the polarisation state of the electromagnetic field, greater throughput can be obtained by designing a horn structure that supports the propagation of more than one mode: few-mode or multi-mode horns in which each additional mode can be thought of, very loosely, as an additional unit of power received. The additional modes give rise to a departure of the beam pattern from Gaussian, and the simulation of these assemblies becomes ever more computationally intensive as more and more modes are included. The computational cost of simulation can be further exacerbated by the design of the assemblies and by poor programing.

In this chapter a concise conceptual overview of the simulation of electromagnetic field propagation and scattering and corrugated horns is given: the method of mode matching, the theory and numerical implementation of which will be the dominant theme for chapters 3, 4, 5 and 6. Thereafter follow sections on the results of modelling the Planck 100 GHz single-mode horn assemblies and the ultra-Gaussian 150 GHz ClOVER horns with a comparison of the evolution of phase and beam shape within the two horns. Section 2.4 presents results from the modelling of the Planck multi-mode horn beam patterns and, where available, comparison with the preliminary beam pattern measurements derived from the planet scans from L2. All of the Planck horns exhibit resonance due to their structure. That resonance leads to extreme sensitivity to manufacturing tolerances, issues that are discussed in sections 2.5 and 2.6 in which it is shown how performance prediction requires attention to manufacturing tolerances as well as to system design. The chapter ends with conclusions addressing the modelling of horns, with particular emphasis on Planck.

2.1 An overview of mode matching

The numerical simulation of corrugated horns by the method of mode matching has been in common use at least since an outline of the method, as applied to single-mode horns and waveguides, was published in [47] building upon earlier work in [13]. The methodology was developed at NUI Maynooth by Murphy for the analysis of multi-mode horns, and mode matching code written and applied to the design of both single and multi-mode horns for CMB applications, in particular for Planck [45] [14] [25] and QUaD [46].

Before proceeding something needs to be said about the general mathematical framework into which the method of mode-matching fits. All necessary mathematical definitions are given in appendix A on page 227. The mathematical setting is functional analysis, specifically Hilbert space theory, and concerns linear spaces of functions or spaces of operators on those functions. The spaces of functions are always those functions describing the electromagnetic fields within a waveguide, the appropriate function space being determined by the waveguide geometry and the boundary conditions. The operators describe the scattering and propagation within a function space or between pairs of spaces, or they are operators acting on a space formed from a pair of function spaces. In all cases the structure of the function spaces allows for the operators to be represented as infinite by infinite matrices that decompose naturally into sub-blocks. Each sub-block is itself an infinite by infinite matrix. Because this thesis only addresses single input, single output situations the structure of the operators can be summarised as follows: The operator is a linear mapping on the direct sum of the input and output spaces; it decomposes into four blocks, each of which is an operator that is a linear contraction on either input or output space, or a linear contraction between input and output, or *vice verca*. Each of those sub-operators is further decomposed into sub-operator blocks that map the electric field components to electric field components, magnetic to magnetic, or between electric and magnetic. In all cases the structure of the operator follows naturally from the structure of the function spaces, and the structure of the function spaces follows naturally form the electromagnetics. The formalism extends to N input, M output systems, but that will not be addressed here as it is not relevant to the applications that have driven this work.

Conceptually the method of mode-matching is simple: Physics enters the picture in the form of wave propagation along the guide axis, described by Helmholtz equation, waveguide impedances, the requirement that power be finite and conserved, and in the conductivity at the waveguide walls that determines boundary conditions; the rest is Sturm-Liouville theory and general Fourier analysis and is geometric in character. The boundary conditions and Sturm-Liouville theory imply the existence of denumerable orthonormal bases for the function space over the waveguide cross-section and, in any such basis, any electromagnetic field satisfying the boundary conditions can be described. The function space will be complete with norm determined by an inner product, and therefore the spaces will be Hilbert with complete basis. All such bases are related by unitary transformations of the Hilbert space. In any such basis the transverse electromagnetic field can be expanded, each basis function, or mode, describing an elementary field particular to the basis.

At a junction in the waveguide the elementary fields are partially reflected (attenuated and changed in phase but not scattered into other modes) and partially scattered into the fields in the next section of the guide with the total power, measured by the L^2-norm (see Appendix A for definitions) of the reflected and scattered field, being conserved. The impedance of each elementary orthogonal electric-magnetic field pair determines the phase shift of that field pair as it propagates a unit distance along a parallel section of the waveguide, and it is the changing relative phase of these basis elements that describes the evolution of the total field as it propagates. The impedance change across the junction for each elementary field determines the reflection amplitude while description of the scattering process across a junction is a purely geometric matter described entirely by the geometry and an appropriate generalised Fourier analysis. As in all Fourier analyses the Fourier coefficients are given by the inner product for the function space, and a mode on one side of a junction is expanded in terms of (the pull-back of) the modes on the other side of the junction with the expansion coefficients interpreted as scattering amplitudes. If the waveguide cross-section at either side of the junction has the same shape, then the function spaces can be taken to be the same but with different domains of definition and normalisation factors, but are trivially isomorphic; if the waveguide cross-sections are different (for example the scattering from a rectangular to a circular guide) then the function spaces are necessarily different, but whatever the basis sets or function spaces may be the mathematical formalism describing the physical process remains the same.

For a single-mode horn there is a single, denumerable, basis for the space of transverse fields in each section of the horn. For a multi-mode system there is a denumerable set of denumerable bases. Since each basis is denumerable the reflection and scattering processes within that basis can be simply described by a matrix representation of the operators. Strictly speaking these operator matrices describing the scattering of modes within a basis are infinite by infinite complex arrays, and every basis element on one side of the junction scatters some power into every basis element on the other side. Conservation of power means that *essentially* all of the power is contained in a finite subset of such modes. Mathematically this means that a judicious choice of a finite dimensional subspace of the function spaces can be made which, from a practical point of view, will give an adequate model of the process with finite matrices as operators. This is simple to describe: let the L^2-norm of the field $\boldsymbol{a}$ be $\| \boldsymbol{a} \|_2$ and its coordinate representation in the basis be $\boldsymbol{a} = (a_1, a_2, \ldots, a_n, a_{n+1}, \ldots)$. Define a sequence $\{\boldsymbol{a}_k\}_{k\in\mathbb{N}}$ by $\boldsymbol{a}_1 = (a_1, 0, 0, \ldots)$, $\boldsymbol{a}_2 = (a_1, a_2, 0, 0, \ldots)$, $\boldsymbol{a}_n = (a_1, a_2, \ldots, a_n, 0, 0, \ldots)$. Then $\boldsymbol{a}_k \to \boldsymbol{a}$ as $k \to \infty$ and the difference in power is given by $\| \boldsymbol{a} \|_2^2 - \| \boldsymbol{a}_k \|_2^2 = \sum_{j=k+1}^{\infty} \boldsymbol{a}_j^2$, which tends to zero. Thus, for any chosen $\varepsilon > 0$ there is some $K_\varepsilon \in \mathbb{N}$ for which $\| \boldsymbol{a} \|_2^2 - \| \boldsymbol{a}_k \|_2^2 < \varepsilon$ for all $k \geq K_\varepsilon$. This number K_ε is the minimum dimension of the model space required to account for a fraction of at least $1 - \varepsilon$ of the power at the junction in the chosen basis, and for given ε different K_ε may be required at different junctions in the same waveguide. But it has to be noted that however large the model space, some power will be scattered out of it at every junction and information lost. Nevertheless, with care and the comparison of the results from different sized models, a suitable dimension for a reliable model can be found. The mathematical description of the conservation of power is, for input vector $\boldsymbol{a}_I$ in space $(\mathcal{H}_A, \| \cdot \|_A)$ reflected at, and scattered across, a junction into space $(\mathcal{H}_B, \| \cdot \|_B)$, we must have $\| \boldsymbol{a}_I \|_A^2 = \| \boldsymbol{a}_R \|_A^2 + \| \boldsymbol{a}_S \|_B^2$, where the norms are 2-norms. The finite dimensional numerical model will give $\| \boldsymbol{a}_R \|_A^2 + \| \boldsymbol{a}_S \|_B^2 + \varepsilon = \| \boldsymbol{a}_I \|_A^2$, but there can be no physical or mathematical justification for forcing conservation of power in the model by rescaling; ε just has to be kept within reasonable bounds. In the particular construction used for the Planck horns the convergence of the model, and hence the reliability of the predictions, required particular attention.

Details will be given in the following chapters, but to summarise: at the junction between two waveguide sections the electromagnetic fields to the left and right of the junction are described by a pair of vectors in a pair of (usually isomorphic) Hilbert spaces, $\mathcal{H}_L$ and $\mathcal{H}_R$, both separable with complete, denumerable bases. Each of these two Hilbert spaces is naturally decomposed into the (orthogonal) direct sum (see A.1.5, page 228) of two isomorphic spaces – the space containing the electric component of the fields and the space containing the magnetic component: $\mathcal{H} = \mathscr{E} \oplus \mathscr{M}$. For a single-mode system that is the end of the decomposition, but for a multi-mode system each of these spaces is itself a direct sum of Hilbert spaces. From a practical point of view this decomposition is finite giving

$$\mathcal{H} = (\mathcal{E}_1 \oplus \cdots \oplus \mathcal{E}_n) \oplus (\mathcal{M}_1 \oplus \cdots \oplus \mathcal{M}_n) = (\mathcal{E}_1 \oplus \mathcal{M}_1) \oplus \cdots \oplus (\mathcal{E}_n \oplus \mathcal{M}_n) \tag{2.1.1}$$

for some $n \in \mathbb{N}$.

The reflection at the junction is determined by an operator determined by the vector of waveguide impedances of the basis elements, while the scattering across the junction is described by an operator mapping between the Hilbert spaces on either side of the junction, and these combine to form a matrix operator that is an endomorphism, S, of the direct sum of the two spaces $\mathcal{H}_L \oplus \mathcal{H}_R$, each of which has the structure given in equation (2.1.1). Because of this structure the operator S decomposes: for each pair $\mathcal{E}_i \oplus \mathcal{M}_i$ on the left and $\mathcal{E}_j \oplus \mathcal{M}_j$ on the right, there is a component S_{ij} of S mapping $(\mathcal{E}_i \oplus \mathcal{M}_i) \oplus (\mathcal{E}_j \oplus \mathcal{M}_j)$ onto itself. For perfectly aligned waveguide sections meeting at a junction this will always be a zero operator unless $i = j$, in which case we can write $S_i = S_{ii}$ and S as the direct sum of all S_i. Whether the sections be aligned or not, because of this structure of the Hilbert spaces, S decomposes into matrix blocks. This structure will be described in detail in later chapters.

The entries in the operator matrix S are the scattering amplitudes, determined by enforcing conservation of total power across the junction and given by integrals over the waveguide cross-section of the type

$$\sum_{\nu\in\mathbb{N}} \left(\bar{A}_\nu - \bar{B}_\nu\right) \int_S \left(\boldsymbol{e}^L_\mu \times \bar{\boldsymbol{h}}^L_\nu\right) \cdot \mathrm{d}S = \sum_{\kappa\in\mathbb{N}} \left(\bar{D}_\kappa - \bar{C}_\kappa\right) \int_S \left(\boldsymbol{e}^L_\mu \times \bar{\boldsymbol{h}}^R_\kappa\right) \cdot \mathrm{d}S \tag{2.1.2-A}$$

$$\sum_{\mu\in\mathbb{N}} \left(A_\mu + B_\mu\right) \int_S \left(\boldsymbol{e}^L_\mu \times \bar{\boldsymbol{h}}^R_\kappa\right) \cdot \mathrm{d}S = \sum_{\nu\in\mathbb{N}} \left(D_\nu + C_\nu\right) \int_S \left(\boldsymbol{e}^R_\nu \times \bar{\boldsymbol{h}}^R_\kappa\right) \cdot \mathrm{d}S \tag{2.1.2-B}$$

for all $\mu \in \mathbb{N}$ and $\kappa \in \mathbb{N}$ respectively with appropriate complex coefficients A, B, C and D, the bar over a symbol denoting complex conjugation, and $\boldsymbol{e}$ and $\boldsymbol{h}$ are elements of a basis for the electric and magnetic fields, respectively. Details are given in equations (3.2.5) and (3.2.6) of section 3.2. It is this matching of power across junctions to determine the scattering coefficients that gives rise to the name "mode matching".

Propagation between junctions is described by a diagonal phase slippage operator matrix, V, that is simply an endomorphism on the space of fields in the waveguide section through which they are propagating. It, too, is a scattering operator, but of a particularly simple form. If S_1 and S_2 are any two scattering operators for any two adjoining sections or sub-units of the guide, with $S_1 \in \mathrm{End}\,(\mathcal{H}_A \oplus \mathcal{H}_B)$ and $S_2 \in \mathrm{End}\,(\mathcal{H}_B \oplus \mathcal{H}_C)$, then the total scattering between the two ends of the concatenated sections is described by a "scattering product" between the operators, $S_2 \odot S_1 \in \mathrm{End}\,(\mathcal{H}_A \oplus \mathcal{H}_C)$, the details of which are given in equation (3.2.13) on page 64. Since the Hilbert spaces naturally embed in the operator spaces, the field at the output can be read from the columns of the "system operator" that is the product of all scattering and phase slippage of the component operators. Thus, labelling the space of fields in the input guide by $\mathcal{H}_1$, the space of fields in the second section $\mathcal{H}_2$, etc. and the scattering operator $S_{j,j+1} : \mathcal{H}_j \oplus \mathcal{H}_{j+1} \to \mathcal{H}_j \oplus \mathcal{H}_{j+1}$, the entire electromagnetic field scattering process for the system of N sections is a matrix operator

$$S_{N1} = V_N \odot S_{N,N-1} \odot V_{N-1} \odot \cdots \odot V_2 \odot S_{21} \odot V_1. \tag{2.1.3}$$

Thus, the operator for the total system describes the scattering between the direct sum of the spaces of all possible input and output fields and its detail (the numerical values of the operator matrix entries) are determined by the entire scattering and phase slippage process through the whole structure. The development of this view of the process and the efficient and accurate numerical computation of the component operators and their products is the main theme of this work.

Consider the system illustrated in figure 2.1 overleaf, the analysis of which is presented in section 2.2. The scattering product is associative; consequently, the system can be described by a product of operators, each of which describes a section of the system that it is computationally or conceptually convenient to treat as a unit, giving a total system operator of the following type:

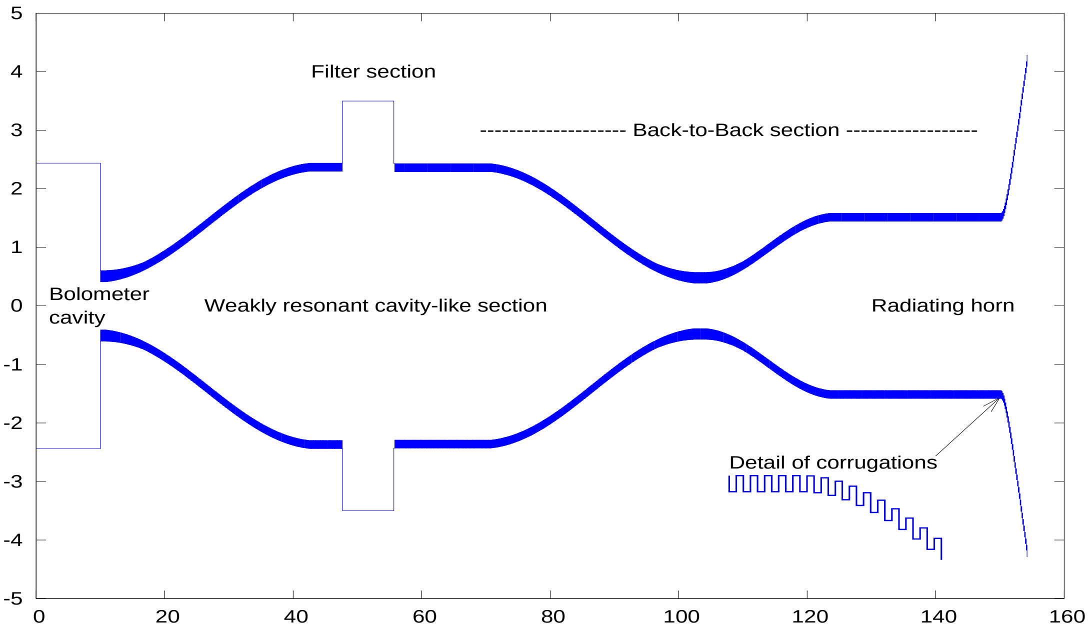


Figure 2.1: Profile of the Planck 545 GHz horn. All of the Planck HFI horns have essentially this same profile, whether they be single or multi-mode. The presence of the cavity-like section gives rise to resonances that considerably complicate the modelling, and greatly increases the computation time required. The thick blue lines are the corrugated sections, not resolved in the plot of the horn assembly, but shown in the detail of the transition to the aperture flare. The scales are millimetres.

$$S = \underbrace{T_5 \odot W_4 \odot T_4}_{\text{radiating horn}} \odot \underbrace{\overbrace{W_3 \odot T_3 \odot W_2^R}^{\text{back horn}} \odot F \odot \overbrace{W_2 \odot T_2 \odot W_1 \odot T_1}^{\text{bolometer feed horn}}}_{\text{feed section forming weakly resonating cavity}} \tag{2.1.4}$$

Here each of the operators S, W, T and F are operators of the type in equation (2.1.3) for a section of the horn assembly: W is a section of smooth walled or parallel sided corrugated waveguide, T any tapered section or section of varying radius and F a filter, free space or lens section. W_2^R means the component is identical to a part of W_2, but reversed. The operator S is to be read with T_1 corresponding to the short input (profiled) taper section at the left hand side of figure 2.1 where the horn joins the bolometer cavity, and so on through to the radiating aperture on the output side of the flare, T_5, at the right hand side of the figure. This is quite a common construction and is similar to the construction of the Planck horns [39]. The model space for the feed to the horn assembly from the bolometer cavity is the domain of of the operator T_1, and the output of fields that radiate to the sky is the codomain of T_5; in between, the domain of each operator is the codomain of the operator to its right. Advantage can be taken of the symmetries in the scattering operators: within the waveguide sections the symmetry across pairs of junctions is exploited to reduce the determination of W to the calculation of powers of a single symmetric waveguide-junction-waveguide-junction-waveguide operator for which the scattering matrix

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} : \mathcal{H}_1 \oplus \mathcal{H}_2 \to \mathcal{H}_1 \oplus \mathcal{H}_2$$

has sub-matrix components $S_{21} = S_{12}$ and $S_{11} = S_{22}$. The work done in calculating the matrix and in forming the scattering product powers of such matrices is thereby halved; this is discussed in detail in section 3.6. The codomain of T_2 is the same as the domain of T_3, and therefore W_2 and W_2^R are powers of the same symmetric operator. Exploitation of such symmetries can greatly reduce the overall computational burden in the analysis of the system. When simulating manufacturing tolerances as outlined in section 2.6 these symmetries cannot be exploited.

The presence of a cavity-like section results in trapped power and significant contributions to the field from evanescent modes. This significantly increases the size of the operator matrices required to give an adequate model of the section $W_3 \odot T_2 \odot W_2^R \odot F \odot W_2 \odot T_2 \odot W_1 \odot T_1$, or any cavity-like section. It is not necessary to maintain the same model dimension throughout the assembly but, as discussed in section 6.4 on the perturbing of models, it can be advantageous to do so.

2.2 Numerical implementation of mode matching

The details of the numerical scheme adopted for the implementation of mode matching will be described in detail in chapter 5, and the efficient representation of fields in the radiating aperture for Physical Optics propagation through the telescope in chapter 6; here only a brief introduction will be given so that the statements made about the relative ease or difficulty of modelling various horn configurations, made later in this chapter, will be comprehensible.

The horn is a described by the scattering product of the operators for all junctions and waveguide sections. Because of the boundary conditions in the guide the function spaces in which the fields are expanded have denumerable bases, and therefore the operators are representable as operator matrices. These, being infinite square matrices that describe linear mappings between two function spaces, must be approximated for numerical work by finite dimensional complex matrices that describe the linear mappings between two finite dimensional complex vector spaces. With the approximation made, the matrices may be handled using the conventional methods of numerical linear algebra. In general these matrices will be rectangular if the field structure in the waveguide to the right and left of a junction are adequately described by complex vector spaces of different dimensions, but for the sake of simplicity it will hereafter be assumed that the spaces are of the same dimension and the matrices therefore square.

There are two inextricable issues that are critical to the reduction of the infinite operator matrices to finite complex matrices: the adequate accounting for power scattered across a junction and the size of matrices that leaves a numerical problem that is solvable in a reasonable amount of time with the computing resources available. It will be seen below that for a relatively simple horn such as the ClOVER horns the size of the operators that give rise to a stable (fully converged) numerical model of the horn is modest and allows for reliable and accurate modelling. The same applies to the back-to-back section of all of the Planck HFI horns (see figure 2.1 on page 12) that had previously been modelled by Murphy, Gleeson, Colgan and co-workers at the band centre and at a couple of frequencies either side of centre, [25] [14]. However, the Planck horns are not used with just the back-to-back section; the back-to-back section is coupled to the bolometer cavity by a second horn. This arrangement, in which the bolometer horn faces the back horn, forms a resonant ‘cavity’ that traps power and gives rise to changes in the beam pattern of the horn that are not predicted by modelling any subsection of the horn assembly. Trapped power is accounted for by evanescent modes (section 3.1.1) and, in the models of the full Planck horn assemblies, leads to very large numerical models.

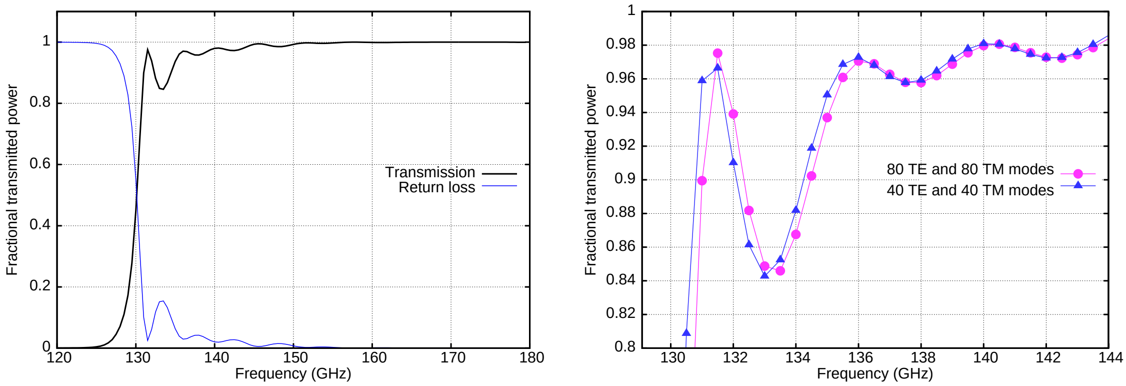
With the code developed for this work and described in chapter 5, the models of any of the complete *single-mode* Planck HFI horns can be run on a conventional PC to give a fully converged model across the operating band within a few minutes. For the *multi-mode* horns the models, supported by measurements made at Cardiff, suggest that the horns exhibit strong resonances minutely influenced by the manufacturing tolerances. This behaviour means that the models can only give a reliable prediction of performance over a band of frequencies, and the power transmission and beam profile at a single frequency should not be taken as an accurate prediction of what would be measured from a real horn made to the design that has been modelled. This is not a practical handicap because the horns are only used broad band, but it does lead to problems with size of the model and consequent run times.

The key to the tractability of mode matching as a numerical modelling method is an efficient and accurate implementation of the scattering product given in equation (3.2.13) on page 64. Chapter 5 gives the general form and method used in this work, particular attention being given to the case of aligned circular cross-section waveguides applicable to Planck. A small gain in overall numerical efficiency, but great gains in stability and accuracy, is obtained by the careful coding of the scattering amplitude equations; that is covered for the circular case in chapters 4 and 5. In section 4.4 the equations for rectangular waveguides are given in a form equivalent to, but different from, those in the literature, along with comments on the coding.

2.3 The C*l*OVER and Planck single-mode horns

The comparison of C*l*OVER ultra-Gaussian and Planck single-mode horns is interesting, both at the field level and at operator matrix level, and is discussed qualitatively here. These two high performance single mode horns are discussed here to introduce some of the basic concepts used in the evaluation of horn performance. They will be returned to later when, in chapter 5, the numerical implementation of mode-matching is discussed. There it will be shown that the shape of the beam is seen in the purity of the mode structure as represented in the sparsity of the scattering operator matrix that describes the scattering from the bolometer cavity to the radiating aperture (see figure 5.1 on page 141); here the evolution of the phase front and Gaussian coupling efficiency as the field propagates towards the aperture is presented – results obtained from the numerical models run using the code described in 5.

The ClOVER horns are designed to give highly Gaussian aperture field distributions in the co-polar component of the field [2]. The desired field distribution is attained by generating $HE11$ and $HE12$ hybrid modes in a cosine-squared profile horn and bringing the two modes into phase at the horn aperture in a long parallel section of corrugated waveguide. This was a development due to Graham Smith of St. Andrew's University, Scotland, that post dates the design and development of the Planck single-mode horns. It utilises the mode dependence of the waveguide impedance to achieve the phasing. When the Planck horns were designed a very similar Gaussian field distribution was achieved in a more complex horn profile. Although the aperture fields of the two horns are similar in that they both have high coupling efficiencies to a Gaussian across a 30% band, the evolution of the propagating field as the beam propagates towards the aperture is noticeably different.



(a) 40 TE and 40 TM mode model in 0.5 GHz steps from 120 to 180 GHz. Computation rate: 665 junction scatterings per second.

(b) Comparison of the 40 TE and 40 TM mode model with an 80 TE and 80 TM mode model over the cut-on region.

Figure 2.2: C*l*OVER 150 GHz horn transmission *vs.* frequency. Modelled in SKITTER with 40 TE and 40 TM modes and with 80 TE and 80 TM modes. The greater accuracy is seen only in the fine detail the cut-on region, not in the overall performance. The increase in computational time – a factor of four – is not an issue given the speed of the code, but the beam shape predicted by the two models is not discernibly changed. (Compare with [2].)

A radiating horn is fed from a bolometer or some other power detector such as a waveguide probe (the details of which will not be discussed in this work) via a waveguide structure of some kind. The field that radiates from the aperture of the horn to illuminate the telescope or the sky is, in the aperture itself, a coherent superposition or a family of coherent superpositions of waveguide ‘modes’ or basis function for the space of solutions the Sturm-Liouville problem in the aperture. For a single-mode horn there is only one such superposition which, generally speaking, is desired to have a power distribution as close to a Gaussian distribution as possible with a phase front as flat as possible at the waist of the horn, the waist being the plane in which the phase front is flattest.

Anything in the design or manufacture of the horn assembly that alters the structure (the relative phase or amplitude) of the modes as they feed into the radiating horn will alter the aperture field and therefore the radiation pattern of the horn in a completely deterministic way. From the mathematical perspective the operator that describes the radiating horn is acting on a different vector or set of vectors in its domain and, since the operator is linear, the output field is different. This effect can be used constructively such as to tune the input field to the ClOVER horns, [2], or it can have a deleterious effect, and this becomes a recurring theme throughout the discussion of the Planck horns.

The physically meaningful measure of the capacity of the horn to absorb power from the electromagnetic field incident upon the aperture is the power coupling efficiency. The power coupling efficiency between two beams is described by the square of the L^2-inner product of the two field functions, each field having been normalised to unity over the domain: given two complex functions ψ and φ on a domain A, the coupling efficiency is given by

$$\text{Power coupling efficiency:}\quad \varepsilon_P = \left\langle \hat{\psi} | \hat{\varphi} \right\rangle^2 = \frac{\left| \int_A \psi\varphi^* \, \mathrm{d}A \right|^2}{\int_A |\psi|^2 \, \mathrm{d}A \int_A |\varphi|^2 \, \mathrm{d}A} \tag{2.3.1}$$

in which $\hat{\varphi} = \varphi \cdot \left[\int_A \varphi\varphi^* \, \mathrm{d}A\right]^{-\frac{1}{2}}$ is the function normalised to have unit total power in A. The power coupling efficiency of the function to itself takes the value 1 and approaches 0 for fields that exchange essentially no power; thus $0 \leq \varepsilon_P \leq 1$. Since the integrals are linear functionals, so is ε_P, and if the waveguide spectrum (or whatever spectrum is appropriate) for the two fields are known, then ε_P is known: in the case of a waveguide the boundary renders the spectrum discrete and ε_P is simply the sum of the products of the individual mode amplitudes.

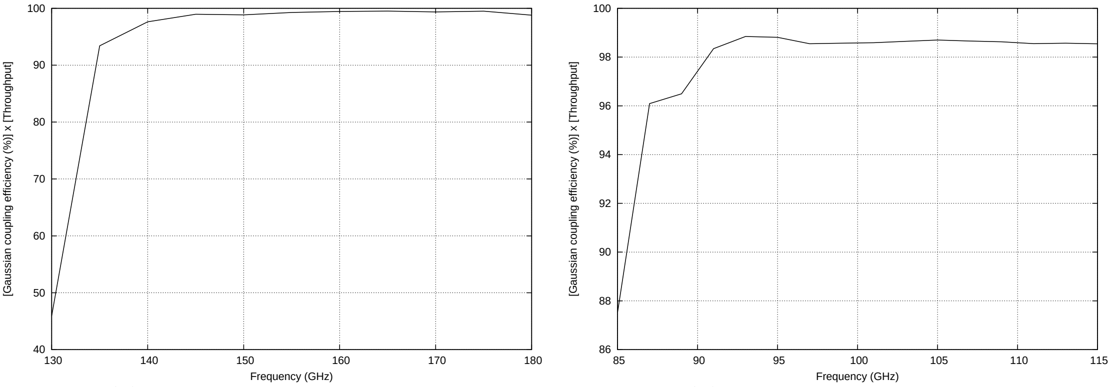
The profile of a Gaussian beam with propagation axis z, at a distance R from its waist of radius w_0, is given by the equation (see [26])

$$\varphi(r, z, w_0) = \frac{\sqrt{2}}{\sqrt{\pi} w(z)} e^{-r^2/w(z)^2} e^{-j\left(2\pi z/\lambda + \pi r^2/\lambda R(z) - \varphi_0(z)\right)}, \tag{2.3.2}$$

in which the phase radius of curvature, $R(z, w_0)$, and phase shift, $\varphi_0(z)$ are given by

$$R_z = z\left(1 + \left(\frac{w_0^2\pi}{z\lambda}\right)^2\right), \text{ and } \quad \varphi_0(z, w_0) = \arctan\left(\frac{z\lambda}{w_0^2\pi}\right)$$

respectively. When fitting a Gaussian beam profile to the horn the Gaussian parameters are varied (the waist radius w_0, the waist position (x_0, y_0, z_0) measured *relative* to the beam data plane, and the angle, (α, β), between the Gaussian beam propagation direction and the aperture plane) to maximise ε_P. (Note that in equation (2.3.2) the propagation distance, z, to any point (x, y, z) in the plane of the data φ is the distance from the waist centre (x_0, y_0, z_0) to the plane through (x, y, z) at angle (α, β). Thus, $r = \sqrt{x^2 + y^2}$ is the length of the perpendicular dropped to the Gaussian beam axis.)

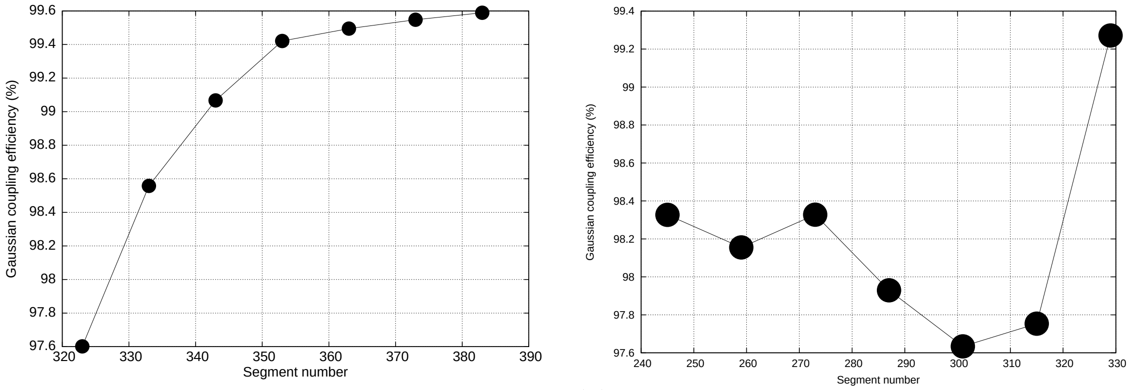


(a) ClOVER 150 GHz horn.

(b) Planck 100 GHz horn.

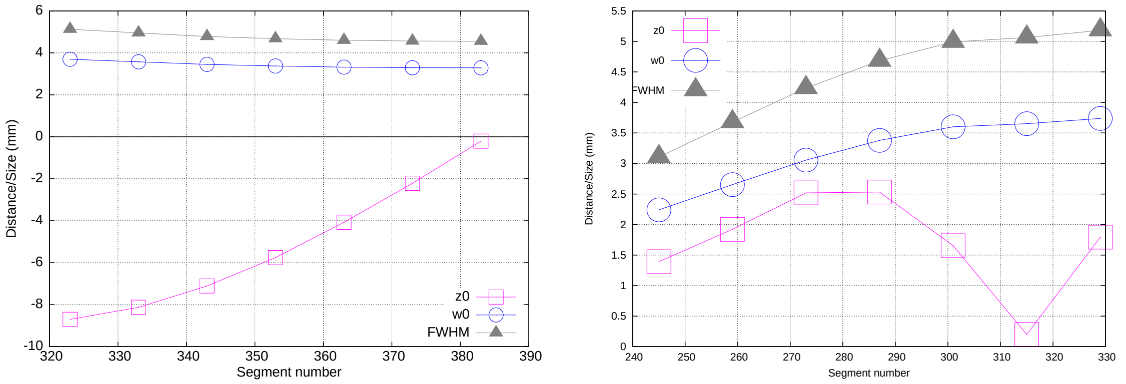
Figure 2.3: Gaussian coupling efficiency (%) as a function of frequency, weighted by the frequency dependent transmission of the horn, of (a) the ClOVER 150 GHz horn and (b) the Planck 100 GHz horn as a function of frequency.

The far field semi-divergence angle is given by $\theta_0 = \arctan(\lambda/w_0\pi) \sim \lambda/w_0\pi$, and the confocal distance by $z_c = w_0^2\pi/\lambda$. These are all frequency dependent parameters and the viable bandwidth of the horn, coupling to a Gaussian source, is determined by the variation in $\varepsilon_P(\lambda)$ and $w_0(\lambda)$.



(a) Gaussian coupling efficiency of the evolving field for the ClOVER 150 GHz horn.

(b) Gaussian coupling efficiency of the evolving field for the Planck 100 GHz horn.

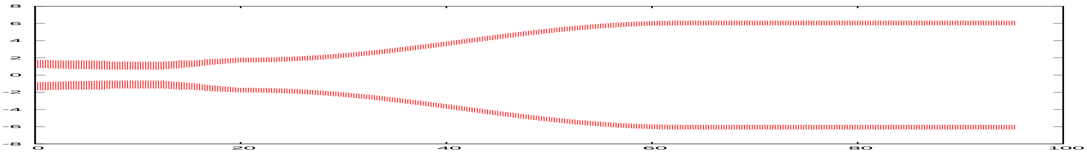


(c) Evolution of the ClOVER 150 GHz horn field parameters, w_0, z_0 and the FWHM.

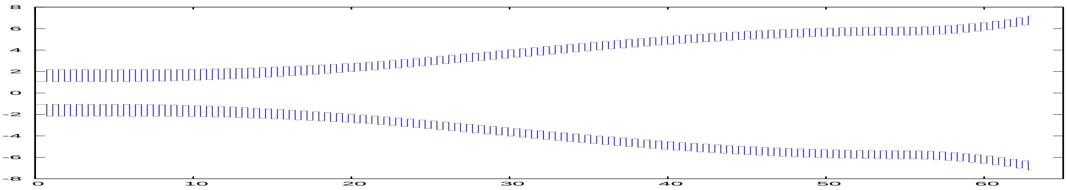
(d) Evolution of the Planck 100 GHz horn field parameters, w_0, z_0 and the FWHM.

Figure 2.4: Gaussian coupling efficiency for the evolving field determined by fitting a complex fundamental Gaussian mode to the co-polar component of the electric field: (a) ClOVER 150 GHz at segment 323 and in steps of 10 segments out to the aperture at junction 383, and (b) Planck 100 GHz at segment 245 to the aperture at segment 339 in steps of 14 segments. Evolution of the horn characteristic parameters: $z0$ – the distance from the horn aperture to the beam waist; $w0$ – the beam waist radius determined by the best fit Gaussian; and $FWHM$ – the full width of the beam measured to the half power level at the waist: (c) ClOVER 150 GHz horn, and (d) Planck 100 GHz horn. Note: whereas the ClOVER horn is parallel sided waveguide over the section modelled here, the Planck horn is gently tapered, but the taper does not account for the increasing value of w_0 and FWHM as the aperture is approached.

For an essentially Gaussian beam profile the quality of the beam is assessed by computing the Gaussian coupling efficiency and thus the waist radius, and from that the full width half maximum (FWHM) and the far field semi-divergence angle. Fig-



(a) Profile of the ClOVER 150 GHz horn.

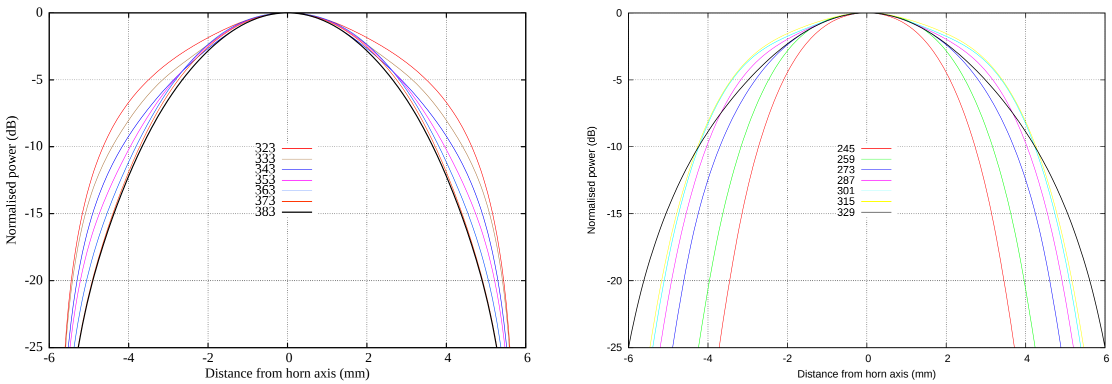


(b) Profile of the radiating section of the Planck 100 GHz horn.

Figure 2.5: ClOVER 150 GHz and Planck 100 GHz horn profiles. The lengths of x axes in figures (a) and (b) are 100 mm and 70 mm respectively; in both plots the y-axis range is ±8 mm.

ure 2.4 above shows how these parameters evolve as the field propagates in the final section of the ClOVER 150 GHz and Planck 100 GHz horns illustrated in figure 2.5.

For the coupling to a telescope or other optical system the $\psi(\lambda)$ in equation 2.3.1 is the complex electric field distribution in the horn aperture plane due to the instrument, while $\varphi(\lambda)$ is the horn aperture field with the aperture as the integration domain for all three integrals. (Strictly, it is the horn radiation field at the aperture that is used, thus accounting for reflection at the aperture, but in a horn of aperture diameter 5λ or larger the impedance step to free space is negligible). There is no computational or mathematical difficulty in fitting a spherical or conic phase front to the phase distribution over the aperture (see [49] for the method). Phase fitting gives a true measure of the phase curvature, but it does not account for the power distribution: it gives equal weights to all points in the field and is therefore not a useful measure of the viable band width of the horn. Modelling the horn aperture field at the centre of the band and at regular wavelength intervals either side of the centre, then calculating the optimal $\varepsilon_P(\lambda)$ for each model, gives a variation in Gaussian coupling for the horn design. These are plotted for the ClOVER 150 GHz horn and the Planck 100 GHz horn in figure 2.3 above where the ε_P for each modelled frequency is weighted by the throughput of the horn predicted by the mode matching model. If a measured or modelled bolometer cavity coupling was available, along



(a) Evolution of the C*l*OVER 150 GHz beam shape from the start of the parallel waveguide section to the aperture at segment 383.

(b) Evolution of the Planck 100 GHz beam shape from section 245 to the aperture at segment 329.

Figure 2.6: Evolution of the beam shape in the C*l*OVER 150 GHz horn and the Planck 100 GHz horn at the band centre. Unlike the C*l*OVER horn, the Plank horn is tapered, which accounts for the significant change in beam width. Both plots are E-plane cuts.

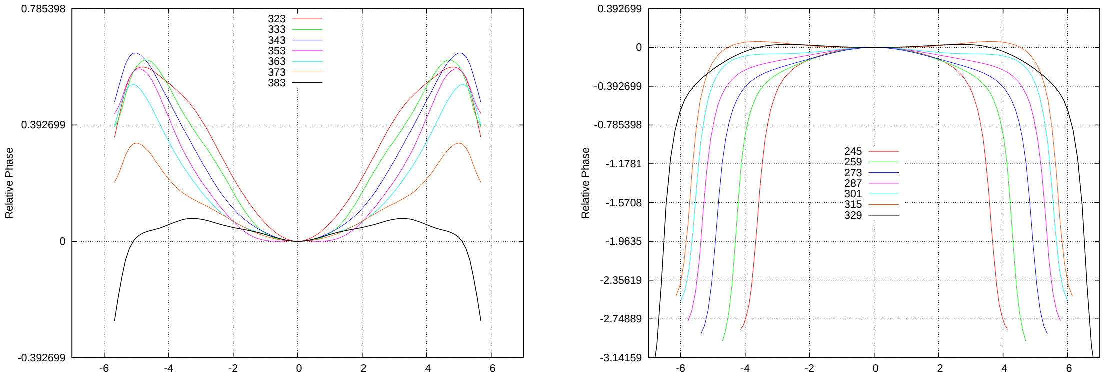
with filter transmissions and other ‘efficiencies’, these, too, would be included in the bandwidth assessment. These should be compared with the phase distribution cross sections in figure 2.8 on page 24.

For a well designed single-mode horn the fundamental mode Gaussian beam formalism gives a very useful measure of the form of the beam on the sky because the power distribution at the beam waist is essentially Gaussian and the Fourier transform of the Gaussian beam waist field is again Gaussian. An optical system is, in the first approximation, a linear system that acts as a position to angle transforming device. Thus, a telescope transforms the Gaussian distribution of the beam waist at its focus to a Gaussian angular distribution on the sky. (Ignoring the diffraction effects of finite apertures, aberration, and so on, the telescope performs an optical Fourier transform of the field distribution at its focus with the angle to length scale being determined by the ‘plate scale’: 206265/(Focal length) arc seconds per mm).

For multi-mode systems the power coupling to a fundamental Gaussian is not a good measure, though it does give a crude indication of the far-field beam pattern if the fundamental mode dominates. The field in the aperture has a discrete Fourier-Bessel series expansion because its domain is a closed disc, or other discrete Fourier series expansion for a rectangle or other closed subset of the plane; the free-space field must have a continuous radial Fourier spectrum while retaining the discrete azimuthal spectrum, continuous because the radial coordinate is $[0, \infty)$. In the case of the disc, appropriate to the ClOVER and all of the Planck horns, the basis functions for the Fourier expansion in the aperture are the discrete set indexed by $(n, m) \in \mathbb{Z}_{\geq 0} \times \mathbb{N}$

$$\Psi_{nm} = \frac{1}{\sqrt{N_{nm}}} J_n(k_{nm} r)\Phi_n(\varphi),$$

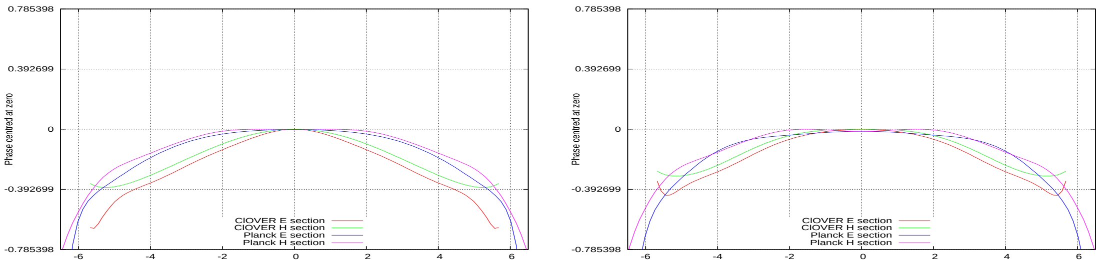
where the N_{nm} are normalisation factors appropriate to the field type and the Φ_n are sines or cosines fitting n times onto the unit circle. Details will be given in chapter 4. The Fourier transform of each mode is a continuous spectrum whether it is calculated in polar or Cartesian coordinates. If the free space spectrum of each mode is calculated, then the far field pattern can be reconstructed for each independent aperture field from the sum of all spectra with their phases, and the total power pattern is then the incoherent sum of all the independent field contributions. The matter of Fourier transforms is relegated to section A.4 in the appendix, and in the final three sections of the appendix the Fourier transform is related to the Gauss-Hermite beam modes.



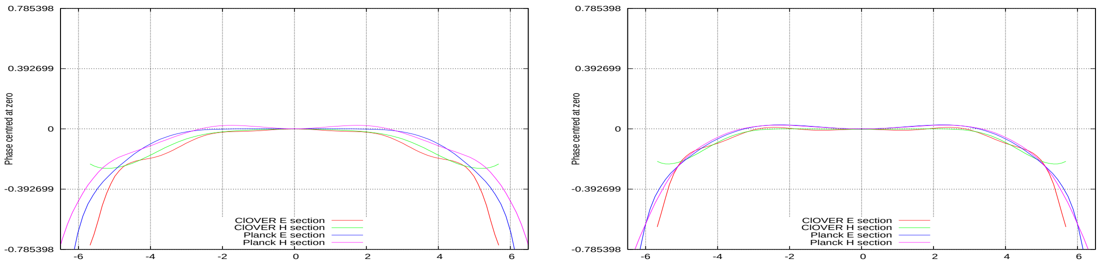
(a) Evolution of the ClOVER 150 GHz phase from the start of the parallel waveguide section to the aperture at segment 383. E-plane cut; phase scale $[-\pi/8, \pi/4)$.

(b) Evolution of the Planck 100 GHz phase in the radiating part of the horn from section 245 to the aperture at segment 329. E-plane cut; phase scale $[-\pi, \pi/8)$.

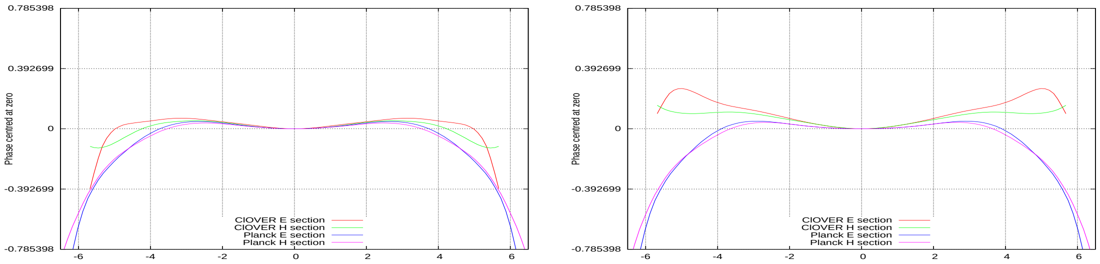
Figure 2.7: Evolution of the phase front in the ClOVER 150 GHz horn and the Planck 100 GHz back-back horn, both at the band centre. Unlike the ClOVER horn, the Plank horn is tapered, which accounts for the significant change in beam width.



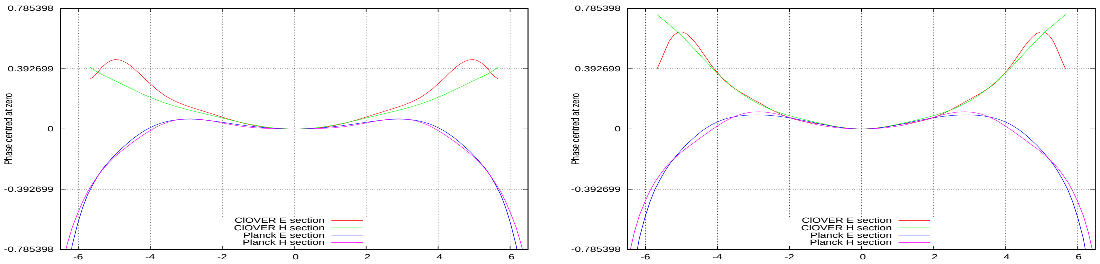
(a) Frequency: 87% of band centre. (b) Frequency: 90% of band centre.



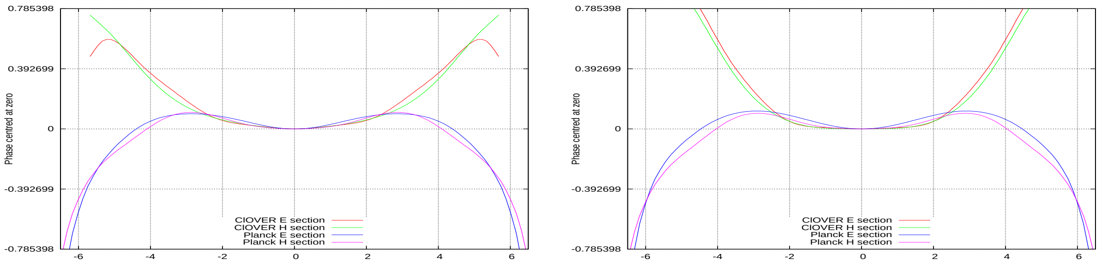
(c) Frequency: 93% of band centre. (d) Frequency: 97% of band centre.



(e) Frequency: band centre. (f) Frequency: 103% of band centre.



(g) Frequency: 107% of band centre. (h) Frequency: 110% of band centre.



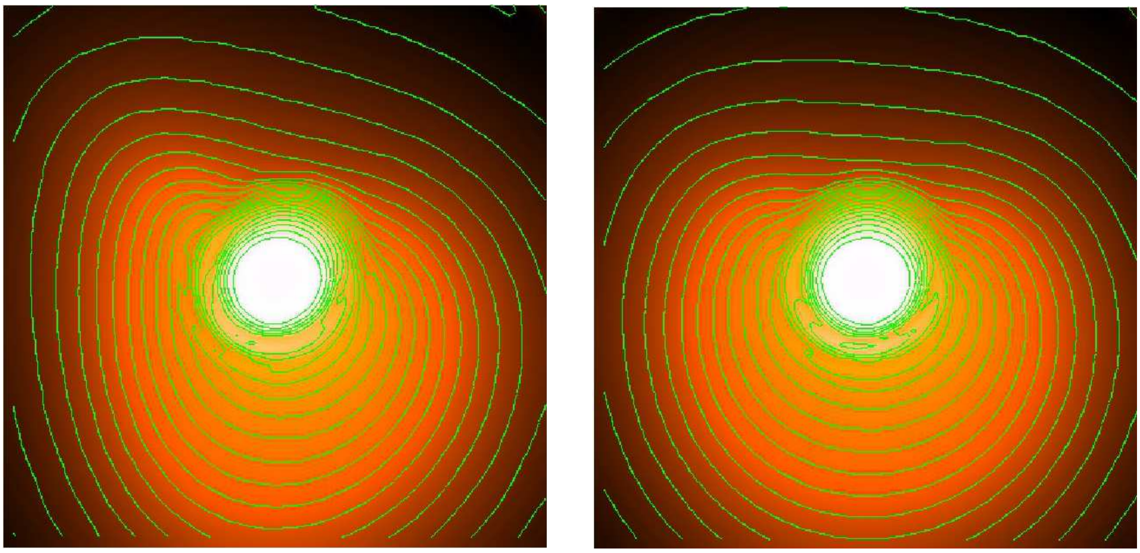
(i) Frequency: 113% of band centre. (j) Frequency: 116% of band centre.

Figure 2.8: Comparison of the changing phase fronts in E and H planes in the apertures of the C*l*OVER 150 GHz and Planck 100 GHz single mode horns. Red and green curves: C*l*OVER ; blue and violet curves: Planck. The phase in each plot has been offset to zero on the horn axis. At band centre, plot (e), it is seen that the phase front of the C*l*OVER horn is flatter than for the Planck horn, but that the Planck horn phase front shape is the less frequency dependent. Phase range in all plots $[-\pi/4 : \pi/4]$.

2.4 The Planck multi-mode horns

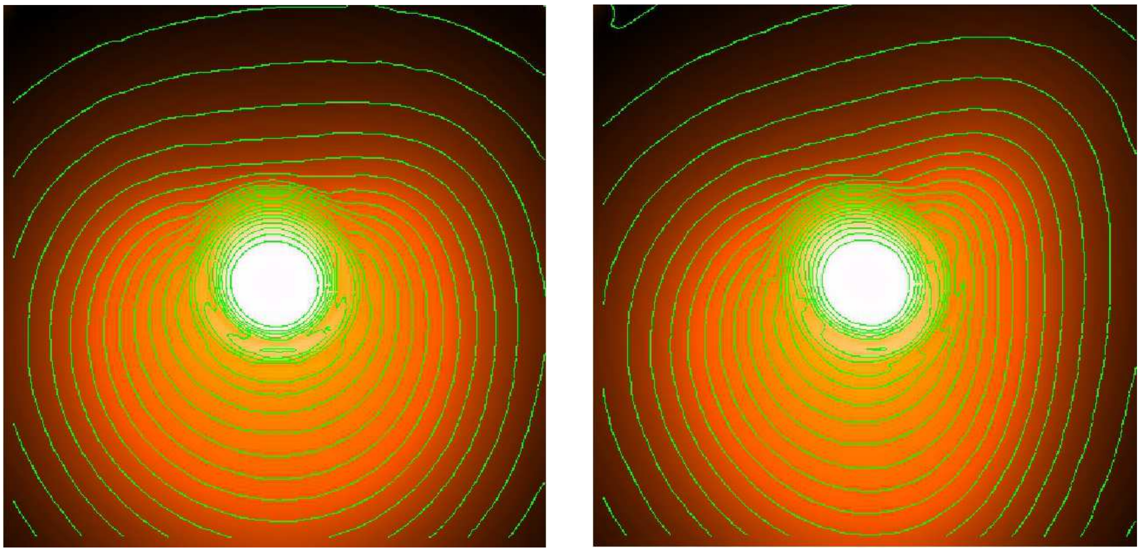
There are eight multi-mode pixels on the Planck telescope, four at 545 GHz and four at 857 GHz. The modelling of the Planck multi-mode pixel assemblies was undertaken on behalf of the HFI core team in the year prior to the launch of Planck and Herschel. The intention was to arrive at reliable predictions for the broad-band beams that would be observed on the sky by each pixel when observing. Prior to this only spot frequency models of the beams that would have resulted from an ideal version of the telescope (as designed rather than as built) illuminated by the radiation pattern that would be emitted by a horn comprising only the back-to-back sections of the horns, had been attempted. The propagation of the beam from the horn aperture, through the telescope model to the sky, was to be performed using the Physical Optics modelling package GRASP9 [56]. To undertake the broad band modelling it was necessary to develop mode-matching code that could derive the required aperture fields at many frequencies over the band accurately and in a short time, and to write the aperture field information into files in a format for source fields for the GRASP9 telescope models, and to run the many resulting cases in batch mode. The telescope modelling process will be described in chapter 8; here the discussion is restricted to the beam patterns derived and comparison with the first attempt, by Brendon Crill at CalTech, to derive beam maps for the multi-mode pixels form the preliminary calibrations scans of Jupiter. At the time of writing no definitive beam maps were available with which to compare the models, but what data was made available on beam widths is tabulated in chapter 8 along with the model beams widths. It was found that, because of the resonant nature of the horn assemblies, large numbers of waveguide modes were required to model the scattering.

Once the aperture fields had been derived and expressed in transverse waveguide modes, a computationally efficient presentation of the aperture fields was needed. It would have been computationally inefficient, though mathematically and physically correct, to propagate every individual aperture field derived by mode-matching through the telescope and onto the sky, and there to assemble the individual mode beam patterns into a total beam pattern. What was required was a way to find the subspace of the space of fields that was actually spanned by the aperture fields, to express the aperture fields in a basis for that space, and to propagate the resulting basis vectors to the sky. That subspace is frequency dependent and is peculiar to the modal structure of the field (equivalently the components of the S_{21} scattering operator) at that frequency. These operators are non-hermitian, and to find the minimal subspace the concept of Schmidt vectors was taken from the mathematical field of meromorphic approximation, a readable outline of which can be found in [74]. These were adapted to give a “Schmidt field” representation of the scattering operator. These ideas are described in some detail in chapter 6, and all beam patterns presented in this thesis were obtained form such field representations.



(a) Pixel H-857-1

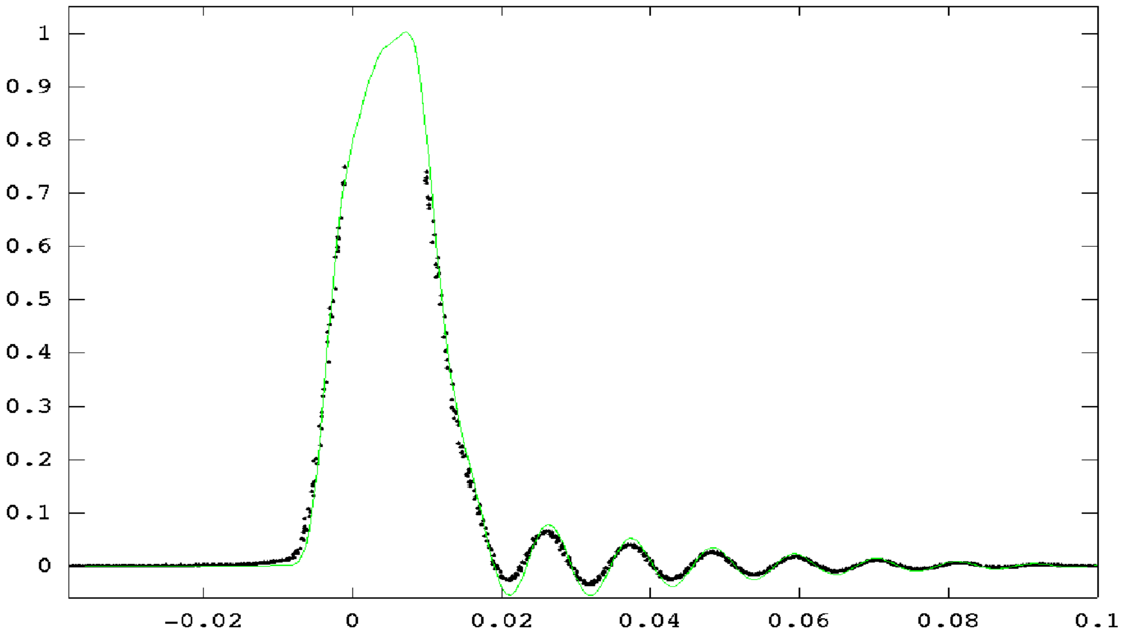
(b) Pixel H-857-2



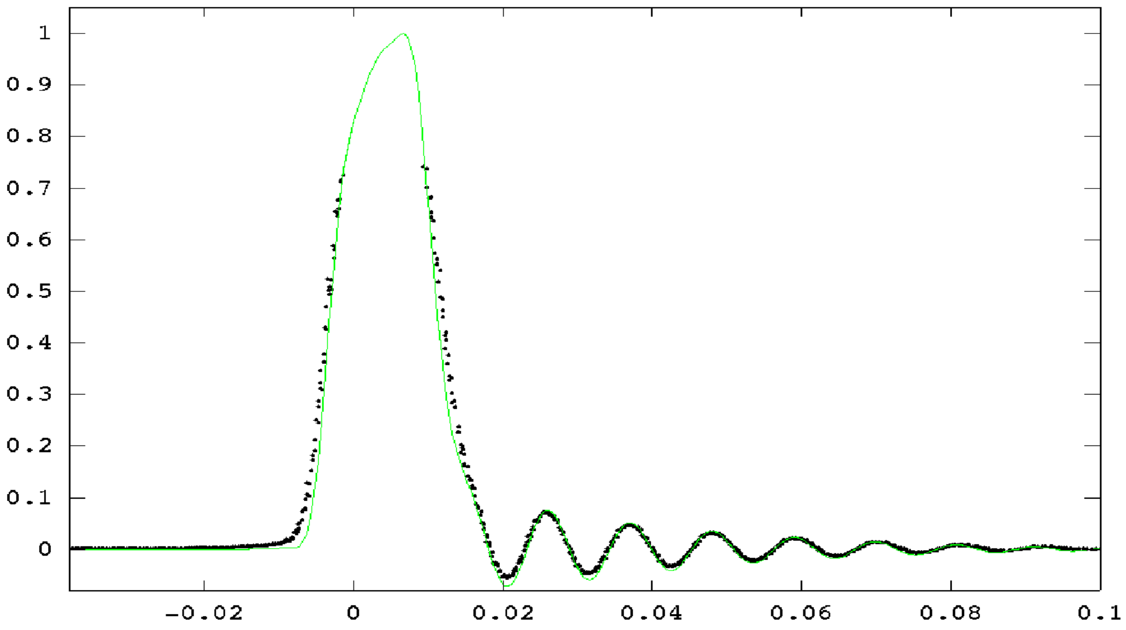
(c) Pixel H-857-3

(d) Pixel H-857-4

Figure 2.9: Decibel plots of the broad-band modelled beam power patterns for the Planck 857 GHz beam on the sky, 730 – 990 GHz in 52 frequency steps; data normalised to a peak power of one. Plot area: $30' \times 30'$; contours -3 dB to -69 dB in -3 dB steps. The patterns exhibit the aberration characteristic the off-axis quasi-Gregorian telescope configuration close to axis: 3^{rd} order coma. This is the 'as built' telescope model that includes the best available pre-launch information on the reflector shapes and the construction of the telescope.

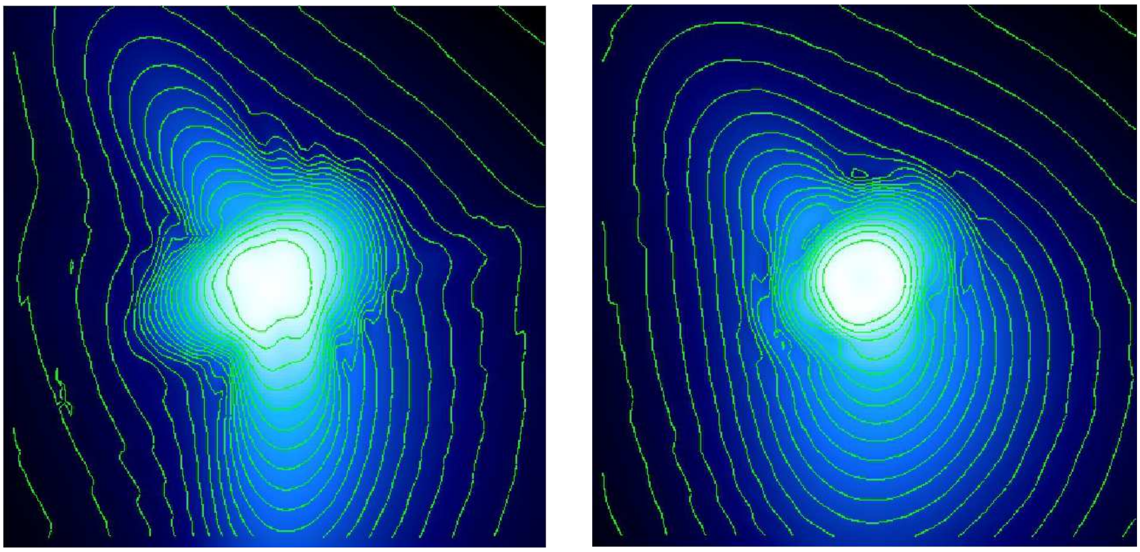


(a) Planck H-857-1 pixel signal on Jupiter.



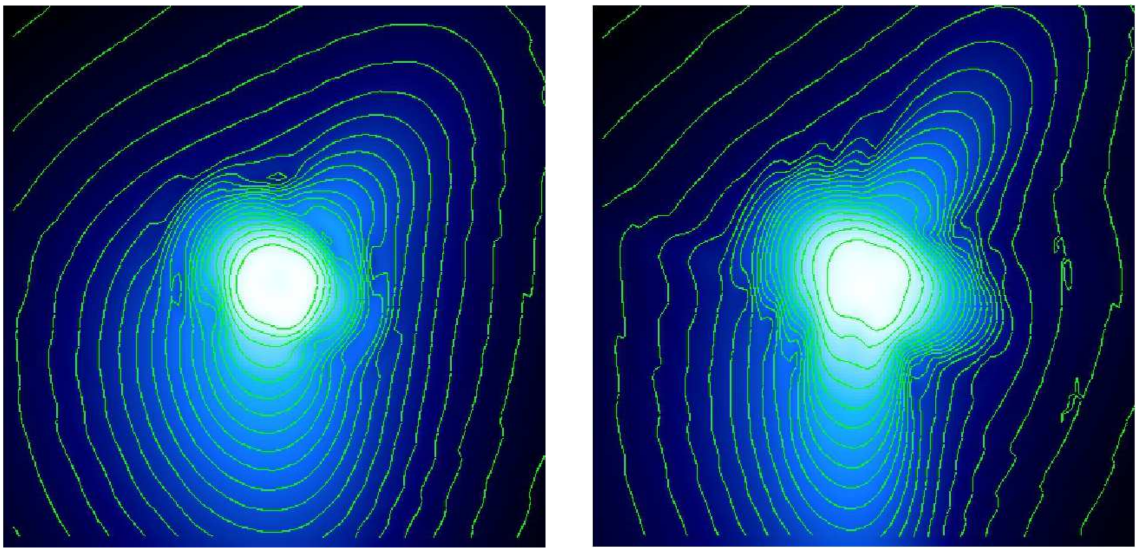
(b) Planck H-857-2 pixel signal on Jupiter.

Figure 2.10: Normalised signal *vs.* time in seconds plots for the Planck 857 GHz pixels H-857-1 and H-857-2. This is early-stage analysis of data taken during the first calibration scan of Jupiter. Green line: broad band model convolved with estimated transfer function, black dots: received signal. The received signal has saturated the pixel above 0.74, so the data is missing. Data processed by Brendan Crill and the Planck HIFI data processing team. Model: full pixel assembly, 52 frequencies over 730 – 990 GHz, pre-launch 'as built' telescope. The dip in the curves below zero energy indicates that the analysis techniques required further development.



(a) Pixel H-545-1

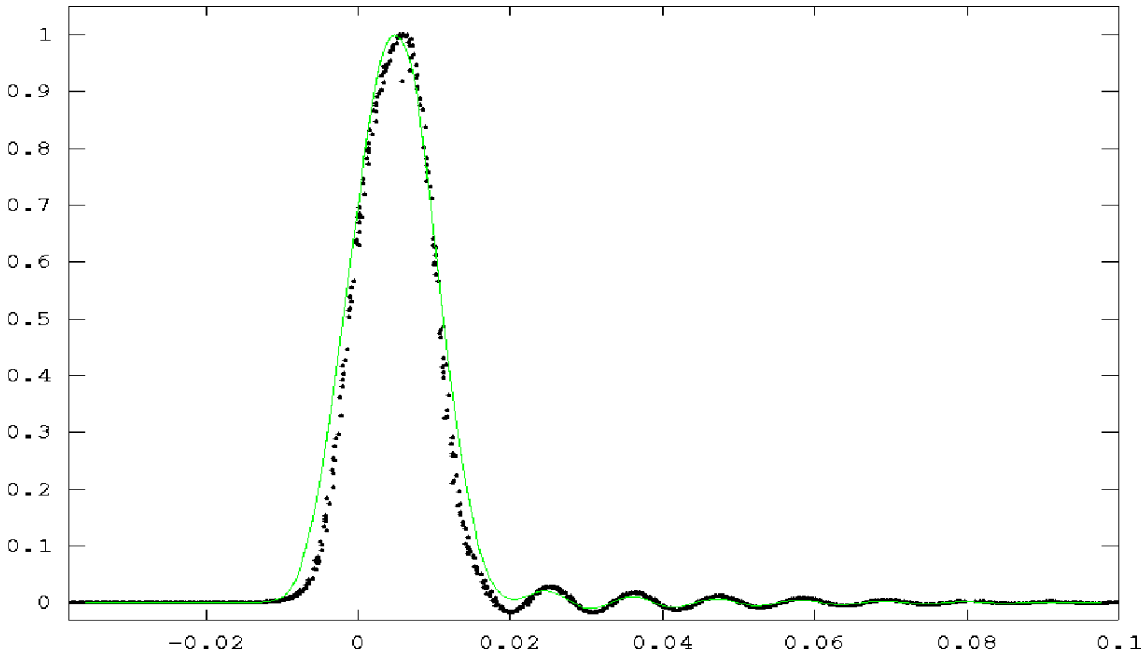
(b) Pixel H-545-2



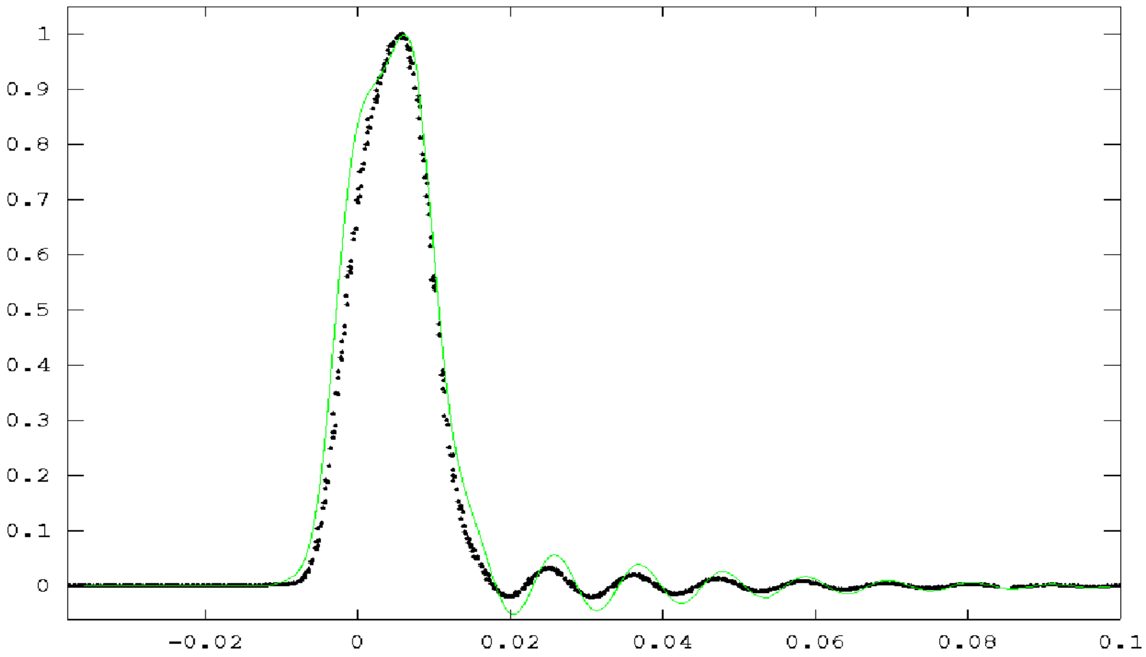
(c) Pixel H-545-3

(d) Pixel H-545-4

Figure 2.11: Decibel plots of the broad-band modelled beam power patterns for the Planck 545 GHz beam on the sky, 460 – 630 GHz in 64 frequency steps; data normalised to a peak power of one. Plot area: $30' \times 30'$; contours: -3 dB to -63 dB in -3 dB steps. There is left-right asymmetry most noticeable between H-545-1 and H-545-4; it arises form slight imperfections in the reflectors and alignment of the telescope, but is not optically significant. This asymmetry is more marked than in the 857 GHz beams because the pixels are further from the telescope axis. The distance of the pixels from the axis is also responsible for the strongly evident distortion with astigmatism dominating over coma, particularly in the outermost pixels, 1 and 4.



(a) Planck H-545-1 pixel signal on Jupiter.



(b) Planck H-545-2 pixel signal on Jupiter.

Figure 2.12: Normalised signal *vs.* time in seconds plots for the Planck 545 GHz pixels H-545-1 and H-545-2. This is early-stage analysis of data taken during the first calibration scan of Jupiter. Green line: broad band model convolved with estimated transfer function, black dots: received signal. Data processed by Brendan Crill and the Planck HIFI data processing team. Model: full pixel assembly, 64 frequencies over 460 – 630 GHz, pre-launch 'as built' telescope. The dip in the curves below zero energy indicates that the analysis techniques required further development.

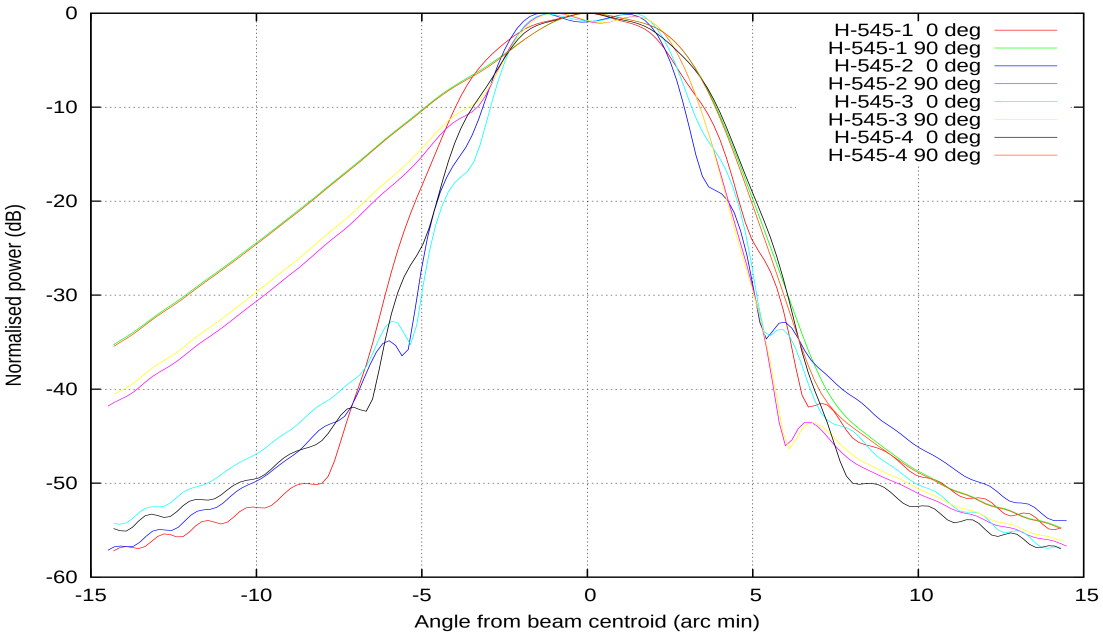


Figure 2.13: Pre-launch broad-band models of the beams on the sky for the Planck multi-mode channels: 545 GHz pixel beam patterns; orthogonal cuts through all beams. The high 'tails' to the left of the main beam are the cuts running vertically downward through the beams as illustrated in figure 2.11. All pixels exhibit marked asymmetry due to optical aberration, particularly the outermost pair of pixels, H-545-1 and H-545-4. Model: full pixel assembly, 64 frequencies over 460 – 630 GHz.

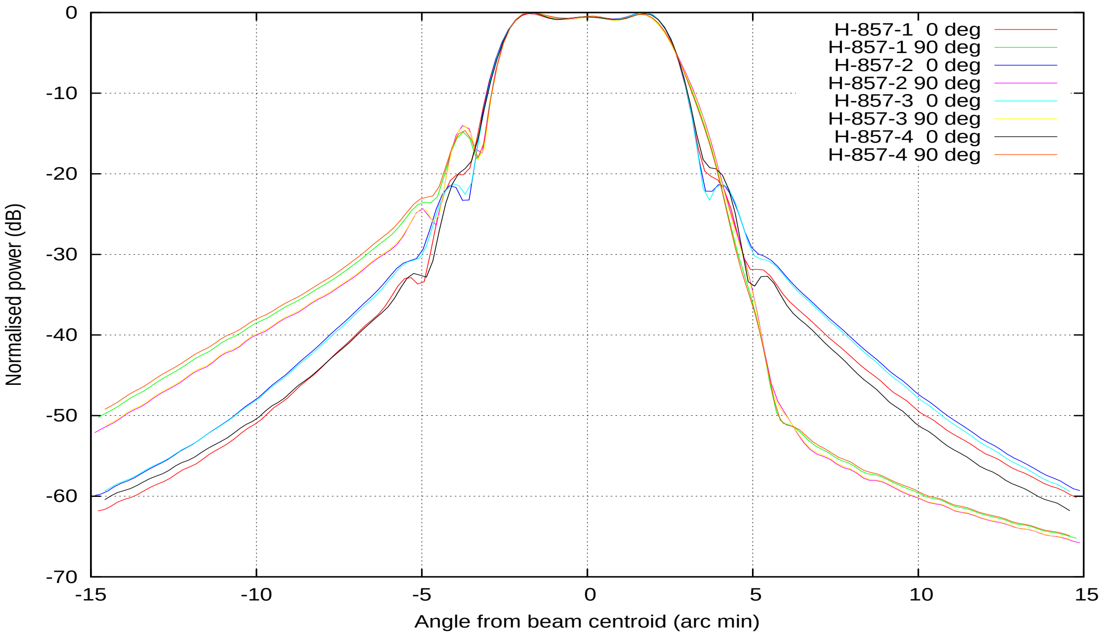
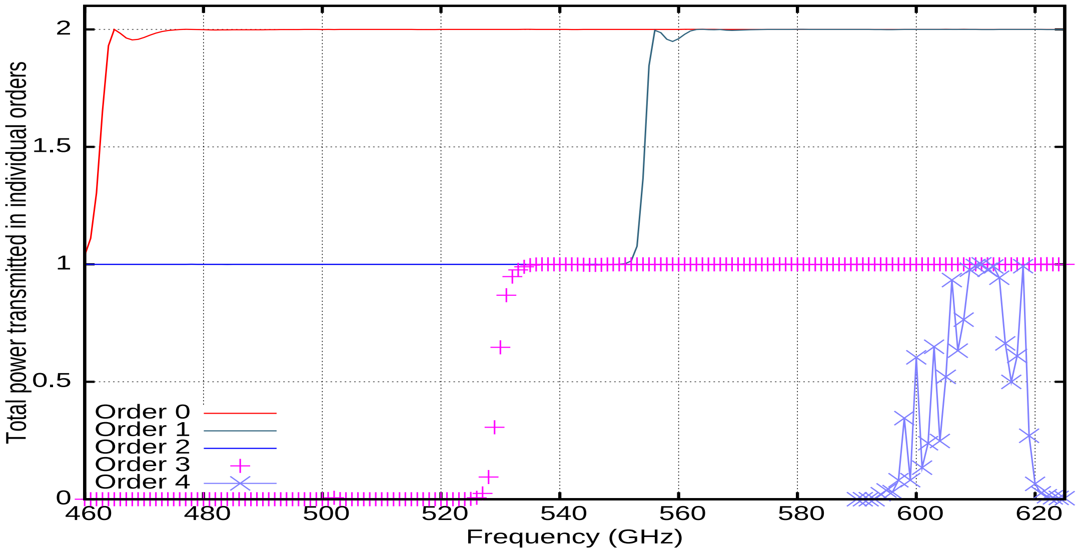


Figure 2.14: Pre-launch broad-band models of the beams on the sky for the Planck multi-mode channels: 857 GHz pixel beam patterns; orthogonal cuts through all beams illustrated in figure 2.9. The outermost pair of pixels, H-857-1 and H-857-4, exhibit marked asymmetry due to coma. Model: full pixel assembly, 52 frequencies 730 – 990 GHz and the as built telescope.

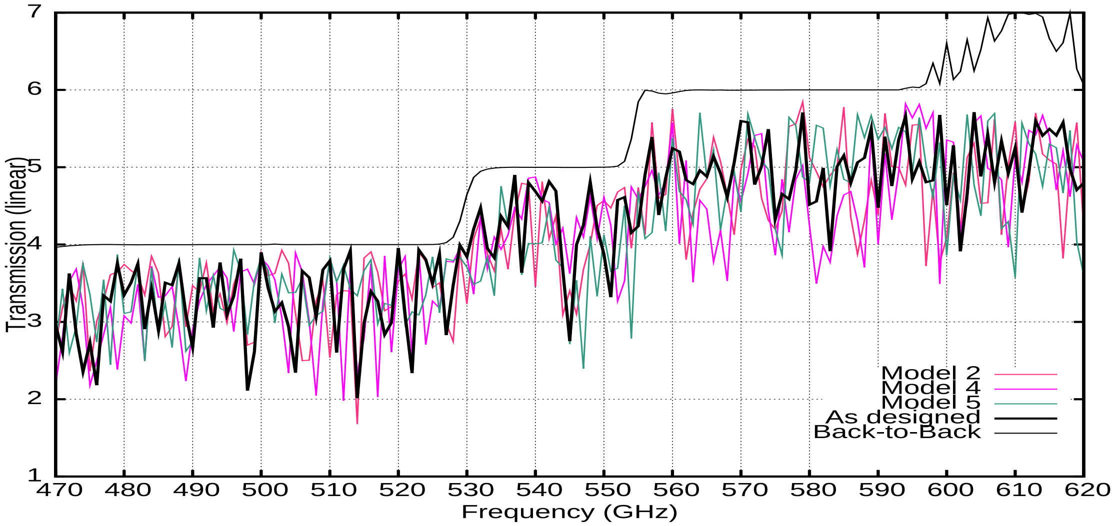
Figure 2.15 (a) on the following page illustrates the frequency dependence of the modal structure of the back-to-back 545 GHz horn as it has been designed. Modes of azimuthal order 2 contribute a constant power level across the band. The order 1 contribution doubles as additional modes cut in at mid band and the order 3 modes cut in at about 530 GHz. The order 4 contribution is a leakage contribution in the back-to-back horn that is not excited in the full pixel assembly. The sum of all these transmitted power contributions gives the total transmission envelope of a single polarisation for the back-to-back section of the horn that is the upper black line in figure 2.15 (b). It forms an upper bound to the very irregular and unpredictable total power transmission of the full pixel assembly, arising because of resonances in the 'cavity' section of the assembly, and discussed further in section 2.5. The modelled power transmission of the as designed pixel, and three tolerance models, are shown below the back-to-back transmission curve. None of these curves are spectrally weighted to account for the presence of the filter stack.

The transmission curve of the 'as designed' full pixel assembly in figure 2.15 (b), when spectrally weighted by the filter stack transmission in figure 2.17 (b) on page 35, gives the modelled transmission curve in figure 2.17 (a). The measured transmission curves for four of the Planck 545 GHz pre-flight corrugated horn assemblies are plotted along with the curve from the model. These Fourier Transform Spectrometer measurements, of filter transmission and of pixel transmission, were provided by P. Ade, G. Savini, B. Maffei and R. Sudiwala, School of Physics and Astronomy, University of Cardiff.

The model assumes a perfect black-body response for the bolometer, no other information being available. The measurements show a marked local maximum at around 530 GHz – close to where the azimuthal order 3 modes cut in, and just below the local minimum of the filter transmission. The cut-in of this mode should not contribute more power than



(a) Excitation of individual azimuthal orders in the Back-to-Back horn model: Power contribution to the radiated field from each azimuthal order *vs* frequency. The plot clearly shows cut-on of the modes of order 3 at 530 GHz. Order 2 modes contribute a constant power across the band while order 1 starts of making a contribution of 1 unit before doubling in power contribution at 555 GHz. The order 4 contribution is not excited in the full horn assembly.



(b) The modelled Back-to-Back horn total power transmission is plotted over the total power transmission of the complete horn assembly models (the sum of all five curves in (a) above). It provides a clear upper bound. The irregular transmission of the complete assembly models is due to resonances in the section of the assembly between the bolometer feed horn and the backward facing section of the Back-to-Back horn.

Figure 2.15: Modelled power transmission (a) trough the back-to-back section of the Planck 545 GHz pixels on an azimuthal order by order basis, and (b) the total power transmission through the back-to-back and the trough the complete pixel assembly model. The resonance effects between the cavity feed horn and the back-horn is a clear impediment to free propagation at all frequencies.

is suppressed by drop in filter transmission. (That is seen from the model curve that is approximately flat on average over the cut on region from around 485 GHz to 545 GHz.) This suggests that either the bolometer cavities have a strong peak in responsivity in the region 520 – 540 GHz, or that, being resonant systems coupling to a resonant system, the resonances of the entire bolometer cavity plus horn assembly give strong throughput in this range, or that the free-space transmission of the filters is not an accurate representation of the transmission when mounted within the cavity. It may also be a feature of the FTS setup. Whatever the case may be, no two pixels, though nominally built to the same design, has the same spectral response.

Figure 2.16 illustrates the broad band beam pattern of the 545 GHz horn model over the main beam superimposed upon the beam pattern at five spot frequencies covering the full spectral band of the horn assembly. All beams are individually normalised for comparison of beam shape. To reiterate: all broad-band multi-mode systems show a changing beam pattern across the band due to the cutting in of additional modes as the frequency increases across the band, but the resonant nature of the horn assembly means that, while a reasonable broad and average beam can be predicted, neither the level nor the exact shape can be predicted at any single frequency.

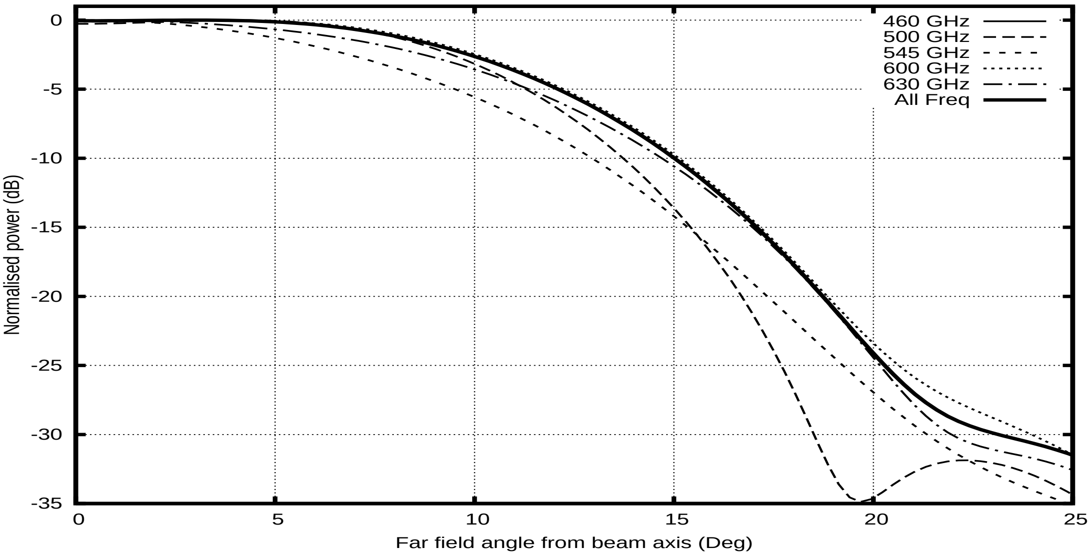
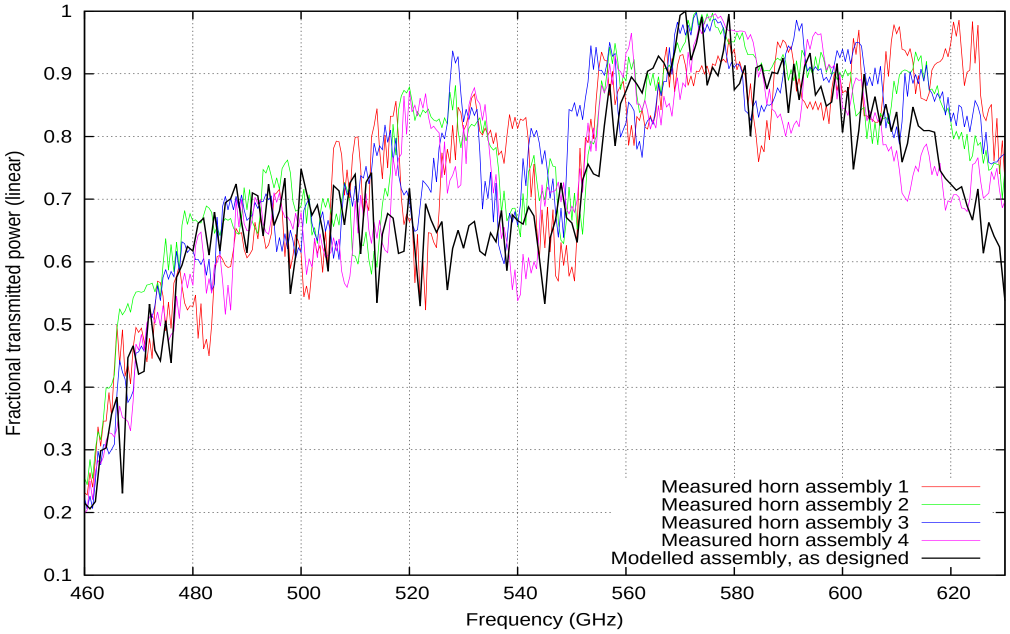
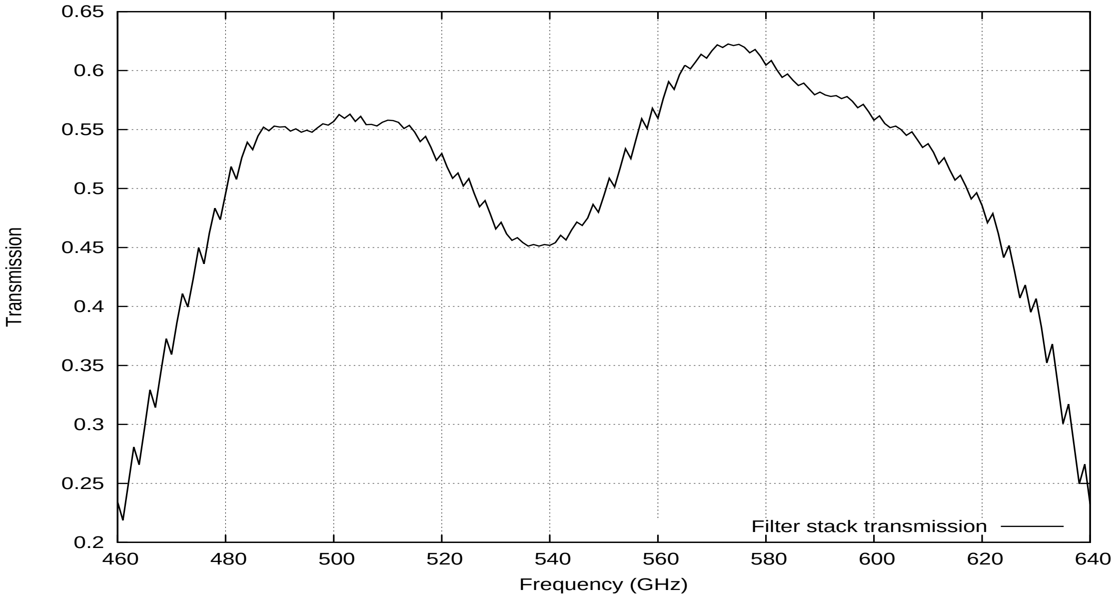


Figure 2.16: Simulated far field beam pattern derived from the aperture fields, model as in figure 2.17. Five frequencies across the band and broad-band estimate based upon 64 frequencies from 460 GHz to 630 GHz.



(a) Measured and modelled 545 GHz horn transmission.



(b) Filter stack transmission model derived from FTS measurements.

Figure 2.17: (a) Comparison of measured power transmission for four Planck 545 GHz horn assemblies and the SKITTER modelled horn with spectrally weighted transmission using a composite model of the filter stack, plotted in (b), built from the measured filter transmissions. There is a strong local maximum in the measured transmission at about 535 GHz, close to the local minimum in the filter stack transmission. This suggests that either (i) the bolometer is tuned to that frequency rather than to the middle of the band, or (ii) that the filter response when mounted within the waveguide is not accurately represented by its free space response, or both of these. FTS measurement data courtesy of P. Ade, G. Savini, B. Maffei and R. Sudiwala, School of Physics and Astronomy, University of Cardiff.

2.5 The problem of resonances

With the mode matching code developed for this study of the Planck multi-mode horns running at approximately sixty times the speed of the mode matching code used for the earlier studies it became possible to make detailed studies of the frequency dependence of the beam shape and power transmission of the Planck horns. Critically, the code produces estimates of the lowest accuracy achieved in the calculation of any scattering product in the model as it is run. There are two run-time indicators of the accuracy and reliability of the results: the minimum reciprocal pivot growth factor (MRPGF) and the minimum reciprocal condition number (MRCN). If the MRPGF is much less than 1 then the results are questionable while the MRCN can be used to estimate the number of reliable decimal places in the calculation. These matters are discussed in chapter 5, but are referred to here.

It has already been stated that the back-to-back sections of the Planck horns exhibit no resonances; furthermore, modelling only that section reduces the number of scattering operations by roughly one third. Consequently modelling them is simple, fast, and the MRPGF and MRCN indicate that the results are reliable. They are not, however, models of the HFI horn assemblies as they are in operation, but are simplified idealisations. Once the bolometer horn is included in the model the system starts to behave in a manner reflecting what has been measured at Cardiff as illustrated in figure 2.17 (a) for, although it is still not a complete system model, it does exhibit some of the qualitative performance characteristics of the real system.

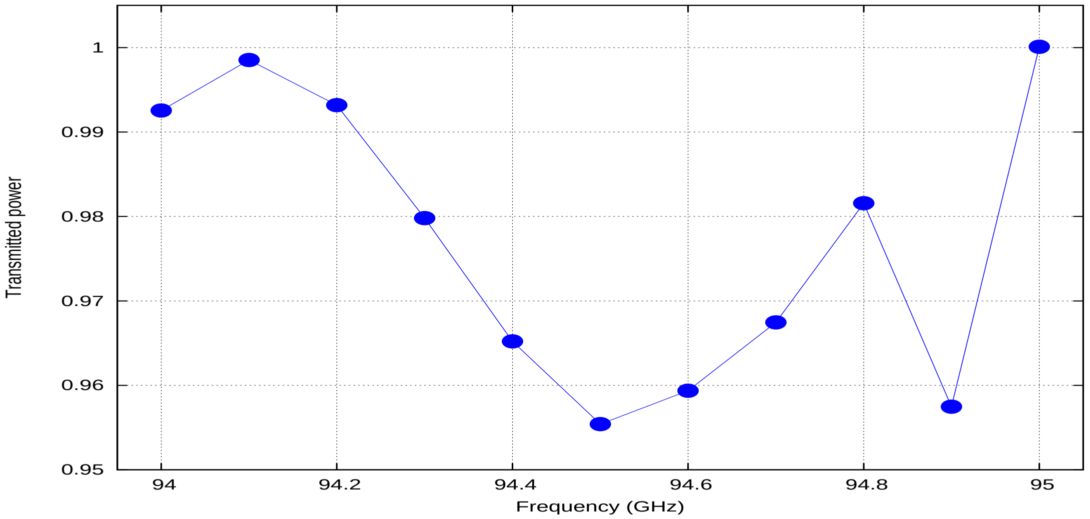
The problems of modelling the resonant behaviour of three systems have been studied in detail: the Planck 100 GHz horn, the Planck 545 GHz horn and W band systems based upon a frequency scaled version of the Planck 857GHz horn.

2.5.1 The Planck 100 GHz model horn assembly

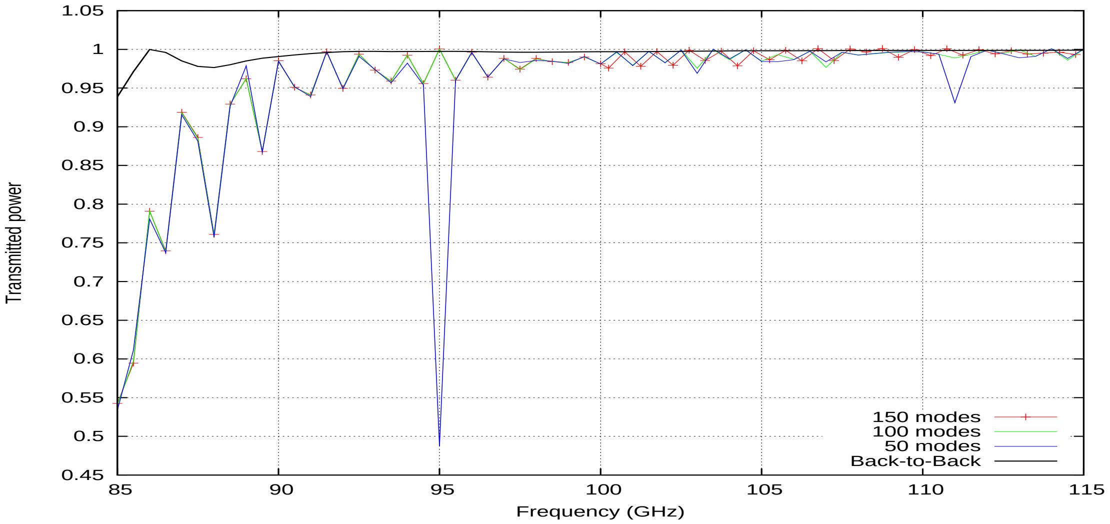
The 100 GHz system is computationally tractable with reliable results obtained on a single core of a PC in short time frames. For the multi-mode systems the problems are far greater and the results less reliable. This difference is accounted for by three factors: the number of azimuthal orders (one or a few), the number of radial modes required to obtain a fully converged model, and the number of scattering junctions.

Put simply, if you have 2000 scattering junctions and 5 azimuthal orders in your model rather than 500 junctions and 1 azimuthal order, then at best you will get less accurate results for the multi-mode system taking roughly 18 times as long for models with the same sized scattering matrices. But it turns out that the situation is far more complex than this.

The power transmission models of the full pixel assembly model of the Planck 100 GHz horn are shown in figure 2.18 on the following page. In this assembly there are only 490 scattering junctions. Models where run at 0.5 GHz steps across the full band with 40, 50, 100 and 150 radial orders (that is, the S_{11}, S_{21} *etc.* matrices are 40×40 to 150×150 complex arrays) and the convergence of the models examined for all frequencies. The lower frequencies converged for smaller numbers of radial modes. All models showed stable numerical behaviour and high precision as indicated by the worst case MRCN and MRPGF and, to within numerical differences of 0.02, the 100 and 150 mode models agreed across the entire band. Therefore, in modelling the system, it would be acceptable to use 100×100 arrays, and such a model takes of the order 75 seconds to run, per frequency, with the single threaded SKITTER code. If complete convergence of the model was required, then 150 modes would be used, with a run time of 217 seconds per frequency. Consequently the horn aperture fields and far field beam pattern predictions, and consequently the Planck beams on the sky, can be considered reliable indicators of what would be measured, though subject to the bolometer response being close to ideal and subject to the models of the telescope being accurate representations of what was built.



(a) 100 radial mode model of the power transmission of the complete Planck 100GHz pixel in 0.1 GHz steps over 94 to 95 GHz. The model has fully converged over this band.



(b) Modelled frequency dependence of power transmission of the complete Planck 100GHz pixel in 0.5 GHz steps. The models show convergence with increasing numbers of radial modes. Solid black curve: back-to-back section of the horn only (no 'cavity').

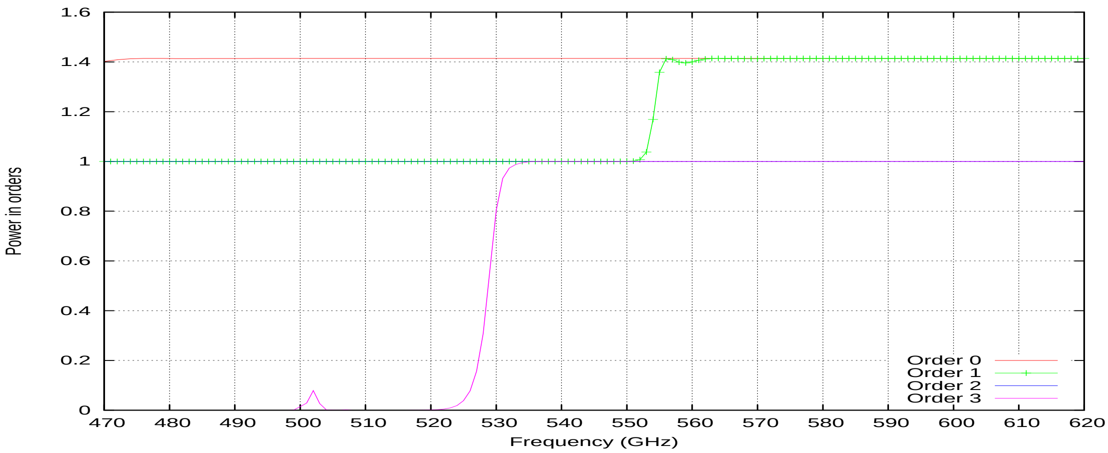
Figure 2.18: The influence of the 'cavity' section in the Planck 100 GHz model single-mode horn on total transmitted power. In (b) the solid black curve shows the predicted power transmission for the back-to-back section of the horn in which there are no resonance effects. The other curves show the convergence of the model as the number of radial modes is increased. With only 20 radial modes (not shown) the model is very erratic, at 100 modes the model has converged at frequencies almost up to the band centre. At 150 modes and above the model has fully converged. This illustrates firstly that modelling the back-to-back section of the horn for beam pattern prediction is inadequate even in a single-mode horn, and secondly that it is essential that sufficient modes be used in the model to adequately account for the power in evanescent modes (see section 3.1.1, page 56).

2.5.2 The Planck 545 GHz model pixel assembly

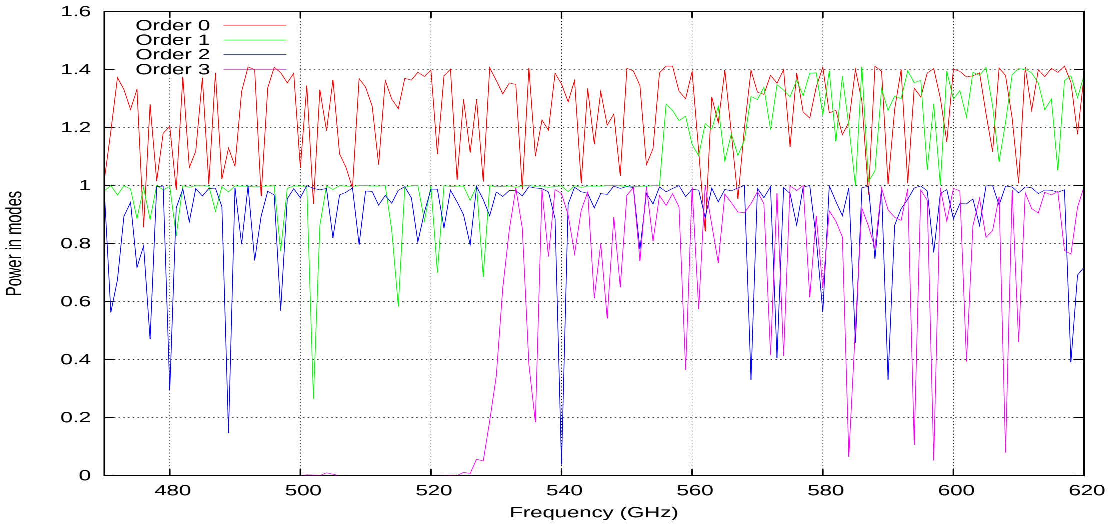
The next detailed study was of the resonant behaviour in the Planck 545 GHz model. These models have 1982 scattering junctions and, whereas the back-to-back model of the horn will pass azimuthal orders 0 to 4 at the top end of the band, the full assembly passes orders 0 to 3. For the beam pattern predictions on the sky models with 64×64 arrays were used. This was done as a compromise between accuracy and feasibility. The pixels are unpolarised and broad band, so it was assumed that, though the point-wise power transmission of the assembly was not reliably known, the broad-band beam pattern prediction would be reasonably reliable since, provided sufficiently large numbers of beams across the frequency band were used in the pattern prediction, the overall error in the broad band power pattern would be small because the mean error would be close to zero.

Figure 2.19 overleaf shows the per-azimuthal order power transmission for the back-to-back section of the horn and for the full pixel model. This particular model uses 100×100 arrays and shows marked resonance in all azimuthal orders. The model has not converged, but both run-time and convergence of the models are significant numerical problems. The comparison the total power transmission of the measured and model horns, plotted in figure 2.17, page 35, suggests that the models are at least qualitatively reasonable. The measurements of the four nominally identical pixels suggests that qualitative agreement is the best that can be hoped for.

Figures 2.20 on page 41 and 2.21 on page 42 plot the MRPGF and MRCN, the run times and the transmitted power predictions for the azimuthal order 3 fields for models with S_{ij} array sizes of 40×40 to 380×380 at 597 GHz. The run times roughly follow the curve $(N/14)^{2.87}$ seconds for $S_{ij} \in \mathrm{M}(N, \mathbb{C})$. Since azimuthal orders 1, 2 and 3 all have the same run times, and order 0 about one half, the full model at N=380 would take about 11 hours 35 minutes per frequency. Therefore, to model the source files at the 64 frequencies of the broad-band model would have taken of the order thirty two days rather than the four and a half hours it actually took. The real problem is not run time, but reliability of the result. It will be shown in the next section that, even with a fully converged model, the system is so sensitive to manufacturing tolerances that a single frequency performance is impossible to predict precisely, so only local average power transmissions and beam profiles should be studied. Nevertheless, to derive those beams some indication of the reliability of the stop frequency models is needed.



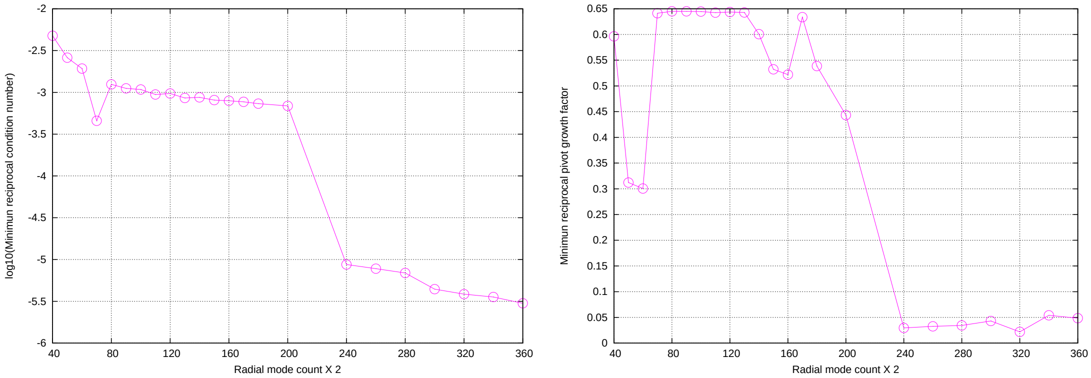
(a) Plank 545 GHz model back-to-back section of the horn assembly.



(b) Plank 545 GHz model of the complete horn assembly

Figure 2.19: Per-azimuthal order transmission through the Plank 545 GHz horn showing the influence of the 'cavity' section on the mode content. In (a) the model uses 50 radial orders and the transmission is that of a simple corrugated horn exhibiting smooth cut-on and no resonance. In (b) the model uses 100 radial orders and the presence of resonant behaviour is clear for all azimuthal orders across the entire band. Because the back-to-back section is free of resonance it is also numerically stable and the output of the model can be taken to be an accurate representation of the radiated field; with the 'cavity' section present the model is no-longer a reliable indication, at any single frequency, of the radiated field.

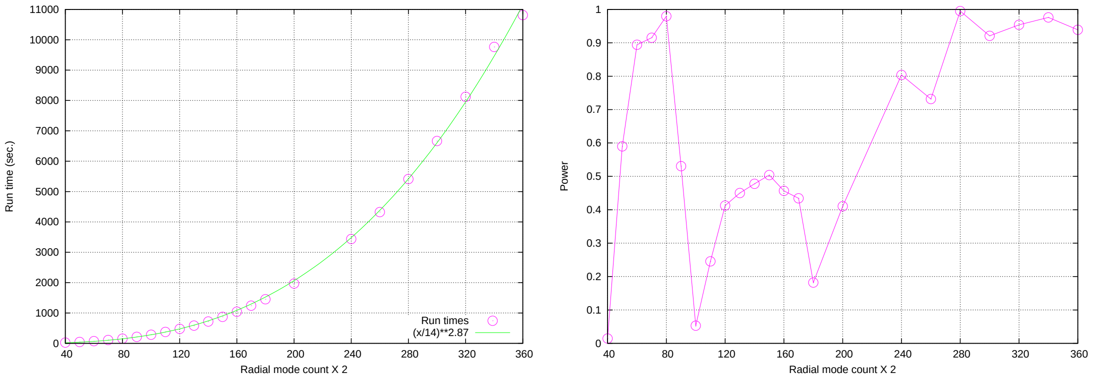
Figure 2.21 (b) overleaf shows that the power transmission model has not converged even with 380×380 arrays. Critically, plots (a) and (b) in Figure 2.20 indicate a reduction in confidence in the results and the number of reliable significant figures when the model is run with arrays larger than 200×200. Approximately, the reduction in significant figures for the worst case solution to the scattering equations is from 12 to 10 decimal places. Consequently, the results become less reliable for very large arrays. Observations of the cash usage during the runs suggests that the large arrays do more swapping of data in memory between different levels, and this might account for the loss of precision. The way forward seems to be to write very carefully constructed code to run on parallel processors, code that is highly tuned for large array sizes on the particular architecture on which it is run.



(a) Minimum reciprocal condition number *vs.* number of radial modes.

(b) Minimum reciprocal pivot growth factor *vs.* number of radial modes.

Figure 2.20: Models of the Planck 545GHz horn assembly with increasing numbers of radial modes, at 597 GHz for the modes of azimuthal order 3. While the numerical stability indicators (a) the minimum reciprocal condition number, and (b) the minimum reciprocal pivot growth factor (see chapter 5) show that all models are numerically stable, the predicted power transmitted varies substantially with the number of radial orders used in the model. Plot (a) indicates that the results of the scattering matrix calculations are accurate to approximately 12 decimal places at worst at any junction for up to 200 modes; thereafter there is a loss of precision to approximately 10 decimal places. Both the minimum reciprocal condition number and the minimum reciprocal pivot growth factor indicate that the reliability of the model falls off above 200 modes. Consequently, although the modelled transmitted power in figure 2.21 (b) on the following page appears to converge towards about 0.95, these results are increasingly unreliable.



(a) Model run times in seconds *vs.* number of radial modes.

(b) Predicted transmission for the azimuthal order 3 fields *vs.* number of radial modes.

Figure 2.21: Continuation of the results plotted in figure 2.20 above: Model runtimes follow the curve $(N/12.4)^{2.72}$ to 200 modes and $(N/14)^{2.87}$ generally but less accurately. Note that the apparent convergence of the power transmission towards about 0.95 i (b) is unreliable due to numerical precision issues.

2.6 The influence of manufacturing tolerances on performance

Manufacturing tolerances result in the radii and lengths of all corrugations deviating randomly from the ideal with some statistical distribution determined by the manufacturing process. As a result there can be no two identical components and the system has to be treated as a whole. The many numerical simulations and measurements that have been performed over the years have shown that a horn without cavity-like sections is sufficiently tolerant of manufacturing errors that beam patterns predicted by modelling the ideal horn conform closely to measurement. When cavity-like sections are present in the system simulations show both throughput and beam pattern varying unpredictably, variations that are attributable to trapped power in the cavity-like sections. Beam pattern variation in a multi-mode horn is due to variation in the distribution of power between the modes (see figure 2.19), and in the power and phase in radial orders within a mode. In a single-mode horn there is only one azimuthal order present and the redistribution of power by a cavity-like section is only between modes of that order. single-mode horns therefore do not exhibit as marked a change in beam shape as multi-mode horns, nor such severe frequency dependent variations in transmission, except possibly in the cut-on region at the lower end of the band.

The results of all of the modelling of multi-mode horns conducted for this work supports an approach to beam pattern prediction that is a stochastic modelling of a number of horns with appropriate randomly assigned errors on corrugation radii and lengths. The method is extremely simple.

2.6.1 Stochastic model generation

Given the ideal system model (the ‘as designed’ system) a choice is made for the statistical distribution of manufacturing errors. This will take the form of a radial offset error due to lathe set-up error plus corrugation-by-corrugation random radial and width errors that are small on the scale of the corrugation dimensions. The overall length error should be close to zero on a digitally encoded CNC lathe, and that must be taken into account in the total distribution of corrugation width errors. For the study presented here the error distribution was uniform.

With the radial offset, the corrugation radius error limit and the width limit decided a probability distribution function is chosen. If N models are to be run and there are M corrugations in the model, then the repetition length (the period of the base pseudo-random number generator beneath the distribution, [37]) of the algorithm should exceed $N \times M$. Typical base pseudo-random number generators have periods exceeding 2^{50}, so repetition within a set of models ought to be no problem regardless of the base generator used: for the Planck multi-mode horns with of the order 2000 corrugations many millions of statistically independent random models could be generated from one call to the random number generator. The real issue is to vary the seed (to set the initial state of the base pseudo-random number generator) so that if a new set of models is generated on another occasion, they are statistically independent from the first set, unless the original models have been lost and need to be regenerated. In that case the use of a base pseudo-random number generator allows the same sequence to be generated given the original seed; consequently the models can be recovered provided the seed has been recorded.

Given a sequence of errors (the pseudo-random numbers of the chosen distribution centred on zero) the errors are added to the as designed model to generate the simulation of the manufactured horn. The overall length of the perturbed model is then scaled to bring the total length error to, or close to, zero. This is how all of the randomised models presented in this work were generated.

In light of the model convergence problems discussed above and the numbers of models to be run the issue of radial mode numbers needs to be addressed. The contention held is that the models do not need to be very large: for the Planck multimode horns array sizes of 64×64 are sufficient for broad-band pattern prediction, and as shown in the next section, and as indicated by the Cardiff measurements in figure 2.17, there is no point in trying to predict narrow band performance precisely for this type of horn assembly.

2.6.2 Modelling results

Figure 2.22 on page 45 illustrates 15 GHz moving averages in the variation in total transmission of six models of the Planck 545 GHz horn assembly with its cavity-like section over a 170 GHz band width centred on 545 GHz. The broadband beam patterns for all of these horn model were found to be smooth and predictable, all models giving essentially the same total power pattern. That means that the beam pattern prediction given in section 2.4 can be taken to be a reliable indication of what would be measured. However, over narrow bands only beam shape is reasonably predictable, not throughput, and over very narrow bands and at spot frequencies neither beam shape nor power throughput can be reliably predicted because of the unpredictable influence of the manufacturing errors on the mode content of the fields. Simulations indicate that the pattern of ringing is sensitive to manufacturing tolerances at the level of $\lambda/1000$, but this does not affect overall quality of the beam pattern. If the bolometer cavity could be used to illuminate the radiating multi-mode horn directly the system does not exhibit this acute sensitivity to manufacturing tolerances and beam pattern prediction over narrow bands would be a simple matter.

On the next page figure 2.23 shows the results of simulating the total power transmission for the as designed and for five randomly perturbed models of the 545 GHz horn at 0.1 GHz intervals over 535 GHz to 555 GHz. The errors are uniformly distributed in the range ±2.5 μm. In this range all azimuthal orders contribute power.

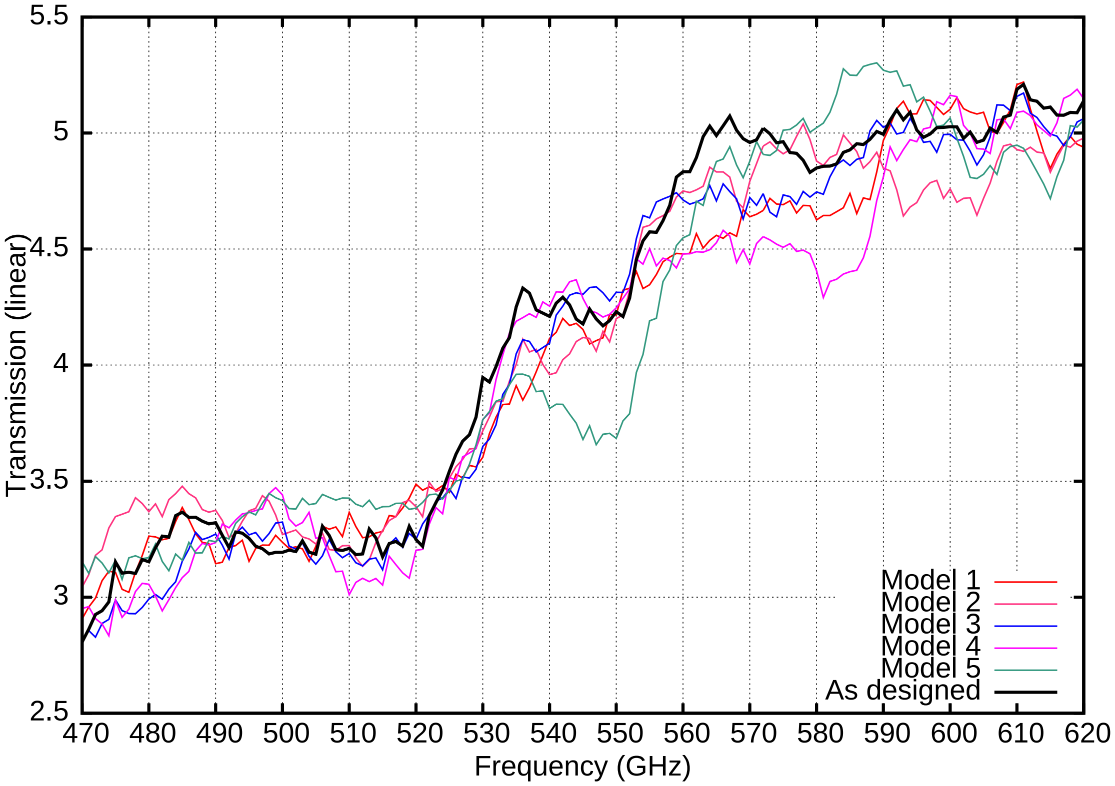
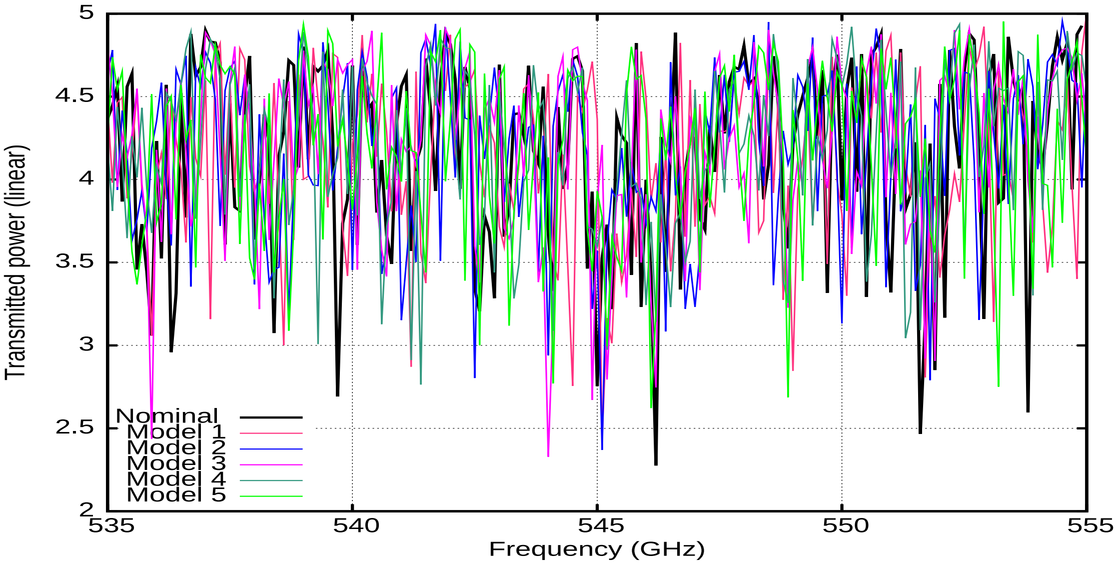
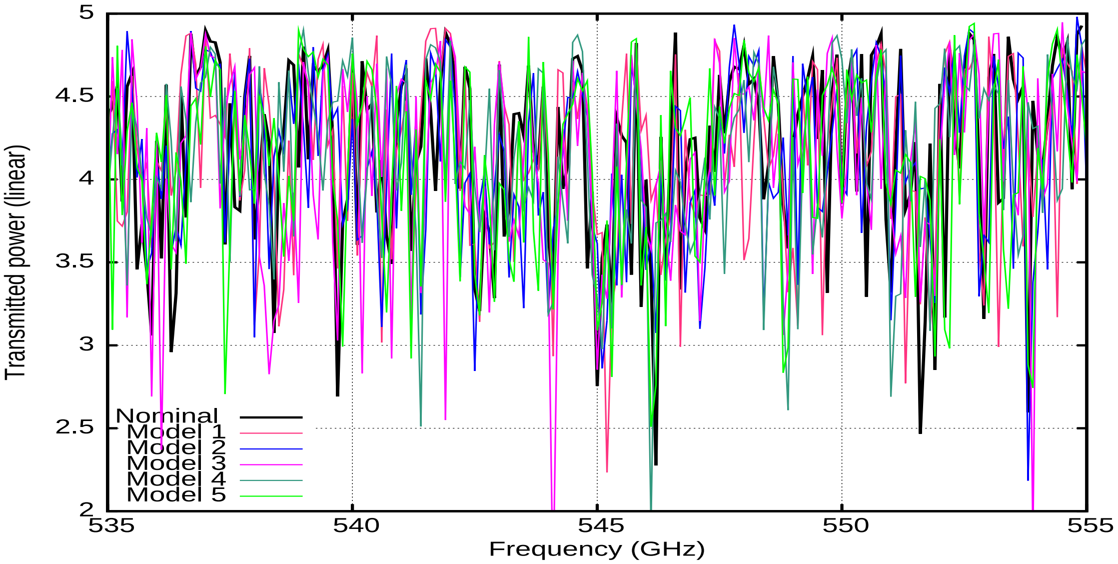


Figure 2.22: 15 GHz moving averages of the power transmitted through the Planck 545 GHz horn assembly, the form of the scattering operator for which is given in equation (2.1.4). The 'as designed' curve is the simulated transmission through the ideal system. The other curves represent simulations of the same system with random manufacturing tolerances of the order ±2.5 μm applied to the segment lengths and radii. All of these models are without spectral weighting for the filters, so the variation is purely due to the influence of these very small radial and length tolerances on the resonances of the assembly. Observe that the models exhibit a power transmission variation of approximately 12.5% at the mid-band cut-in of the additional modes, and this variation is supported by the measured curves in figure 2.17 (a). This suggests that the design exhibits quite severe sensitivity to very small manufacturing errors.



(a) Modelled power transmission at 0.1 GHz steps over the range 535 to 555 GHz for the nominal 545 GHz horn design and for five tolerance models with section length errors uniformly distributed over the range ±2.5μm. The total length of the three model sub-assemblies (cavity feed horn, back horn and radiating horn) are constrained to be exactly as designed to reflect cumulative machining error. The section radii are as in the nominal system.



(b) Modelled horns as in (a) but with the manufacturing error on the radii of the sections, the section lengths being as in the nominal system.

Figure 2.23: Detailed models of the effects of section radius and length tolerances on the resonances in the transition region where azimuthal orders 1 and 3 cut in. The models exhibit great sensitivity to random uniformly distributed manufacturing errors in the range ±2.5μm. The plots indicate that the qualitative effects of length and radial tolerances are essentially the same, but the exact power transmission of a horn cannot be predictable because the 'cavity' renders the horn assembly so sensitive to manufacturing tolerances.

In plot (a) the errors are on the length of the corrugations only, in (b) they are on the radii only. The plots indicate that the effects of radial and of corrugation length errors are qualitatively the same. The plots clearly show that, in so far as the models can be taken as indicative of true behaviour, the spot frequency performance of the Planck multi-mode horns could never be predicted.

2.7 Conclusions

The work presented in this chapter shows that, while a simple waveguide or corrugated horn such as the ClOVER single-mode horns, or the back-to-back section of the Planck single-mode or multi-mode horns can be modelled easily and accurately with a model that will run quickly, this is not true of the multi-mode systems when the full pixel assembly is included. In that case the system becomes resonant due to the cavity-like section formed by the bolometer horn facing the back horn, and the model becomes acutely sensitive both to the exact waveguide section dimensions (to the manufacturing errors) and to the size of the arrays used to model the scattering processes. The model must include large numbers of evanescent modes to reasonably account for trapped power in the 'cavity' section. The convergence of the models becomes a critical issue because, though the current version of the SKITTER mode matching code has been carefully constructed to give accurate results, the run-time indicators of accuracy and reliability suggest that when the size of the arrays goes beyond 200×200 the precision drops to give (at worst) 10 decimal places of accuracy in the solutions to the systems of linear equations at the heart of the scattering product calculations. For small systems such as the 100 GHz horn assemblies such a level of accuracy is more than sufficient, but the Planck multi-mode horns have of the order 2000 scattering junctions, and the accumulation of rounding errors becomes an important consideration.

The Planck multi-mode horns in particular are very sensitive to manufacturing tolerances. Despite this sensitivity, multi-mode systems of the type considered here are suited to broad-band use such as CMB observation and the broad-band beam patterns presented here and delivered to the HFI consortium are reliable indicators of telescope system performance. This is discussed further in chapter 8 on the reverse engineering of the telescope. Two and a half years after the time that this work was completed there were still reliability issues with the HFI calibration scans of the planets, and ±10% swings in measured beam widths were being reported for the multi-mode beams between scans. These swings are believed to be due to a lack of understanding of the individual pixel responses. Once these data reduction issues have been resolved beam maps will become available to the consortium with which the models can be reliably compared. To date only the plots in figures 2.10 on page 28 and 2.12 on page 30 have been made available, in addition to the Cardiff FTS measurements of the power transmission of the 545 GHz and 857 GHz horn assemblies.

The modelling indicates that it should not be expected that precise broad-band beam patterns can be predicted for the horns in the way that can easily be done for traditional single-mode horns or for multi-mode horns not exhibiting resonant behaviour. Over narrow bands beam pattern predictions should be considered indicative only, and it would be advisable to produce several beam pattern predictions from stochastically generated variations on the design, and thus arrive at an indication of the variation that is possible.

It has been found that to get a realistic performance prediction the broad band beam on the sky must be simulated as the incoherent sum of beams at many frequencies. The number of frequencies should be as large as possible so that the random errors in the beam patterns, both mode content and total power transmission, can reasonably be expected to average out. The matter of generating very large numbers of horn aperture fields very quickly is addressed in section 6.4 on page page 164. Though they have not been discussed here, the same methods of analysis and conclusions apply to the much simpler case of single-mode horn assemblies.

Chapter 3

Scattering in waveguides

The fields in a tubular waveguide of simply connected cross section with perfectly conducting walls are described by transverse electric and magnetic fields that decompose naturally into orthogonal 'modes' – the elements of an orthogonal basis for the set of functions satisfying the boundary conditions determined by the physics. In this chapter the waveguides of primary interest are the cylindrical waveguides, the Planck horns being corrugated cylindrical horns. All of the Planck High Frequency Instrument horns are electrically large; in the case of the two multi-mode channels – 545 GHz and 857 GHz – the complete horn assemblies have of the order 2000 scattering junctions and support from three to five azimuthal orders of modes. The problem then became how to model the entire assemblies and arrive at broad-band radiated field pattern predictions for the horns that could then be used as inputs to the GRASP9 models of the telescope, and so arrive at realistic beam pattern prediction for the telescope in operation at L2. Prior to the work described here, models of the beam patterns were derived by modelling only the back-to-back section of the horns at one to five frequencies across the band and propagating then through idealised models of the telescope. The scattering software that had previously been used to generate the horn patterns was the mode-matching code SCATTER developed at Maynooth [43, 14, 25].

To study the frequency dependence of the radiated field pattern in the multi-mode horns demanded very fast, but reliable, code. To that end the algebraic structure of the problem was studied in detail. That lead to the development of code with significantly different, and more complex, structure than the SCATTER code. The numerical implementation of the mode-matching method that was developed will be presented in chapter 5; the beam pattern predictions derived with it have been described in chapter 2. In this chapter it is the scattering theory in stepped waveguides that will be developed. The basic assumptions will be that each section of guide between junctions will be parallel sided, of simply connected cross-section, and have perfectly conducting walls and be homogeneously filled. (Simple connectivity is not a requirement of the theory, but it is all that is required for the horns of interest here. Waveguide cross-sections of arbitrary connectivity are handled in the same way, only the bases for the spaces of functions changes. For annular sections – connectivity 1 – with perfectly conducting walls the function spaces are particularly simple, for connectivity 2 and more the function spaces become complicated, though conformal mapping could help.) In section 3.1 to 3.6 the development makes no assumptions other than these: there is no reason to assume that adjoining sections have the same cross-section. The presentation in sections 3.1 and 3.2 is expository and the material well known, but the presentation is not entirely conventional. In the next chapter section 4.1 develops the algebraic analysis of mode-matching for waveguides of circular section waveguides in detail.

3.1 Fields in cylindrical waveguides

The fundamental assumption is that Maxwell's equations describe the electromagnetic field in the waveguide so that the field propagation for a pure frequency component is described by Helmholtz equation. The particular form of the solution is determined by the boundary conditions and the cross-sectional geometry of the guide. Assume that the guide walls are perfectly conducting, then Dirichlet conditions apply to the electric field and Neumann conditions to the magnetic field. The general formalism outlined here is applicable to guides of any cross-section, but the application will be restricted to guides of cylindrical section, cylindrical meaning of connected piece-wise differentiable boundary with contractible interior, but more particularly, the waveguide section is conformal to a closed disc except possibly at a finite set of discrete points on the boundary (a rectangle, for example).

Denote the cross-section of the guide by S and assume that S is contractible and work in general curvilinear coordinates adapted to the guide: position vector $\boldsymbol{r} = \boldsymbol{r}(u, v, z)$. Define the electric and magnetic Hertz potentials (polarisation potentials) Π_E and Π_M by their relations to the more familiar vector and scalar potentials by $\boldsymbol{A} = \mu_0 \nabla\times \Pi_M + \mu\partial\Pi_E/\partial t$ and $\Phi = -\nabla\cdot\Pi_E/\varepsilon$, as in references [11][33]. Throughout this thesis the notation for electric and magnetic fields and potentials will take the general form X_E^n, X_E, X_M^n or X_M, notation used consistently in the following sense: The superscript denotes the azimuthal order; for the *axial* potentials or fields the subscripts E and M refer to the axial electric and magnetic potential or field; for the *transverse* fields that the axial potentials induce the subscripts E and M are shorthand for *transverse* electric and magnetic fields respectively. Thus, the transverse fields E_E and E_M denote the transverse fields E_{TE} and E_{TM} respectively and a scattering operator S, from E_E to E_M, would be denoted S_{ME} rather than S_{TM-TE}.

In terms of *axial* Hertz potentials $\boldsymbol{\Pi}_M = \hat{\boldsymbol{z}}\Pi_M$ and $\boldsymbol{\Pi}_E = \hat{\boldsymbol{z}}\Pi_E$, the *transverse* electric and magnetic fields are given by [15]

$$\boldsymbol{E}_E = -j\omega\mu_0 \nabla\times\boldsymbol{\Pi}_M, \qquad \boldsymbol{H}_E = \nabla\nabla\cdot\boldsymbol{\Pi}_M + k_0^2\boldsymbol{\Pi}_M, \tag{3.1.1-A}$$

$$\boldsymbol{H}_M = j\omega\varepsilon_0 \nabla\times\boldsymbol{\Pi}_E, \qquad \boldsymbol{E}_M = \nabla\nabla\cdot\boldsymbol{\Pi}_E + k_0^2\boldsymbol{\Pi}_E, \tag{3.1.1-B}$$

respectively. By assumption these axial Hertz potentials satisfy Helmholtz equation, and their axial and transverse coordinate dependence means they must both take the general form $\boldsymbol{\Pi}_F(u, v, z) = \hat{\boldsymbol{z}}\Psi_F(u, v)e^{\pm j\gamma_F z}$ for some scalar function Ψ_F and propagation constant γ_F, where F is one of the fields E or M, as appropriate. It is important to observe that the function Ψ_F is simply a point in the function space $L^2(S)$ – square integrable functions on the domain S that is the waveguide cross section on which the (u, v)-coordinate system is defined. The physics determines the boundary conditions that determines the subspace of $L^2(S)$ in which Ψ resides, as well as γ_F. In describing the scattering of fields at a waveguide junction, and in the numerical implementation of the scattering, this split between a purely geometric (function-analytic) aspect and a physical aspect of the representation of the fields will be exploited.

Since axial $\boldsymbol{\Pi}_F$ satisfies Helmholtz equation, Ψ_F must satisfy the scalar Helmholtz equation point-wise in the (u, v) plane. Thus $\nabla_T^2\, \Psi_F + \kappa_F^2 \Psi_F = 0$, with $\kappa_F^2 \overset{\text{def}}{=} k_0^2 - \gamma_F^2$ and solving for the fields in terms of the potentials from equations (3.1.1) gives

$$\boldsymbol{H}_E = \pm\, \gamma_M\, \nabla_T\, \Psi_M e^{\pm j\gamma_M z}, \quad H_z = \kappa_M^2 \Psi_M e^{\pm j\gamma_M z}, \quad \boldsymbol{E}_E = \pm\, Z_E \hat{\boldsymbol{z}} \times \boldsymbol{H}_E, \tag{3.1.2-A}$$

$$\boldsymbol{E}_M = \pm\, \gamma_E\, \nabla_T\, \Psi_E e^{\pm j\gamma_E z}, \quad E_z = \kappa_E^2 \Psi_E e^{\pm j\gamma_E z}, \quad \boldsymbol{H}_M = \mp\, Y_M \hat{\boldsymbol{z}} \times \boldsymbol{E}_M \tag{3.1.2-B}$$

for the transverse electric and transverse magnetic field components and the axial field components. The coefficients $Z_E = Z_0 k_0/\gamma_M$ and $Y_M = Y_0 k_0/\gamma_E$ are the guide impedance and admittance, respectively. The notation Y_M and Z_E has been chosen to agree with the subscript labelling of the transverse fields and ought not to cause confusion since the γ_E and γ_M are associated with the axial fields which induce TM and TE fields respectively.

From equations (3.1.2) it is immediate that the fields in the waveguide can be completely and conveniently described by expanding the functions $\Psi_E(u, v)$ and $\Psi_M(u, v)$ in terms of any appropriate basis functions for $L^2(S)$ that satisfy the Dirichlet and Neumann boundary conditions, respectively. This is the analytical framework for the mode-matching method for the description of the transverse fields in a cylindrical guide in terms of ‘modes’ – basis functions spanning an infinite dimensional, denumerable and separable complex Hilbert space. Scattering that occurs at a waveguide junction is described by an operator matrix acting on the direct sum of the two Hilbert spaces at either side of the junction. With identical waveguide sections at either side of the junction the two Hilbert spaces are trivially isomorphic, but there are three factors that give rise to the scattering operator: the purely geometric effect of the step change in guide section, and the physical effects of the step change in impedance and the enforcement of the boundary conditions on the step flange. The idealised physical assumption of perfectly conducting waveguide walls means that there is no tangential component to the electric field and no axial component of the magnetic field on the radial flange at the step. Basis functions (modes) on either side of the junction have to be matched: the modes in the larger section are expanded in terms of the modes in the smaller section. This gives rise to a purely real operator matrix that describes the scattering across the junction. The impedance depends upon the basis function and is real for a propagating mode and pure imaginary for an evanescent mode. These impedances scale the components of the real scattering operator giving rise to a generally complex operator.

Let $\Psi = \sum \Psi_k$ denote the expansion of either Ψ_E or Ψ_M in terms of modes $\{\Psi_k\}$ and let $\{\gamma_k\}$ be the corresponding set of propagation constants. Helmholtz equation gives

$$\begin{aligned} 0 &= (\Psi_i \nabla_T^2 \Psi_j + (k_0^2 - \gamma_j^2)\Psi_i\Psi_j) = (\Psi_j \nabla_T^2 \Psi_i + (k_0^2 - \gamma_i^2)\Psi_j\Psi_i) \\ \implies (\gamma_i^2 - \gamma_j^2)\int_S \Psi_i\Psi_j \,\mathrm{d}S &= \int_S \left(\Psi_i \nabla_T^2 \Psi_j - \Psi_j \nabla_T^2 \Psi_i\right) \mathrm{d}S \\ &= \int_{\partial S} \left(\Psi_i \frac{\partial \Psi_j}{\partial \hat{\boldsymbol{n}}} - \Psi_j \frac{\partial \Psi_i}{\partial \hat{\boldsymbol{n}}}\right) \mathrm{d}l, \end{aligned} \tag{3.1.3}$$

where the equality between the surface integral and the boundary integral is Green's second identity (A.3.2), and $\hat{\boldsymbol{n}}$ is the normal to the boundary. Since the Dirichlet conditions imply that $\Psi_E|_{\partial S} \equiv 0$, and the Neumann conditions imply that $\partial\Psi_M/\partial\hat{\boldsymbol{n}}|_{\partial S} \equiv 0$, the integrals in equation (3.1.3) are identically zero as long as the non-degeneracy condition $\gamma_i \neq \gamma_j$ holds. In all cases considered in this thesis this condition will hold, but in general if $\gamma_{i_1} = \gamma_{i_2} = \cdots = \gamma_{i_n}$ holds, the familiar Gram-Schmidt process of orthonomalisation can be applied to the set $\{\Psi_{i_k}\}_{k=1}^n$, n possibly infinite. If the Ψ_{i_k} spans a subspace $H \subset L^2(S)$, then so too does the resulting orthonormalised set $\{\hat{\Psi}_k\}_{k=1}^n$, and the result will be a denumerable, orthonormal, basis set for the space of axial E_z and M_z fields over S.

From equations (3.1.2-A and B) the transverse fields are expanded in terms of the Laplacians of the basis functions Ψ_i as $\boldsymbol{E}_M = \pm\sum_i \gamma_{E,i} \nabla_T \Psi_{E,i} e^{j\gamma_{E,i}z}$, etc. Orthogonality of the modes requires that the appropriate one of the three inner products

$$\langle \nabla_T \Psi_i, \nabla_T \Psi_j \rangle\,,\ \langle \hat{\boldsymbol{z}} \times \nabla_T \Psi_i, \hat{\boldsymbol{z}} \times \nabla_T \Psi_j \rangle\,,\ \text{or}\ \langle \nabla_T \Psi_i, \hat{\boldsymbol{z}} \times \nabla_T \Psi_j \rangle$$

be zero. These inner products are defined as the integrals over the pointwise scalar product of the two components over the guide cross section, S. Pointwise, $(\hat{\boldsymbol{z}} \times$

$\nabla_T \Psi_i) \cdot (\hat{\boldsymbol{z}} \times \nabla_T \Psi_j) = \nabla_T \Psi_i \cdot \nabla_T \Psi_j$ so that the first and second inner products both become

$$\begin{aligned}\int_S \nabla_T \Psi_i \cdot \nabla_T \Psi_j \,\mathrm{d}S &= \int_{\partial S} \Psi_i \frac{\partial \Psi_j}{\partial \hat{\boldsymbol{n}}} \,\mathrm{d}l - \int_S \Psi_i \nabla_T^2 \Psi_j \,\mathrm{d}S \\ &= k_0^2 \int_S \Psi_i \Psi_j \,\mathrm{d}S \propto \delta_{ij}\end{aligned} \tag{3.1.4}$$

by Green's first identity (A.3.1), the application of the appropriate boundary conditions to eliminate the boundary integral, using Helmholtz equation, and applying the previous result. Thus, constituent modes for the field pairs of type (E_E, E_E), (H_E, H_E) (E_M, E_M), (H_M, H_M), (E_M, H_M) and (E_M, H_E) are pairwise orthogonal because the axial functions that determine them, up to mode dependent factors, are orthogonal.

For the third inner product, which applies to (E_E, E_M) and (H_E, H_M) pairs only,

$$\begin{aligned}\langle \nabla_T \Psi_{Ei}, \hat{\boldsymbol{z}} \times \nabla_T \Psi_{Mj} \rangle &= \int_S \nabla_T \Psi_{Ei} \cdot (\hat{\boldsymbol{z}} \times \nabla_T \Psi_{Mj}) \,\mathrm{d}S \\ &= -\int_S \nabla_T \cdot (\hat{\boldsymbol{z}} \Psi_{Mj} \times \nabla_T \Psi_{Ei}) \,\mathrm{d}S \\ &= \int_{\partial S} \Psi_{Mj} (\hat{\boldsymbol{z}} \times \nabla_T \Psi_{Ei}) \cdot \hat{\boldsymbol{n}} \,\mathrm{d}l.\end{aligned} \tag{3.1.5}$$

However the boundary conditions give $(\hat{\boldsymbol{z}} \times \nabla_T \Psi_{E,i}) \cdot \hat{\boldsymbol{n}} \equiv 0$, from which the orthogonality follows. Reversing the rôles of the Ψ_E and Ψ_M it is the $\Psi_E|_{\partial S} \equiv 0$ that gives the orthogonality. The final cases are (E_E, H_E) and (E_M, H_M), but these two have the same general form as for equations (3.1.5) giving the orthogonality of any E_E mode to any H_E mode and any E_M mode to any H_M mode in the same section of the guide. This will be used below for the decomposition of the spaces of TE and of TM fields into direct sums of electric and magnetic components.

Denote the Hilbert space of transverse electric and transverse magnetic fields over S by $\mathcal{H}_E$ and $\mathcal{H}_M$ respectively; then the completeness of the spaces and denumerability and orthogonality of the bases means that the spaces have complete, orthonormal bases, upon normalisation with respect to the inner product. It is in these bases, the modes of the waveguide section, that the fields and scattering operators are to be expanded.

Given two linearly independent solutions, $\boldsymbol{E}$, $\boldsymbol{H}$ and $\boldsymbol{E}'$, $\boldsymbol{H}'$ to Maxwell's equations, $0 = -j\mu\omega(\boldsymbol{H}\cdot\boldsymbol{H}' - \boldsymbol{H}'\cdot\boldsymbol{H}) = \boldsymbol{H}'\cdot(\nabla\times\boldsymbol{E}) - \boldsymbol{H}\cdot(\nabla\times\boldsymbol{E}')$ and $0 = -j\varepsilon\omega(\boldsymbol{E}\cdot\boldsymbol{E}' - \boldsymbol{E}'\cdot\boldsymbol{E}) = \boldsymbol{E}'\cdot(\nabla\times\boldsymbol{H}) - \boldsymbol{E}\cdot(\nabla\times\boldsymbol{H}')$, which, upon addition of the right hand sides gives

$$0 = \nabla\cdot(\boldsymbol{E}\times\boldsymbol{H}' - \boldsymbol{E}'\times\boldsymbol{H}) = \nabla_T\cdot(\boldsymbol{E}\times\boldsymbol{H}' - \boldsymbol{E}'\times\boldsymbol{H}) + \hat{\boldsymbol{z}}\partial_z\cdot(\boldsymbol{E}\times\boldsymbol{H}' - \boldsymbol{E}'\times\boldsymbol{H}). \tag{3.1.6}$$

With the decomposition of the fields into transverse and axial components with axial dependence of the form $e^{-j\gamma z}$ this becomes

$$0 = \nabla_T\cdot(\boldsymbol{E}\times\boldsymbol{H}' - \boldsymbol{E}'\times\boldsymbol{H}) - j(\gamma+\gamma')\hat{\boldsymbol{z}}\cdot(\boldsymbol{E}_T\times\boldsymbol{H}_T' - \boldsymbol{E}_T'\times\boldsymbol{H}_T).$$

But then the integral of (3.1.6) over S is identically zero and, since the first term on the right gives an integral around the boundary, under the assumption of a perfectly conducting boundary $0 = \hat{\boldsymbol{n}}\cdot(\boldsymbol{E}\times\boldsymbol{H}) = \boldsymbol{H}\cdot(\hat{\boldsymbol{n}}\times\boldsymbol{E})|_{\partial S}$, it follows that

$$(\gamma+\gamma')\int_S(\boldsymbol{E}_T\times\boldsymbol{H}_T' - \boldsymbol{E}_T'\times\boldsymbol{H}_T)\cdot\,\mathrm{d}\boldsymbol{S} = 0. \tag{3.1.7}$$

With the axial dependence $e^{-j\gamma z}$ the transverse fields have the form

$$\begin{aligned}\boldsymbol{H}_T(u,v,z) &= \boldsymbol{h}(u,v)e^{-j\gamma z} = \sum_m \boldsymbol{h}_m(u,v)e^{-j\gamma_m z},\\ \boldsymbol{E}_T(u,v,z) &= \boldsymbol{e}(u,v)e^{-j\gamma z} = \sum_n \boldsymbol{e}_n(u,v)e^{-j\gamma_n z},\end{aligned}$$

and equation (3.1.7) becomes

$$(\gamma+\gamma')\int_S(\boldsymbol{e}\times\boldsymbol{h}' - \boldsymbol{e}'\times\boldsymbol{h})\cdot\,\mathrm{d}\boldsymbol{S} = 0. \tag{3.1.8}$$

The fields propagating in the reverse direction have z-dependence $e^{j\gamma z}$ and give

$$(\gamma-\gamma')\int_S(-\,\boldsymbol{e}\times\boldsymbol{h}' - \boldsymbol{e}'\times\boldsymbol{h})\cdot\,\mathrm{d}\boldsymbol{S} = 0. \tag{3.1.9}$$

Adding equation (3.1.8) to (3.1.9) and subtracting equation (3.1.9) form (3.1.8) gives the required orthogonality of linearly independent modes:

$$\int_S \boldsymbol{e}_m\times\boldsymbol{h}_n\cdot\mathrm{d}\boldsymbol{S} = \int_S \boldsymbol{e}_n\times\boldsymbol{h}_m\cdot\mathrm{d}\boldsymbol{S} \propto \delta_{nm} \tag{3.1.10}$$

This is an inner product measuring *power* coupling, and by orthogonality of the basis:

$$\boldsymbol{e}_n\cdot\boldsymbol{e}_m = \boldsymbol{e}_n\cdot(\boldsymbol{h}_m\times\hat{\boldsymbol{z}})Z_m = (\boldsymbol{e}_n\times\boldsymbol{h}_m)\cdot\hat{\boldsymbol{n}}Z_m = \delta_{nm}.$$

For $n = m$ the inner product (3.1.10) is the integral of the Poynting vector over the waveguide section. Physically, two fields $(\boldsymbol{e}_1, \boldsymbol{h}_1)$, $(\boldsymbol{e}_2, \boldsymbol{h}_2)$ in a waveguide are orthogonal over a transverse section, S, if the $n$et power, measured as the integral over the plane of the axial components of $\boldsymbol{e}_1 \times \boldsymbol{h}_2 = \boldsymbol{e}_2 \times \boldsymbol{h}_1$ crossing the plane, is zero at all times. In a lossless guide the real Poynting vector can be replaced with the complex Poynting vector, $\boldsymbol{e}_1 \times \boldsymbol{h}_2^*$, and that will be the case for the Planck horns and all other waveguide structures considered hereafter, all being treated as having perfectly conducting walls and no dielectric filling anywhere in the horn assembly.

3.1.1 Power flow

With the fields expanded in terms of orthonormal modes, the time-averaged power flow across a section of the guide is given by the integral of the Poynting vector over the guide section

$$\begin{aligned} P &= \frac{1}{2}\mathrm{Re}\int_S \boldsymbol{E} \times \boldsymbol{H}^* \cdot \mathrm{d}\boldsymbol{S} = \frac{1}{2}\mathrm{Re}\int_S \boldsymbol{E}_T \times \boldsymbol{H}_T^* \cdot \mathrm{d}\boldsymbol{S} \\ &= \frac{1}{2}\sum_{n=0}^{\infty} A_n \mathrm{Re}\int_S \boldsymbol{e}_n \times \boldsymbol{h}_n^* \cdot \mathrm{d}\boldsymbol{S}. \end{aligned} \tag{3.1.11}$$

where the $\boldsymbol{e}_n$ and $\boldsymbol{h}_n$ are now normalised over the guide section, the A_n are the expansion coefficients, and the cross terms have been eliminated using equation (3.1.10).

From equations (3.1.2) the general form of the transverse fields for both TE and TM modes is $\boldsymbol{F}_n = \pm\gamma \nabla_T \Psi_n e^{j\gamma z}$, $\boldsymbol{G}_n = \pm\zeta \hat{\boldsymbol{z}} \times \boldsymbol{F}_n$, where $\boldsymbol{F}$ and $\boldsymbol{G}$ stand for the electric or magnetic field as appropriate, γ is the propagation constant, and ζ the appropriate impedance or admittance. In either case, the integrand in equation (3.1.11) takes the form

$$\begin{aligned} \zeta\gamma^2 \hat{\boldsymbol{z}} \cdot [(\hat{\boldsymbol{z}} \times \nabla_T \Psi_n) \times \nabla_T \Psi_n] &= -\zeta\gamma^2 \nabla_T \Psi_n \cdot \nabla_T \Psi_n + (\hat{\boldsymbol{z}} \cdot \nabla_T \Psi_n)^2 \\ &= -\zeta\gamma^2 \nabla_T \Psi_n \cdot \nabla_T \Psi_n. \end{aligned}$$

The power in the n^{th} propagating mode is then obtained from Green's first identity as

$$\begin{aligned}P_n &= \frac{1}{2}\zeta\gamma^2 \int_S \nabla_T \Psi_n \cdot \nabla_T \Psi_n \,\mathrm{d}S = -\frac{1}{2}\zeta\gamma^2 \int_S \Psi_n \nabla_T^2 \Psi_n \,\mathrm{d}S + \frac{1}{2}\zeta\gamma^2 \int_{\partial S} \Psi_n \frac{\partial \Psi_n}{\partial \hat{\boldsymbol{n}}}\,\mathrm{d}l \\ &= \frac{1}{2}\zeta\gamma^2\kappa^2 \int_S \Psi_n^2 \,\mathrm{d}S, \\ &= \begin{cases} \frac{1}{2}Z_0 k_0 \gamma_{Mn}\kappa_{Mn}^2 \|\Psi_{Mn}\|_S^2 & : \quad \text{for the TE case,} \\ \frac{1}{2}Y_0 k_0 \gamma_{En}\kappa_{En}^2 \|\Psi_{En}\|_S^2 & : \quad \text{for the TM case,} \end{cases}\end{aligned} \tag{3.1.12}$$

where κ is as on page 52, and Z_0, Y_0 and k_0 are the free space impedance, admittance and wavenumber. Here the boundary integral gives zero in both the TE and TM cases by the boundary assumptions, and $\|\Psi_n\|_S^2$ is the squared $L^2(S)$ norm of the basis function over the waveguide cross section.

For non-evanescent fields, $\kappa = k_0\sqrt{1-\gamma^2/k_0^2}$ is real, this integral is real and the time averages of the electric and magnetic field energies, w_e and w_m, are equal (see [11], [15]) and the time averaged power flow is $P = 2w_e v_g$, which determines the group velocity, v_g. This is the regime in which Helmholtz equation is a wave equation.

When $\kappa \in i\,\mathbb{R}$ Helmholtz equation is a diffusion type equation, second order in axial distance, z. Within a section of guide of constant cross section the general form for the time averaged power stored between transverse planes at $z = z_0$ and $z = z_0 + \delta_z$ for TE modes is given by

$$\begin{aligned}2i\omega(w_m - w_e) &= \frac{1}{2} i Z_0 k_0 \kappa_M^2 |\gamma_M| e^{-2|\gamma_M|z_0}(1 - e^{-2|\gamma_M|\delta_z})\|\Psi_M\|_S^2 \\ \Longrightarrow\; w_m - w_e &= \frac{1}{4\omega} Z_0 k_0 \kappa_M^2 |\gamma_M| e^{-2|\gamma_M|z_0}(1 - e^{-2|\gamma_M|\delta_z})\|\Psi_M\|_S^2\end{aligned}$$

so that for TE modes $w_m > w_e$ in the length of waveguide. For TM modes the equation becomes

$$w_e - w_m = \frac{1}{4\omega} Y_0 k_0 \kappa_E^2 |\gamma_E| e^{-2|\gamma_E|z_0}(1 - e^{-2|\gamma_E|\delta_z})\|\Psi_E\|_S^2$$

and $w_e > w_m$ in the length of waveguide. For a numerical model these relations need to be observed. There must be sufficient evanescent modes in the model that these conditions hold in all sections of the guide.

3.2 Scattering at a waveguide junction

For the purposes of modelling and analysis the fields are expanded in terms of functions in $L^2(S)$. Physically this must be the correct space because there must be finite power in the fields, and the power is related to the L^2 norm by equation (3.1.12). The assumption of perfectly conducting boundaries selects orthogonal subspaces of $L^2(S)$ within which the TE and TM fields can be expanded. From the previous section it is seen that the fields are described in terms of an L^2 function multiplied by terms related to the physics of wave propagation in the guide – the frequency, impedances and dielectric properties. In what follows it will be shown that the scattering of the modes across a junction is described by an expansion of the modes on the larger guide section in terms of the modes in the smaller section – a Fourier series expansion, in the general sense – and multiplication by an appropriate impedance term.

Denote the guide cross section to the left of the junction by S_L and that to the right by S_R. Without loss of generality assume that the guide is smaller on the left than on the right. Whatever the two sections may be, and whatever the chosen coordinate systems on S_L and S_R, the physical junction determines (is described by) an injective mapping $\pi : S_L \hookrightarrow S_R$ that will be a C^∞ isometric embedding. Usually in analysis of scattering in waveguides this mapping, and what follows from it, is ignored, but in the analysis of misaligned guide sections it is critical. Furthermore, for the analysis to be formally correct, it has to be used. The functions in $\mathcal{H}(S_R)$ pull back to $\mathcal{H}(S_L)$ via the linear pull-back induced by the mapping π:

$$\pi^* : \mathcal{H}(S_R) \to \mathcal{H}(S_L), \quad (\pi^* f)(x) = f(\pi(x)), \quad \text{for all } x \in \pi(S_L). \tag{3.2.1}$$

Both of the spaces spanned by the TE fields and by the TM fields are representable as functions in $\mathcal{H}(S)$. Denote these two spaces by $\mathscr{E}$ and $\mathscr{M}$ respectively; the total fields to the left and right of the junction are contained in the spaces $(\mathscr{E} \oplus \mathscr{M})_L$ and $(\mathscr{E} \oplus \mathscr{M})_R$. Then π^* gives the pull-back $\pi^* : (\mathscr{E} \oplus \mathscr{M})_R \to (\mathscr{E} \oplus \mathscr{M})_L$ and, since the constituent spaces all have denumerable bases and π^* is linear, it determines infinite operator matrix, to be denoted by P, of the form

$$\begin{bmatrix} P_{EE} & P_{ME} \\ P_{EM} & P_{MM} \end{bmatrix} : (\mathscr{E} \oplus \mathscr{M})_L \longrightarrow (\mathscr{E} \oplus \mathscr{M})_R \, . \tag{3.2.2}$$

Each sub-matrix in this operator matrix is an infinite operator matrix, the P_{mk} component of which maps the k^{th} TE or TM basis vector on the left to the m^{th} basis vector in the expansion of TE or TM on the right of the junction being $P_{mk} = \langle \psi_k | \pi^* \psi'_m \rangle$. Thus, this operator P, and its adjoint $P^\dagger$, describe the geometric aspects of the scattering – those aspects of the scattering that relate purely to the embedding $\pi : S_L \to S_R$ and the choice of bases in $\mathcal{H}(S_L)$ and $\mathcal{H}(S_R)$, the particular structure of P being determined purely by the geometry of S_L and S_R and the embedding π. For geometries such as the discs and rectangles considered below, and for regular polygons generally, this structure will be particularly simple provided the sections are perfectly coaxial and aligned. In that situation the matrix P will be found to take a block diagonal form; effectively a direct sum of arrays, $P = \bigoplus P_n$. If the alignment is not perfect the mapping π becomes critical to the numerical implementation and the operator matrix P will be dense. (For circular guides the index n will label the azimuthal orders, and if a horn is constructed from section not perfectly aligned there will be scattering between azimuthal orders with a resulting change in beam structure and loss of efficiency.) From hereon, except when misaligned guides are considered, π will be the identity mapping and all reference to it will usually be dropped and the distinction between the domains of definition of the functions in the integrands will be ignored.

There must also be endomorphisms that account for the reflection of modes at the junction due to the step change in impedance:

$$R : (\mathscr{E} \oplus \mathscr{M})_L \longrightarrow (\mathscr{E} \oplus \mathscr{M})_L \, ,$$
$$Q : (\mathscr{E} \oplus \mathscr{M})_R \longrightarrow (\mathscr{E} \oplus \mathscr{M})_R \, .$$

It is immediate from the orthogonal decomposition of the fields and the integrals of the previous section that these operator matrices will be diagonal with terms $\langle \psi_i | \psi_j \rangle \propto \delta_{ij}$; physically the reflection operators will be determined solely by the impedance step across the junction and will not scatter power between modes.

The scattering at the junction will then be described by an operator

$$S = \begin{bmatrix} S_{LL} & S_{LR} \\ S_{RL} & S_{RR} \end{bmatrix} \in \mathrm{End}\left((\mathscr{E}\oplus\mathscr{M})_L \oplus (\mathscr{E}\oplus\mathscr{M})_R\right) \tag{3.2.3}$$

that is to be constructed from the operators P, $P^\dagger$, Q and R.

At any junction in the guide there will be scattering of fields travelling towards the junction from both the left and the right, both across the junction and back from the junction. Choosing some ordering for the modes and expanding the fields on either side of the junction in terms of the orthonormal TE and TM modes, the TE fields are spanned by a basis $\{\boldsymbol{e}_{E_n}, \boldsymbol{h}_{E_n}\}_{n\in\mathbb{N}}$ for $\mathscr{E}$ and the TM fields are spanned by a basis $\{\boldsymbol{e}_{M_m}, \boldsymbol{h}_{M_n}\}_{n\in\mathbb{N}}$ for $\mathscr{M}$.

Denote the μ^{th} electric and magnetic modes by $\boldsymbol{e}_\mu$ and $\boldsymbol{h}_\mu$ respectively. Denote the complex coefficients of the modes in $\mathcal{H}_L$ by A_μ and by B_μ, and those in $\mathcal{H}_R$ by C_μ and D_μ. Then all electric and magnetic fields to the left and right of the junction are superpositions of modes travelling to the left and to the right, and take the form

$$\begin{bmatrix} \boldsymbol{E}_L \\ \boldsymbol{H}_L \end{bmatrix} = \sum_{\mu\in\mathbb{N}} \left[A_\mu e^{jk_\mu^L z} \pm B_\mu e^{-jk_\mu^L z}\right] \begin{bmatrix} \boldsymbol{e}_\mu^L \\ \boldsymbol{h}_\mu^L \end{bmatrix}, \tag{3.2.4-A}$$

$$\begin{bmatrix} \boldsymbol{E}_R \\ \boldsymbol{H}_R \end{bmatrix} = \sum_{\mu\in\mathbb{N}} \left[D_\mu e^{jk_\mu^R z} \pm C_\mu e^{-jk_\mu^R z}\right] \begin{bmatrix} \boldsymbol{e}_\mu^R \\ \boldsymbol{h}_\mu^R \end{bmatrix}, \tag{3.2.4-B}$$

with $+z$ being the positive propagation direction, and k^L and k^R denoting the propagation constants in the waveguide to the left and right of the junction. The expansion coefficient vectors A, B, C and D will, in general, be complex. That these fields are also solutions to Helmholtz equation is immediate from the linearity of the operator $\boldsymbol{\nabla}_T^2 + k^2$ and its independence of the axial coordinate, z.

With reference to the general description of the fields given in equation (3.2.4), consider the magnetic fields to the left and right of the boundary. For the μ^{th} electric field mode $\boldsymbol{e}_\mu^L \in \mathscr{E}_L$, continuity of the magnetic fields across the junction gives equality between the time averaged power at the junction from either side:

$$\int_{S_L} \left[\boldsymbol{e}_\mu^L \times \sum_{\nu\in\mathbb{N}} (\bar{A}_\nu - \bar{B}_\nu)\bar{\boldsymbol{h}}_\nu^L\right] \cdot \mathrm{d}\boldsymbol{S} = \sum_{\nu\in\mathbb{N}} (\bar{A}_\nu - \bar{B}_\nu) \int_{S_L} \left[\boldsymbol{e}_\mu^L \times \bar{\boldsymbol{h}}_\nu^L\right] \cdot \mathrm{d}\boldsymbol{S}$$

$$
\begin{aligned}
&\stackrel{\text{def}}{=} \sum_{\nu\in\mathbb{N}} (\bar{A}_\nu - \bar{B}_\nu) R_{\mu\nu} && \text{(3.2.5-A)}\\
&= \int_{S_L} \left[\boldsymbol{e}^L_\mu \times \sum_{\kappa\in\mathbb{N}} (\bar{D}_\kappa - \bar{C}_\kappa) \bar{\boldsymbol{h}}^R_\kappa \right] \cdot \mathrm{d}\boldsymbol{S}\\
&= \sum_{\kappa\in\mathbb{N}} (\bar{D}_\kappa - \bar{C}_\kappa) \int_{S_L} \left[\boldsymbol{e}^L_\mu \times \bar{\boldsymbol{h}}^R_\kappa \right] \cdot \mathrm{d}\boldsymbol{S}\\
&\stackrel{\text{def}}{=} \sum_{\kappa\in\mathbb{N}} (\bar{D}_\kappa - \bar{C}_\kappa) P_{\mu\kappa} && \text{(3.2.5-B)}
\end{aligned}
$$

for all $\mu \in \mathbb{N}$. Here the over-bar denotes complex conjugation, and the integrals involving $\boldsymbol{h}^R$ should strictly be written as integrals of $\boldsymbol{e}^L \times \pi^* \boldsymbol{h}^R$ over S_L, but the natural identification of $\pi(S_L) \subset S_R$ with S_L and the form of π^* in equation (3.2.1) results in the given form. The form of π^* has the physical interpretation that, for a perfectly conducting waveguide wall, the tangential component of the electric field is zero on the junction flange. If the guide wall is not a perfect conductor equations 3.2.5 will not determine the scattering amplitudes and the formalism breaks down.

In like manner, continuity of the transverse fields across the junction gives

$$
\begin{aligned}
\int_{S_L} \left[\sum_{\mu\in\mathbb{N}} (A_\mu + B_\mu)\, \boldsymbol{e}^L_\mu \times \bar{\boldsymbol{h}}^R_\kappa \right] \cdot \mathrm{d}\boldsymbol{S} &= \sum_{\mu\in\mathbb{N}} (A_\mu + B_\mu) \int_{S_L} \left[\boldsymbol{e}^L_\mu \times \bar{\boldsymbol{h}}^R_\kappa \right] \cdot \mathrm{d}\boldsymbol{S}\\
&\stackrel{\text{def}}{=} \sum_{\mu\in\mathbb{N}} P_{\mu\kappa} (A_\mu + B_\mu) && \text{(3.2.6-A)}\\
&= \int_{S_L} \left[\sum_{\nu\in\mathbb{N}} (D_\nu + C_\nu)\, \boldsymbol{e}^R_\nu \times \bar{\boldsymbol{h}}^R_\kappa \right] \cdot \mathrm{d}\boldsymbol{S}\\
&= \sum_{\nu\in\mathbb{N}} (D_\nu + C_\nu) \int_{S_L} \left[\boldsymbol{e}^R_\nu \times \bar{\boldsymbol{h}}^R_\kappa \right] \cdot \mathrm{d}\boldsymbol{S}\\
&\stackrel{\text{def}}{=} \sum_{\nu\in\mathbb{N}} Q_{\nu\kappa} (D_\nu + C_\nu) && \text{(3.2.6-B)}
\end{aligned}
$$

It follows from equations (3.2.5) and (3.2.6) that the field coefficient vectors $A \pm B = [A_\mu \pm B_\mu] \in \mathcal{H}_L$ and $D \pm C = [D_\mu \pm C_\mu] \in \mathcal{H}_R$ are related by the operator matrix equations

$$P(A + B) = Q(D + C) \qquad \text{(3.2.7-A)}$$

$$\bar{R}(A - B) = P^\dagger (D - C) \qquad \text{(3.2.7-B)}$$

Here the adjoints of equations (3.2.5-A) and (3.2.5-B) have been taken. The scattering problem is to solve these simultaneous equations for the elements of the vectors

A, B, C and D. The situation is illustrated by the diagrams

$$\begin{tikzcd} A \arrow[r, "P"] \arrow[d, "\bar{R}"'] & D \\ B & C \arrow[l, "P^\dagger"] \arrow[u, "Q"'] \end{tikzcd} \qquad \begin{tikzcd} \mathcal{H}_L \arrow[r, bend left, "P"] \arrow[loop below, "\bar{R}"] & \mathcal{H}_R \arrow[l, bend left, "P^\dagger"] \arrow[loop below, "Q"] \end{tikzcd} . \tag{3.2.8}$$

From (3.2.3) the scattering operator for the junction is the matrix operator

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} : \mathcal{H}_L \oplus \mathcal{H}_R \longrightarrow \mathcal{H}_L \oplus \mathcal{H}_R \tag{3.2.9}$$

with entries that are operator matrices $S_{11} \in \operatorname{End}(\mathcal{H}_L)$, $S_{12} \in \mathscr{L}(\mathcal{H}_R, \mathcal{H}_L)$, $S_{21} \in \mathscr{L}(\mathcal{H}_L, \mathcal{H}_R)$ and $S_{22} \in \operatorname{End}(\mathcal{H}_R)$. (Use of indices 1 and 2 rather than L and R is to conform to standard usage.) Clearly, the power of the input and the scattered fields, measured by the L^2 norm, are bounded, and $\|S_{ij}F\| \leq \|S_{ij}\| \cdot \|F\| \leq \|S\| \cdot \|F\|$, so the operators are all contractions (see [23, 74] for definition). In an appropriate norm, $\|S\|$ would represent the total scattered time-averaged power, and conservation of power would require $\|S\| = 1$.

The problem at hand is to find the scattering operator S for the junction in terms of P, Q and R from equations (3.2.7). The values of the operator elements p_{ij}, q_{ij} and r_{ij} will then be derived from the particular representations of the bases for $\mathcal{H}_L$ and $\mathcal{H}_R$ using equations (3.2.5) and (3.2.6). Formal manipulation of the simultaneous equations (3.2.7) to eliminate output B gives

$$\begin{aligned} &2PA = \left[\left(P\bar{R}^{-1}P^\dagger + Q\right) D - \left(P\bar{R}^{-1}P^\dagger - Q\right) C\right] \\ \Longrightarrow\ &D = 2\left[\left(P\bar{R}^{-1}P^\dagger + Q\right)^{-1} P\right] A - \left[\left(P\bar{R}^{-1}P^\dagger + Q\right)^{-1} \left(Q - P\bar{R}^{-1}P^\dagger\right)\right] C. \end{aligned}$$

Similarly, elimination of the output D gives

$$B = 2\left[\left(\bar{R} + P^\dagger Q^{-1} P\right)^{-1} P^\dagger\right] C + \left[\left(\bar{R} + P^\dagger Q^{-1} P\right)^{-1} \left(\bar{R} - P^\dagger Q^{-1} P\right)\right] A.$$

Then, since S maps the inflowing fields $A \in \mathcal{H}_L$ and $C \in \mathcal{H}_R$ to the outflowing fields $B \in \mathcal{H}_L$ and $D \in \mathcal{H}_R$, the entries in the scattering operator (3.2.9) for the process (3.2.7) are

$$S_{11} = [\bar{R} + P^\dagger Q^{-1} P]^{-1} [\bar{R} - P^\dagger Q^{-1} P] \tag{3.2.10-A}$$

$$S_{12} = 2[\bar{R} + P^\dagger Q^{-1} P]^{-1} P^\dagger \tag{3.2.10-B}$$

$$S_{21} = 2[Q + P\bar{R}^{-1}P^{\dagger}]^{-1}P \tag{3.2.10-C}$$

$$S_{22} = -[Q + P\bar{R}^{-1}P^{\dagger}]^{-1}[Q - P\bar{R}^{-1}P^{\dagger}] \tag{3.2.10-D}$$

giving, as the scattering operator matrix for a single junction with inputs A and C (power flow towards the junction from left and right) and outputs B and D (power flow away from the junction)

$$\begin{bmatrix} B \\ D \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} A \\ C \end{bmatrix}. \tag{3.2.11}$$

Given two adjacent junctions in a waveguide separating sections of guide with functions spaces $\mathcal{H}_1$, $\mathcal{H}_2$ and $\mathcal{H}_3$, the total scattering operator must be a product of the two separate junction S-matrices in some appropriate sense. The situation is represented in the following scattering diagram:

$$\begin{tikzcd}
A \arrow[r, "A_{21}"] \arrow[d, "A_{11}"'] & D \arrow[r, "B_{32}"] \arrow[d, shift left, "B_{22}"] & E & & A \arrow[r, "S_{31}"] \arrow[d, "S_{11}"'] & E \\
B & C \arrow[l, "A_{12}"] \arrow[u, shift left, "A_{22}"] & F \arrow[l, "B_{23}"] \arrow[u, "B_{33}"'] & & B & F \arrow[l, "S_{13}"] \arrow[u, "S_{33}"']
\arrow[from=1-3, to=2-5, phantom, "="]
\end{tikzcd} \tag{3.2.12}$$

The product will be written $B \odot A = S$ with $A \in \operatorname{End}(\mathcal{H}_1 \oplus \mathcal{H}_2)$, $B \in \operatorname{End}(\mathcal{H}_2 \oplus \mathcal{H}_3)$ and $S \in \operatorname{End}(\mathcal{H}_1 \oplus \mathcal{H}_3)$. This is a simultaneous equation derived from

$$\begin{bmatrix} B \\ D \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} A \\ C \end{bmatrix}, \quad \begin{bmatrix} C \\ E \end{bmatrix} = \begin{bmatrix} B_{22} & B_{23} \\ B_{32} & B_{33} \end{bmatrix} \begin{bmatrix} D \\ F \end{bmatrix}$$

The algebraic approach to the solution is the following: Eliminating C and D from $D = A_{21}A + A_{22}C$ and $C = B_{22}D + B_{23}F$ and solving for B in terms of the left side and right side inputs A and F gives

$$\begin{aligned} C &= B_{22}\left(A_{21}A + A_{22}C\right) + B_{23}F \\ &= (I - B_{22}A_{22})^{-1}\left[B_{22}A_{21}A + B_{23}F\right], \\ \implies B &= A_{11}A + A_{12}C \\ &= \left[A_{11} + A_{12}(I - B_{22}A_{22})^{-1}B_{22}A_{21}\right]A + \left[A_{12}(I - B_{22}A_{22})^{-1}B_{23}\right]F. \end{aligned}$$

Likewise, solving for E in terms of the inputs A and F gives

$$E = \left[B_{32}(I - A_{22}B_{22})^{-1}A_{21}\right]A + \left[B_{33} + B_{32}(I - A_{22}B_{22})^{-1}A_{22}B_{23}\right]F.$$

Thus, the component operators of the cascaded scattering operator matrices are

$$S_{11} = A_{11} + A_{12}(I - B_{22}A_{22})^{-1}B_{22}A_{21} \tag{3.2.13-A}$$

$$S_{13} = A_{12}(I - B_{22}A_{22})^{-1}B_{23} \tag{3.2.13-B}$$

$$S_{31} = B_{32}(I - A_{22}B_{22})^{-1}A_{21} \tag{3.2.13-C}$$

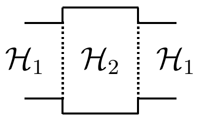
$$S_{33} = B_{33} + B_{32}(I - A_{22}B_{22})^{-1}A_{22}B_{23} \tag{3.2.13-D}$$

These equations represent the scattering of power between the left-most domain, $\mathcal{H}_1$, and the right-most domain $\mathcal{H}_3$, $S_{11} \in \operatorname{End}(\mathcal{H}_1)$, $S_{13} \in \mathscr{L}(\mathcal{H}_3, \mathcal{H}_1)$, $S_{31} \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_3)$, and $S_{33} \in \operatorname{End}(\mathcal{H}_3)$; the product is clearly associative. It cannot be commutative except in the situation where $S_1 = S_3$. That situation does occur in the case of a parallel corrugated waveguide. In that situation the waveguide is a concatenation of symmetric units comprising pairs of junctions, except, possibly, for the addition of a single junction on one end, and for the system of symmetric units the scattering is symmetric and it is immediate that the operator reduces to

$$S_{11} = S_{22} = A_{11} + A_{12}A_{22}(I - A_{22}^2)^{-1}A_{21} \tag{3.2.14-A}$$

$$S_{12} = S_{21} = A_{12}(I - A_{22}^2)^{-1}A_{21} \tag{3.2.14-B}$$

which is illustrated by the following diagram of a basic unit from which the corrugated waveguide is constructed, and the associated scattering diagram for the scattering operator across this unit, $S \in \operatorname{End}(\mathcal{H}_1 \oplus \mathcal{H}_1)$:



$$\begin{tikzcd}
\bullet \arrow[r, "A_{21}"] \arrow[d, "A_{11}"'] & \bullet \arrow[r, "A_{12}"] \arrow[d, shift right, "A_{22}"'] & \bullet \\
\bullet & \bullet \arrow[l, "A_{12}"] \arrow[u, shift right, "A_{22}"'] & \bullet \arrow[l, "A_{21}"] \arrow[u, "A_{11}"']
\end{tikzcd}
=
\begin{tikzcd}
\bullet \arrow[r, "S_{21}"] \arrow[d, "S_{11}"'] & \bullet \\
\bullet & \bullet \arrow[l, "S_{12}"] \arrow[u, "S_{22}"']
\end{tikzcd} \tag{3.2.15}$$

Once the appropriate phase slippage has been included into the scattering operators, this will represent the scattering through a basic unit of the guide. It means that, in calculating the propagation through a corrugated waveguide, or any section of a horn that forms a corrugated waveguide, the amount of calculation needed is almost halved. The exploitation of this fact, and of the algebraic properties of these operators, is part of the key to efficient and accurate numerical simulation that will be addressed in section 3.6.

The system of equations (3.2.13) can be arrived at formally by chasing around the diagram (3.2.12) following all possible paths that the diagram permits – follow the arrows in the direction that they point. To see this, write

$$(I - B_{22}A_{22})^{-1} = I + B_{22}A_{22} + (B_{22}A_{22})^2 + \cdots\cdots$$

which describes the scattering of power back and forth across the junction. Thus, with reference to diagram 5.3.2-B, to get from A to B sum A_{11} and all paths $A_{12}(B_{22}A_{22})^n B_{22}A_{21}$ giving $S_{11} = A_{11} + A_{12}\sum_{n=0}^{\infty}(B_{22}A_{22})^n B_{22}A_{21} = A_{11} + A_{12}(I - B_{22}A_{22})^{-1}B_{22}A_{21}$. Likewise the formal structure of all S_{NM} for all scattering products for the concatenation of any number of junctions are found from the diagrams. Diagram chasing works for a K-fold junction: at the common space $\mathcal{H}_2$ there are K arrows with one arrow pointing down and $K-1$ pointing up; scattering from any input to any output must follow all paths taking the one downward arrow and any possible upward arrow. For a T-junction the diagram gives scattering $(B_{22}A_{22})^n(C_{22}A_{22})^m$ within the space common at the T, hence $S_{11} = A_{11} + A_{12}(I - B_{22}A_{22})^{-1}(I - C_{22}A_{22})^{-1}(C_{22} + B_{22})A_{21}$, and so on for the mappings S_{13}, S_{14}, S_{34} etc. The only technical point is to remember the exact equivalence of all spaces when drawing the diagrams; therefore $K-1$ diagrams are needed. For a horn $K = 2$ and the equivalence is trivial. This does not address the problem of matching the modes at the common junction to solve for the scattering coefficients, but it does give the structure of the operator on $\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n$.

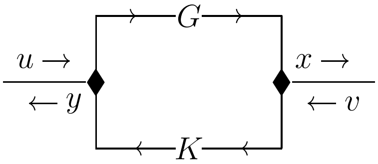
Two points have been glossed over. The first is the rather obvious point that since the scattering is scattering of power, conservation of power must be observed. In terms of the norms of the operator that means that the S-matrix of a length of guide would be of norm $\|S\| = 1$ if there were no evanescent modes present or the guide was of zero length. The S-matrix of a single junction must, therefore, be of norm 1 as an endomorphism of $\mathcal{H}_L \oplus \mathcal{H}_R$. But if that endomorphism arises as a true step discontinuity there will be evanescent modes induced and when propagation takes place there is amplitude decay because the Helmholtz equation becomes a diffusion equation of second order in the axial coordinate for any mode with imaginary propagation coefficient. The second point is that for the scattering product to be defined the terms of the form $(I - B_{22}A_{22})^{-1}$ to be defined the operators must satisfy

$\|B_{22}A_{22}\| < 1$. But that is immediate since all operators A_{22} and B_{22} must both be strict contractions except where there is perfect reflection – a short – and in that case $P \equiv 0$ in equations (3.2.10) and the entire process becomes trivial: $S \equiv I$.

3.3 An alternative description of scattering

The form of the component operators in the scattering operator given in equations (3.2.10) is formally correct for any pair of waveguide sections S_L and S_R for which the corresponding spaces of fields admit denumerable bases. They are not, however, expressed in a way that will lead to efficient computation. There are several forms in which equations (3.2.10) can be rewritten, but the one given here seems particularly suitable for computational purposes.

The motivation comes from the following diagram that is common in the elimentary theory of control for a single input, single output system with feedback:



What follows is not a true presentation of a control problem, it is simply the presentation of the idea that gave rise to the alternative form that scattering operators given in equations (3.3.1), (3.3.2) and (3.3.3) that is the basis for the numerical computation in chapter 5.

Let G represent the 'transfer', or throughput of some kind of system with input u from the left. The output in the absence of a 'control', or feedback, K is $x = Gu$, so the input to the system with control is $u - Kx$. Then formal manipulation gives

$$x = G(u - Kx) = G(I + KG)^{-1}u \implies 2x = 2(I + GK)^{-1}Gu \overset{\text{def}}{=} S_{21}u,$$

and define $S_{11}u \overset{\text{def}}{=} u - 2Kx = (I+KG)^{-1}(I-KG)u$. Imagine G and K are rectified, then for input v from the right the rôles of G and K are reversed and set $y = Kv$. Then $S_{12}v = 2(I + KG)^{-1}Kv$ and $S_{22}v = (I + GK)^{-1}(GK - I)$ and note that $S_{12} = 2(I+KG)^{-1}K = (I+KG)^{-1}(I+KG+I-KG)K = (I+S_{11})K$. Rearranging the equations gives various representations of the component S-matrices, amongst which are

$$S_{11} = (I + KG)^{-1}(I - KG) = I - S_{12}G \tag{3.3.1-A}$$

$$S_{12} = 2(I + KG)^{-1}K = (I + S_{11})K \tag{3.3.1-B}$$

$$S_{21} = 2(I + GK)^{-1}G = (I - S_{22})G \tag{3.3.1-C}$$

$$S_{22} = (I + GK)^{-1}(GK - I) = S_{21}K - I. \tag{3.3.1-D}$$

As a result, and after re-expressing S_{21} and S_{22}, the S-matrix for a junction can always be written in terms of S_{11}, I, G and K:

$$S = \begin{bmatrix} S_{11} & (I + S_{11})K \\ G(I + S_{11}) & G(I + S_{11})K - I \end{bmatrix}. \tag{3.3.2}$$

Here it is assumed that, in the finite sized approximation to S being used for computation of the scattering, S_{11} is $N \times N$ and S_{22} is $M \times M$ with $N \leq M$; was $N > M$, then an equivalent rearrangement can be made in which the rôles of S_{11} and S_{22} are reversed, so minimising the total size of the computation task in either case.

To exploit the computational simplicity that results from the scheme above the waveguide junction scattering operators need to be presented in this form: Define the operators $G = Q^{-1}P : \mathcal{H}_L \to \mathcal{H}_R$ as a 'transfer' operator and $K = \bar{R}^{-1}P^\dagger = (PR^{-1})^\dagger : \mathcal{H}_R \to \mathcal{H}_L$ as the 'control' operator. From equations (3.2.10) write S_{11} as

$$\begin{aligned} S_{11} &= [\bar{R} + P^\dagger Q^{-1}P]^{-1}[\bar{R} - P^\dagger Q^{-1}P] \\ &= [I + \bar{R}^{-1}P^\dagger Q^{-1}P]^{-1}\bar{R}^{-1}\bar{R}[I - \bar{R}^{-1}P^\dagger Q^{-1}P] \\ &= (I + KG)^{-1}(I - KG) \end{aligned} \tag{3.3.3}$$

and the equations (3.3.1) and matrix (3.3.2) follow. Alternatively all of these equations are derived directly from equations (3.2.7). The particulars of K and G depend upon the waveguide cross-section geometry, the boundary conditions and the dielectric properties of any filling, and that will dictate the particulars of any scheme for computing the modes and scattering products, but since these are completely general expressions it does not matter what the guide section may be – the form of the scattering operators can always be expressed in this form.

There are several advantages to computing junction scattering in this way. It is in the form of equations (3.3.2) that the matrices representing the scattering operators in a finite model of the systems will be computed. The computational scheme will will be described in some detail in chapter 5.

3.4 $S^2 \equiv I$ and $S \odot S \equiv I$ at a junction

It has already been observed at the end of section 3.2 that, at a junction, S must be of norm 1. Here it will be shown that there is are formal algebraic conditions $S^2 \equiv I$ and $S \odot S \equiv I$ which, along with $\|S\|_F = 1$, give potentially valuable tests on the accuracy of computation over and above the numerical run-time conditioning tests that are discussed in chapter 5: at a junction the scattering operator must satisfy

$$\begin{aligned} S_{11}^2 + S_{12}S_{21} &= I_L, \quad & S_{22}^2 + S_{21}S_{12} &= I_R, \\ S_{11}S_{12} + S_{12}S_{22} &= 0, \quad & S_{21}S_{11} + S_{22}S_{21} &= 0. \end{aligned} \tag{3.4.1}$$

For a finite size numerical model none of these conditions can hold exactly, but they must hold to high accuracy if the model is to be accurate.

From the definition $S^{-1}S = I$, straightforward manipulation of the operator sub-matrices S_{ij} gives the formulæ

$$(S^{-1})_{11} = (S_{11} - S_{12}S_{22}^{-1}S_{21})^{-1} \tag{3.4.2-A}$$

$$(S^{-1})_{12} = -(S_{11} - S_{12}S_{22}^{-1}S_{21})^{-1}(S_{12}S_{22}^{-1}) \tag{3.4.2-B}$$

$$(S^{-1})_{21} = -(S_{22}^{-1}S_{21})(S_{11} - S_{12}S_{22}^{-1}S_{21})^{-1} \tag{3.4.2-C}$$

$$(S^{-1})_{22} = S_{22}^{-1} - (S_{22}^{-1}S_{21})(S_{11} - S_{12}S_{22}^{-1}S_{21})^{-1}(S_{12}S_{22}^{-1}) \tag{3.4.2-D}$$

where $(S^{-1})_{ij}$ refers to the ij^{th} block matrix of the operator matrix S^{-1}. With reference to equations (3.3.1) we have $S_{11}^{-1} = (I - KG)^{-1}(I + KG)$ and can substitute for all S_{ij} into equation (3.4.2-A) in terms of K and G. Formal power series manipulation gives $(I \pm XY)^{\pm 1}X = X(I \pm YX)^{\pm 1}$ and $(I \pm XY)^{\pm 1}(I \mp XY) = (I \mp XY)(I \pm XY)^{\pm 1}$, from which equation (3.4.2-A) gives

$$(S^{-1})_{11} = \left[(I + KG)^{-1}(I - KG) + 4(I + KG)^{-1}K(I - GK)^{-1}G\right]^{-1}$$

$$= \left[(I - KG)^2 + 4KG\right]^{-1} (I - KG)(I + KG) = S_{11}.$$

Substituting this result into (3.4.2-B to D) gives the identities

$$(S^{-1})_{12} = S_{12}, \quad (S^{-1})_{21} = S_{21} \text{ and } (S^{-1})_{22} = S_{22}.$$

Thus, $S^{-1} = S$ and $S^2 = I$, giving the result that the scattering operator S at a junction is a square root of the identity on $\mathcal{H}_L \oplus \mathcal{H}_R$.

S_{11} and S_{22} are morphisms of $\mathcal{H}_L$ and $\mathcal{H}_R$ respectively so that there is no difficulty over the domain and codomain when they are interpreted as component operators of S^{-1}. Likewise both S_{12} and $(S^{-1})_{12}$ are mappings $\mathcal{H}_R \to \mathcal{H}_L$ and S_{21} and $(S^{-1})_{21}$ are mappings $\mathcal{H}_L \to \mathcal{H}_R$. Thus there is no algebraic difficulty with these equations; furthermore, substitution of S for both A and B into equations (3.2.13) and use of equations (3.4.1) shows that $S \odot S = I$, so this is not just a matrix inverse, but a scattering inverse equation. Nevertheless, it has *no* physical meaning at all: there is no scattering diagram from which these equations follow. A moment's consideration of the mechanics of the junction will show that this equation can relate to no physical junction except the trivial junction where the step size is zero – the join between two waveguide sections of identical cross section where the components of S reduce to $S_{11} = S_{22} = 0$ and $S_{12} = S_{21} = I$. The relations (3.4.1) and (3.4.2) are purely algebraic, they say nothing about the physics of scattering at a junction; there is no physical process inverting a scattering process. None of the operators $(S^{-1})_{IJ}$, nor the inverses S_{IJ}^{-1} of the component operators S_{IJ} correspond to physical processes in the sense that it is not possible build a structure that will generate the inverse scattering.

3.5 Phase slippage

It is immediate from the scattering product equations (3.2.13) that the scattering processes of phase slippage along a waveguide section followed by scattering at a junction, and the scattering at a junction followed by phase slippage, are represented by the matrix products

$$S \odot V = \begin{bmatrix} V_1 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} V_1 & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} V_1 S_{11} V_1 & V_1 S_{12} \\ S_{21} V_1 & S_{22} \end{bmatrix}, \tag{3.5.1-A}$$

$$V \odot S = \begin{bmatrix} I & 0 \\ 0 & V_2 \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & V_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} V_2 \\ V_2 S_{21} & V_2 S_{22} V_2 \end{bmatrix}, \tag{3.5.1-B}$$

where V_1 and V_2 are diagonal matrices that are functions of the section length, d, cross sectional geometry and impedance of the guide taking the form

$$V = \begin{bmatrix} V_E & 0 \\ 0 & V_M \end{bmatrix}, \quad (V_E)_{ij} = \delta_{ij} \exp\left(-id\kappa_{E,j}\right), \quad (V_M)_{ij} = \delta_{ij} \exp\left(-id\kappa_{M,j}\right). \tag{3.5.2}$$

Since S is representable as an operator matrix in $M_2[\mathcal{H}_1 \oplus \mathcal{H}_2] = M_2[(\mathscr{E} \oplus \mathscr{M})_1 \oplus (\mathscr{E} \oplus \mathscr{M})_2]$, the components of each of the four operator sub-matrices $V_I \odot S_{IJ} \odot V_J$, for $I,\ J \in \{1, 2\}$, are given by

$$[V_I \odot S_{IJ} \odot V_J]_{mn} = \exp\left(-id_I \kappa_{In}\right) \exp\left(-id_J \kappa_{Jm}\right) (S_{IJ})_{mn}, \tag{3.5.3}$$

where d_I and d_J are the section lengths and κ_{In} and κ_{Jm} are the propagation coefficients of the appropriate waveguide section at either side of the junction. The indices I and J label the codomain and domain respectively from the set $\{\mathscr{E}_1,\ \mathscr{M}_1,\ \mathscr{E}_2,\ \mathscr{M}_2\}$.

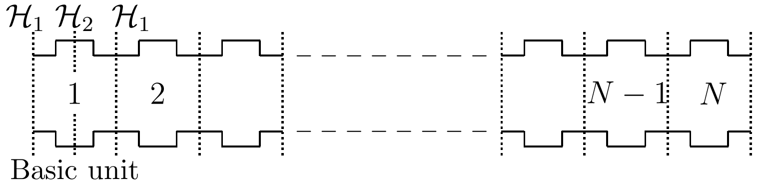
It is immediately apparent that the diagonal phase slippage operators and their finite dimensional matrix approximations for an N mode model lie in a copy of $\mathbb{C}^N$ in the space of operators and matrices. If, in the model there are K propagating and $N - K$ evanescent modes, the phase slippage of propagating modes occupy a Cartesian product of K copies of the unit circle (a K-torus) while the evanescent mode slippage coefficients lie in an $(N-K)$-fold Cartesian product $(0, 1] \times \cdots \times (0, 1]$. This gives a mapping

$$(\mathbb{R}, +) \to (V, \cdot), \quad d \mapsto \operatorname{diag}\{e^{-id\kappa_{E1}}, \cdots, e^{-id\kappa_{Mn}}\},$$

from the reals to the diagonal phase slippage matrices *in any one section*. It follows that each section length can be split into sub-lengths for the purposes of propagation, with the computational convenience of negative length waveguide sections allowed.

3.6 Propagation in parallel corrugated waveguides

Equations (3.2.14) of section 3.2 used the symmetry in an elementary section of a parallel corrugated waveguide to give the scattering, $\mathcal{H}_1 \oplus \mathcal{H}_1 \to \mathcal{H}_1 \oplus \mathcal{H}_1$, across the basic unit illustrated here:



The scattering between the two ends of the entire guide of N basic units is the $(N-1)$-fold scattering power of the right-hand side of equations (3.2.14).

Let d_1 be the half length of the narrow section in the basic unit and d_2 be the half length of the wide section, so that the total length of the basic unit is $2(d_1+d_2)$. Let V_1 and V_2 be the phase slippage matrix images of d_1 and d_2 and junction scattering be denoted A_{ij}; then the combination of the junction scattering and the phase slippage along the half sections gives the scattering for the entire basic unit as

$$\begin{aligned} S_{11} &= V_1 \left[A_{11} + \frac{1}{2} A_{12} V_2^2 A_{22} V_2^2 \left((I - A_{22}V_2^2)^{-1} + (I + A_{22}V_2^2)^{-1} \right) A_{21} \right] V_1 \\ &= V_1 \left[A_{11} + A_{12} V_2^2 B_{22} (I - B_{22}^2)^{-1} A_{21} \right] V_1, \qquad (3.6.1\text{-A}) \\ S_{12} &= \frac{1}{2} V_1 \left[A_{12} V_2^2 \big((I - A_{22}V_2^2)^{-1} + (I + A_{22}V_2^2)^{-1} \big) A_{21} \right] V_1 \\ &= V_1 A_{12} V_2^2 (I - B_{22}^2)^{-1} A_{21} V_1, \qquad (3.6.1\text{-B}) \end{aligned}$$

where $B_{22} \overset{\text{def}}{=} A_{22}V_2^2$. The total scattering $\mathcal{H}_1 \oplus \mathcal{H}_1 \to \mathcal{H}_1 \oplus \mathcal{H}_1$ along all N basic units is $(N-1)$-fold product $S \odot S \odot \cdots \odot S$ which, since N can be written as $N = 2^{n_1} + 2^{n_2} + \cdots 2^{n_k}(+1)$ for suitable constants n_j, will factor into powers and products of S. For example, a waveguide segment of $1024 = 2^{10}$ basic units requires ten scattering products: form $S \odot S$, square to get $(S \odot S) \odot (S \odot S)$, then keep squaring successive results; one of $121 = 2^6 + 2^5 + 2^4 + 2^3 + 1$ units requires nine products once S has been formed, while a guide comprising $81 = 2^6 + 2^4 + 1$ basic units requires five products. The cost in scattering products follows from the binary representation of the number of basic units. The computational cost is in the extra storage needed to temporarily retain some partial results when the number of units is not a power of two.

In addition to great efficiency this decomposition of a parallel waveguide section into concatenations of a basic unit leads to improved numerical accuracy over brute force modelling because the accumulation of numerical errors is greatly reduced in long sections. This matter will be addressed in chapter 5. In the above example of a waveguide of 1024 base units the total computational effort is the following: form A_{11} and A_{12} for a single junction then form S_{11} and S_{12} as in equations (3.6.1), then perform the ten scattering products as in equation (3.2.14) – a total of twenty two scattering products to describe the entire waveguide. If the same system was modelled naïvely there would be four scattering operations at each of the 2048 junctions, plus four scattering operations at 2048 sections between adjacent junctions, a total of 16384 scattering operations. The method described is approximately 745 times faster than a simple approach given the same scattering product algorithm; for the 121 unit guide the speed-up would be of the order 40. Coding the process is relatively complex, but the efficiency and the accuracy resulting form the great reduction in error accumulation makes the effort worth while.

3.7 The transmission operators at a junction

The scattering operator $S : \mathcal{H}_L \oplus \mathcal{H}_R \to \mathcal{H}_L \oplus \mathcal{H}_R$ at a junction has associated with it a transmission operator $T : \mathcal{H}_L \oplus \mathcal{H}_L \to \mathcal{H}_R \oplus \mathcal{H}_R$ and, if the system of equations is to be algebraically consistent, each operator must determine the other. Here the physical interpretation of $\mathcal{H}_X \oplus \mathcal{H}_X$ is of fields travelling in both directions in a waveguide section, while the mathematical interpretation is simply as the direct sum of two identical copies of the space of transverse fields supported by the guide. It is the scattering operator that is the primary and natural operator that describes the physical process because the process of field propagation in a waveguide is one of scattering simultaneously from the input at both ends to the output at both ends; the transmission operator is an algebraic object derived from it that can be constructed only from the components of the scattering operator.

The transmission operators seem to be appealing as objects in that they appear to give information about the transmission of fields through a waveguide structure that, it would be hoped, would be useful to model the 'transmission' from the bolometer cavity to the horn or waveguide aperture. Here it is argued that they are purely algebraic objects which have no real physical significance or rôle in waveguide analysis. The physically significant operators are the subcomponents of the scattering operators and, though it is possible to derive transmission operators and products from them, nothing is gained either in terms of understanding or in computational effort.

The diagrams associated with S and T, and with the fictitious component-wise inverses for S (as distinct from the components of the fictitious S^{-1} of section 3.4) are, respectively:

$$\begin{tikzcd}
A \arrow[r, "S_{21}"] \arrow[d, "S_{11}"'] & D \\
B & C \arrow[l, "S_{12}"] \arrow[u, "S_{22}"']
\end{tikzcd}
\qquad
\begin{tikzcd}
B \arrow[r, "S_{12}^{-1}"] \arrow[d, "S_{11}^{-1}"'] & C \\
A & D \arrow[l, "S_{21}^{-1}"] \arrow[u, "S_{22}^{-1}"']
\end{tikzcd}
\qquad
\begin{tikzcd}
A \arrow[r, "T_{11}"] \arrow[d, "T_{21}"'] & D \\
C & B \arrow[l, "T_{22}"] \arrow[u, "T_{12}"']
\end{tikzcd}$$

Given the transmission operator T, the action on the fields can be written as a transformation between leftward travelling and rightward travelling fields on either side of the junction, subsystem or system, to give

$$\begin{bmatrix} l_2 \\ r_2 \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} l_1 \\ r_1 \end{bmatrix} \tag{3.7.1}$$

Whereas the scattering operators map the inputs $A \oplus C$ onto the outputs $B \oplus D$, the transmission operators map the input-output on the left, $A \oplus B$, to the output-input on the right, $D \oplus C$, which are the input-output for the next transmission operator in the chain. Therefore the transmission operators for concatenated waveguide sections are modelled by simple operator matrix multiplication of the component operators.

Algebraic manipulation of equation (3.7.1) gives

$$\begin{aligned}
r_2 &= T_{21} l_1 + T_{22} r_1 \implies r_1 = -(T_{22}^{-1} T_{21}) l_1 + T_{22}^{-1} r_2, \\
l_2 &= T_{11} l_1 + T_{12} r_1 = (T_{11} - T_{12} T_{22}^{-1} T_{21}) l_1 + (T_{12} T_{22}^{-1}) r_2,
\end{aligned}$$

from which the components of the scattering operator, expressed in terms of the components of the transmission operator, are

$$S = \begin{bmatrix} -T_{22}^{-1}T_{21} & T_{22}^{-1} \\ T_{11} - T_{12}T_{22}^{-1}T_{21} & T_{12}T_{22}^{-1} \end{bmatrix} : A \oplus C \longrightarrow B \oplus D \tag{3.7.2}$$

This operator S, associated with T, can also be obtained by chasing around the diagrams above and using the second row of equation (3.4.1).

The process of obtaining S from T can be reversed to obtain T from S by algebraic manipulation. The resulting transmission matrix is

$$T = \begin{bmatrix} S_{21} - S_{22}S_{12}^{-1}S_{11} & S_{22}S_{12}^{-1} \\ -S_{12}^{-1}S_{11} & S_{12}^{-1} \end{bmatrix} : A \oplus B \to D \oplus C. \tag{3.7.3}$$

S cannot be obtained from T if T_{22} is singular, nor can T be obtained from S if S_{12} is singular. The expression for S in (3.7.2) is in terms of the sub-components of T^{-1}. Using equations (3.7.3) and (3.4.2) the inverse transmission operator can be written

$$T^{-1} = \begin{bmatrix} S_{21}^{-1} & -S_{21}^{-1}S_{22} \\ S_{11}S_{21}^{-1} & S_{12} - S_{11}S_{21}^{-1}S_{22} \end{bmatrix} \tag{3.7.4}$$

so that S_{21} must also be non-singular. Since T_{22}, S_{12} and S_{21} map fields on the left to fields on the right, for these operators to be non-singular the domain and co-domain must have the same dimension and the operator matrix representations cannot be row or column degenerate. Therefore, in the finite dimensional models, the number of modes must be the same on either side of a junction or section of horn or waveguide. It follows that a numerical implementation that varies the number of modes from section to section is not algebraically consistent.

If such an operator corresponded to a physical process in a waveguide it would be possible to build two waveguides structures, one for T and one for T^{-1}, and by joining them end to end produce a perfect transmission system that transferred the fields at one of the waveguide to the other end of the waveguide structure without any net change to the fields. That is an absurdity; in section 3.4 is was stated that S_{21}^{-1} is a purely algebraic object corresponding to no physical process, that is the why T is purely an algebraic object. Indeed, since $S_{12} = 2(I + KG)^{-1}K$ and $K = \bar{R}^{-1}P^{\dagger}$, for T to be physical would require K^{-1} to exist. That requires $P^{\dagger}$ to correspond to an invertible physical process, but P is a strict contraction so for $P^{\dagger}$ to be invertible power would have to be created.

With the reservation that the transmission operators are purely formal algebraic objects, the relations between the algebraic properties of scattering and transmission can be pursued a little further. Denote the the set of all scattering operators by $\mathcal{S}$ and the set of all transmission operators by $\mathcal{T}$. Both $(\mathcal{S}, \odot)$ and $(\mathcal{T}, \cdot)$ are both closed and the products are associative, the product on $\mathcal{T}$ being simply matrix multiplication when the transmission operators are written as in (3.7.1).

Let $\sigma_\tau : \mathcal{T} \to \mathcal{S}$ denote the mapping that takes a transmission operator to the scattering operator associated given by equation (3.7.2), and denote the opposite mapping of equation (3.7.3) by $\tau_\sigma : \mathcal{S} \to \mathcal{T}$. A tedious but simple calculations shows that the scattering product $\mathcal{S} \times \mathcal{S} \to \mathcal{S}$, and the transmission product $\mathcal{T} \times \mathcal{T} \to \mathcal{T}$ are related by

$$(\sigma_\tau T_2) \odot (\sigma_\tau T_1) = \sigma_\tau (T_2 \cdot T_1), \tag{3.7.5-A}$$

$$\tau_\sigma S_2 \cdot \tau_\sigma S_1 = \tau_\sigma (S_2 \odot S_1), \tag{3.7.5-B}$$

for transmission operators T_1 and T_2 and scattering operators S_1 and S_2. Furthermore, the composite operators $\sigma_\tau \circ \tau_\sigma = 1_{\mathcal{S}}$, the identity operator on the set $\mathcal{S}$, and $\tau_\sigma \circ \sigma_\tau = 1_{\mathcal{T}}$, the identity operator on the set $\mathcal{T}$. The scattering identity operator has already been noted in section 3.4 to have $S_{11} = S_{22} = 0$ and $S_{21} = S_{12} = I$, but corresponds only to the trivial junction, and its transmission image is simply the identity operator matrix. Thus it is seen that the mappings σ_τ and τ_σ are homomorphisms of sets that respect the products, but nothing more.

On the matter of computational complexity it should be observed that there is no gain to be derived from avoiding the calculation of the scattering products by calculating the transmission operator from pairs of scattering operators and performing a matrix product. The net computational cost is slightly higher than working entirely within the framework of the scattering operators.

Chapter 4

Circular and rectangular waveguides

The motivation for the development of the particular approach to the scattering models and methods that are described in this thesis was the need to model the Planck horns in a reasonable time frame. Those horns were all of circular, simply connected, cross section. The equations for the fields and the scattering amplitudes are presented in this chapter in the form in which they were used in the code development. The mathematical formalism is that presented in chapter 3 and section 4.1 develops the equations for perfectly aligned sections from first principles. The scattering amplitude formulæ are presented in section 4.2 followed by a discussion of the radial dependence of the amplitudes in section 4.3. In section 4.4 the equations for the scattering amplitude in rectangular waveguides are presented, followed by a discussion of the breakdown of mode orthogonality due to finite wall conductivity in section 4.5. The final section looks at the scattering between modes and azimuthal orders that arises at imperfectly aligned junctions.

4.1 Modes in a circular cylindrical waveguide

The homogeneous Helmholtz equation for a transverse field, F, of wavenumber $k = \omega/c$, in cylindrical polar coordinates is

$$\left[\boldsymbol{\nabla}^2 + k^2\right] F = \left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial\varphi^2} + k^2\right] F = 0. \tag{4.1.1}$$

Ignoring time dependence that is assumed to be a Fourier component, $e^{j\omega t}$, of the time signal, rearranging and seeking a separable solution $F(r,\varphi,z) = R(r)\Phi(\varphi)Z(z)$ gives

$$\begin{aligned}\frac{\Phi Z}{r}\frac{\mathrm{d}}{\mathrm{d}r}\left(r\frac{\mathrm{d}R}{\mathrm{d}r}\right) + \frac{RZ}{r^2}\frac{\mathrm{d}^2\Phi}{\mathrm{d}\varphi^2} + R\Phi\frac{\mathrm{d}^2Z}{\mathrm{d}z^2} + k^2R\Phi Z &= 0\\ \Longrightarrow \frac{1}{rR}\frac{\mathrm{d}}{\mathrm{d}r}\left(r\frac{\mathrm{d}R}{\mathrm{d}r}\right) + \frac{1}{\Phi r^2}\frac{\mathrm{d}^2\Phi}{\mathrm{d}\varphi^2} + k^2 &= -\frac{1}{Z}\frac{\mathrm{d}^2Z}{\mathrm{d}z^2}.\end{aligned} \tag{4.1.2}$$

Independence of the left and right-hand sides of equation (4.1.2) gives

$$\frac{\mathrm{d}^2Z}{\mathrm{d}z^2} = -\gamma^2 Z, \quad \gamma \in \mathbb{C}, \tag{4.1.3}$$

whereupon, setting $\kappa^2 = k^2 - \gamma^2$ and rearranging gives

$$\frac{r}{R}\frac{\mathrm{d}}{\mathrm{d}r}\left(r\frac{\mathrm{d}R}{\mathrm{d}r}\right) + r^2\kappa^2 = -\frac{1}{\Phi}\frac{\mathrm{d}^2\Phi}{\mathrm{d}\varphi^2}. \tag{4.1.4}$$

Again, independence gives a solution

$$\frac{\mathrm{d}^2\Phi}{\mathrm{d}\varphi^2} = -n^2\Phi, \quad n \in \mathbb{C}\,. \tag{4.1.5}$$

Substituting from equation (4.1.5) into (4.1.4) gives the radial function $R(r)$ as any solution to Bessel's equation

$$r\frac{\mathrm{d}}{\mathrm{d}r}\left(r\frac{\mathrm{d}R}{\mathrm{d}r}\right) + (r^2\kappa^2 - n^2)R = 0. \tag{4.1.6}$$

Both equation (4.1.5) and (4.1.6) can be rearranged and put in self-adjoint Sturm-Liouville form [8],

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)\frac{\mathrm{d}u(x)}{\mathrm{d}x}\right] + \left[\lambda\rho(x) - q(x)\right]u(x) = 0,$$

and it is a characteristic of all such systems that the solutions can be expanded in terms of the eigenfunctions of the equation.

Equation (4.1.5) is defined on the interval $[-\pi, \pi]$ with periodic endpoint conditions, $p(\varphi) = \rho(\varphi) = 1$, $q(\varphi) = 0$ and $\lambda = n^2$ giving rise to Fourier series. Equation (4.1.6) with $p(r) = \rho(r) = r$, $q(r) = n^2/r$ and $\lambda = \kappa^2$ is defined on the half-open interval $(0, a]$ to avoid the vanishing of $p(r)$ on the waveguide axis that would give rise to a singular system. (The functions ρ become the weight functions for the orthogonality of the eigenfunctions in the mode expansion of the fields.) The extension of the solutions (the field equations) to the waveguide axis is by continuity.

To obtain a particular solution and give a system of equations that can be used in modelling a waveguide, appropriate boundary conditions and restrictions must be imposed to force compatibility with the physical system, under whatever simplifying assumptions are made.

First of all the functions $\Phi(\varphi)$ must be single valued and periodic on the unit circle. As well as being physically necessary, this gives a separable solution space to the azimuthal functions indexed by $n \in \mathbb{Z}$. Consequently equation (4.1.6) is forced to be of integral order. The eigenfunctions of (4.1.6) are then the Bessel functions of integer order of first kind, J_n, of second kind, N_n, and the modified Bessel functions of first and second order: I_n and K_n respectively. The physical assumption that the fields be everywhere bounded eliminates K_n which increase without bound as $r \to 0$. On the same physical grounds the solution N_n is dismissed for being unbounded below on the guide axis.

For the modelling of the Planck horns the physical assumption made is that the horn walls, being gold coated copper, are perfectly conducting so that the pure Dirichlet and Neumann boundary conditions and the analysis of section 3.1 apply without modification. These assumptions eliminate I_n as a possibility since both $I_n(r)$ and $I_n'(r)$ are strictly positive on $(0, \infty)$ and the boundary conditions could not be matched. That leaves only the J_n for physically plausible eigenfunctions to (4.1.6) under the assumed boundary conditions. Since $J_{-n}(r) = (-1)^n J_n(r)$, the indexing set for the solution space is reduced to the non-negative integers, $n \in \mathbb{Z}_{\geq 0}$. (From a mathematical point of view the physically correct eigenfunctions are found by passing from the non-singular to the singular Sturm-Liouville system with bounded solutions as $r \to 0$.)

Denote the radius of the waveguide by a. The boundary conditions on the wall for the electric field require $J_n(\kappa a) = 0$, and for the magnetic field $J_n'(\kappa a) = 0$ with $\kappa^2 = k^2 - \gamma^2$ to be determined for each n in each case, independently. Thus, in each case, κa is either a zero of the Bessel function or of its derivative. Let q_{nm} denote the m^{th} root of J_n' and and p_{nm} denote the m^{th} root of J_n with $m \in \mathbb{N}$. With reference to the notation of equation (3.1.2), page 52, the boundary conditions will be satisfied with $\kappa^2_{E_{nm}} = k^2 - \gamma^2_{E_{nm}} = (p_{nm}/a)^2$ for the electric case and $\kappa^2_{M_{nm}} = k^2 - \gamma^2_{M_{nm}} = (q_{nm}/a)^2$ for the magnetic case. Rearranging these equations, define the guide wavenumber for the m^{th} radial mode of the n^{th} azimuthal order for the transverse electric and magnetic fields respectively to be

$$\gamma_{E_{nm}}(a) = k\sqrt{1 - \left(\frac{p_{nm}}{ka}\right)^2}, \qquad \gamma_{M_{nm}}(a) = k\sqrt{1 - \left(\frac{q_{nm}}{ka}\right)^2}.$$

Depending upon the wavenumber, k, these radicals may be real or complex. In free space, $k = k_0 \in \mathbb{R}$ so that the axial dependence of the field must, from equation (4.1.3), take the form $e^{\pm j\gamma z}$. The real roots correspond to modes that satisfy Helmholtz equation and therefore propagate in the waveguide; the imaginary roots correspond to evanescent modes with axial dependence $e^{\mp|\gamma|z}$ that satisfy the diffusion equation that is second order in z and, as discussed in subsection 3.1.1, are critical to the correct accounting for power in the scattering system. In the general case $k \in \mathbb{C}$, but the boundary conditions at the waveguide walls must still hold. In that case $\gamma_{E_{nm}} = k\left[1 - (p_{nm}/|k|a)^2 e^{-2j\arg(k)}\right]^{\frac{1}{2}}$, with the equivalent expression for $\gamma_{M_{nm}}$.

From equation (4.1.5) with the periodic endpoint conditions $\Phi(-\pi) = \Phi(\pi)$ and $\Phi'(-\pi) = \Phi'(\pi)$, the azimuthal dependence will be given by the eigenfunction solutions 1, $\cos(n\varphi)$ or $\sin(n\varphi)$ corresponding to the eigenvalues n^2. Thus, for each $n > 0$, there will be two linearly independent eigenfunctions. The general form of the axial fields in a lossless waveguide of circular cross-section with perfectly conducting walls will therefore take the general form

$$\Psi_{E_{nm},z}(r, \varphi, z) \propto A_{nm} J_n\left(\frac{p_{nm}}{a} r\right) \begin{Bmatrix} \cos(n\varphi) \\ \sin(n\varphi) \end{Bmatrix} e^{\pm j(\omega t - \gamma_{E_{nm}} z)}, \tag{4.1.7-A}$$

$$\Psi_{M_{nm},z}(r,\varphi,z) \propto B_{nm} J_n\left(\frac{q_{nm}}{a} r\right) \begin{Bmatrix} \sin(n\varphi) \\ \cos(n\varphi) \end{Bmatrix} e^{\pm j(\omega t - \gamma_{M_{nm}} z)}, \tag{4.1.7-B}$$

for some constants A_{nm} and B_{nm} that depend upon the radial order, m, and the azimuthal order n, of the mode. Then, using these expressions for the axial fields in equations (3.1.2-A and B), the transverse electric and magnetic components of the fields take the form

$$\boldsymbol{H}_{E_{nm}} \propto J_n'\left(\frac{q_{nm}}{a} r\right) \begin{Bmatrix} \sin(n\varphi) \\ \cos(n\varphi) \end{Bmatrix} \hat{\boldsymbol{r}} + \frac{Na}{q_{nm} r} J_n\left(\frac{q_{nm}}{a} r\right) \begin{Bmatrix} \cos(n\varphi) \\ -\sin(n\varphi) \end{Bmatrix} \hat{\boldsymbol{\varphi}}, \tag{4.1.8-A}$$

$$\boldsymbol{E}_{E_{nm}} \propto \frac{na}{q_{nm} r} J_n\left(\frac{q_{nm}}{a} r\right) \begin{Bmatrix} \cos(n\varphi) \\ -\sin(n\varphi) \end{Bmatrix} \hat{\boldsymbol{r}} - J_n'\left(\frac{q_{nm}}{a} r\right) \begin{Bmatrix} \sin(n\varphi) \\ \cos(n\varphi) \end{Bmatrix} \hat{\boldsymbol{\varphi}}, \tag{4.1.8-B}$$

$$\boldsymbol{E}_{M_{nm}} \propto J_n'\left(\frac{p_{nm}}{a} r\right) \begin{Bmatrix} \cos(n\varphi) \\ \sin(n\varphi) \end{Bmatrix} \hat{\boldsymbol{r}} + \frac{na}{p_{nm} r} J_n\left(\frac{p_{nm}}{a} r\right) \begin{Bmatrix} -\sin(n\varphi) \\ \cos(n\varphi) \end{Bmatrix} \hat{\boldsymbol{\varphi}}, \tag{4.1.8-C}$$

$$\boldsymbol{H}_{M_{nm}} \propto \frac{na}{p_{nm} r} J_n\left(\frac{p_{nm}}{a} r\right) \begin{Bmatrix} \sin(n\varphi) \\ -\cos(n\varphi) \end{Bmatrix} \hat{\boldsymbol{r}} + J_n'\left(\frac{p_{nm}}{a} r\right) \begin{Bmatrix} \cos(n\varphi) \\ \sin(n\varphi) \end{Bmatrix} \hat{\boldsymbol{\varphi}}. \tag{4.1.8-D}$$

The normalisation of these fields is the power normalisation of equation (3.1.10) on page 55. The general form of this normalisation expression for both TE and TM fields is the same: integrate the cross product $\boldsymbol{E} \times \boldsymbol{H}^*$ over the disc. (It was noted in the previous chapter that the both E and H are real up to a possibly imaginary scalar determined by the impedance of the guide for the mode, and it follows that the integrand is real up to a possibly imaginary factor that is independent of the integration variables, so that the integral is real.) The integrals are separable and the sine and cosine terms give a factor of π while the radial integrals for the TM fields reduce to

$$\int_0^a \left(J_n'\left(\frac{p_{nm}}{a} r\right)^2 + (1+\delta_0^n)\left(\frac{na}{p_{nm} r} J_n\left(\frac{p_{nm}}{a} r\right)\right)^2 \right) r\,\mathrm{d}r.$$

Using the well known recurrence relations for Bessel functions and their derivatives

$$\frac{n}{x} J_n(x) = \frac{1}{2}\left(J_{n-1}(x) + J_{n+1}(x)\right), \qquad J_n'(x) = \frac{1}{2}\left(J_{n-1}(x) - J_{n+1}(x)\right), \tag{4.1.9}$$

and the symmetry $J_{-1} = -J_1$ the integral becomes, upon multiplying by the azimuthal factor π,

$$P_{M_{nm}}^{-2} \stackrel{\text{def}}{=} \pi \frac{(1+\delta_0^n)}{2} \int_0^a \left(J_{n-1}^2\left(\frac{p_{nm}}{a} r\right) + J_{n+1}^2\left(\frac{p_{nm}}{a} r\right) \right) r\,\mathrm{d}r.$$

This integral is a sum of two Lommel integrals of the second kind [6], the general solution to such integrals being, for $n \in \mathbb{Z}_{\geq 0}$,

$$\int_0^a J_n^2(\alpha r) r \, \mathrm{d}r = \frac{a^2}{2} \left\{ \left(J_n'(a\alpha)\right)^2 + \left(1 - \frac{n^2}{a^2\alpha^2}\right) J_n^2(a\alpha) \right\}.$$

Rearranging (4.1.9) to get $J'_{n\pm1}$ in terms of J_n and $J_{n\pm1}$ and observing that $J_{n-1}(p_{nm}) = -J_{n+1}(p_{nm})$, then substituting into the Lommel integral, expanding and cancelling, then adding the contributions from the two integrands gives

$$P_{M_{nm}}^{-2} \overset{\mathrm{def}}{=} \frac{a^2\pi(1+\delta_0^n)}{2} J_{n+1}^2(p_{nm}),$$

from which the power normalisation coefficient for the nm^{th} transverse magnetic modes are

$$P_{M_{nm}} = \frac{\sqrt{2-\delta_0^n}}{a\sqrt{\pi}|J_{n+1}(p_{nm})|}. \tag{4.1.10}$$

Equivalent manipulations give the radial power normalisation factor for the TE fields as

$$P_{E_{nm}}^{-2} = \pi\frac{(1+\delta_0^n)}{2} \int_0^a \left(J_{n-1}^2\left(\frac{q_{nm}}{a}r\right) + J_{n+1}^2\left(\frac{q_{nm}}{a}r\right)\right) r \, \mathrm{d}r.$$

Again, using the Lommel integral and relations (4.1.9) and using $J_n'(q_{nm}) = 0$ gives the power normalisation coefficient for the nm^{th} transverse electric mode as

$$P_{E_{nm}} = \frac{\sqrt{2-\delta_0^n}}{a\sqrt{\pi}|J_n(q_{nm})|} \cdot \left[1 - \frac{n^2}{q_{nm}^2}\right]^{-\frac{1}{2}}. \tag{4.1.11}$$

The system of equations (4.1.8) with normalisations (4.1.10) and (4.1.11) give, with reference to (3.1.2), the TE and TM modes in cylindrical polar coordinates in the form

$$\frac{\boldsymbol{e}_{E_{nm}}}{P_{E_{nm}}} = \frac{na}{q_{nm}r} J_n\left(\frac{q_{nm}}{a}r\right) \begin{Bmatrix} \cos(n\varphi) \\ -\sin(n\varphi) \end{Bmatrix} \hat{\boldsymbol{r}} - J_n'\left(\frac{q_{nm}}{a}r\right) \begin{Bmatrix} \sin(n\varphi) \\ \cos(n\varphi) \end{Bmatrix} \hat{\boldsymbol{\varphi}}, \tag{4.1.12-A}$$

$$Z_{E_{nm}} \frac{\boldsymbol{h}_{E_{nm}}}{P_{E_{nm}}} = J_n'\left(\frac{q_{nm}}{a}r\right) \begin{Bmatrix} \sin(n\varphi) \\ \cos(n\varphi) \end{Bmatrix} \hat{\boldsymbol{r}} + \frac{na}{q_{nm}r} J_n\left(\frac{q_{nm}}{a}r\right) \begin{Bmatrix} \cos(n\varphi) \\ -\sin(n\varphi) \end{Bmatrix} \hat{\boldsymbol{\varphi}}, \tag{4.1.12-B}$$

$$\frac{\boldsymbol{e}_{M_{nm}}}{P_{M_{nm}}} = J_n'\left(\frac{p_{nm}}{a}r\right) \begin{Bmatrix} \cos(n\varphi) \\ \sin(n\varphi) \end{Bmatrix} \hat{\boldsymbol{r}} + \frac{na}{p_{nm}r} J_n\left(\frac{p_{nm}}{a}r\right) \begin{Bmatrix} -\sin(n\varphi) \\ \cos(n\varphi) \end{Bmatrix} \hat{\boldsymbol{\varphi}}, \tag{4.1.12-C}$$

$$Z_{M_{nm}}\frac{\boldsymbol{h}_{M_{nm}}}{P_{M_{nm}}} = \frac{na}{p_{nm}r}J_n\left(\frac{p_{nm}}{a}r\right)\begin{Bmatrix}\sin(n\varphi)\\ -\cos(n\varphi)\end{Bmatrix}\hat{\boldsymbol{r}} + J_n'\left(\frac{p_{nm}}{a}r\right)\begin{Bmatrix}\cos(n\varphi)\\ \sin(n\varphi)\end{Bmatrix}\hat{\boldsymbol{\varphi}}, \tag{4.1.12-D}$$

in which the transverse electric and magnetic mode impedances are

$$Z_{E_{nm}} = \frac{Z_0}{\gamma_{M_{nm}}} = \frac{Z_0}{\sqrt{1-(q_{nm}/k_0a)^2}}$$

and

$$Z_{M_{nm}} = Z_0\gamma_{E_{nm}} = Z_0\sqrt{1-(p_{nm}/k_0a)^2}$$

in conformity with the notation on page 52. The TE and TM magnetic fields in equations (4.1.12-A to D) are related to the electric field components through the off-axis block diagonal matrices

$$[\boldsymbol{h}_{E_{nm}}] = \begin{bmatrix}0 & -Y_{E_{nm}}\\ Y_{E_{nm}} & 0\end{bmatrix}[\boldsymbol{e}_{E_{nm}}] \quad\text{and}\quad [\boldsymbol{h}_{M_{nm}}] = \begin{bmatrix}0 & -Y_{M_{nm}}\\ Y_{M_{nm}} & 0\end{bmatrix}[\boldsymbol{e}_{M_{nm}}].$$

The Cartesian versions of these equations are obtained from these by the application of a formal rotation through an angle of φ: for the electric field modes

$$\begin{bmatrix}\boldsymbol{e}_x\\ \boldsymbol{e}_y\end{bmatrix} = \begin{bmatrix}\cos(\varphi) & -\sin(\varphi)\\ \sin(\varphi) & \cos(\varphi)\end{bmatrix}\begin{bmatrix}\boldsymbol{e}_r\\ \boldsymbol{e}_\varphi\end{bmatrix}, \tag{4.1.13}$$

and equivalently for the magnetic modes.

4.2 Formulæ for the scattering amplitudes

The scattering across a junction is described by the weighted expansion of the pull-back of the basis functions for the TE and TM fields in the larger section of the waveguide at the junction in terms of the basis functions on the narrower side of the junction. For perfectly aligned guides we can ignore the pull-back. Formally, the fields in the wider section, of radius b, are expanded as a Fourier series of the fields in the narrower section, of radius $a < b$; call these the right and left hand sides respectively. Since the fields on the right are not required to match the boundary conditions that apply to the left it is to be expected that the expansion of the series will require both TE and TM modes to provide the expansion for the fields scattered across the boundary. Equally obviously the back-scattered fields – left to left or right to right – cannot scatter into new modes, but must simply scale by a factor not greater than one. The problem at hand is to find the scattering coefficients to fill the operator matrices P, Q and R of section 3.2.

The matrix Q is associated with the scattering operator $Q \in \operatorname{End}(\mathscr{E}_R \oplus \mathscr{M}_R)$ which, from the orthogonality of the modes, must reduce to a direct sum $Q_E \oplus Q_M \in \operatorname{End}(\mathscr{E}_R) \oplus \operatorname{End}(\mathscr{M}_R)$ of two diagonal matrices, the components of which will simply be the waveguide admittances for each of the modes in the guide to the right of the junction. Likewise $R = R_E \oplus R_M \in \operatorname{End}(\mathscr{E}_L) \oplus \operatorname{End}(\mathscr{M}_L)$ will be diagonal with components the waveguide admittances of the modes to the left of the junction.

Since k_0 and the waveguide radius, a, and the roots of the Bessel functions and their derivatives are all strictly positive and finite, $0 < k_0 a/q_{nm}$, and as $k_0 \to q_{nm}/a$ from below, $\left[1 - (k_0 a/q_{nm})^2\right] \to \infty$, the impedance $Z_{E_{nm}}$ becoming infinite at $k_0 a/q_{nm} = 1$. This is the cut-off condition for propagation at angular frequency $\omega_{nm} = c q_{nm}/a$ below which the impedance becomes purely imaginary and Helmholtz equation is of diffusion type. Likewise, $Y_{M_{nm}} \to \infty$ as $k_0 \to p_{nm}/a$, $Y_{M_{nm}}$ and $Z_{M_{nm}}$ becoming purely imaginary for $k_0 > p_{nm}/a$. For each $n \geq 0$ the matrix operators Q and R then take the form

$$\begin{bmatrix} Q_{E_n} & 0 \\ 0 & Q_{M_n} \end{bmatrix}, \qquad \begin{bmatrix} R_{E_n} & 0 \\ 0 & R_{M_n} \end{bmatrix}$$

where the sub-matrices are the diagonal matrices of admittances with $Q_{ij} = R_{ij} = 0$ for $i \neq j$ and

$$\begin{aligned} (Q_{E_{nm}})_{mm} &= Y_{E_{nm}}(b), & \qquad (Q_{M_{nm}})_{mm} &= Y_{M_{nm}}(b), \\ (R_{E_{nm}})_{mm} &= Y_{E_{nm}}(a), & \qquad (R_{M_{nm}})_{mm} &= Y_{M_{nm}}(a). \end{aligned}$$

Define the symbols

$$\forall\, n \geq 1,\ \alpha_k^n = \mathtt{sgn}(J_n(q_{nk})) = \begin{cases} 1 & : \ k \equiv 1 \mod (2) \\ -1 & : \ k \equiv 0 \mod (2) \end{cases} \tag{4.2.1-A}$$

$$\alpha_k^0 = -\alpha_k^1 \tag{4.2.1-B}$$

$$\forall\, n,\ \beta_k^n = -\,\mathtt{sgn}(J_{n+1}(p_{nk})) = \begin{cases} -1 & : \ k \equiv 1 \mod (2) \\ 1 & : \ k \equiv 0 \mod (2) \end{cases} \tag{4.2.1-C}$$

Denote the ratio of smaller to the larger waveguide radii at the step by $\rho = a/b$, then $P = Q^*\Pi$, in which Π is a real matrix with components as given in the following two sets of equations, (4.2.3) and (4.2.4) at the foot of this page. The scattering amplitudes arises from substitution of equations (4.1.12) into the integrals of the various $\boldsymbol{e}_L \times \boldsymbol{h}_R^*$ which, ignoring the mode power normalisation factors $P_{E_{nm}}$ and $P_{M_{nm}}$ and the impedances, all of which pull out of the integrand, and using the relations (4.1.9), give rise to integrands that are a sum of terms with the general form $J_{n\pm1}(\alpha r)J_{m\pm1}(\beta r)\mathrm{cs}(n\varphi)\mathrm{cs}(m\varphi)$, in which the terms $\mathrm{cs}(n\varphi)\mathrm{cs}(m\varphi)$ are either pairs of sine or cosine functions. The integrands are then separable and the orthogonality of the sines and cosines forces $n = m$ for non-zero integrals. That eliminates the cross terms in which $n \pm 1 = m \mp 1$ leaving a sum of Lommel integrals of the first kind, the general form and solution of which is

$$\int_0^a J_n(\alpha r)J_n(\beta r) r\,\mathrm{d}r = \frac{a}{\alpha^2 - \beta^2} \cdot \left[\beta J_n(a\alpha)J_n'(a\beta) - \alpha J_n'(a\alpha)J_n(a\beta)\right]. \tag{4.2.2}$$

Solving the integral gives, for azimuthal order $n = 0$, the scattering amplitudes

$$(\Pi_{EE}^0)_{mk} = 2\,\alpha_k^0 \frac{\rho}{q_{0m}^2} \cdot \frac{J_1(q_{0m}\rho)}{|J_0(q_{0m})|}\left[\rho^2 - \left(\frac{q_{0k}}{q_{0m}}\right)^2\right]^{-1}, \tag{4.2.3-A}$$

$$(\Pi_{EM}^0)_{mk} = 0, \tag{4.2.3-B}$$

$$(\Pi_{ME}^0)_{mk} = 0, \tag{4.2.3-C}$$

$$(\Pi_{MM}^0)_{mk} = 2\,\beta_k^0 \frac{\rho}{p_{0m}} \cdot \frac{J_0(p_{0m}\rho)}{|J_1(p_{0m})|}\left[\rho^2 - \left(\frac{p_{0k}}{p_{0m}}\right)^2\right]^{-1}, \tag{4.2.3-D}$$

while for all azimuthal orders $n > 0$

$$(\Pi_{EE}^n)_{mk} = \frac{2\,\alpha_k^n \rho}{q_{nm}^2} \cdot \frac{J_n'(q_{nm}\rho)}{|J_n(q_{nm})|} \cdot \left[\rho^2 - \frac{q_{nk}^2}{q_{nm}^2}\right]^{-1}\left[\left(1 - \frac{n^2}{q_{nk}^2}\right)\left(1 - \frac{n^2}{q_{nm}^2}\right)\right]^{-\frac{1}{2}}, \tag{4.2.4-A}$$

$$(\Pi_{EM}^n)_{mk} = 0, \tag{4.2.4-B}$$

$$(\Pi^n_{ME})_{mk} = \frac{\alpha^n_k\, n}{p_{nm}\cdot q_{nk}} \cdot \frac{J_n\,(p_{nm}\rho)}{|J_{n+1}\,(p_{nm})|} \cdot \left[1 - \frac{n^2}{q^2_{nk}}\right]^{-\frac{1}{2}}, \tag{4.2.4-C}$$

$$(\Pi^n_{MM})_{mk} = 2\,\beta^n_k \frac{\rho}{p_{nm}} \cdot \frac{J_n\,(p_{nm}\rho)}{|J_{n+1}\,(p_{nm})|} \cdot \left[\rho^2 - \frac{p^2_{nk}}{p^2_{nm}}\right]^{-1}. \tag{4.2.4-D}$$

Note also that as the step size increases $\rho \longrightarrow 0_+$ and, for a waveguide terminating with its aperture in an infinite perfectly conducting plane (a step to free space in a ground plane) the scattering amplitudes in equations (4.2.3) and (4.2.4) all tend to zero. This implies a breakdown in the formalism and the step to free space cannot be handled by treating the mouth of the horn as a wide flange. The reason is that the enforced boundary conditions on the free space side of the plane are unphysical. Indeed, a look at the derivation of the normalisation coefficients $P^{-2}_{E_{nm}}$ and $P^{-2}_{M_{nm}}$ shows that the integrals are either zero or they are unbounded, depending upon n and nonsense is to be expected.

The equations (4.1.12) for the basis elements of the TE and the TM fields in the waveguide make clear that, in the circular cylinder case, the space of transverse fields in one section of the guide has a basis set indexed by pairs $(n, m) \in \mathbb{Z}_{\geq 0} \times \mathbb{N}$, and from the preceding discussion it is clear that modes of differing azimuthal order, n, cannot scatter into each other at a junction provided the guide sections are coaxial. From this it is immediate that the P operator matrices, as well as those for Q and R, must decompose into direct sums $P = \bigoplus_{n=0}^{\infty} P_n$ as discussed on page 59. When the guide sections at either side of the junction are not coaxial this decomposition fails because the pull-back function $(\pi^* J_n)(q_{kl}r/b) = J_n(q_{kl}\pi(r)/b)$ is an infinite sum of products of Bessel functions. As a result the operator matrix P is no longer block diagonal, but becomes dense. (By dense is meant that the system of operator matrices is, strictly speaking, an infinite array of infinite operator matrices, though it will be seen that most of the power is scattered into ‘nearby’ azimuthal orders, thus rendering the array of operator matrices essentially band diagonal.) This situation is discussed in section 4.6.

All equations (4.2.3-A and D) and (4.2.4-A, C and D) tend uniformly to zero as either $\rho \longrightarrow 1_-$ or as $m \longrightarrow \infty$. The difference between the case $n = 0$ and $n \neq 0$ can be exploited in numerical work resulting in a computational complexity and time saving factor of two when calculating the azimuthal order 0 case for multi-mode horns.

Numerical stability of the model requires, for some machine dependent $\varepsilon > 0$, that $|p_{nm}\rho - p_{nk}| \gg \varepsilon$ and $|q_{nm}\rho - q_{nm}| \gg \varepsilon$ for all $n,\ k,\ m$ at all junctions. Equations (4.2.4-A and C) are always bounded since $n < q_{nm}$; however, $(1 - n^2/q_{nm}^2) \xrightarrow{m\to\infty} 0$ which could lead to numerical instability if the equations where programmed as written. Reformulations of the equations (4.2.3) and (4.2.4) for numerical use in any situation where small step sizes occur are given in section 5.2.

4.3 Scattering amplitude dependence on radial index m and k

For the modelling of the operator Π the development of the scattering amplitudes as the indices m and k vary has to be understood so that the reduction from an infinite dimensional scattering operator to a finite dimensional matrix model retains sufficient information to give a reliable model and account adequately for the power in evanescent modes.

In section 5.2 it will be shown that, writing $\varepsilon = p_{nm}/b - p_{nk}/a$, equations (4.2.3-D) and (4.2.4-D) can be written using (5.2.3) in the alternative form

$$\frac{-2\beta_k^n}{|J_{n+1}(p_{nm})|} \cdot \frac{p_{nk} + a\varepsilon}{2p_{nk} + \varepsilon} \cdot \sum_{\nu\in\mathbb{N}} \frac{(a\varepsilon)^{\nu-1}}{\nu!} J_n^{(\nu)}(p_{nk})$$

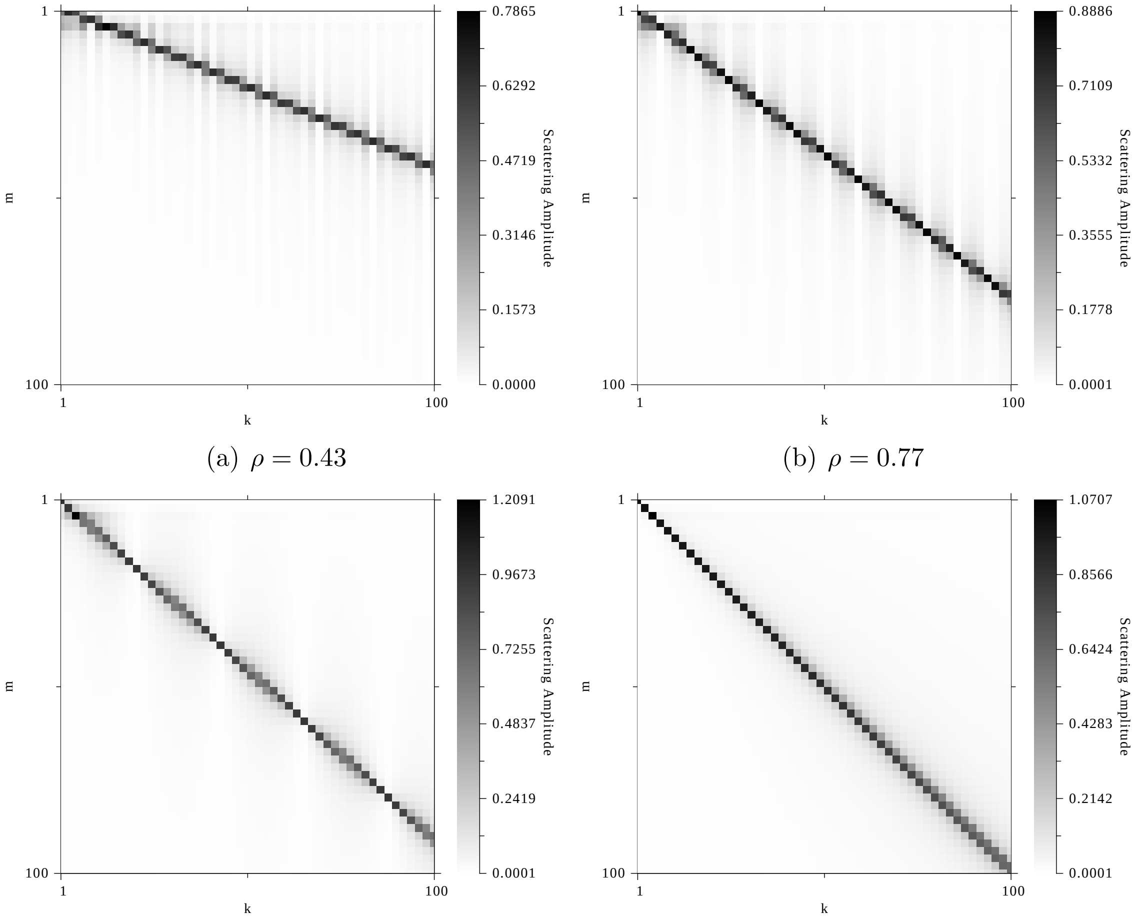
which, using the relations (4.1.9) and the definition of β_k^n in equation (4.2.1-C), is seen to have the limit $|J_{n+1}(p_{nk})| \,/\, |J_{n+1}(p_{nm})|$ as $\varepsilon \to 0$. For TM to TM scattering this is the maximum possible contribution to the scattering amplitude across a junction from a radial mode index of k on the left of the junction to mode of index m on the right. For $m > k$ the roots satisfy $p_{nm} > p_{nk}$ and for $p_{nm}\rho > p_{nk}$

$$\begin{aligned}(\Pi^n_{MM})_{mk} &= 2\beta^n_k \frac{J_n(p_{nm}\rho)}{p_{nm}\rho\,|J_{n+1}(p_{nm})|} \cdot \left[1 - \left(\frac{p_{nk}}{p_{nm}\rho}\right)^2\right]^{-1} \\ &< 2\beta^n_k \frac{J_n(p_{nm}\rho)}{p_{nm}\rho\,|J_{n+1}(p_{nm})|} \xrightarrow{m\to\infty} 0_+ .\end{aligned} \tag{4.3.1}$$

For $m < k$ the equivalent equation for $p_{nk} > p_{nm} > p_{nm}\rho$ again shows that

$$(\Pi^n_{MM})_{mk} \xrightarrow{m\to 1} 2\beta^n_k \frac{p_{n1}\,\rho\, J_n(p_{n1}\rho)}{p^2_{nk}\,|J_{n+1}(p_{n1})|\,\left[\,(p_{n1}\rho/p_{nk})^2 - 1\right]}$$

and there are similar equations for Π_{EE} with equivalent trends.



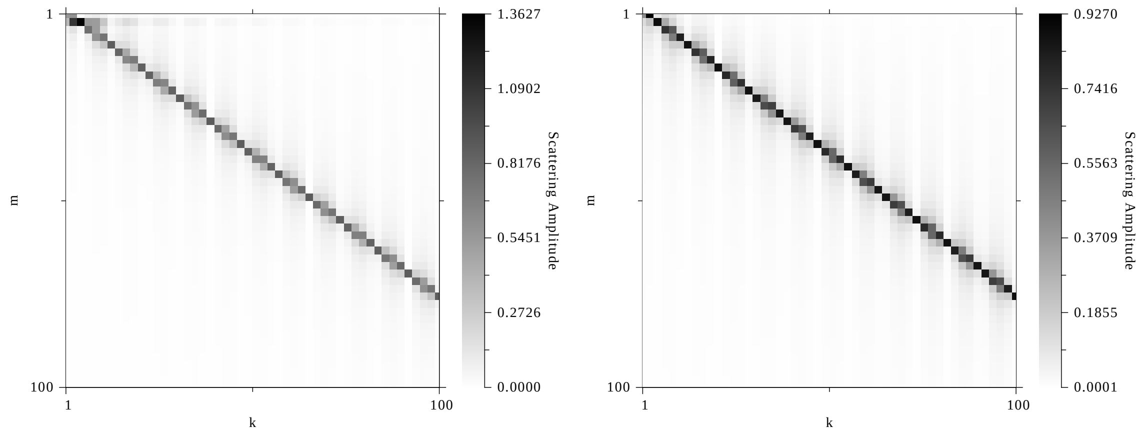
(a) $\rho = 0.43$

(b) $\rho = 0.77$

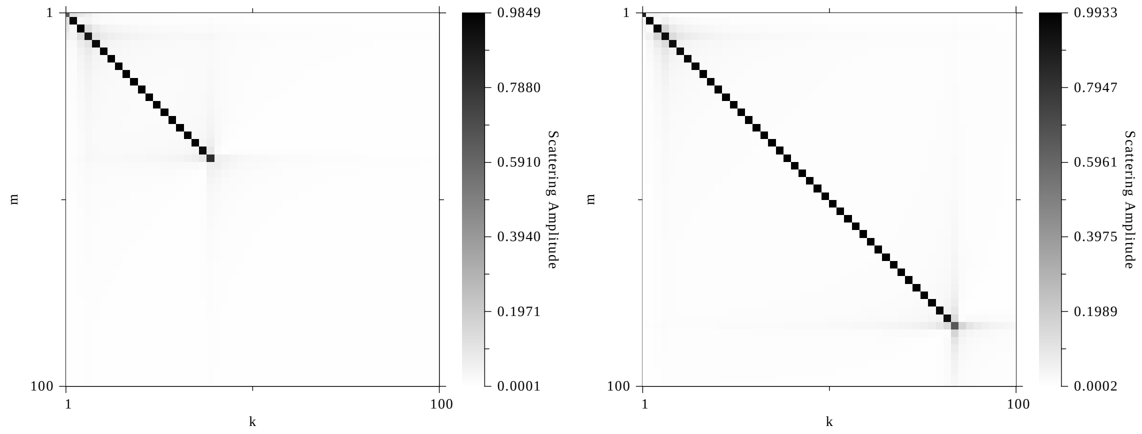
(c) $\rho = 0.91$

(d) $\rho = 0.99$

Figure 4.1: Illustrations of the variation in the form of the scattering amplitude for radial orders k into m in sub-matrices Π^0_{EE} with ρ. The arrays are rotated through 90°. In (a) ρ is typical of the filter section in a single-mode horn; in (b) it is in the taper section. In (c) ρ is typical of the parallel section of a single-mode ultra-Gaussian horn, and in (d) ρ represents the almost diagonal condition for the step size in a model of a smooth walled horn. For fixed ρ the arrays of all azimuthal order have the same essential form, only the amplitudes differ, ρ determining the slope of the dominant scattering amplitude band.



(a) Π^5_{EE} for $\rho = 0.77$ (b) Π^5_{MM} for $\rho = 0.77$



(c) The operator KG^1_{EE} for $\rho = 0.40$ (d) The operator KG^1_{EE} for $\rho = 0.83$

Figure 4.2: Illustration of the (a) TE to TE, and (b) TM to TM scattering amplitudes for radial orders k into m and azimuthal order 5 with $\rho = 0.77$. Figures (c) and (d) illustrate the sparsity of the operator product $KG = \bar{R}^{-1}\Pi^T\bar{Q}\Pi$ of equation (3.3.1) which gives rise to sparse sub-arrays of S, here the azimuthal order is 1. In (c) the TE_{11} field is propagating in the filter section of the horn whereas in (d) the field is propagating in the broadening section of the horn. By the time the field is in the wide section of the horn $\rho \sim 0.95$ and KG will be essentially a diagonal operator. Both (c) and (d) are modelled at the same frequency (500 GHz).

Thus it is seen that the maximum scattering amplitude in the k^{th} column of Π occurs in the m^{th} row where m minimises $\varepsilon = \left|p_{nm}/b - p_{nk}/a\right|$ and falls off in amplitude in the rows above and below. Typically, in a standard corrugated horn the corrugation depths are of the order $0.25\lambda \lesssim d \lesssim 0.5\lambda$ and $0.43 \lesssim \rho \lesssim 0.95$, the smaller values for ρ applying in the neck of a single-mode horn where the mode selection is made and the largest values applying close to the radiating aperture.

It follows that the maximum scattering amplitudes lie off the diagonal, but the operators become increasingly diagonal as $\rho \to 1$; this is illustrated in figure 4.1. The value of ρ determines the slope of the high amplitude band and the structure of these operator arrays means that the L^2 norm is determined largely by the values in the the region immediately surrounding this band. The same applies in the finite size model for the Frobenius norm.

From the equations on page 67, chapter 3, and the above it is seen that the operators $K = \bar{R}^{-1}\Pi^T Q$ and $G = \bar{Q}Q^{-1}\Pi$, so the derived operators in equations (3.3.1) and (3.3.2) are obtained from Π or Π^T by multiplication by diagonal operators. G therefore has the same structure as Π, rows being multiplied by the squares of the argument of the complex conjugates of the mode impedances; likewise K has essentially the same structure as Π. Both KG and GK become diagonal dominant. It follows that the structure of the operators in the S-matrix is essentially sparse, and so sparse methods could be exploited for highly efficient preliminary design and analysis of corrugated waveguides; see figures 4.1 and 4.2. See also figure 5.1, page 141 for the complete S_{21} scattering operators for single and multi-mode horns. In particular the operators $(I \pm KG)^{\pm 1}$ and $(I \pm GK)^{\pm 1}$ must be essentially diagonal, and that closeness to diagonality determines the structure S_{11} and S_{22} and would allow the solution to $S_{11} = (I + KG)^{-1}(I - KG)$ as $(I + KG)S_{11} = (I - KG)$ to be solved by guessing a diagonal solution $s_{ii} = (1 - (KG)_{ii})/(1 + (KG)_{ii})$ and resorting immediately to iterative refinement of the trial solution. For smooth walled horns in particular the step size needs to be kept very small; consequently the operators become very nearly diagonal. Therefore, for such horns sparse methods could be exploited for the detailed design and analysis of the horns without significant loss of accuracy but with potentially great savings in computation time.

4.4 Scattering in rectangular waveguides

Given a rectangular waveguide of dimension $a \times b$, set up a rectangular coordinate system centred on the axis of the guide with x axis parallel to the edges of length

a and y axis parallel to the side of length b, with orientation is such that $+z$ is the direction of forward propagation.

Let k denote the wavenumber in free space, m, $n \in \mathbb{N} \cup \{0\}$ the mode numbers in the x and y directions respectively. Then the modal wavenumbers and attenuation in the guide are

$$k_m = \frac{\pi m}{a} \qquad k_n = \frac{\pi n}{b} \tag{4.4.1-A}$$

$$k_{mn} = \sqrt{k_m^2 + k_n^2} \qquad \beta_{mn}^* = \sqrt{k^2 - k_{mn}^2}. \tag{4.4.1-B}$$

With normalisation coefficient A_{mn}, the electric and magnetic transverse electric field vectors for the mn^{th} mode are [42]

$$\begin{aligned}(\boldsymbol{E}_E)_{mn} = j\frac{\omega\mu A_{mn}}{k_{mn}^2}\Bigg[&k_n \cos\Big(k_m\big(x+\frac{a}{2}\big)\Big)\sin\Big(k_n\big(y+\frac{b}{2}\big)\Big)\hat{\boldsymbol{x}}\\ &- k_m \sin\Big(k_m\big(x+\frac{a}{2}\big)\Big)\cos\Big(k_n\big(y+\frac{b}{2}\big)\Big)\hat{\boldsymbol{y}}\Bigg],\end{aligned} \tag{4.4.2-A}$$

$$\begin{aligned}(\boldsymbol{H}_E)_{mn} = j\frac{\beta_{mn} A_{mn}}{k_{mn}^2}\Bigg[&k_m \sin\Big(k_m\big(x+\frac{a}{2}\big)\Big)\cos\Big(k_n\big(y+\frac{b}{2}\big)\Big)\hat{\boldsymbol{x}}\\ &+ k_n \cos\Big(k_m\big(x+\frac{a}{2}\big)\Big)\sin\Big(k_n\big(y+\frac{b}{2}\big)\Big)\hat{\boldsymbol{y}}\Bigg].\end{aligned} \tag{4.4.2-B}$$

With normalisation coefficient B_{mn}, the electric and magnetic transverse magnetic field vectors for the mn^{th} mode are [42]

$$\begin{aligned}(\boldsymbol{E}_M)_{mn} = -j\frac{\beta_{mn} B_{mn}}{k_{mn}^2}\Bigg[&k_m \cos\Big(k_m\big(x+\frac{a}{2}\big)\Big)\sin\Big(k_n\big(y+\frac{b}{2}\big)\Big)\hat{\boldsymbol{x}}\\ &+ k_n \sin\Big(k_m\big(x+\frac{a}{2}\big)\Big)\cos\Big(k_n\big(y+\frac{b}{2}\big)\Big)\hat{\boldsymbol{y}}\Bigg],\end{aligned} \tag{4.4.3-A}$$

$$\begin{aligned}(\boldsymbol{H}_M)_{mn} = j\frac{\omega\varepsilon B_{mn}}{k_{mn}^2}\Bigg[&k_n \sin\Big(k_m\big(x+\frac{a}{2}\big)\Big)\cos\Big(k_n\big(y+\frac{b}{2}\big)\Big)\hat{\boldsymbol{x}}\\ &- k_m \cos\Big(k_m\big(x+\frac{a}{2}\big)\Big)\sin\Big(k_n\big(y+\frac{b}{2}\big)\Big)\hat{\boldsymbol{y}}\Bigg].\end{aligned} \tag{4.4.3-B}$$

It would be exceedingly inefficient to use these equations in the form given here in numerical code. Both speed and accuracy can be improved by making a few elementary observations and it will be shown that, rather than the $16nm$ trigonometric calls, only two calls to cosine and two calls to sine are required. With reference to equation (4.4.2-A), observe that the term in square brackets evaluates to one of the following four cases:

$$\begin{aligned}
&\alpha_1[k_m \sin(k_m x)\cos(k_n y)\hat{\boldsymbol{x}} + k_n \cos(k_m x)\sin(k_n y)\hat{\boldsymbol{y}}] && : \text{case } m \text{ odd, } n \text{ odd,}\\
&\alpha_2[k_m \sin(k_m x)\sin(k_n y)\hat{\boldsymbol{x}} - k_n \cos(k_m x)\cos(k_n y)\hat{\boldsymbol{y}}] && : \text{case } m \text{ odd, } n \text{ even,}\\
&\alpha_3[k_m \cos(k_m x)\cos(k_n y)\hat{\boldsymbol{x}} - k_n \sin(k_m x)\sin(k_n y)\hat{\boldsymbol{y}}] && : \text{case } m \text{ even, } n \text{ odd,}\\
&\alpha_4[k_m \cos(k_m x)\sin(k_n y)\hat{\boldsymbol{x}} + k_n \sin(k_m x)\cos(k_n y)\hat{\boldsymbol{y}}] && : \text{case } m \text{ even, } n \text{ even,}
\end{aligned}$$

in which the coefficients α_j take the values

$$\alpha_1 = \begin{cases} -1 & : \quad ((m \equiv 1 \mod (4)) \wedge (n \equiv 1 \mod (4))) \vee \\ & \qquad ((m \equiv 3 \mod (4)) \wedge (n \equiv 3 \mod (4))), \\ 1 & : \quad ((m \equiv 1 \mod (4)) \wedge (n \equiv 3 \mod (4))) \vee \\ & \qquad ((m \equiv 3 \mod (4)) \wedge (n \equiv 1 \mod (4))), \\ 0 & : \quad \text{otherwise.} \end{cases}$$

$$\alpha_2 = \begin{cases} -1 & : \quad ((m \equiv 1 \mod (4)) \wedge (n \equiv 0 \mod (4))) \vee \\ & \qquad ((m \equiv 3 \mod (4)) \wedge (n \equiv 2 \mod (4))), \\ 1 & : \quad ((m \equiv 1 \mod (4)) \wedge (n \equiv 2 \mod (4))) \vee \\ & \qquad ((m \equiv 3 \mod (4)) \wedge (n \equiv 0 \mod (4))), \\ 0 & : \quad \text{otherwise.} \end{cases}$$

$$\alpha_3 = \begin{cases} 1 & : \quad ((m \equiv 0 \mod (4)) \wedge (n \equiv 1 \mod (4))) \vee \\ & \qquad ((m \equiv 2 \mod (4)) \wedge (n \equiv 3 \mod (4))), \\ -1 & : \quad ((m \equiv 0 \mod (4)) \wedge (n \equiv 3 \mod (4))) \vee \\ & \qquad ((m \equiv 2 \mod (4)) \wedge (n \equiv 1 \mod (4))), \\ 0 & : \quad \text{otherwise.} \end{cases}$$

$$\alpha_4 = \begin{cases} 1 & : \quad ((m \equiv 0 \mod (4)) \wedge (n \equiv 0 \mod (4))) \vee \\ & \qquad ((m \equiv 2 \mod (4)) \wedge (n \equiv 2 \mod (4))), \\ -1 & : \quad ((m \equiv 0 \mod (4)) \wedge (n \equiv 2 \mod (4))) \vee \\ & \qquad ((m \equiv 2 \mod (4)) \wedge (n \equiv 0 \mod (4))), \\ 0 & : \quad \text{otherwise.} \end{cases}$$

The equations for $\boldsymbol{H}_E$, $\boldsymbol{E}_M$ and $\boldsymbol{H}_M$ can be similarly rewritten.

Thus the α_j and the appropriate sine and cosine terms are simply cycled through in a double loop over the mode indices. Furthermore, with the equations expressed in this way and using the recurrence relations

$$\begin{aligned} \cos(k_2 x) &= 2\cos^2(k_1 x) - 1, \\ \cos(k_{m+1} x) &= 2\cos(k_1 x)\cos(k_m x) - \cos(k_{m-1} x) \quad \text{for } m \geq 3, \end{aligned}$$

likewise for $\cos(k_n y)$, and for the sines

$$\sin(k_{m+1} x) = 2\cos(k_1 x)\sin(k_m x) - \sin(k_{m-1} x) \quad \text{for } m \geq 1,$$

it is seen that it is only necessary to make a single call to evaluate $\sin(x\pi/a)$, $\cos(x\pi/a)$, $\sin(y\pi/b)$ and $\cos(y\pi/b)$ at each junction, and no other calls to sine or cosine routines are required to evaluate each $\cos(k_{m+1}x)$ etc. For that reason the calculation of the fields in a rectangular waveguide can be written in an exceedingly efficient and computationally accurate way, far more efficient than calculations of the Bessel functions for circular waveguides.

If the computational procedure follows the above scheme, then it would always be best to use n and m as multiples of four. If that was deemed undesirable, then the recurrence routines can be used in equations (4.4.2) and (4.4.3) directly but, whichever of the two approaches is taken, more than two calls to the cosine and two to the sine routine is unnecessary.

Scattering across junctions is determined exactly as in chapter 3 from the continuity of the magnetic fields and integration of the complex Poynting vector over the open junction, giving

$$\begin{aligned} P &= \int_S (\boldsymbol{e}_L \times \boldsymbol{h}_R^*) \cdot \, \mathrm{d}\boldsymbol{S} \\ &= \int_{-\frac{a}{2}}^{\frac{a}{2}} \int_{-\frac{b}{2}}^{\frac{b}{2}} \left(E_{x,L} H_{y,R}^* - E_{y,L} H_{x,R}^* \right) \, \mathrm{d}x \, \mathrm{d}y, \end{aligned} \tag{4.4.4}$$

and equivalent equations for Q and R. These integrals are clearly separable, and using the impedance relations $E_x/H_y = \omega\mu/\beta = Z_E$, etc. the normalisation coefficients are derived by setting

$$\int_{-\frac{a}{2}}^{\frac{a}{2}} \int_{-\frac{b}{2}}^{\frac{b}{2}} \left(E_x^2 + E_y^2\right) \mathrm{d}x \, \mathrm{d}y = 1, \tag{4.4.5}$$

and these separable integrals give the cases

$$A_{m0} =: A_m = \sqrt{\frac{2}{ab}} \cdot \frac{k_m}{\omega\mu} \quad : \quad m \neq 0,\, n = 0, \tag{4.4.6-A}$$

$$A_{0n} =: A_n = \sqrt{\frac{2}{ab}} \cdot \frac{k_n}{\omega\mu} \quad : \quad m = 0,\, n \neq 0, \tag{4.4.6-B}$$

$$A_{mn} = \sqrt{2(A_m^2 + A_n^2)} = \frac{2k_{mn}}{\omega\mu\sqrt{ab}} \quad : \quad m \neq 0,\, n \neq 0, \tag{4.4.6-C}$$

$$B_{mn} = \frac{2}{\sqrt{ab}} \cdot \frac{k_{nm}}{|\beta_{mn}|} = \omega\mu\frac{A_{mn}}{|\beta_{mn}|} \quad : \quad \begin{cases} m \neq 0,\, n \neq 0, \\ \text{otherwise undefined.} \end{cases} \tag{4.4.6-D}$$

Label the modes in the waveguide section to the right of the junction with $M,\ N \in \mathbb{N} \cup \{0\}$ and the dimensions of the guide by A and B. Then all of the equations above give the equations for the right-hand section by substitution of the symbols M for m, N for n, A for a and B for b. Denote the electric-electric scattering coefficient for the mode mn to mode MN by $(P_{EE})_{MN,mn}$, and so on. Solving the integrals of type (3.1.10) gives the junction scattering equations between modes at the junction:

$$(R_E)_{mn} = \frac{1}{Z^*_{E,mn}}, \qquad (R_M)_{mn} = \frac{1}{Z^*_{M,mn}}, \tag{4.4.7-A}$$

$$(Q_E)_{MN} = \frac{1}{Z^*_{E,MN}}, \qquad (Q_M)_{MN} = \frac{1}{Z^*_{M,MN}}. \tag{4.4.7-B}$$

In the reduction of the scattering coefficient equations to the form (4.4.9) below, the functions $\tau(m, M, a, A)$ and $\tau(n, M, b, B)$ are defined for $M,\ N > 0$ by

$$\begin{aligned} \tau(m, M, a, A) &= \sin\left[k_M \frac{A-a}{2}\right] - (-1)^m \sin\left[k_M \frac{A+a}{2}\right] \\ &= \begin{cases} -2\sin\left(\dfrac{k_M a}{2}\right)\cos\left(\dfrac{k_M A}{2}\right) & : \quad m \equiv 0 \mod (2), \\ 2\sin\left(\dfrac{k_M A}{2}\right)\cos\left(\dfrac{k_M a}{2}\right) & : \quad m \equiv 1 \mod (2). \end{cases} \end{aligned} \tag{4.4.8}$$

Again, for numerical work the evaluation of sines and cosines is avoided by use of the recurrence relations indexed by M and N.

Integral (4.4.4) gives the coefficients of the scattering matrix, P, which, scaled by the waveguide impedances are as follows: define $\delta = 1/\sqrt{abcd}$, $\beta = \arg(\beta_{MN})$ for $|\beta_{MN}| > 0$, $\beta = 0$ otherwise, and the functions τ as in equation (4.4.8) above, then

$$Z^*_{E,0N}(P_{EE})^{0n}_{0N} = \frac{2\,\delta\,a\,k_n}{(k_n^2 - k_N^2)}\,\tau(n,N,b,B)$$

$$: \quad m = 0,\, n \neq 0,\, M = 0,\, N \neq 0, \tag{4.4.9-A}$$

$$Z^*_{E,M0}(P_{EE})^{m0}_{M0} = \frac{2\,\delta\,b\,k_m}{(k_m^2 - k_M^2)}\,\tau(m,M,a,A)$$

$$: \quad m \neq 0,\, n = 0,\, M \neq 0,\, N = 0, \tag{4.4.9-B}$$

$$Z^*_{E,MN}(P_{EE})^{0n}_{MN} = \frac{-2\sqrt{2}\,\delta\,k_n k_N}{k_M k_{MN}(k_n^2 - k_N^2)}\,\tau(0,M,a,A)\,\tau(n,N,b,B)$$

$$: \quad m = 0,\, n \neq 0,\, M \neq 0,\, N \neq 0, \tag{4.4.9-C}$$

$$Z^*_{E,MN}(P_{EE})^{m0}_{MN} = \frac{-2\sqrt{2}\,\delta\,k_m k_M}{k_N k_{MN}(k_m^2 - k_M^2)}\,\tau(m,M,a,A)\,\tau(0,N,b,B)$$

$$: \quad m \neq 0,\, n = 0,\, M \neq 0,\, N \neq 0, \tag{4.4.9-D}$$

$$Z^*_{E,MN}(P_{EE})^{mn}_{MN} = \frac{4\,\delta\,k_M k_N\,k_{mn}}{k_{MN}(k_m^2 - k_M^2)(k_n^2 - k_N^2)}\,\tau(m,M,a,A)\,\tau(n,N,b,B)$$

$$: \quad m \neq 0,\, n \neq 0,\, M \neq 0,\, N \neq 0, \tag{4.4.9-E}$$

$$(P_{EM})^{mn}_{MN} = 0 \qquad : \quad \forall\, m,\, n,\, M, \text{ and } N, \tag{4.4.9-F}$$

$$Z^*_{M,MN}(P_{ME})^{0n}_{MN} = \frac{-2\sqrt{2}\,\delta\beta\,k_n}{k_{MN}(k_n^2 - k_N^2)}\,\tau(0,M,a,A)\,\tau(n,N,b,B)$$

$$: \quad m = 0,\, n \neq 0,\, M \neq 0,\, N \neq 0, \tag{4.4.9-G}$$

$$Z^*_{M,MN}(P_{ME})^{m0}_{MN} = \frac{2\sqrt{2}\,\delta\beta\,k_m}{k_{MN}(k_m^2 - k_M^2)}\,\tau(m,M,a,A)\,\tau(0,N,b,B)$$

$$: \quad m \neq 0,\, n = 0,\, M \neq 0,\, N \neq 0, \tag{4.4.9-H}$$

$$Z^*_{M,MN}(P_{ME})^{mn}_{MN} = \frac{-4\,\delta\beta\,(k_m^2 k_N^2 - k_n^2 k_M^2)}{k_{mn} k_{MN}(k_m^2 - k_M^2)(k_n^2 - k_N^2)}\,\tau(m,M,a,A)\,\tau(n,N,b,B)$$

$$: \quad m \neq 0,\, n \neq 0,\, M \neq 0,\, N \neq 0, \tag{4.4.9-I}$$

$$Z^*_{M,MN}(P_{MM})^{mn}_{MN} = \frac{4\,\delta\,k_m k_n k_{MN}}{k_{mn}(k_m^2 - k_M^2)(k_n^2 - k_N^2)}\,\tau(m,M,a,A)\,\tau(n,N,b,B)$$

$$: \quad m \neq 0,\, n \neq 0,\, M \neq 0,\, N \neq 0. \tag{4.4.9-J}$$

For all cases not listed the scattering coefficients equations (4.4.9) are identically zero, there being no scattering between those mode combinations. Note that the right hand sides of these equations are determined entirely by the geometry.

Define the ratios $Z^*_{E,mn}/Z^*_{E,MN}$ and $Z^*_{M,mn}/Z^*_{M,MN}$ at a junction to be to be

$$(C_E)^{mn}_{MN} = \frac{Z^*_{E,mn}}{Z^*_{E,MN}} = \frac{\beta^*_{MN}}{\beta^*_{mn}} = \frac{Z^*_{M,mn}}{Z^*_{M,MN}} = (C_M)^{mn}_{MN},$$

so define the ratio of impedances

$$C^{mn}_{MN} = \sqrt{\frac{1-(k_{MN}/k)^2}{1-(k_{mn}/k)^2}}. \tag{4.4.10}$$

Likewise $Z^*_{E,mn}/Z^*_{M,MN} = C^{mn}_{MN} \cdot \mu/\varepsilon$.

Writing the scattering equations in terms of $G = Q^{-1}P$ and $K = \bar{R}^{-1}P^\dagger$ as in section 3.3, page 66 the angular frequency, permeability and permittivity cancel and the impedances do not occur explicitly. In that case all of the $1/Z^*$ terms are replaced to give the entries of $K^\dagger$ and G in terms of C^{mn}_{MN} and the right-hand sides of equations (4.4.9)

$$(K^\dagger_{EE})^{mn}_{MN} = (P_{EE}R^{-1}_E)^{mn}_{MN} = C^{mn}_{MN}Z^*_{E,MN}(P_{EE})^{mn}_{MN}, \tag{4.4.11-A}$$

$$(K^\dagger_{ME})^{mn}_{MN} = (P_{ME}R^{-1}_E)^{mn}_{MN} = \frac{\mu}{\varepsilon}C^{mn}_{MN}Z^*_{M,MN}(P_{ME})^{mn}_{MN}, \tag{4.4.11-B}$$

$$(K^\dagger_{MM})^{mn}_{MN} = (P_{MM}R^{-1}_M)^{mn}_{MN} = C^{mn}_{MN}Z^*_{M,MN}(P_{MM})^{mn}_{MN}, \tag{4.4.11-C}$$

$$(G_{EE})^{mn}_{MN} = (Q^{-1}_E P_{EE})^{mn}_{MN} = C^{mn}_{MN}Z^*_{E,MN}(P_{EE})^{mn}_{MN}, \tag{4.4.11-D}$$

$$(G_{ME})^{mn}_{MN} = (Q^{-1}_M P_{ME})^{mn}_{MN} = Z^*_{M,MN}(P_{MM})^{mn}_{MN}, \tag{4.4.11-E}$$

$$(G_{MM})^{mn}_{MN} = (Q^{-1}_M P_{MM})^{mn}_{MN} = C^{mn}_{MN}Z^*_{M,MN}(P_{MM})^{mn}_{MN}. \tag{4.4.11-F}$$

It is these terms that determine whether or not the fields are real (propagating) or imaginary (evanescent), in particular, form equation (4.4.11-E), matrix G_{ME} is always real.

In the calculation of the operator matrices P, G and K there are five general cases to be considered: $(m = M = 0) \wedge (n, N > 0)$, $(m, M > 0) \wedge (n = N = 0)$, $(m = 0, M > 0) \wedge (n, M > 0)$, $(m, M > 0) \wedge (n = 0, N > 0)$ and $(m, n, M, N > 0)$. Corresponding to these five cases there are the mappings:

$$(m = M = 0) \wedge (n, N > 0) \ : \ E_{0N} = \left((P_{EE})^{0n}_{0N}\right) E_{0n}, \tag{4.4.12-A}$$

$$(m, M > 0) \wedge (n = N = 0) \ : \ E_{M0} = \left((P_{EE})^{m0}_{M0}\right) E_{m0}, \tag{4.4.12-B}$$

$$(m = 0, M > 0) \wedge (n, M > 0) \ : \begin{cases} E_{MN} = ((P_{EE})^{0n}_{MN})\, E_{0n}, \\ H_{MN} = ((P_{ME})^{0n}_{MN})\, E_{0n}, \end{cases} \tag{4.4.12-C}$$

$$(m, M > 0) \wedge (n = 0, N > 0) \ : \begin{cases} E_{MN} = ((P_{EE})^{m0}_{MN})\, E_{m0}, \\ H_{MN} = ((P_{ME})^{m0}_{MN})\, E_{m0}, \end{cases} \tag{4.4.12-D}$$

$$(m, n, M, N > 0) \ : \begin{cases} E_{MN} = ((P_{EE})^{mn}_{MN})\, E_{mn}, \\ H_{MN} = ((P_{ME})^{mn}_{MN})\, E_{mn} \\ \qquad\qquad + ((P_{MM})^{mn}_{MN})\, H_{mn}. \end{cases} \tag{4.4.12-E}$$

4.5 Mode coupling and attenuation due to finite wall conductivity

The working assumption that the walls of the waveguide are perfectly conducting is the main reason why the scattering analysis is analytically tractable leading, for circular waveguides, to the Lommel integrals of section 4.1, Bessel series and ordinary Fourier series. In short, because of the assumed physical boundary conditions the system separates cleanly into Sturm-Liouville series with either periodic, Dirichlet or Neumann boundary conditions. The approximation is good because the conductivity is very high and the skin depth very small at the operating frequencies, but if the finite value of the guide walls is taken into account the model breaks down at two levels. The first is that the propagation of the modes derived in section 4.1 are no longer completely independent as is assumed; rather, they are coupled via the currents induced in the guide walls, and those currents induce a component of the electric field parallel to the walls which must, in turn, be the boundary values of fields over the waveguide section that must scatter at the junction. The second is that the magnetic fields over the flanges at the junctions induce electric fields parallel to the flanges.

This is the first level of correction; deeper levels can be investigated, but they are corrections of higher order. Thus, the Sturm-Liouville system no longer has the tidy separation of Dirichlet and Neumann boundary conditions. Nevertheless, it is clear that the integrands over the waveguide cross section will remain separable so that, in the case of the circular cylindrical waveguide, there will be no scattering between different azimuthal orders in a perfectly aligned waveguide. The simple form of the scattering amplitudes in equations (4.2.3) and (4.2.4) will hold only as a zeroth order approximation.

The basic theory of first order approximation to the the fields proximate to a good, but finite, conductivity wall is given by Jackson [33]. The basic idea is as follows: assume that the perfect conductor theory is a good approximation and make successive orders of correction. Only the first order of correction will be considered here where the aim is only to be able to derive estimates of the relative magnitude of these corrections.

Let $\boldsymbol{H}_{\|}$ and $\boldsymbol{E}_{\|}$ denote components of the magnetic fields at the guide wall that are everywhere parallel to the tangent plane to the wall; let $\boldsymbol{E}_{\perp}$ and $\boldsymbol{H}_{\perp}$ denote the orthogonal components and $\boldsymbol{E}_c$ and $\boldsymbol{H}_c$ the fields within the conductor. For a perfect conductor a surface current, $K = \hat{\boldsymbol{n}} \times H_{\|}$, is required to cancel the fields that would otherwise be within the conductor. Assume that the fields outside the conductor are as for the perfect conductor model: $\boldsymbol{E} \to \boldsymbol{E}_{\perp}$ and $\boldsymbol{H} \to \boldsymbol{H}_{\|}$ as the walls are approached. The model is essentially a quasi-static one and so it is assumed that the electric displacement is essentially zero. From Ohm's Law and Ampére's equation and Faraday's equation the approximations

$$\boldsymbol{E}_c \sim \frac{1}{\sigma} \nabla\times \boldsymbol{H}_c, \qquad \boldsymbol{H}_c \sim -\frac{i}{\mu\omega} \nabla\times \boldsymbol{E}_c \tag{4.5.1}$$

follow for the fields within the conductor.

Assume that, because the conductor is good and the fields must therefore decay very rapidly within the conductor in the direction $-\boldsymbol{n}$ (into the conductor), the rate of change of the fields orthogonal to the wall is much greater than parallel to it, then the nabla operator is well approximated by $\nabla \sim -\hat{\boldsymbol{n}}\, \partial/\partial r$ where r increases into the guide wall. Equations 4.5.1 then combine to give the two equations

$$\left(\frac{\partial^2}{\partial r^2} + \frac{2i}{\delta^2}\right) (\hat{\boldsymbol{n}} \times \boldsymbol{H}_c) \sim 0, \qquad \hat{\boldsymbol{n}} \cdot \boldsymbol{H}_c \sim 0,$$

where $\delta = \sqrt{2/\mu\omega\sigma}$ is the skin depth. Immediately $\boldsymbol{H}_c$, the field within the conductor, is everywhere parallel to the wall and the solution has exponential decay in the $-\hat{\boldsymbol{n}}$ direction so that $\boldsymbol{H}_c$ has the form $\boldsymbol{H}_c \sim \boldsymbol{H}_\| \, e^{-r(1-i)/\delta}$ to match the boundary value in the first order correction. Then the first of equations (4.5.1) gives

$$\begin{aligned} \boldsymbol{E}_c &\sim -\frac{1}{\sigma}\hat{\boldsymbol{n}} \times \frac{\partial \boldsymbol{H}_c}{\partial r} = -\frac{1}{\sigma}\hat{\boldsymbol{n}} \times \frac{\partial}{\partial r}\left(\boldsymbol{H}_\| \, e^{-r(1-i)/\delta}\right) \\ &= \frac{(1-i)}{\delta}(\hat{\boldsymbol{n}} \times \boldsymbol{H}_\|)e^{-r(1-i)/\delta} \\ &= (1-i)\sqrt{\frac{\omega\mu}{2\sigma}} \cdot e^{-r(1-i)/\delta} \, \hat{\boldsymbol{n}} \times \boldsymbol{H}_\| \end{aligned} \tag{4.5.2}$$

which is, once again, a field vector parallel to the walls of the waveguide. $\boldsymbol{E}_c$ is orthogonal to, out of phase with, and proportional to, $\boldsymbol{H}_c$, with a complex proportionality coefficient that goes to zero as $\sigma \to \infty$, decaying like H_c with depth of penetration, though out of phase by $-\pi/4$.

The boundary condition $\hat{\boldsymbol{n}} \times (\boldsymbol{E} - \boldsymbol{E}_c) = 0$ requires that there is a component of the electric field *parallel* to the guide wall immediately outside the wall which must match the field within the wall in the limit $r \to 0$. Thus

$$\boldsymbol{E}_\| = (1-i)\sqrt{\frac{\omega\mu}{2\sigma}} \, \hat{\boldsymbol{n}} \times \boldsymbol{H}_\| = \frac{\sqrt{2}}{\sigma\delta} \, e^{-i\pi/4} \, \hat{\boldsymbol{n}} \times \boldsymbol{H}_\| \, . \tag{4.5.3}$$

From the Ohm's Law relation $\boldsymbol{J} = \sigma \, \boldsymbol{E}_c$ the ohmic losses are derived, and in this first order correction to the model it is this electric field component that gives rise to coupling between modes that these currents can couple to. Denote by $\boldsymbol{J}_n$ the current induced by the n^{th} waveguide mode $\boldsymbol{h}_n$ (of either electric or magnetic type). Let $\boldsymbol{e}_m$ be an electric field mode of either type. The coupling of the modes within a waveguide section is determined by the integral

$$\left\langle \boldsymbol{e}_{m\|} \,|\, \boldsymbol{J}_n \right\rangle = \int_{\partial D} \boldsymbol{e}_{m\|} \cdot \boldsymbol{J}_n^* \, \mathrm{d}s \tag{4.5.4}$$

over the boundary of the guide, the purely radial function $\exp(-2r/\delta)$ having been integrated out as in equation (4.5.6) below. In the case of the circular cylinder $\partial D \cong S^1$, the unit circle, scaling the arc length measure, $\mathrm{d}s$, appropriately. This is the standard inner product on $L^2(S^1)$ and, ignoring possible frequency dependence of the conductivity and magnetic permeability, equation (4.5.4) evaluates to

$$\begin{aligned}\left\langle \boldsymbol{e}_{m\|} \,\middle|\, \boldsymbol{J}_n \right\rangle = \sigma \left\langle \boldsymbol{e}_{m\|} \,\middle|\, \boldsymbol{e}_{n\|} \right\rangle &= \Delta\omega\mu \int_{S^1} (\hat{\boldsymbol{n}} \times \boldsymbol{h}_{m\|}) \cdot (\hat{\boldsymbol{n}} \times \boldsymbol{h}^*_{n\|})\, \mathrm{d}s \\ &= \Delta\omega\mu \int_{S^1} \boldsymbol{h}_{m\|}\, \boldsymbol{h}^*_{n\|}\ \mathrm{d}s = \Delta\omega\mu \left\langle \boldsymbol{h}_{m\|} \,\middle|\, \boldsymbol{h}_{n\|} \right\rangle\end{aligned} \tag{4.5.5}$$

where Δ is an integral involving δ. Equation (4.5.5) means that whenever the $L^2(S^1)$ inner product between two magnetic field modes is non-zero, there will be coupling between the modes and the propagation of those modes will not be independent.

Evaluating equations (4.1.12-B and D) at the waveguide wall and substituting into equation (4.5.5) gives, for electric-electric field coupling,

$$\begin{aligned}\left\langle \boldsymbol{e}_{m\|} \,\middle|\, \boldsymbol{J}_n \right\rangle = \omega\mu \cdot J_m(q_{mk}) J_n(q_{nl}) \cdot \frac{mn}{q_{mk} q_{nl}} \cdot \frac{P_{E^m_k}}{Z_{E^m_k}} \cdot \frac{P_{E^n_l}}{Z^*_{E^n_l}} \cdot \int_{r=0}^{\infty} e^{-2r/\delta}\, \mathrm{d}r \\ \times \int_{\varphi=0}^{2\pi} \begin{Bmatrix} \cos(m\varphi) \\ \sin(m\varphi) \end{Bmatrix} \begin{Bmatrix} \cos(n\varphi) \\ \sin(n\varphi) \end{Bmatrix} \mathrm{d}\varphi.\end{aligned} \tag{4.5.6}$$

Here, and throughout, it is assumed that making the approximation that the waveguide walls are infinitely thick is reasonable because they are sufficiently thick that the fields within them have effectively decayed away before the outside is reached. The integral of the exponential then contributes $\delta/2$. From the orthogonality properties of the trigonometric integrands there is only coupling if $m = n$, and then the integrals give coupling factors π in the sine-sine case for $n > 0$, and 0 for $n = 0$, while for the cosine-cosine case the factors are π when $n > 0$ and 2π for $n = 0$. Thus, within this approximation, different azimuthal orders do not couple, but within an azimuthal order there are coupling factors

$$\left\langle \boldsymbol{e}_{n\|} \,\middle|\, \boldsymbol{J}_n \right\rangle = \frac{J_n(q_{nk}) J_n(q_{nl})}{\sigma\delta} \cdot \frac{n^2}{q_{nk} q_{nl}} \cdot \frac{P_{E^n_k}}{Z_{E^n_k}} \cdot \frac{P_{E^n_l}}{Z^*_{E^n_l}} \begin{cases} \pi & : n > 0,\ \sin - \sin \\ \pi & : n > 0,\ \cos - \cos \\ 2\pi & : n = 0,\ \cos - \cos \\ 0 & : \text{all other cases} \end{cases} \tag{4.5.7}$$

which, for all n, tends to zero as $k,\, l \to \infty$ and as $\sigma \to \infty$. This coupling coefficient is only real in the case $k = l$; generally the waveguide impedance renders it complex. The mathematical perspective of orthogonality between sine and cosine terms in equation (4.5.6) corresponds with the physical independence of orthogonal polarisations of the fields.

It is worth noting here that the orthogonality expressed above is a purely geometrical property – it has nothing to do with the frequency. In the general case one could consider mode coupling across the spectral band of operation. In the general case the coefficient in equation (4.5.7) would not be real; rather with frequencies $\omega_1 \neq \omega_2$ we would not necessarily have $\sigma(\omega_1) = \sigma(\omega_2)$, or even $\mu(\omega_1) = \mu(\omega_2)$, though generally $\mu = \mu_0$ for the materials from which the guide would be manufactured. The coefficient would then become

$$\left\langle \boldsymbol{e}_{n\|} \,\middle|\, \boldsymbol{J}_n \right\rangle \propto \frac{J_n(q_{nk})J_n(q_{nl})}{2\sigma(\omega_1)} \left(\frac{1}{\delta(\omega_1)} + \frac{1}{\delta(\omega_2)} \right) \cdot \frac{n^2}{q_{nk}q_{nl}} \cdot \frac{P_{E_k^n}}{Z_{E_k^n}} \cdot \frac{P_{E_l^n}}{Z^*_{E_l^n}}.$$

In the conventional TE/TM waveguide modes this expression, and equation (4.5.7), would be real and positive for propagating modes, purely positive imaginary for the coupling between propagating and evanescent modes, and negative real for evanescent to evanescent coupling. In an arbitrary basis, obtained from the standard basis by a unitary transformation, the coefficients would be a general complex number.

At first sight the possibility of mode coupling between different frequencies seems problematic – likewise for pure imaginary and negative coupling coefficients – and it is natural to ask if it can be physical. There are matters that are overlooked here and not amenable to analytic treatment. The first is that the guide walls are treated as being perfectly smooth. A real guide has some level of surface micro-roughness that will depend upon the manufacturing technique, and as the skin depth approaches the micro-roughness depth the theory must break down and attenuation be greater than predicted due to higher equivalent surface resistance. This has been confirmed by experiments reported in [5].

Equation (4.5.3) is a function on the boundary of the disc. We can assume that it is holomorphic. Then by Cauchy's theorem of complex analysis [7] it determines a field over the entire disc that is holomorphic in the two transverse coordinates. This contribution to the transverse field must scatter into the electric and magnetic field components at the other side of a junction. It would be interesting to pursue these induced fields, but the mathematics is inappropriately difficult.

4.6 Scattering between misaligned circular guides

It is common practise in the construction of horns to build the horn as an assembly of separately manufactured units. There are two reasons for doing this: ease of manufacture and the requirement to fit filters into the assembly. The Planck horns required blocking filters in the assembly and the length to diameter ratio meant that it would not have been possible to manufacture the back to back horn as a single piece. If the alignment of sections is not perfect, scattering no longer takes place only from modes of one azimuthal only into modes of the same azimuthal order at the join, but there is a scattering amplitude to all azimuthal orders. Furthermore, the separation of polarisations breaks down – orthogonal polarisations scatter into each other across an out of alignment junction – and there will be scattering of magnetic into electric fields which does not take place in the aligned waveguides.

In the case of rectangular waveguides, if two waveguide sections are slightly offset relatively to each other by some (ξ, η), but without relative rotation, then the scattering amplitude integral (4.4.4) remains analytically tractable, all be it splitting into four integrals rather than one. This is simply a consequence of writing $\cos\big(k_N(\pi^* x + A/2)\big) = \cos\big(k_N(x + A/2)\big)\cos(k_N\xi) - \sin\big(k_N(x + A/2)\big)\sin(k_N\xi)$, and so on, expanding the products and observing that the basic form of the four new scattering amplitude integrals defining scattering between modes remains the same as for the aligned case, though now with mixed terms and scale factors of the type $\cos(k_M\xi)\sin(k_N\eta)$, and so on. The mathematics remains as simple as in the aligned case. For the circular waveguide no such tractable, analytical solutions exist, and it is the circular case that is examined.

The matter is of interest to horns constructed like the Planck horns in which the cavity horn is separated from the back-to-back horn by a filter and/or lens section that gives scope for imperfect mechanical alignment of the sections. Intuitively it would then be expected that, since the functions defining the modes on either side of the junction are no longer defined on the same coordinate system, the Fourier expansions will become very complicated, this complexity being a mathematical

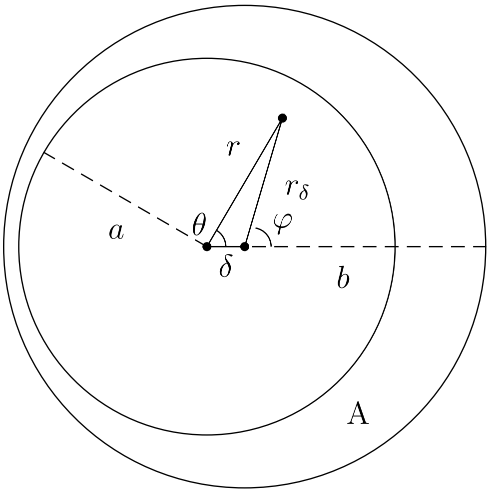


Figure 4.3: Diagram showing the coordinate systems over the domain of integration for calculation of the scattering amplitudes for misaligned waveguides. In general the displacements need not be aligned to the polarisation axes. In that case the coordinate system in the larger waveguide section is rotated and there is a mixing of polarisations.

reflection of complex inter-mode and inter-order scattering in the horn.

In this section the consequences of the misalignment of two joined guides is considered. First an analytic expression for the solution is sought; then a numerical approach is taken. The analytic expression gives an understanding of the redistribution of power between azimuthal orders, but the expressions are complex because the integrals are not separable: they are mixed in both polar coordinates. The sine and cosine terms over the displaced larger section are simple to express in terms of the angular coordinate in the smaller guide, and that expression already indicates the complexity of the scattering, but the main problem lies in the Bessel functions in the larger guide: they are now functions of radius, guide offset, and polar angle. To establish the form of the integrand for the scattering across the junction the cross product is formed to find the appropriate Poynting vector as in section 4.1. Denoting the mode of the electric field to the left by $\boldsymbol{e} = \boldsymbol{e}_r \,\hat{\boldsymbol{r}} + \boldsymbol{e}_\theta \,\hat{\boldsymbol{\theta}}$ and the magnetic mode to the right of the junction by $\boldsymbol{h} = \boldsymbol{h}_{r_\delta}\,\hat{\boldsymbol{r}}_\delta + \boldsymbol{h}_\varphi\,\hat{\boldsymbol{\varphi}}$, the cross product becomes, on the twice punctured disc $\{(r,\theta) \,:\, 0 < r \leq a,\ 0 \leq \theta < 2\pi\} - \{(\delta, 0)\}$

$$\boldsymbol{e} \times \boldsymbol{h} = \frac{1}{r_\delta}\left[(r - \delta\cos\theta)(\boldsymbol{e}_r\,\boldsymbol{h}_\varphi - \boldsymbol{e}_\theta\,\boldsymbol{h}_{r_\delta}) + \delta\sin\theta(\boldsymbol{e}_r\,\boldsymbol{h}_{r_\delta} + \boldsymbol{e}_\theta\,\boldsymbol{h}_\varphi)\right]. \tag{4.6.1}$$

This expression clearly has the correct limit as $\delta \to 0$, but will give rise to a far more complex integrand than in the aligned case. The need to exclude the point $(\delta, 0)$ from the domain is simply a reflection of the fact that polar coordinates are defined only on $\mathbb{R}^n - \{0\}$ and the coordinate transformation required to derive (4.6.1) is not defined on the axis of the displaced guide. The apparent pole in the equation is not an analytic reality, but careful structuring of the integrand is required to avoid numerical singularities.

Consider the case of magnetic to magnetic scattering. (The $T_E - T_E$ scattering is formally identical to the $T_M - T_M$ scattering described here, requiring only the formal substitution of symbols and the solution of equivalent integrals. The $T_E - T_M$ and also $T_M - T_E$ scattering integrals are also mathematically equivalent and handled by the same numerical approach. These other cases will not be described.) Expanding equation (4.6.1) gives, up to a constant $P_{M_{nm}}(a) P_{M_{nm}}(b) / Z_{M_{nm}}(b)$,

$$\begin{aligned}
\boldsymbol{e}_M \times \boldsymbol{h}_M = \frac{r - \delta \cos\theta}{r_\delta} & \Bigg[J_n'\left(\frac{p_{nk}}{a} r\right) J_m'\left(\frac{p_{ml}}{b} r_\delta\right) \begin{Bmatrix} \cos(n\theta)\cos(m\varphi) \\ \sin(n\theta)\sin(m\varphi) \end{Bmatrix} \\
& \quad + \frac{na}{p_{nk} r} \frac{mb}{p_{ml} r_\delta} J_n\left(\frac{p_{nk}}{a} r\right) J_m\left(\frac{p_{ml}}{b} r_\delta\right) \begin{Bmatrix} \sin(n\theta)\sin(m\varphi) \\ \cos(n\theta)\cos(m\varphi) \end{Bmatrix} \Bigg] \\
+ \frac{\delta \sin\theta}{r_\delta} & \Bigg[\frac{mb}{p_{ml} r_\delta} J_n'\left(\frac{p_{nk}}{a} r\right) J_m\left(\frac{p_{ml}}{b} r_\delta\right) \begin{Bmatrix} \cos(n\theta)\sin(m\varphi) \\ -\sin(n\theta)\cos(m\varphi) \end{Bmatrix} \\
& \quad + \frac{na}{p_{nk} r} J_n\left(\frac{p_{nk}}{a} r\right) J_m'\left(\frac{p_{ml}}{b} r_\delta\right) \begin{Bmatrix} -\sin(n\theta)\cos(m\varphi) \\ \cos(n\theta)\sin(m\varphi) \end{Bmatrix} \Bigg]
\end{aligned} \tag{4.6.2}$$

where $r_\delta = r_\delta(r, \delta, \theta)$ and $\varphi = \varphi(r, \delta, \theta)$ are functions of both integration variables, and each of the pairs of Bessel functions and derivatives, for the case $m \geq 2$, reduces to a sum of products of the six terms

$$\frac{p_{ml}}{4b} J_{n\pm1}\left(\frac{p_{nk}}{a} r\right) J_{m\pm2}\left(\frac{p_{ml}}{b} r_\delta\right) \quad \text{and} \quad \frac{p_{ml}}{4b} J_{n\pm1}\left(\frac{p_{nk}}{a} r\right) J_m\left(\frac{p_{ml}}{b} r_\delta\right)$$

up to constant factors. The cases $m = 0$ and $m = 1$ must be dealt with separately.

Observe that $(r-\delta\cos\theta)/r_\delta = \sqrt{1-\delta^2\sin^2\theta/(r^2+\delta^2-2r\delta\cos\theta)}$ and that

$$\delta^2\sin^2\theta/(r^2+\delta^2-2r\delta\cos\theta) \xrightarrow{r\to\delta_\pm} (1+\cos\theta)/2 = \cos^2(\theta/2)$$

Thus, $(r-\delta\cos\theta)/r_\delta \to \sin(\theta/2)$ and $\delta\sin\theta/r_\delta \to \cos(\theta/2)$ smoothly on the circle of radius δ. Furthermore since the terms involving J_m and J_m' all tend to finite limits on the same circle for all m, it follows that the scattering integrand has no singularity in its domain. Expressing the integrand in a form that is *numerically* non-singular across the entire integration domain is the dominant problem for developing a successful quadrature scheme.

An indication of what happens to the scattering is the following: For mechanical realism we can suppose that the offset is $\delta < b-a \ll a$ so that the offset is smaller than the flange width. With reference to the diagram 4.3, the usual boundary conditions and mathematical formulation apply in the guides on either side of the junction, each in the appropriate centred coordinate system. In the smaller guide of radius a the radial fields are expanded in the $J_n(\alpha r/a)$ while in the larger guide they are to be expanded in $J_m(\beta r_\delta/b)$, where α and β are roots of the Bessel function. To estimate the scattering coefficients it would be convenient to start from the series expansion of $r_\delta = \sqrt{r^2+\delta^2-2r\delta\cos\theta}$, and then to substitute the truncated series into the Bessel function, and expand that as a series. However, the series for r_δ cannot just be truncated at order δ^2 say, because while δ^2 may be small, $p_{mk}\delta^2/b$ might be large. Herein lies the main difficulty in the attempt to obtain analytic estimates: in theory it can be done, but the result is a nested sequence of infinite series in both δ^n and in $\cos^n\theta$. Abandoning that approach as impractical, the next step is to consider the Bessel functions, and then the functions $\sin(m\varphi)$ and $\cos(m\varphi)$.

The term $1/r_\delta$ in equation 4.6.1 is a potential problem in a numerical procedure, but use can be made of the product expansion (see [34])

$$J_m\left(\frac{p_{mk}}{b}r_\delta\right) = \frac{1}{m!}\left[\frac{p_{mk}}{2b}r_\delta\right]^m \prod_{l\in\mathbb{N}}\left(1-\left(\frac{p_{mk}r_\delta}{bp_{ml}}\right)^2\right)$$

as will be shown, so that the product $J_m(\cdot)\cos(m\varphi)$ has no terms in $1/r_\delta$. The scattering is dominated by integral equations of the form (4.6.4-A) on the next page which, using the above equation and equation (4.6.7-A) on page 107, take the form

$$\begin{aligned}\int_{\theta=0}^{2\pi}\int_{r=0}^{a} & J_{n\pm1}\left(\frac{p_{nj}}{a}r\right) J_{m\pm2}\left(\frac{p_{mk}}{b}r_\delta\right) r^2\cos(n\theta)\cos(m\varphi)\,\mathrm{d}r\,\mathrm{d}\theta \\ &= \frac{1}{(m\pm2)!}\left[\frac{p_{mk}}{2b}\right]^{m\pm2}\int_{\theta=0}^{2\pi}\int_{r=0}^{a} J_{n\pm1}\left(\frac{p_{nj}}{a}r\right)\cos(n\theta) \\ &\quad\times\prod_{l\in\mathbb{N}}\left(1-\left(\frac{p_{mk}}{bp_{(m\pm2)l}}\right)^2(r^2+\delta^2-2r\delta\cos\theta)\right) \\ &\quad\times(r^2+\delta^2-2r\delta\cos(\theta))^{\pm1}\sum_{i=0}^{m}(-1)^i\binom{m}{i} \\ &\quad\times r^{m-i+2}\delta^i\cos((m-i)\theta)\,\mathrm{d}r\,\mathrm{d}\theta,\end{aligned}\tag{4.6.3}$$

with equivalent integrals for the $\sin(n\theta)$ and $\sin(m\varphi)$ forms. This expression is not practical from the perspective of numerical evaluation because the product initially diverges rapidly before converging slowly. However, it does illustrate the extent of the scattering: For any m and n there are integrals of all possible moments of $J_n(p_{nj}r/a)$ and powers of $\cos\theta$. For all $\nu\in\mathbb{N}$ and $0\le i\le m$ there are terms $\cos^\nu\theta\cos(n\theta)\cos((m-i)\theta)$, where $\cos^\nu\theta=a_0+a_1\cos\theta+\cdots+a_\nu\cos(\nu\theta)$ for some rational coefficients a_j. (For even ν the odd indexed a_j will be zero while for odd ν the even indexed a_j will be zero.) For any n, m and i there will be some ν for which one of the possible terms $j-n\pm(m-i)$, for $0\le j\le\nu$, will be zero, and for that j the integral over θ in equation (4.6.3) will give π. With some thought it is seen that, for $\nu=0$, this is just $J_m\left(\frac{p_{mk}}{b}\sqrt{r^2+\delta^2}\right)$, while for $\nu\ne0$ it takes the form

$$J_{m\pm2}\left(\frac{p_{mk}}{b}\sqrt{r^2+\delta^2}\right)\cdot(2r\delta)^\nu\sum_{j=0}^{\nu}\left[a_j\cos(j\theta)\sum_{l_1\cdots l_\nu}\frac{c_{l_1}\cdots c_{l_\nu}}{\alpha_{l_1}\cdots\alpha_{l_\nu}}\right],$$

where the inner sum is over all possible multi-indices $l_1\cdots l_\nu$, $l_i\ne l_j$, and

$$\frac{c_l}{\alpha_l}\overset{\text{def}}{=}\frac{(p_{mk}/bp_{nl})^2}{(1-p_{mk}^2(r^2+\delta^2)/(bp_{nl})^2)}\xrightarrow{l\to\infty}0_+.$$

Thus, the integral (4.6.3) takes the form of a sum over j of integrals

$$\begin{aligned}&\frac{2^j\pi}{(m\pm2)!}\left[\frac{p_{mk}}{2b}\right]^{m\pm2}\sum_{i=0}^{m}(-1)^i\delta^{j+i}\binom{m}{i} \\ &\qquad\times\int_{\theta=0}^{2\pi}\int_{r=0}^{a}J_n\left(\frac{p_{nj}}{a}r\right)J_{m\pm2}\left(\frac{p_{mk}}{b}\sqrt{r^2+\delta^2}\right)\sum_{l_1\cdots l_\nu}\frac{r^{j+m-i+2}}{P_{l_1\cdots l_\nu}(r)}\,\mathrm{d}r\,\mathrm{d}\theta.\end{aligned}$$

where the $P_{l_1\cdots l_\nu}(r)$ are the polynomials of order 2ν in r given by the c_l/α_l products above. Therefore, since none of the moments of the Bessel functions are zero [6], it follows that there is scattering from all possible azimuthal orders n into all possible azimuthal orders m. Having obtained the qualitative understanding that was sought, the problem is to achieve a presentation of the integrand of practical use for numerical modelling.

To obtain estimates for the scattering amplitudes for misaligned circular waveguides, the problem becomes one of solving non-separable integrals of the following three general types obtained by expanding the terms of equation (4.6.2):

$$\int_{\theta=0}^{2\pi}\int_{r=0}^{a} J_{n\pm1}\left(\frac{p_{nk}}{a}r\right) J_{m\pm2}\left(\frac{p_{ml}}{b}r_\delta\right) r^2\, cs_n(\theta)cs_m(\varphi)\,\mathrm{d}r\,\mathrm{d}\theta \tag{4.6.4-A}$$

$$\delta\int_{\theta=0}^{2\pi}\int_{r=0}^{a} J_{n\pm1}\left(\frac{p_{nk}}{a}r\right) J_{m\pm2}\left(\frac{p_{ml}}{b}r_\delta\right) r\, cs(\theta)cs_n(\theta)cs_m(\varphi)\,\mathrm{d}r\,\mathrm{d}\theta \tag{4.6.4-B}$$

$$\delta\int_{\theta=0}^{2\pi}\int_{r=0}^{a} J_{n\pm1}\left(\frac{p_{nk}}{a}r\right) J_{m}\left(\frac{p_{ml}}{b}r_\delta\right) r\, cs(\theta)cs_n(\theta)cs_m(\varphi)\,\mathrm{d}r\,\mathrm{d}\theta \tag{4.6.4-C}$$

where the functions $cs(\theta)$, $cs_n(\theta)$ and $cs_m(\varphi)$ are the appropriate sines and cosines. As will be seen shortly, the functions $J_m(p_{ml}r_\delta/b)cs_m(\varphi)$ and $J_{m+2}(p_{ml}r_\delta/b)cs_m(\varphi)$ can be written in a form that is numerically stable and, though algebraically complex, reasonably straightforward to compute. However, the remaining function, $J_{m-2}(p_{ml}r_\delta/b)cs_m(\varphi)$, has a second order pole at the centre of the offset waveguide that arises from the term $J_0(p_{ml}r_\delta/b)/r_\delta^2$ (for all $m > 0$) occurring in the presentation of the function. This is physically meaningless as well as numerically unstable and the quadrature scheme needs to handle this in a simple, stable and non-disruptive manner. Devising a presentation of the integrand that is completely stable represents the main challenge for work in this section. A method is presented below.

In the particular case $m = 0$ only the derivatives of $J_0(p_{0k}r_\delta/b)$ occur in equation (4.6.2), the terms $\cos(0\varphi)$ and $\sin(0\varphi)$ give 1 and 0 respectively, and the following substitution can be made:

$$\frac{J_0'\left(\dfrac{p_{0k}}{b}r_\delta\right)}{r_\delta} = -\frac{p_{0k}}{2b}\left(J_0\left(\frac{p_{0k}}{b}r_\delta\right) + J_2\left(\frac{p_{0k}}{b}r_\delta\right)\right).$$

This gives numerically stable integrals for the scattering amplitudes for the azimuthal order $m = 0$ in the right hand waveguide, the lower line in equation (4.6.2) giving zero, while the upper line can be rearranged to give the integral

$$\begin{aligned}\frac{p_{0k}}{4b}\int_{\theta=0}^{2\pi}\int_{r=0}^{a}&\left(J_0\left(\frac{p_{0k}}{b}r_\delta\right)+J_2\left(\frac{p_{0k}}{b}r_\delta\right)\right)\\ &\times\Bigg[J_{n+1}\left(\frac{p_{nl}}{a}r\right)(r\cos(n\theta)-\delta\cos((n-1)\theta))\\ &\qquad -J_{n-1}\left(\frac{p_{nl}}{a}r\right)(r\cos(n\theta)-\delta\cos((n+1)\theta))\Bigg]\,r\,\mathrm{d}r\,\mathrm{d}\theta\end{aligned}\tag{4.6.5}$$

From figure 4.3 it is clear that for $r_\delta > 0$, $\cos\varphi = (r\cos\theta - \delta)/r_\delta$ and $\sin\varphi = (r\sin\theta)/r_\delta$. Then, for the other particular case, $m = 1$, observe that

$$\frac{r-\delta\cos\theta}{r_\delta}\cos\varphi = \cos\theta - \frac{\sin^2\theta}{2}\left(\frac{r^2+\delta^2}{2r\delta}-\cos\theta\right)^{-1},\tag{4.6.6-A}$$

$$\frac{r-\delta\cos\theta}{r_\delta}\sin\varphi = \sin\theta\left(1-\left[1+\frac{r(r-\delta\cos\theta)}{\delta(\delta-r\cos\theta)}\right]^{-1}\right),\tag{4.6.6-B}$$

$$\frac{\delta\sin\theta}{r_\delta}\cos\varphi = -\sin\theta\left(1-\left[1+\frac{\delta(\delta-r\cos\theta)}{r(r-\delta\cos\theta)}\right]^{-1}\right),\tag{4.6.6-C}$$

$$\frac{\delta\sin\theta}{r_\delta}\sin\varphi = \frac{\sin^2\theta}{2}\left(\frac{r^2+\delta^2}{2r\delta}-\cos\theta\right)^{-1}.\tag{4.6.6-D}$$

These will be numerically stable as $r_\delta \to 0$, approaching the correct limits (one for (4.6.6-A) and zero for equations B to D) smoothly and monotonically, and the exceptional point on the offset axis, $r_\delta = 0$, will be handled provided the compiler interprets, or can be instructed to interpret, 0/0 as 0, or a switch or IEEE exception handles the overflow. Then with these forms used in equation (4.6.2) the computation will proceed stably across the entire disc.

The general case, $m \geq 2$ has now to be dealt with, and to handle the general case requires extensive rewriting of the functions under the integral. The orthogonality of the modes of different azimuthal orders in aligned guides is expressed by the orthogonality of the sine and cosine terms. Here the situation is complicated by the fact that the azimuthal components of the basis functions over the smaller guide are expanded in terms of $\cos(\theta)$, $\sin(\theta)$, $\cos(n\theta)$ and $\sin(n\theta)$, while those over the larger section involve the terms $\cos(m\varphi)$ and $\sin(m\varphi)$. By induction on m, these become

$$\cos(m\varphi) = \frac{1}{r_\delta^m}\sum_{k=0}^{m}(-1)^k\binom{m}{k}r^{m-k}\delta^k\cos((m-k)\theta),\tag{4.6.7-A}$$

$$\sin(m\varphi) = \frac{1}{r_\delta^m}\sum_{k=0}^{m-1}(-1)^k\binom{m}{k}r^{m-k}\delta^k\sin((m-k)\theta),\tag{4.6.7-B}$$

Here it has been assumed that the offset is aligned to the polarisation axis, as in figure 4.3 above. More generally it will be offset at $(\delta\cos(\theta_0), \delta\sin(\theta_0))$ for some angle θ_0. The result is that φ is replaced by $\varphi+\theta_0$, $\cos(\varphi+\theta_0) = r\sin(\theta-\theta_0)/r_\delta - \delta\sin(\theta_0)/r_\delta$, and so on, so that δ in equation (4.6.7) and in all that follows would replaced by $\delta\sin(\theta_0)$. Since that is a trivial change to the computation and makes no difference to the physics it is ignored in what follows. Note that it is convenient to extend the upper limit in the sum for $\sin(m\varphi)$ from $m-1$ to m, which just adds zero.

The problematic term $1/r_\delta^m$ that, at first sight, gives rise to a pole of order m on the axis of the offset waveguide, is handled by observing that for the Bessel function of order s in r_δ

$$\begin{aligned}
\frac{1}{r_\delta^m} J_s\left(\frac{p_{sk}}{b} r_\delta\right) &= \frac{r_\delta^{s-m}}{r_\delta^s} J_s\left(\frac{p_{sk}}{b} r_\delta\right) \\
&= \frac{r_\delta^{s-m}}{r_\delta^{s-1}} \cdot \frac{p_{sk}}{2b} \cdot \frac{1}{s}\left[J_{s-1}\left(\frac{p_{sk}}{b} r_\delta\right) + J_{s+1}\left(\frac{p_{sk}}{b} r_\delta\right)\right] \\
&= \frac{r_\delta^{s-m}}{r_\delta^{s-2}} \cdot \left[\frac{p_{sk}}{2b}\right]^2 \cdot \frac{(s-2)!}{(s+1)!}\Big[(s+1) J_{s-2}\left(\frac{p_{sk}}{b} r_\delta\right) \\
&\qquad + 2s J_s\left(\frac{p_{sk}}{b} r_\delta\right) + (s-1) J_{s+2}\left(\frac{p_{sk}}{b} r_\delta\right)\Big] \\
&\qquad \vdots \\
&= \left[\frac{p_{sk}}{2b}\right]^s \frac{r_\delta^{s-m}}{(2s-1)!} \cdot \sum_{l=0}^{s} A_{sl} J_{2l}\left(\frac{p_{sk}}{b} r_\delta\right),
\end{aligned} \tag{4.6.8}$$

where the terms A_{sl} are positive integer coefficients derived by induction in the s steps required to eliminate the factor $1/r_\delta^s$ at the apparent singularity. Pre-calculating the coefficients in the last line of equation (4.6.8) outside the quadrature loop, the standard recursion algorithm for the Bessel function $J_n(x)$ for $n \geq 2$ can be modified to take the coefficient vector as an additional argument and to calculate the entire right hand expression in equation (4.6.8) with only m additional floating point additions and multiplications. Consequently, although the expression looks complex, from the algorithmic point of view it is scarcely more complex than calculating the left hand side of equation (4.6.8), but is numerically stable everywhere.

From equations (4.6.7) and (4.6.8) we then have the product of the Bessel function in r_δ times the sines or cosine of φ under the integral in equations (4.6.4-A to

C) taking the form

$$\begin{aligned} J_s\left(\frac{p_{mk}}{b}r_\delta\right) cs_m(\varphi) = \left[\frac{p_{mk}}{2b}\right]^s \frac{r_\delta^{s-m}}{(2s-1)!} \cdot \sum_{l=0}^{s} A_{sl} J_{2l}\left(\frac{p_{mk}}{b}r_\delta\right) \\ \times \sum_{i=0}^{m}(-1)^i \binom{m}{i} r^{m-i}\delta^i\, cs_{m-i}(\theta). \end{aligned} \tag{4.6.9}$$

The integral now has a non-singular presentation and can be evaluated stably providing $s \geq m$. From the numerical perspective the difficulty arises only for the case $m-2$. Re-writing equation (4.6.9) for this case gives

$$\begin{aligned} J_{m-2}\left(\frac{p_{mk}}{b}r_\delta\right) cs_m(\varphi) &= \left[\frac{p_{mk}}{2b}\right]^{m-2} \frac{1}{(2m-5)!} \cdot \sum_{l=0}^{m-2} \frac{A_{m-2,l}}{r_\delta^2} J_{2l}\left(\frac{p_{mk}}{b}r_\delta\right) \\ &\quad \times \sum_{i=0}^{m}(-1)^i \binom{m}{i} r^{m-i}\delta^i\, cs_{m-i}(\theta) \\ &= \left[\frac{p_{mk}}{2b}\right]^{m} \frac{1}{(2m-5)!} \Bigg\{ \sum_{l=1}^{m-2} \frac{A_{m-2,l}}{(2l+1)(2l)(2l-1)} \Bigg[(2l+1)J_{2l-2}\left(\frac{p_{mk}}{b}r_\delta\right) \\ &\quad + 2lJ_{2l}\left(\frac{p_{mk}}{b}r_\delta\right) + (2l-1)J_{2l+2}\left(\frac{p_{mk}}{b}r_\delta\right)\Bigg] \\ &\qquad + \frac{4b^2}{p_{nk}^2}\, \frac{A_{m-2,0}}{r_\delta^2} J_0\left(\frac{p_{mk}}{b}r_\delta\right) \Bigg\} \cdot \sum_{i=0}^{m}(-1)^i \binom{m}{i} r^{m-i}\delta^i\, cs_{m-i}(\theta), \end{aligned} \tag{4.6.10}$$

and from the last line it is seen that the pole arises only from the term $J_0\left(\frac{p_{m,k}}{b}r_\delta\right)/r_\delta^2$ which is not present for analytic reasons. Because of this term the otherwise stable presentation of the integrand will give rise to numerical overflow. It will be removed in a simple way below.

When $m \geq 2$ the terms in equation (4.6.2) can be expanded, reorganised and gathered and, using equations (4.6.7) and for $n \geq 0$, $|m| \geq 2$ and $k \in \{0,1\}$, to write

$$c(n,m,k,\theta) \overset{\text{def}}{=} \sum_{i=0}^{m}(-1)^i \binom{|m|}{i} r^{|m|-i}\delta^i \cos\left((n+m+k-i)\theta\right),$$

then using equations (4.6.9) and (4.6.10) six integrands are obtained:

$$\begin{aligned} I_{m+2}^{\mp} = \frac{r}{8(m+1)} \left[\frac{p_{mk}}{2b}\right]^{m+2} \frac{r^2+\delta^2-2r\delta\cos\theta}{(2m+3)!} \sum_{l=0}^{m+2} A_{m+2,l} J_{2l}\left(\frac{p_{mk}}{b}r_\delta\right) \\ \times \left\{ J_{n-1}\left(\frac{p_{nj}}{a}r\right) \left[r\, c(n,-m,0,\theta) - \delta\, c(n,-m,1,\theta)\right] \right. \end{aligned}$$

$$\mp J_{n+1}\left(\frac{p_{nj}}{a}r\right)\left[r\,c(n,m,0,\theta)-\delta\,c(n,m,1,\theta)\right]\Big\} \tag{4.6.11-A}$$

$$\begin{aligned}I_m^{\mp} &= \frac{r}{4(m-1)}\left[\frac{p_{mk}}{2b}\right]^m\frac{1}{(2m-1)!}\sum_{l=0}^{m}A_{m,l}J_{2l}\left(\frac{p_{mk}}{b}r_\delta\right)\\ &\times\Big\{J_{n-1}\left(\frac{p_{nj}}{a}r\right)\left[r\,c(n,-m,0,\theta)-\delta\,c(n,-m,1,\theta)\right]\\ &\qquad\mp J_{n+1}\left(\frac{p_{nj}}{a}r\right)\left[r\,c(n,m,0,\theta)-\delta\,c(n,m,1,\theta)\right]\Big\}\end{aligned} \tag{4.6.11-B}$$

$$\begin{aligned}I_{m-2}^{\mp} &= \frac{r}{8(m-1)r_\delta^2}\left[\frac{p_{mk}}{2b}\right]^{m-2}\frac{1}{(2m-5)!}\sum_{l=1}^{m-2}A_{m-2,l}J_{2l-2}\left(\frac{p_{mk}}{b}r_\delta\right)\\ &\times\Big\{J_{n-1}\left(\frac{p_{nj}}{a}r\right)\left[r\,c(n,-m,0,\theta)-\delta\,c(n,-m,1,\theta)\right]\\ &\qquad\mp J_{n+1}\left(\frac{p_{nj}}{a}r\right)\left[r\,c(n,m,0,\theta)-\delta\,c(n,m,1,\theta)\right]\Big\}\end{aligned} \tag{4.6.11-C}$$

These integrands are then summed as $I^- = I^-_{m+2}+I^-_m+I^-_{m-2}$ and $I^+ = I^+_{m+2}+I^+_m+I^+_{m-2}$ to give the upper and lower polarisation lines in equation (4.6.2).

Some comments need to be made about these equations: Firstly, in (4.6.11-A) it is not strictly necessary to reduce $J_{m+2}(p_{mk}r_\delta/b)$ to a sum of terms $J_{2l}(p_{mk}r_\delta/b)$, $l=0,\ldots,m+2$; instead the same procedure could be halted after m steps, sufficient to remove the $1/r_\delta$ factor and giving a reduction to a sum of terms $J_{2l}(p_{mk}r_\delta/b)$, $l=1,\ldots,m$ and so avoid the r_δ^2. To do so would require developing a second algorithm for no marked gain in computational efficiency. Secondly, all six integrands are now numerically stable as both as $r\to 0$ and as $r_\delta \to 0$ provided that, in $I^{\mp}_{m-2}$ r/r_δ is written as

$$\frac{r}{r_\delta^2} = \begin{cases}\dfrac{r}{r^2+\delta^2-2r\delta\cos\theta}\xrightarrow{r\to 0} 0 & : \quad r\le \delta/2,\\[2ex] \dfrac{1}{\delta/r+r-2\delta\cos\theta}\xrightarrow{r_\delta\to 0}\dfrac{1}{1-\delta} & : \quad r>\delta/2.\end{cases}$$

Calculation of the coefficients A_{ml} is performed inductively in m steps for each m following, in outline, the sequence of equation (4.6.8). It is a common problem with factorial and related numerical integer arithmetic problems such as the calculation of the A_{ml} that the values rapidly exceed the representable integers. If 32 bit arithmetic is used then the maximum azimuthal index that can be handled is $m=9$; beyond that overflow occurs unless 64 bit integers are used, but even then $m=14$ is all that can be achieved without overflow. For the Planck multi-mode horns, where the maximum azimuthal order is four, 32 bit arithmetic would be sufficient for the investigation of alignment tolerances. Table 4.1 on page 117, gives the coefficients A_{ml} for the first ten azimuthal orders. Table 4.2 gives the corresponding coefficients

$$C_{ml} \overset{\text{def}}{=} \frac{m!}{2^m(2m-1)!} A_{ml}. \tag{4.6.12}$$

but the original, singular, non-separable integral has been replaced by a sum of non-separable integrals over a sum: from equation (4.6.8)

$$\begin{aligned}
&\int_{\theta=0}^{2\pi}\int_{r=0}^{a} J_n\left(\frac{p_{nj}}{a}r\right) J_m\left(\frac{p_{mk}}{b}r_\delta\right) r\cos(n\theta)\cos(m\varphi)\,\mathrm{d}r\,\mathrm{d}\theta \\
&= \begin{cases}
\displaystyle\int_{\theta=0}^{2\pi}\cos(n\theta)\int_{r=0}^{a} J_n\left(\frac{p_{nj}}{a}r\right) J_0\left(\frac{p_{mk}}{b}r_\delta\right) r\,\mathrm{d}r\,\mathrm{d}\theta & : \ m=0 \\[2ex]
\displaystyle\frac{p_{mk}}{2b}\int_{\theta=0}^{2\pi}\cos(n\theta)\int_{r=0}^{a} J_n\left(\frac{p_{nj}}{a}r\right)\left(J_0\left(\frac{p_{mk}}{b}r_\delta\right)+J_2\left(\frac{p_{mk}}{b}r_\delta\right)\right) & : \ m=1 \\
\qquad \times\, r\left(r\cos(\theta)-\delta\right)\,\mathrm{d}r\,\mathrm{d}\theta & \\[2ex]
\displaystyle\left[\frac{p_{mk}}{2b}\right]^m \frac{1}{(2m-1)!}\sum_{i=0}^{m}(-1)^i\binom{m}{i}\delta^i\int_{\theta=0}^{2\pi}\cos(n\theta)\cos((m-i)\theta) & \\
\qquad \displaystyle\times\int_{r=0}^{a} r^{m-i+1}J_n\left(\frac{p_{nj}}{a}r\right)\sum_{l=0}^{m}A_{ml}J_{2l}\left(\frac{p_{mk}}{b}r_\delta\right)\,\mathrm{d}r\,\mathrm{d}\theta & : \ m\geq 2.
\end{cases}
\end{aligned} \tag{4.6.13}$$

The normalisation factors $P_{M_j^n}$ and $P_{M_k^m}$ from equation (4.1.10) have been omitted in the above integral. It is immediate from equations (4.6.7) and (4.6.8) that the right hand expressions are analytically correct in the limit $\delta \to 0$. Therefore, the validity and accuracy of any quadrature scheme written to evaluate (4.6.13) can be tested by evaluation at $\delta = 0$ and comparing with the known analytic value (in particular, for $a = b$, it should return the value one). There is an equivalent expression for the sine-sine scattering and, unlike in the case of the aligned waveguides, also one for sine-cosine scattering because there is no orthogonality of the integrands.

With numerically stable forms for the integrals described, the final problem is to develop a suitable quadrature scheme for their evaluation. First a standard Gaussian quadrature scheme was coded. Such schemes are relatively straightforward to encode, but the sampling of the integration domain follows the coordinate system on the domain and does not take into account the fluctuations and values of the integrand. In this case the domain is the half disc (see figure 4.4 below). This is particularly the case for Gaussian quadrature in polar coordinates where the sampling is very fine at the centre of the disc and where, for these integrals, the integrand is essentially zero. The amplitude of the integrand is mostly very close to zero except in a few well defined regions, and it is also oscillatory in both polar coordinates. Consequently a standard quadrature scheme spends most of the time evaluating and summing very small contributions to the scattering amplitude. Thus, though the Gaussian quadrature scheme worked, it was inefficient in its use of memory and very slow, particularly for higher azimuthal and radial orders. It was therefore abandoned in favour of an adaptation of a markedly different approach that had been developed by the author to overcome similar sampling issues in the evaluation of integrals occurring in Physical Optics analysis.

When Physical Optics is used for the analysis of optical systems integrands of the type given in equations (6.3.3) on page 162 have to be solved, integrating the source function for each field point, $\boldsymbol{x}'$, at which the resulting field is to be found. The integrand there is a function of the field point as well as the source, and so it changes with $\boldsymbol{x}'$. Consequently, even in the simple situation of a single-mode Gaussian horn aperture field as source, which upon first sight is perfectly suited to Gaussian quadrature in polar coordinates, Gaussian quadrature is inefficient for calculating the field at points off the horn axis. In [54] an self-adaptive quadrature scheme was developed that used the integrand values and/or derivatives to determine the appropriate local sampling density for the quadrature. Essentially, if the integrand is changing rapidly in a region of the domain, and there is power in the field there, then sampling is fine; where the rate of change is low coarse sampling is used. In Physical optics the situation is complicated by the phase, but for the evaluation of the integrals in equation (4.6.4) there is no phase information and all that matters are the local gradient and absolute value of the integrand.

The evaluation of the integrals of type (4.6.13) was done in the following way. Since the purpose of the procedure is to investigate, quantitatively, the effects of inter-modal scattering for misaligned waveguide junctions, assume for the sake of simplicity that the displacement is along the x-axis of the coordinate system (polar angle $\theta = 0$). Then the integration domain is reduced to a half disc, $D = \{(r, \theta) : 0 \leq r \leq a,\ 0 \leq \theta \leq \pi\}$, and the quadrature procedure begins with a coarse Delaunay triangulation $\mathcal{T}$ of constant scale over a point set P in the closure of D. (By coarse is meant that the distances between adjacent points of P is roughly four times the minimum distance between zeroes of the highest order Bessel function occurring in the integrand divided by the radius of the waveguide, though the exact value is not very important; by constant scale is meant that the local mesh scale is everywhere approximately the same, subject to the constrains of domain geometry.) A Delaunay triangulation is a triangulation of a point set, P, in which the circumcircle of no triangle contains a point of P. For technical details see [17] or texts on computational geometry. In particular the centre of the disc $(0, 0)$ and the two points $(a, 0)$ and (a, π) are in P. Each triangle will become a quadrature sub-domain, the final integral being the sum of the sub-integrals. Let there be N triangles in $\mathcal{T}$ and denote the set of all triangles in $\mathcal{T}$ by $\{T_n\}_{n=1}^{N}$. Because the triangulation is Delaunay the triangles are all roughly of the same shape and size at this stage. Let T_S be the standard triangle of side length one subdivided into 16 regular sub-triangles (i.e. its vertices is set $\{(0, 0),\ (1, 0),\ (0, 1)\}$ with each side subdivided into four equal lengths by the addition of three points, and the sub-triangles are obtained by drawing lines between the nine added points and adding an additional point at the intersections of the lines. The points will hereafter be referred to as the nodes). Let $\tau : T_n \to T_S$ be the natural mapping, then each T_n is sub-triangulated by pulling back the triangulation from T_S. This ensures that the preliminary quadrature on all adjoining triangles shares common vertices on their common boundary.

With the preliminary triangulation established the next step is to establish a quadrature order for the set $\mathcal{T}$: evaluate the integrand at each of the 12 nodes in each $T_n \in \mathcal{T}$ and integrate over each T_n using Simpson's rule for each of the 16 sub-triangles. If I_n is the approximate integral over T_n, then impose an order in $\{T_n\}$, re-indexing the triangles so that $I_1 \geq I_2 \geq \cdots \geq I_N$. In the subsequent quadrature procedure the convergence of the quadrature over the T_n will proceed in this established order.

The quadrature proper over T_1 proceeds by first estimating the local gradient at each of the 12 nodes using finite differences. Number the nodes $n_1, \ldots, n_{12}$ and let the gradient and the absolute value of the integrand at the k^{th} node be g_k and v_k respectively. Let the mean length of all triangle edges meeting at n_k be d_k, then the real number $h(k) = d_k e^{b g_k v_k} \geq d_k$ is the value of a mesh distribution function (MDF), h, over T_n that extends to an MDF over P because the values of h on the common nodes on the boundaries of adjacent T_n agree. The purpose of h is to control the subsequent refinements of the quadrature domains $T_n \in \mathcal{T}$; for technical details see [35]. Note that if either the gradient or the value of the integrand is zero at the k^{th} node of T_n, then $h(k) = d_k$, otherwise $d_k < h(k)$. The real number b in the expression for h is a forcing factor that can be used to give more aggressive refinement of the quadrature domain if desired.

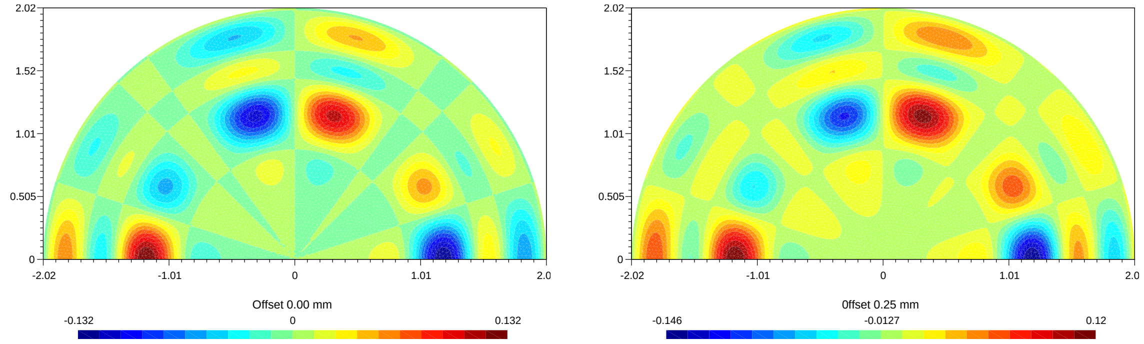
With h established for T_1 a new Delaunay triangulation of T_1 is generated under the control of h. The purpose of the MDF is to ensure that, where there is no change in value of the integrand the triangulation scale remains essentially unchanged, but where the value is changing rapidly the local scale of the triangulation is changed to give a finer triangulation. Thus, the quadrature sub-domains over T_1 will reflect the form of the integrand locally. Let the new triangulation of T_1 be $\{T_{1k}\}_{k=1}^{K}$, then there is a refined quadrature estimate over T_1 given by applying Simpson's rule over each T_{1k}; denote this new estimate $I_1^{(1)}$ and let $\varepsilon_1^{(1)} = |I_1^{(1)} - I_1|$.

Suppose that the overall quadrature error required is to be not greater than ε; then since there are N triangles we require errors $\varepsilon_n \leq \varepsilon/N$ for each sub-quadrature. If $\varepsilon_1^{(1)} \leq \varepsilon/N$ then the quadrature has converged and the process moves on the quadrature over T_2, and so on. Generally it will not have converged and the next refinement $\{T_{1k}\}$ by addition of an additional nodes at the weighted centre of each edge: if e_j is the edge with nodes n_{j1} and n_{j2} at its ends, and $h(n_{j1}) > h(n_{j2})$ then the new node is placed at the centre of mass of the two nodes and will be closer to n_{j1} than to n_{j2}. Repeat the quadrature over this new refinement to get a value $I_1^{(2)}$ and error estimate $\varepsilon_1^{(2)} = |I_1^{(2)} - I_1^{(1)}|$. Proceed making further refinements until at the m^{th} refinement we have $\varepsilon_1^{(m)} = |I_1^{(m)} - I_1^{(m-1)}| \leq \varepsilon/N$; the value $I_1^{(m)}$ becomes the estimate for the integral over T_1. Denoted this final value by I'_1 and its error estimate ε'_1.

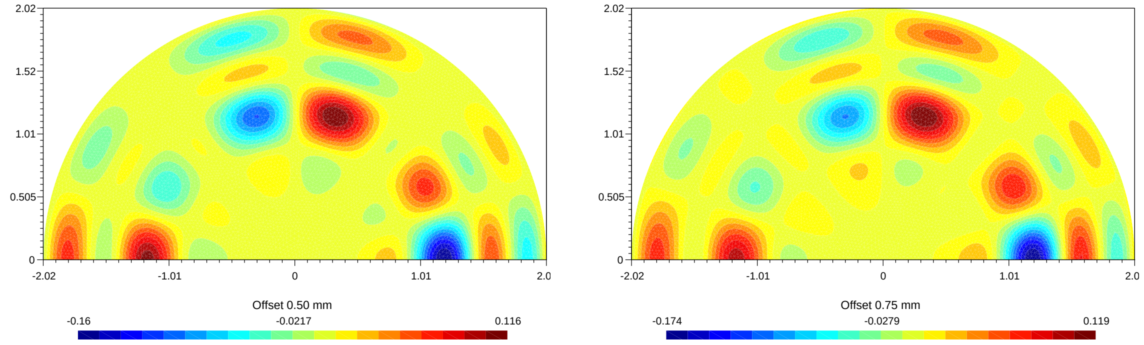
Next proceed to seek convergence over T_2. Since $I_2 \leq I_1$ we can expect convergence in no more than m steps if the integrand is everywhere quite well behaved over T_1 and T_2 and the preliminary sampling scale for P chosen as stated. In any case the convergence procedure will find the appropriate stopping point. Generally, for the quadrature over $T_2, T_3, \ldots, T_N$ convergence can be expected to become faster.

The final estimate for the inter-modal scattering integrand is then $I = \sum_{n=1}^{N} I'_n$ and the estimated error will be $\sum_{n=1}^{N} \varepsilon'_n \leq N \cdot \varepsilon/N = \varepsilon$. This procedure, though relatively complex, has proven to be faster than standard Gaussian quadrature. The main complexity is in the establishment of the preliminary MDF and the triangulation refinement. However, these are all well established procedures in computational geometry and algorithms and code are freely available. It would be interesting to pursue a more sophisticated adaptive meshing procedure using anisotropic meshes generated by the metric $(\det H)^{-1/2(p+n)}H$, where H is the Hessian of the integrand function, as presented in [12]. An anisotropic mesh would be stretched in the direction of least local change in the integrand, thus reducing the overall number of samples in an appropriate way. The triangulation would no longer be Delaunay, but that would not affect the efficacy of the procedure.

Although code has been developed to evaluate the integrals for the TM-TM scattering case this mater has not been taken further in this work. The interest lies in the qualitative investigation of the effects of waveguide section misalignment on mode suppression and excitation as power is scattered between modes of the same azimuthal order with amplitudes different from the aligned case, and scattering between different azimuthal orders that either propagate through to the sky or are reflected back, or are evanescent. In all cases some influence on both throughput and beam shape is to be expected. The full coding and testing of the necessary code, and its incorporation into the mode matching software, would be a more complex undertaking than the development of the mode matching software itself. All that has been done here is to set out an approach to the development of such code, finding a practical method of performing the quadrature essential to obtaining reliable results.



(a) Offset 0.00 (b) Offset 0.25



(c) Offset 0.50 (d) Offset 0.75

Figure 4.4: The amplitude of the integrand over the integration domain in equation (4.6.4) for the scattering of azimuthal order 2, radial order 3 into azimuthal order 5, radial order 3 with the offsets ranging from 0.00 mm to 0.75 mm. These offsets have been set unrealistically large to illustrate the changes of the scattering integrand with increasing offset. The waveguide radii are 2.0 and 3.0 mm respectively. Note that the symmetries of figure 4.4(a) are generated by the reflections in the x and y axes, while the symmetries of the fields with offset waveguides are generated by the reflection in the x axis only, thus limiting the reduction in the integrand over polar angle to $[0, \pi)$; hence the half disc domains. Note: these are maps of the integrands (real functions) and are dependent purely upon the geometry of the waveguide and independent of the frequency.

Table 4.1: The values of the first ten coefficients A_{ml}, for azimuthal orders 0 through to 9, calculable using 32 bit arithmetic without overflow.

1	1	1									
2	3	4	1								
3	20	30	12	2							
4	210	336	168	48	6						
5	3024	5040	2880	1080	240	24					
6	55440	95040	59400	26400	7920	1440	120				
7	1235520	2162160	1441440	720720	262080	65520	10080	720			
8	32432400	57657600	40360320	22014720	9172800	2822400	604800	80640	5040		
9	980179200	1764322560	1283143680	748500480	345461760	123379200	32901120	6168960	725760	40320	
10	33522128640	60949324800	45711993600	28130457600	14065228800	5626091520	1758153600	413683200	68947200	7257600	362880

Table 4.2: The values of the first ten coefficients $C_{ml} \overset{\text{def}}{=} \frac{m!}{2^m(2m-1)!}A_{ml}$. In the integral, these are further scaled by factors of $\delta^i/i!(m-i)!$, so that only the integrals associated with small l and i contribute significantly to the integral.

$m \setminus l$	0	1	2	3	4	5	6	7	8	9	10
1	1.000000E+00	1.000000E+00									
2	5.000000E-01	6.666667E-01	1.666667E-01								
3	2.500000E-01	3.750000E-01	1.500000E-01	2.500000E-02							
4	1.250000E-01	2.000000E-01	1.000000E-01	2.857143E-02	3.571429E-03						
5	6.250000E-02	1.041667E-01	5.952381E-02	2.232143E-02	4.960317E-03	4.960317E-04					
6	3.125000E-02	5.357143E-02	3.348214E-02	1.488095E-02	4.464286E-03	8.116883E-04	6.764069E-05				
7	1.562500E-02	2.734375E-02	1.822917E-02	9.114583E-03	3.314394E-03	8.285985E-04	1.274767E-04	9.105478E-06			
8	7.812500E-03	1.388889E-02	9.722222E-03	5.303030E-03	2.209596E-03	6.798757E-04	1.456876E-04	1.942502E-05	1.214064E-06		
9	3.906250E-03	7.031250E-03	5.113636E-03	2.982955E-03	1.376748E-03	4.916958E-04	1.311189E-04	2.458479E-05	2.892328E-06	1.606849E-07	
10	1.953125E-03	3.551136E-03	2.663352E-03	1.638986E-03	8.194930E-04	3.277972E-04	1.024366E-04	2.410274E-05	4.017123E-06	4.228550E-07	2.114275E-08

Chapter 5

Numerical implementation of mode-matching

Prior to this work the software available for the modelling of the Planck multi-mode horns had been the code developed by Murphy at the National University of Ireland Maynooth, [43], and later versions that retained the same computational core. To make a serious attempt at the broad band modelling of the Planck multi-mode pixels, from the cavity to the radiating aperture, and to be able both to derive realistic broad band beam patterns and to investigate the spectral properties of the scattering within the structure, new and much more efficient software was required. At the most trivial level the formulae for the scattering coefficients used in the original code and given in [43], [14] and [25] needed to be written for efficient computation. Doing so would give improved accuracy, but only minor gains in speed. The form of the scattering coefficient equations that have been used here are those given in equations (4.2.3) and (4.2.4). A significant gain in the speed of computation would only be achieved with a complete restructuring of the code and new algorithms. At the same time it had become evident that whatever code was written it needed to have built into it run-time checks on the results at every scattering step. Checks were needed for numerical stability of the matrix operations because these can become ill conditioned; without stability checks the matrix operations could become close to being numerically singular without the code failing. Further checks were needed to ensure that the model remained physically realistic at every step. A physically realistic model requires a model space of sufficient dimension to account for essentially all power at all scattering steps so that the finite dimension of the model does not lead to significant information loss. In either case, without suitable run-time checks the computer will produce output that can appear to be a valid electric field structure but is physically unrealistic; sometimes wildly unrealistic.

The computational methods described here exploit the relative simplicity of the scattering equations as written in matrix (3.3.2) over the familiar matrix equations (3.2.10). The aim is both speed and accuracy gains over straightforward coding of the matrix components in any of the forms given in equations (3.3.1). The gains fall into two kinds: major gains in both speed and accuracy, and minor gains in speed. Both are discussed because the long term aim (beyond the scope of this work) is to produce code that can be used for broad-band optimisation of complex systems: systems like the Planck multi-mode horns, smooth walled or hybrid equivalents, and more general structures. In all of the sections describing the computational scheme the general case is covered, applicable to all azimuthal orders. For order zero the off diagonal block P_{ME} is identically zero which considerably simplifies the scheme, a simplification that is exploited by writing a special set of algorithms to handle the order zero case. Since the modifications to what follows are obvious they are not discussed further, but with their implementation the azimuthal order zero case takes half the computation time of the other orders.

5.1 Preliminaries: computational error

The major problem with any complex numerical computation is that there are a plethora of sources of error that will accumulate and render the computation invalid if they are not controlled or detected. Furthermore, it is useful to be able to estimate the overall error in the computed result so that validity of the result can be judged. In the modelling of scattering in centred, simply connected cross-sectional corrugated waveguides for which there exist analytic solutions to the expansion of the fields in a waveguide section, the main sources of error are, in addition to the basic problems of finite precision arithmetic and the density of representable numbers, (a) truncation of the operators P, Q, R and the S_{ij} to finite size, (b) the errors inherent in the formation of matrix products and, most particularly, in (c) the solution of linear systems to which the entire computational scattering problem reduces.

Accurate numerical solution of systems of equations such as (3.2.13), (3.3.1) and (3.3.2) is predicated upon the solution of equations of the type $S = A+BC^{-1}D$ with low *relative* error. For the Planck 857 GHz horn there are 2314 scattering junctions in the full assembly model. Assume that the S-matrix at each junction is to be presented in the form (3.3.2). Then if we proceed naïvely, at each junction there will be the following equations to solve in addition to the formation of P, Q and R:

$$G = Q^{-1}P, \quad K = \bar{R}^{-1}P^{\dagger}, \quad KG, \quad S_{11} = (I+KG)^{-1}(I-KG), \quad S_{12} = (I+S_{11})K,$$

and so on. The calculation of G, K, and KG is trivial, as is the calculation of S_{12}, S_{21} and S_{22} once S_{11} is known. However, the computation of S_{11}, and of the many systems of equations of type $A^{-1}B$ at each of the 2313 scattering products, has to be performed with extreme care.

Standard numerical analysis of the errors in solutions to linear systems, [32][36], is as follows: Given matrices $A \in \mathrm{GL}\,(n;\mathbb{C})$ and $B = [\boldsymbol{b}_1 \cdots \boldsymbol{b}_n] \in \mathrm{M}(n;\mathbb{C})$ the problem is to solve the system of n linear equations $A\,\boldsymbol{x}_j = \boldsymbol{b}_j$ simultaneously to obtain $X = [\boldsymbol{x}_1 \cdots \boldsymbol{x}_n] = A^{-1}B$. If the true solution to $A\,\boldsymbol{x} = b$ is $\boldsymbol{x}_t$ and the calculated solution is $\boldsymbol{x}_c$, then the residual is $\boldsymbol{x}_r = A(\boldsymbol{x}_t - \boldsymbol{x}_c)$. Denote compatible matrix and vector norms by $\|\cdot\|$, then the relative error for which estimates are required is defined to be $\|\,\boldsymbol{x}_r\,\|/\|\,\boldsymbol{x}_t\,\|$, while the norm compatibility condition $\|A\,\boldsymbol{x}\,\| \le \|A\|\cdot\|\,\boldsymbol{x}\,\|$ gives

$$\frac{1}{\|A\|\cdot\|A^{-1}\|}\cdot\frac{\|A\,\boldsymbol{x}_c - \boldsymbol{b}\,\|}{\|\,\boldsymbol{b}\,\|} \le \frac{\|\,\boldsymbol{x}_t - \boldsymbol{x}_c\,\|}{\|\,\boldsymbol{x}_t\,\|} \le \|A\|\cdot\|A^{-1}\|\cdot\frac{\|A\,\boldsymbol{x}_c - \boldsymbol{b}\,\|}{\|\,\boldsymbol{b}\,\|}.$$

Defining $\varepsilon = \|A\,\boldsymbol{x}_c - \boldsymbol{b}\,\|/\|\,\boldsymbol{b}\,\|$ and the condition number to be $\kappa(A) = \|A\|\cdot\|A^{-1}\|$, the relative error is bounded by

$$\frac{\varepsilon}{\kappa(A)}\cdot\|\,\boldsymbol{x}_t\,\| \le \|\,\boldsymbol{x}_r\,\| \le \varepsilon\kappa(A)\cdot\|\,\boldsymbol{x}_t\,\|.$$

If $\kappa(A) \sim 1$, then ε would be a true estimate of the relative error, but if $\kappa(A) \gg 1$, then ε may still be small whilst the true relative error is large. In that case $A\,\boldsymbol{x} = \boldsymbol{b}$

may be ill conditioned and A be close to singular. This is a condition that can occur in the modelling of waveguides, particularly smooth walled guides, when small radial steps arise and the scattering coefficients at a junction are calculated with the standard numerical presentation of the equations (for example, using equations (4.2.3-A and D) and (4.2.4-A and D) instead of the recurrence relations that will be given in subsections 5.2.1, 5.2.2 and 5.2.3 below. Note that this is *numerical* singularity, not analytic, and has no physical meaning.

The measure of closeness to singularity that is used in the numerical modelling of the waveguides is the reciprocal condition number

$$\frac{1}{\kappa(A)} = \min\left\{\|A - B\|/\|A\| \ : \ B \text{ singular}\right\}. \tag{5.1.1}$$

With computational errors the attempt to solve $A\boldsymbol{x} = \boldsymbol{b}$ will have lead to the solution of $(A + E)\,\boldsymbol{x} = \boldsymbol{b}$ for some error matrix E and, if the reciprocal condition number is very small, then $A + E$ is close to singular. There is a general rule of thumb [19] that says that if (i) $\eth$ is the number of decimal places in the computation, (ii) $\kappa(A) \sim 10^k$, and (iii) A is correctly scaled, then the result of the computation will be accurate to approximately $\eth - k$ significant figures. (Correctly scaled means solve $(D_r A D_l) D_l^{-1}\boldsymbol{x} = D_r\,\boldsymbol{b}$ for well chosen diagonal matrices D_r and D_l.)

The following lemma is given in [32]: Denoting the floating point operations on two real numbers by $\mathrm{fl}(\,\cdot\,)$, if, for floating point numbers with mantissa $d_1 \cdots d_m$ the floating point inner product (at the core of matrix multiplication)

$$\mathrm{fl}\left[\sum_{i=1}^{n} a_i b_i\right] = \mathrm{fl}\left[\mathrm{fl}\left[\sum_{i=1}^{n-1} a_i b_i\right] + \mathrm{fl}(a_n b_n)\right],$$

is calculated with rounding, and if $n \cdot 10^{1-m} \le 1$, then for any δ such that

$$|\delta a_1| \le n|a_1| \cdot 10^{1-m} \quad \text{and} \quad |\delta a_i| \le (n - i + 2)|a_i| \cdot 10^{1-m}$$

for all $i = 1, \ldots, n$, then the floating point and true values are related by

$$\mathrm{fl}\left[\sum_{i=1}^{n} a_i b_i\right] = \sum_{i=1}^{n}(1 + \delta)a_i b_i.$$

This is useful because it says that the computation cannot distinguish between vectors $\boldsymbol{a}$ and $\boldsymbol{a}_\delta$ with $|\,\boldsymbol{a}\,|/(1 + \delta) \le |\,\boldsymbol{a}_\delta\,| \le (1 + \delta)|\,\boldsymbol{a}\,|$. Thus the $|\delta|/(1 - |\delta|)$ neighbourhood of vectors surrounding the true $\boldsymbol{a}$ give the same matrix products.

5.2 Numerically stable forms of Π_{MM} and Π_{EE}

For small step sizes, used for example in the modelling of smooth walled horns, or more generally if either $p_{nk}/p_{nm} \simeq a/b$ or $q_{nk}/q_{nm} \simeq a/b$ at a junction, then the scattering equations (4.2.3-A and D) and (4.2.4-A and D) will become numerically unstable. In any horn design optimisation process the radii a and b have to be free variables and to fragment the feasible domain for solutions by bounding a/b away from all possible ratios p_{nk}/p_{nm} and q_{nk}/q_{nm} would destroy the optimisation search. Thus, for both routine waveguide analysis and optimisation the Π_{MM} and Π_{EE} component equations need to be rewritten to render all the terms completely numerically stable. Taking the example of Π_{MM}, at the level of the Lommel integrals it is clear that, in the limit as $p_{nm}/b \to p_{nk}/a$, the Lommel integral of the first kind becomes a Lommel integral of the second kind. In the development of the equations for scattering simulations the limit is undefined and the equations for electric-electric and for magnetic-magnetic scattering are numerically unstable close to the numerical singularity.

The limiting value as $p_{nm}/b - p_{nk}/a \to 0$ is found by writing $\varepsilon = \dfrac{p_{nm}}{b} - \dfrac{p_{nk}}{a}$

$$\frac{\rho}{p_{nm}} \frac{J_n\left(\dfrac{p_{nm}a}{b}\right)}{\rho^2 - \left(\dfrac{p_{nk}}{p_{nm}}\right)^2} = \frac{p_{nm}}{ab} \cdot \frac{J_n\left(\dfrac{p_{nm}a}{b}\right)}{\left(\dfrac{p_{nm}}{b}\right)^2 - \left(\dfrac{p_{nk}}{a}\right)^2} = \frac{p_{nm}}{ab} \cdot \frac{J_n(p_{nk} + a\varepsilon)}{\varepsilon\left(\dfrac{2p_{nk}}{a} + \varepsilon\right)}, \qquad (5.2.1)$$

which, since both numerator and denominator tend to zero as $\varepsilon \to 0$, takes the limiting value

$$-\frac{a^2}{2p_{nk}} J_{n+1}(p_{nk}) \qquad (5.2.2)$$

when $p_{nm}/b = p_{nk}/a$, by l'Hôpital's rule. This could also have been obtained from the Taylor series expansion of $J_n(p_{nk}x + x\varepsilon)$ in the Lommel integral of the first kind and taking the limit to obtain the Lommel integral of the second kind. For numerical simulation when $|\varepsilon| \sim 0$, what is required is a Padé approximation or a series expansion of the right hand side of equation (5.2.1) and an efficient means of evaluating it.

The solution adopted here is to expand $J_n(p_{nm}a/b)$ as a Taylor series about

$J_n(p_{nk})$: Since $J_n(p_{nk}) = 0$, write

$$\frac{J_n(p_{nk} + a\varepsilon)}{\varepsilon\left(\dfrac{2p_{nk}}{a} + \varepsilon\right)} = \frac{a^2}{2p_{nk} + a\varepsilon} \cdot \sum_{m\in\mathbb{N}} \frac{(a\varepsilon)^{m-1}}{m!} J_n^{(m)}(p_{nk}). \tag{5.2.3}$$

Two methods of solution have been found, and both are given here: The first is the product of a polynomial with coefficients that are Bessel functions of the first kind and integer order with a polynomial approximation to $J_n(x)/x$ for small $|x|$; the second is a four term recurrence relation. Both converge stably to $-a^2 J_{n+1}(p_{nk})/2p_{nk}$.

5.2.1 Polynomial approximation to Π_{MM}

A very accurate polynomial approximation to P_{MM} is described. The equivalent approximation for P_{EE} will not be given since its derivation is essentially the same.

Consider the generating function for the Bessel functions [1][6]

$$g(x,t) = e^{(t-1/t)(x/2)} = \sum_{n\in\mathbb{Z}} J_n(x)t^n.$$

Differentiating m times with respect to x gives

$$\begin{aligned}\frac{\partial^l}{\partial x^l} g(x,t) &= \left[\frac{1}{2}\left(t - \frac{1}{t}\right)\right]^l g(x,t)\\ &= \sum_{n\in\mathbb{Z}} \frac{1}{2^l}\left(t - \frac{1}{t}\right)^l t^n J_n(x) = \sum_{n\in\mathbb{Z}} t^n J_n^{(l)}(x).\end{aligned}$$

Rewriting this equality as

$$\sum_{n\in\mathbb{Z}} t^n J_n^{(l)}(x) = \frac{1}{2^l}\left(t - \frac{1}{t}\right)^l \sum_{l\in\mathbb{Z}} t^l J_l(x) \tag{5.2.4}$$

and equating powers of t gives the derivatives as polynomials with coefficients that are Bessel functions of the first kind and integer order which can be written out in a table with the following pattern: up to a factor of $1/2^l$,

	-4	-3	-2	-1	0	1	2	3	4
$J_n^{(1)} =$				J_{n-1}		$-J_{n+1}$			
$J_n^{(2)} =$			J_{n-2}		$-2J_n$		J_{n+2}		
$J_n^{(3)} =$		J_{n-3}		$-3J_{n-1}$		$3J_{n+1}$		$-J_{n+3}$	
$J_n^{(4)} =$	J_{n-4}		$-4J_{n-2}$		$6J_n$		$-4J_{n+2}$		J_{n+4}
$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$

and so on, the coefficient for J_{n-l+2k}, $0 \le k \le l$, in the row for $J_n^{(l)}$ being $(-1)^k\binom{l}{k}$.

From equations (5.2.3) and (5.2.4) it follows that the coefficient for the l^{th} row in the table is $(a\varepsilon)^{l-1}/2^l l!$ and since the Bessel functions are to be evaluated at the zero of J_n, the central column entries are all zeros. Furthermore, the coefficients of the terms J_{n+l} and J_{n-l} are related by the factor $(-1)^l$. Observing that for all $l \in \mathbb{N}$ the coefficients of the sum of all terms $[J_{n-l}(p_{nk}) + (-1)^l J_{n+l}(p_{nk})]$ is

$$\sum_{j=0}^{\infty}(-1)^j \frac{(a\varepsilon)^{2j+l-1}}{2^{2j+l}(2j+l)!}\binom{2j+l}{j} = \frac{1}{a\varepsilon}\cdot\sum_{j=0}^{\infty}\frac{(-1)^j}{j!(j+l)!}\left(\frac{a\varepsilon}{2}\right)^{l+2j},$$

it is seen that the sum of derivatives in equation (5.2.3) can be obtained by summing down the columns of the table giving

$$\begin{aligned}\sum_{l\in\mathbb{N}}\frac{(a\varepsilon)^{l-1}}{l!}J_n^{(l)}(p_{nk}) &= \sum_{l\in\mathbb{N}}\Bigg[\Big(J_{n-l}(p_{nk}) + (-1)^l J_{n+l}(p_{nk})\Big)\\ &\qquad\qquad \times\frac{1}{a\varepsilon}\cdot\sum_{j=0}^{\infty}\frac{(-1)^j}{j!(j+l)!}\left(\frac{a\varepsilon}{2}\right)^{l+2j}\Bigg]\\ &= \sum_{l\in\mathbb{N}}\Big(J_{n-l}(p_{nk}) + (-1)^l J_{n+l}(p_{nk})\Big)\frac{J_l(a\varepsilon)}{a\varepsilon}.\end{aligned}$$

As written, this equation is just as numerically unstable as the equation it was designed to replace, but for $a\varepsilon \ll 1$ the term $J_l(a\varepsilon)/a\varepsilon$ can be evaluated as a polynomial approximation to the series by writing

$$\begin{aligned}h_l(ae) = \sum_{j=0}^{\infty}(-1)^j\frac{(a\varepsilon)^{2j+l-1}}{2^{2j+l}j!(j+l)!} &\approx \frac{(a\varepsilon)^{l-1}}{2^l l!}\left[1-\frac{(a\varepsilon)^2}{4(l+1)}\left(1-\frac{(a\varepsilon)^2}{4\cdot 2(l+2)}\left(\cdots\right.\right.\right.\\ &\left.\left.\left.\cdots\left(1-\frac{(a\varepsilon)^2}{4\cdot(j-1)(l+j-1)}\left(1-\frac{(a\varepsilon)^2}{4\cdot(j)(l+j)}\right)\right)\cdots\right)\right)\right].\end{aligned} \tag{5.2.5}$$

Denote a truncation of this series (to any chosen number of terms) by $\bar{h}_l$, then the polynomial approximation is then given by

$$\frac{J_n\left(\dfrac{p_{nm}a}{b}\right)}{\left(\dfrac{p_{nm}}{b}\right)^2-\left(\dfrac{p_{nk}}{a}\right)^2} \approx \sum_{l=1}^{N}\big(J_{n-l}(p_{nk}) + (-1)^l J_{n+l}(p_{nk})\big)\,\bar{h}_l(ae). \tag{5.2.6}$$

In the evaluation of $J_n(x)$ the numerical method requires a downward recurrence of length $N = 2(n + \sqrt{256n}/2)$, where the arithmetic is integer arithmetic. As a conservative estimate for the limit, N, in equation (5.2.6) and the number of terms, j, in equation (5.2.5), the value N just given and the value and $j = 5$ will suffice to give the same accuracy as the Bessel function routine providing $|a\varepsilon| < 1$. Thus, since evaluation of $J_n(x)$ by downward recurrence requires the evaluation of $J_0(x), \cdots, J_N(x)$ (the correction to the first approximation being done at $J_0(x)$ and propagated up to $J_n(x)$, see [57]) the additional cost is only the minor extension of the correction propagation up to $J_N(x)$.

5.2.2 Four term recurrence relation for Π_{MM}

The Bessel functions of the first kind and integer order are defined by the relation

$$x^2 J_n''(x) + x J_n'(x) + (x^2 - n^2) J_n(x) = 0. \tag{5.2.7}$$

Substituting $x = p_{nk}$ gives $J_n''(p_{nk}) = -J_n'(p_{nk})/p_{nk} = J_{n+1}(p_{nk})/p_{nk}$. Repeated differentiation of equation (5.2.7) gives the higher derivatives in terms of the lower:

$$J_n^{(3)}(x) = -\frac{3}{x} J_n''(x) - \left[1 - \frac{(n^2-1)}{x^2}\right] J_n'(x) - \frac{2}{x} J_n(x) \tag{5.2.8-A}$$

and for all $m \geq 0$

$$\begin{aligned} J_n^{(m+4)}(x) = &-\frac{(2(m+2)+1)}{x} J_n^{(m+3)}(x) - \left[1 - \frac{(n^2-(m+2)^2)}{x^2}\right] J_n^{(m+2)}(x) \\ &- \frac{2(m+2)}{x} J_n^{(m+1)}(x) - \frac{(m+1)(m+2)}{x^2} J_n^{(m)}(x). \end{aligned} \tag{5.2.8-B}$$

Substituting $x = p_{nk}$ eliminates the $J_n(x)$ terms in $J_n^{(3)}(p_{nk})$ and $J_n^{(4)}(p_{nk})$. An attempt to find a closed form solution to the sum in equation (5.2.3) as a multiple of $J_{n+1}(p_{nk})/p_{nk}$ gives a leading term of $J_{n+1}(p_{nk})[1 - \ln(1 - a\varepsilon)]$, but there remains an infinite sequence of polynomials, the truncation of which would not lead to good numerical approximation. However, using equations (5.2.8) it is now immediately clear that the sum can be easily evaluated in a loop.

From the coefficients of $J_n^{(m+l)}(x)$ in equation (5.2.8-B), evaluated at $x = p_{nk}$, define the following functions of an integer argument: for azimuthal and radial orders

n and k respectively

$$M^n_{0,k}(m) = \frac{(m+1)(m+2)}{p^2_{nk}}, \qquad M^n_{1,k}(m) = \frac{2m+4}{p_{nk}},$$
$$M^n_{2,k}(m) = \left[1 - \frac{n^2-(m+2)^2}{p^2_{nk}}\right], \qquad M^n_{3,k}(m) = \frac{2m+5}{p_{nk}}.$$

Dividing the sum in equation (5.2.3) by $J_{n+1}(p_{nk})$ gives an infinite order polynomial in p_{nk} and $a\varepsilon$ with real coefficients; so define

$$T^n_{m,k}(a\varepsilon) = \frac{(a\varepsilon)^{m-1}}{(m)!}\frac{J^{(m)}_n(p_{nk})}{J_{n+1}(p_{nk})}, \quad \text{for all } \ m \geq 1.$$

The terms in the loop are given by the sequence of constants and the following four term recurrence relation: evaluating the first four $T^n_{m,k}$ gives

$$T^n_{1,k}(a\varepsilon) = -1, \qquad T^n_{2,k}(a\varepsilon) = \frac{a\varepsilon}{2p_{nk}},$$
$$T^n_{3,k}(a\varepsilon) = \frac{(a\varepsilon)^2}{6}\left(1 - \frac{n^2+2}{p^2_{nk}}\right), \qquad T^n_{4,k}(a\varepsilon) = \frac{(a\varepsilon)^3}{12p_{nk}}\left(\frac{3(n^2+1)}{p^2_{nk}} - 1\right),$$

and for all $m \geq 1$

$$T^n_{m+4,k}(a\varepsilon) = -\frac{a\varepsilon}{m+4}\left\{M^n_{3,k}(m)T^n_{m+3,k}(a\varepsilon) + \frac{a\varepsilon}{m+3}\left[M^n_{2,k}(m)T^n_{m+2,k}(a\varepsilon) + \right.\right.$$
$$\left.\left.\frac{a\varepsilon}{m+2}\left(M^n_{1,k}(m)T^n_{m+1,k}(a\varepsilon) + \frac{a\varepsilon}{m+1}M^n_{0,k}(m)T^n_{m,k}(a\varepsilon)\right)\right]\right\}. \tag{5.2.9}$$

The exact and numerically stable form for the ratio in the components of P^n_{MM}, intended for use when $|a\varepsilon| = |p_{nm}a/b - p_{nk}| \ll 1$ is given by

$$\frac{J_n\left(p_{nm}a/b\right)}{\left(\frac{p_{nm}}{b}\right)^2 - \left(\frac{p_{nk}}{a}\right)^2} = \frac{a^2 J_{n+1}(p_{nk})}{2p_{nk}+a\varepsilon}\cdot\sum_{m\in\mathbb{N}} T^n_{m,k}(a\varepsilon), \quad \text{for } \varepsilon = \frac{p_{nm}}{b} - \frac{p_{nk}}{a}. \tag{5.2.10}$$

Since $p_{nk} > n$ for all n and k, it follows that $T^n_{m,k} \sim \left(1 + \frac{m^2}{p^2_{nk}}\right)\frac{(a\varepsilon)^{m-1}}{m!}$, and the series therefore converges faster than $(a\varepsilon)^{m-1}(m^2+n^2)/[n^2m!]$ and the fifth term (the firsts term in the loop) will be of order 10^{-10} or smaller for $a\varepsilon \leq 10^{-2}$. Convergence is therefore extremely rapid. In practise, with the threshold $a\varepsilon \leq 10^{-2}$ the loop exits after a single evaluation because 16 decimal place precision has already been achieved.

The recurrence relation will converge rapidly at least for $0 < |a\varepsilon| \lesssim 10$ with careful coding, but there is no gain in computational efficiency over direct computation of the right hand side when $|\varepsilon| \gg 0$; it is, however, correct for *all* values.

5.2.3 Four term recurrence relation for Π_{EE}

The same method is used to find the asymptotic behaviour of the terms in P^n_{EE} in the limit as $\varepsilon = q_{nm}/b - q_{nk}/a \to 0$. As above, using l'Hôpital's rule and the recurrence relation (5.2.7),

$$\frac{J_n'(q_{nm}a/b)}{\left(\frac{q_{nm}}{b}\right)^2 - \left(\frac{q_{nk}}{a}\right)^2} = \frac{J_n'(q_{nk}+a\varepsilon)}{\varepsilon\left(\frac{2q_{nk}}{a}+\varepsilon\right)} \xrightarrow{\varepsilon\to 0} -\frac{a^2}{2q_{nk}}\left(1-\frac{n^2}{q_{nk}^2}\right)J_n(q_{nk}). \tag{5.2.11}$$

The left hand side of this equation expands as a Taylor series giving

$$\frac{J_n'(q_{nk}+a\varepsilon)}{\varepsilon\left(\frac{2q_{nk}}{a}+\varepsilon\right)} = \frac{a^2}{(2q_{nk}+a\varepsilon)}\left[\left(\frac{n^2}{q_{nk}^2}-1\right)J_n(q_{nk}) + \sum_{m=2}^{\infty}\frac{(a\varepsilon)^{m-1}}{m!}J_n(q_{nk})^{(m+1)}\right]. \tag{5.2.12}$$

For all $m \in \mathbb{N}$ define the functions $H^n_{m,k}(a\varepsilon) = \dfrac{(a\varepsilon)^{m-1}}{(m)!}\dfrac{J_n^{(m+1)}(q_{nk})}{J_n(q_{nk})}$, and the following functions for all $m \in \mathbb{Z}$:

$$E^n_{0,k}(m) = -\frac{(m+2)(m+3)}{q_{nk}^2}, \qquad E^n_{1,k}(m) = -\frac{2m+6}{q_{nk}},$$
$$E^n_{2,k}(m) = \left[\frac{n^2-(m+3)^2}{q_{nk}^2}-1\right], \qquad E^n_{3,k}(m) = -\frac{2m+7}{q_{nk}}.$$

Substituting from equations (5.2.7) and (5.2.8), the evaluation of the series in the right hand side of equation (5.2.12) is given by a loop over m with initialisation functions and four term recurrence relation

$$H^n_{1,k}(a\varepsilon) = \frac{n^2}{q_{nk}^2} - 1, \tag{5.2.13-A}$$

$$H^n_{2,k}(a\varepsilon) = \frac{a\varepsilon}{2q_{nk}}\left(1-\frac{3n^2}{q_{nk}^2}\right) \tag{5.2.13-B}$$

$$H^n_{3,k}(a\varepsilon) = \frac{a\varepsilon}{3}\left[E^n_{3,k}(-1)H^n_{2,k}(a\varepsilon) + \frac{a\varepsilon}{2}\left[E^n_{2,k}(-1)H^n_{1,k}(a\varepsilon)+E^n_{0,k}(-1)\right]\right], \tag{5.2.13-C}$$

$$H^n_{4,k}(a\varepsilon) = \frac{a\varepsilon}{4}\left[E^n_{3,k}(0)H^n_{3,k}(a\varepsilon) + \frac{a\varepsilon}{3}\left(E^n_{2,k}(0)H^n_{2,k}(a\varepsilon)+\frac{a\varepsilon}{2}E^n_{1,k}(0)H^n_{1,k}(a\varepsilon)\right)\right], \tag{5.2.13-D}$$

and for all $m \geq 1$

$$\begin{aligned} H^n_{m+4,k}(a\varepsilon) = \frac{a\varepsilon}{m+4}\Bigg[&E^n_{3,k}(m)H^n_{m+3,k}(a\varepsilon) + \frac{a\varepsilon}{m+3}\Bigg\{E^n_{2,k}(m)H^n_{m+2,k}(a\varepsilon) \\ &+ \frac{a\varepsilon}{m+2}\left[E^n_{1,k}(m)H^n_{m+1,k}(a\varepsilon) + \frac{a\varepsilon}{m+1}E^n_{0,k}(m)H^n_{m,k}(a\varepsilon)\right]\Bigg\}\Bigg] \end{aligned} \tag{5.2.14}$$

Thus, the exact and numerically stable form for the ratio in the components of P^n_{EE} intended for use when $|a\varepsilon| = |q_{nm}a/b - q_{nk}| \ll 1$ is given by

$$\frac{J'_n\left(q_{nm}a/b\right)}{\left(\frac{q_{nm}}{b}\right)^2 - \left(\frac{q_{nk}}{a}\right)^2} = \frac{a^2 J_n(q_{nk})}{2q_{nk} + a\varepsilon} \cdot \sum_{m\in\mathbb{N}} H^n_{m,k}(a\varepsilon). \tag{5.2.15}$$

Just as the recurrence relation for $J_n(p_{nm}a/b)/\left((p_{nm}/b)^2 - (p_{nk}/a)^2\right)$ is stable for all $a\varepsilon$ and converges rapidly for $0 \leq a\varepsilon \lesssim 10$, equation (5.2.15) is also stable and rapidly convergent over the same range. From the point of view of numerical efficiency, however, the recurrence relations should only be used when $|a\varepsilon| \lesssim 0.01$ to ensure that there is no loss of precision.

5.2.4 Using the recurrence relations in scattering code

There are basically two ways that the recurrence relations of the previous two subsections, or the polynomial approximation of subsection 5.2.1 and its Π_{EE} equivalent can be incorporated into scattering code. The most obvious way is to calculate the ε and to have set a threshold below which the evaluation of the components of P_{MM} or P_{EE} uses the recurrence relations instead of the standard algorithms. An alternative would be to always use the standard algorithm and proceed with the matrix operations as described in this chapter. If the horn geometry (step size ratio) gives rise to numerical instability and any junction the high threshold for numerical stability (the chosen threshold for the reciprocal condition number $1/\kappa(A)$ of equation (5.1.1) and the forward and backward error estimates of subsection 5.4.2) will detect the problem. Then, instead of aborting with a failure report, the code could reevaluate the operator matrices and repeat the calculation. Only then, if there was still numerical instability, would the code abort.

5.2.5 Limiting cases as $a \to b$

For numerical modelling of smoothed walled horns the particular case $a \to b$ must be stable and return the identity matrix in the limit.

The form of equation (4.2.4-C) for the terms $(\Pi^n_{ME})_{mk}$ is numerically stable and tends to zero with both m and k. The same is true for equations (4.2.4-A and D) for indices $m \neq k$ and $|a-b| \ll \min_{m,k} |q_{nm} - q_{nk}|$ or $|a-b| \ll \min_{m,k} |p_{nm} - p_{nk}|$ respectively. As in equation (4.3.1) the limit is determined simply by the limiting values $J_n'(q_{nm}a/b) \to 0$ and $J_n(p_{nm}a/b) \to 0$. Thus, all off diagonal entries in the Π matrix tend stably to zero.

For the case $m = k$ equation (5.2.15), or (5.2.11) with equation (4.2.4-A), gives

$$\begin{gathered}\frac{2\,\alpha_k^n \cdot J_n'\left(\frac{q_{nm}a}{b}\right)}{ab|J_n\left(q_{nm}\right)| \cdot \left[\left(\frac{q_{nm}}{b}\right)^2 - \left(\frac{q_{nk}}{a}\right)^2\right]} \left[\left(1-\frac{n^2}{q_{nk}^2}\right)\left(1-\frac{n^2}{q_{nm}^2}\right)\right]^{-\frac{1}{2}} \\ \overset{\lim}{\xrightarrow{a\to b}} \quad \frac{2\,\alpha_k^n \cdot J_n(q_{nk})}{a^2|J_n(q_{nk})|\left(1-\frac{n^2}{q_{nk}}\right)} \cdot \frac{a^2}{2}\left(1-\frac{n^2}{q_{nk}}\right) = 1\end{gathered} \tag{5.2.16}$$

by definition of α_k^n. Likewise, equation (5.2.10), or (5.2.1) combined with equation (4.2.4) gives, in the limit $p_{nk}/p_{nm} \to a/b$,

$$-\frac{a^2 J_n(p_{nk})}{2\,p_{nk}} \cdot \frac{2\,\beta_k^n\,p_{nk}}{a^2|J_n(p_{nk})|} = -\,\mathtt{sgn}(J_n(p_{nk}))\,\beta_k^n = 1 \tag{5.2.17}$$

by definition of β_k^n. Thus, for all indices m and k, the real part of the matrix Π will converge uniformly and stably to the identity matrix as $a \to b$ when using the recurrence relations (5.2.10) and (5.2.15), as was required, but *not* with the standard forms of the equations.

5.3 Formulation of the coding problem

For efficient modelling of the horns the structure of the operator matrices needs to be exploited. What has to be avoided above all else is the use of matrix multiplication and inversion where not absolutely essential and the use of complex matrices where families of real matrices can be used in their place. It will be shown that, while matrix multiplication cannot be avoided, the size of the problem can be reduced, and that matrix inversion is completely avoidable.

For a system of n_E electric and n_M magnetic modes, a literal approach to the solution of $S_{11} = (\bar{R}+P^\dagger Q^{-1}P)^{-1}(\bar{R}-P^\dagger Q^{-1}P)$ requires $4(n_E+n_M)^3$ 64 bit complex multiplications; treating the matrices K and G as complex matrices and solving $(I+KG)S_{11} = (I-KG)$ reduces this to $2(n_E+n_M)^3$. Since complex multiplication is performed on a machine with fast multiplication with four real multiplications, one addition and one subtraction, if the structure of the matrices K and G can be exploited to eliminate the need for any complex multiplication in forming KG, the size of the problem can be reduced further from $8(n_E+n_M)^3$ real multiplications to $5(n_E+n_M)^3$. To do so requires analysis of the structure of the matrices K and G.

For each azimuthal order writing the matrices P in terms a purely real matrix, Π, and the admittance factors, gives an array of the form

$$\begin{bmatrix} P_{EE} & 0 \\ P_{ME} & P_{MM} \end{bmatrix} = \begin{bmatrix} Q_E^*(b)\Pi_{EE} & 0 \\ Q_M^*(b)\Pi_{ME} & Q_M^*(b)\Pi_{MM} \end{bmatrix}. \tag{5.3.1}$$

Pre-multiplying P by Q^{-1} and $P^\dagger$ by $\bar{R}^{-1}$ gives

$$\begin{bmatrix} G_{EE} & 0 \\ G_{ME} & G_{MM} \end{bmatrix} = \begin{bmatrix} Q_E^{-1}(b)Q_E^*(b)\Pi_{EE} & 0 \\ Q_M^{-1}(b)Q_M^*(b)\Pi_{ME} & Q_M^{-1}(b)Q_M^*(b)\Pi_{MM} \end{bmatrix}, \tag{5.3.2-A}$$

$$\begin{bmatrix} K_{EE} & K_{EM} \\ 0 & K_{MM} \end{bmatrix} = \begin{bmatrix} [Q_E(b)\Pi_{EE}R_E^*(a)]^T & [Q_M(b)\Pi_{ME}R_E^*(a)]^T \\ 0 & [Q_M(b)\Pi_{MM}R_M^*(a)]^T \end{bmatrix}, \tag{5.3.2-B}$$

where M^T denotes the transpose of a matrix M. For free space or loss-less dielectric filled guides the matrix G is purely real because the components, z, of $Q^{-1}Q^*$ have the form $e^{2i\arg(z)}$ which has value $+1$ for propagating modes and -1 for evanescent modes; thus, $(G_{IJ})_{mk} = \pm(\Pi_{IJ})_{mk}$. In the operator matrix K the products of the admittance and impedance gives purely real coefficients where either the admittance to the left and impedance to the right of the junction are both real, or both are pure imaginary; they give a pure imaginary coefficient only where either the left side or right side coefficient is imaginary. The result is that all three of the matrices K_{IJ} have a block structure of the form $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ where the sub-matrices A, B, C and D are generally rectangular and of shape particular to the submatrix of K, with components a_{ij}, $d_{ij} \in \mathbb{R}$ and b_{ij}, $c_{ij} \in i\,\mathbb{R}$, or general complex if the material filling the waveguide at either junction gives a complex ratio Z_a/Z_b. The dimensions and shape of the submatrices in K is determined by the admittance-impedance products in equation (5.3.2-B). Denoting the waveguide material to the left and right of the junction by Z_a and Z_b respectively, these are

$$(Z_E(a))_{mm}(Y_E(b))_{ll} = (Z_a/Z_b)\cdot\left[\left(1-(q_{nl}/kb)^2\right)/\left(1-(q_{nm}/ka)^2\right)\right]^{\frac{1}{2}} \qquad (5.3.3\text{-A})$$

$$(Z_E(a))_{mm}(Y_M(b))_{ll} = (Z_a/Z_b)\cdot\left[\left(1-(q_{nm}/ka)^2\right)\left(1-(p_{nl}/kb)^2\right)\right]^{-\frac{1}{2}} \qquad (5.3.3\text{-B})$$

$$(Z_M(a))_{mm}(Y_M(b))_{ll} = (Z_a/Z_b)\cdot\left[\left(1-(p_{nm}/ka)^2\right)/\left(1-(p_{nl}/kb)^2\right)\right]^{\frac{1}{2}}. \qquad (5.3.3\text{-C})$$

All submatrices in both K and G are therefore dimensionless.

For fixed wavenumber, k, write

$$q(a,m) = |1-(q_{nm}/ka)^2|^{\frac{1}{2}}, \quad p(b,l) = |1-(p_{nl}/kb)^2|^{\frac{1}{2}}$$

etc. Then, using the logical AND symbol, $\wedge$, up to a factor of Z_a/Z_b:

$$(Z_E(a))_{mm}(Y_E(b))_{ll} = \begin{cases} iq(b,l)/q(a,m) & : \quad (ka < q_{nm}) \wedge (kb > q_{nl}) \\ -iq(b,l)/q(a,m) & : \quad (ka > q_{nm}) \wedge (kb < q_{nl}) \\ q(b,l)/q(a,m) & : \quad \text{otherwise} \end{cases} \qquad (5.3.4\text{-A})$$

$$(Z_E(a))_{mm}(Y_M(b))_{ll} = \begin{cases} -1/\left(q(a,m)\cdot p(b,l)\right) & : \quad (ka < q_{nm}) \wedge (kb < p_{nl}) \\ 1/\left(q(a,m)\cdot p(b,l)\right) & : \quad (ka > q_{nm}) \wedge (kb > p_{nl}) \\ -i/\left(q(a,m)\cdot p(b,l)\right) & : \quad \text{otherwise} \end{cases}$$

$$(5.3.4\text{-B})$$

$$(Z_M(a))_{mm}(Y_M(b))_{ll} = \begin{cases} ip(a,m)/p(b,l) & : \quad (ka < p_{nm}) \wedge (kb > p_{nl}) \\ -ip(a,m)/p(b,l) & : \quad (ka > p_{nm}) \wedge (kb < p_{nl}) \\ p(a,m)/p(b,l) & : \quad \text{otherwise} \end{cases} \qquad (5.3.4\text{-C})$$

Note $ka = q_{nm}$ would imply infinite impedance in the waveguide to the left of the junction, while $kb = p_{nl}$ would imply zero impedance in the waveguide to the right of the junction. Usually $Z_a = Z_b$ (*i.e.* the same matrial filling the waveguide at either side of the junction) in which case the factor plays no rôle in the scattering operation. Otherwise the ratio introduces an additional, possibly complex, factor. If the factor is real it presents no impediment to increasing efficiency in the numerical implementation of scattering models; if complex, K will be complex at every locus. Nevertheless, the computation of the S-matrix for the junction can still make full use of the stratagem outlined in what follows, though there will be an additional order N^2 complex multiplication of the array KG by the constant factor Z_a/Z_b.

5.4 Computation of the S-matrix at a junction

The computation of the full S-matrix at a junction will return a complex matrix that is first scattered with the system S-matrix up to the current junction and then phase slipped to the following junction. There are three steps for the computation if the S-matrix once the relatively trivial task of obtaining K and G has been completed: firstly calculation of the products $(I \pm KG)$; secondly the calculation of S_{11}; finally calculation of S_{12}, S_{21} and S_{22}. Note: in the presentation that follows it has been assumed that the number of TE modes is not greater than the number of TM modes in the model; if that is not the case then S_{22} is calculated first and all other S_{IJ} calculated from it in the obvious modifications to the procedure given below.

5.4.1 Computing $(I \pm KG)$

Fix the frequency, hence the wavenumber k, and fix the azimuthal order, n. At any given junction the radii a and b are fixed which determines the critical parameter $\rho = a/b$. The three sets of three equations in (5.3.4) are three-valued multiplication tables on the sets $\{p_{nm}\}$ and $\{q_{nm}\}$ with set product $\cap$ (equivalent to logical $\wedge$): the values ka and kb partition the sets $\{p_{nm}\}$ and $\{q_{nm}\}$ giving a four place table, each product in the table corresponding to a submatrix in $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ with value ± 1 on the diagonal blocks of A and D and $\pm i$ for B and C. The aim here is to replace the complex submatrices in K with real matrices, form real products equivalent to each of $K_{EE}G_{EE}$, $K_{EM}G_{ME}$, $K_{EM}G_{MM}$, $K_{MM}G_{ME}$ and $K_{MM}G_{MM}$, and then to form the complex arrays $I \pm KG$.

To follow the above strategy with minimum operation count allocate memory for the matrices as follows: Define $X = (I + KG)$ and $Y = (I - KG)$ and allocate complex arrays $X_{11}(1 : n_E, 1 : n_E)$, $X_{12}(1 : n_E, 1 : n_M)$, $X_{21}(1 : n_M, 1 : n_E)$ and $X_{22}(1 : n_M, 1 : n_M)$, and the equivalent submatrices for matrix Y. Calculate the size to wavelength scales ka and kb and search for $q_a = \max\{m : ka < q_{nm}\}$, $q_b = \max\{m : kb < q_{nm}\}$ and $p_a = \max\{m : ka < p_{nm}\}$, $p_b = \max\{m : kb < p_{nm}\}$. The multiplication table for equation (5.3.4-A and C) is $\begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}$. Then KG is stored in X by assigning to the j^{th} column of $X_{11} = K_{EE}G_{EE} + K_{EM}G_{ME}$, for all $j \in \{1, \cdots, n_E\}$

for $i = 1, q_a$

$$(X_{11})_{ij} = \left[\sum_{m=1}^{q_b}(K_{EE})_{im}(G_{EE})_{mj} - \sum_{m=1}^{p_b}(K_{EM})_{im}(G_{ME})_{mj}\right] - i\left[\sum_{m=q_b+1}^{n_E}(K_{EE})_{im}(G_{EE})_{mj} - \sum_{m=p_b+1}^{n_M}(K_{EM})_{im}(G_{ME})_{mj}\right] \quad (5.4.1\text{-A})$$

for $i = q_a + 1, n_E$

$$(X_{11})_{ij} = \left[\sum_{m=q_b+1}^{n_E}(K_{EE})_{im}(G_{EE})_{mj} + \sum_{m=p_b+1}^{n_M}(K_{EM})_{im}(G_{ME})_{mj}\right] + i\left[\sum_{m=1}^{q_b}(K_{EE})_{im}(G_{EE})_{mj} + \sum_{m=1}^{p_b}(K_{EM})_{im}(G_{ME})_{mj}\right] \quad (5.4.1\text{-B})$$

The other three submatrices are simpler: for $X_{12} = K_{EM}G_{MM}$ the sums for the columns $j \in \{1, \cdots, n_M\}$ are

for $i = 1, q_a$

$$(X_{12})_{ij} = -\sum_{m=1}^{p_b}(K_{EM})_{im}(G_{MM})_{mj} - i\sum_{m=p_b+1}^{n_M}(K_{EM})_{im}(G_{MM})_{mj} \quad (5.4.2\text{-A})$$

for $i = q_a + 1, n_E$

$$(X_{12})_{ij} = -\sum_{m=p_b+1}^{n_M}(K_{EM})_{im}(G_{MM})_{mj} + i\sum_{m=1}^{p_b}(K_{EM})_{im}(G_{MM})_{mj} \quad (5.4.2\text{-B})$$

Equivalent sums are obtained for $X_{21} = K_{MM}G_{ME}$ and $X_{22} = K_{MM}G_{MM}$.

There follows the $(n_E+n_M)^2$ copies to fill Y: $Y_{11} = -X_{11}$ etc. then the (n_E+n_M) additions for each of $Y \rightarrow Y + I$, and $X \rightarrow X + I$.

The total cost, counted in terms of 64-bit *real* operations, required to fill the arrays $X = (I + KG)$ and $Y = (I - KG)$, given K and G, for the method outlined above and for multiplication of a real matrix G by a complex matrix K, ignoring operations common to both methods, are as follows. For the above method, $2(n_E^3 + 2n_E^2 n_M + n_E n_M^2 + n^3)$ multiplications and $2(n_E^3 + n_E n_M(n_E + n_M - 1) - n_E^2 - n_M^2 + n_M^3)$ additions or subtractions are required; using a complex array K, $4(n_E^3 + 2n_E^2 n_M + n_E n_M^2 + n^3)$ multiplications and $n_E^3 + n_E^2(n_M - 2) + n_M^2(2n_E - 1) - 2n_E n_M + n_M^3$ additions. On a fast floating point machine real addition and real multiplication, performed at the same precision, both take a single clock cycle; on such machines the net gain from using real K is $(n_E + n_M)^2 + n_M^2(1 - n_E)$ clock cycles.

When modelling the Planck 545 GHz and 857 GHz horns the total number of scattering junctions is roughly 2×10^3. For typical model sizes the clock cycle reduction per junction is 7×10^6, so the cycle saving for the horn assembly for each frequency is of the order 1.4×10^{10}. That is about five seconds per frequency on a 3000 MHz processor. The gain would not merit the time spent coding for simple horns and models run at few frequencies, but for complex horn optimisation run at many frequencies and hundreds of times, the coding effort is worthwhile. In studies of the power transmission dependence on frequency in the horns, conducted over one or two hundred frequency samples, the gain was noticeable. (Depending upon how the program has been compiled to handle denormal numbers, the gain could be much greater.)

The second place that using real K and a similar approach to the above to computation of a matrix product is in the formation of $S_{12} = (I + S_{11})K$. There $(I + S_{11})$ is always complex but, at a modest cost in coding complexity, the product $(I + S_{11})K$ can be done entirely with real multiplication with a operation count reduction of almost one half.

5.4.2 Computing the S-matrix

It is at this stage that the greatest gains can be made in the speed of computation and in accuracy, and the procedure outlined in this section, and applied to the computation of the scattering product in the following section, is what made the scattering studies for the Planck multi-mode horn assemblies a tractable problem.

The input are the eight complex submatrices of $X = (I+KG)$ and $Y = (I-KG)$ from page 133 and the first stage output is to be $S_{11} = X^{-1}Y$. As a matter of principle in numerical work with arrays you never calculate $A = B^{-1}C$ by calculating B^{-1} and then performing the multiplication unless that approach cannot be avoided. Here there is no need to calculate any inverses or perform the very large complex matrix multiplication $X^{-1}Y$; rather, the strategy is to solve the system of linear equations $XS_{11} = Y$. To do so naïvely would only halve the time cost of solving for X^{-1} and then $X^{-1}Y$; to do so efficiently gives far greater time savings.

In the following procedure the strategy is to exploit the structure of the arrays X and Y to solve for S_{11}. Alternatives would be to solve for S_{12} from X and K, or to have computed $(I \pm GK)$ and a complex G to compute S_{21}. The strategy would be the same in all cases, but the sequence of calculations described here gives the lowest *total* cost in CPU clock cycles at each junction.

Write X and Y as above and $S_{11} = \begin{bmatrix} \mathfrak{S}_{11} & \mathfrak{S}_{12} \\ \mathfrak{S}_{21} & \mathfrak{S}_{22} \end{bmatrix}$. The systems of equations to be solved are

$$X_{11}\,\mathfrak{S}_{11} + X_{12}\,\mathfrak{S}_{21} = Y_{11} \tag{5.4.3-A}$$

$$X_{11}\,\mathfrak{S}_{12} + X_{12}\,\mathfrak{S}_{22} = Y_{12} \tag{5.4.3-B}$$

$$X_{21}\,\mathfrak{S}_{11} + X_{22}\,\mathfrak{S}_{21} = Y_{21} \tag{5.4.3-C}$$

$$X_{21}\,\mathfrak{S}_{12} + X_{22}\,\mathfrak{S}_{22} = Y_{22}. \tag{5.4.3-D}$$

Then, since X_{11} is square, equations (5.4.3-A and C) give the pair of simultaneous equations in $\mathfrak{S}_{11}$ and $\mathfrak{S}_{21}$:

$$\begin{aligned} X_{21}\,\mathfrak{S}_{11} + (X_{21}X_{11}^{-1}X_{12})\,\mathfrak{S}_{21} &= X_{21}X_{11}^{-1}Y_{11} \\ X_{21}\,\mathfrak{S}_{11} + \qquad\qquad X_{22}\,\mathfrak{S}_{21} &= Y_{21} \\ \Longrightarrow\ (X_{22} - X_{21}X_{11}^{-1}X_{12})\,\mathfrak{S}_{21} &= Y_{21} - X_{21}(X_{11}^{-1}Y_{11}). \end{aligned}$$

Write $H = X_{11}^{-1}X_{12}$ and $A = X_{11}^{-1}Y_{11}$. The first problem is to solve $X_{11}H = X_{12}$ for H and then $X_{11}A = Y_{11}$ for A. The method is to use equilibrated PLU

decomposition of X_{11} followed iterative improvement (see [57]) and then by forward and backward substitution to derive H and A. Here L and U are lower and upper triangular matrices respectively and P is a permutation matrix. Equilibration helps ensure numerical stability and the reciprocal condition number of the equilibrated array, the pivot growth factor, and forward and backward error bounds can be estimated. These checks, explained in the following paragraphs, are performed at every junction and at every scattering product. Thus, the numerical stability of the procedure can be checked at every step for every azimuthal order and the model aborted if it approaches numerical singularity or the reciprocal condition number indicates possible or actual instability.

A detailed account of the meanings of these error checking terms can be found in [32]; a summary follows. Suppose that when solving for X in some linear system $AX = B$ the exact solution is X and the calculated solution is $\bar{X}$; the forward error is defined to be the estimate of the error bound on the columns of $\bar{X}$ given by

$$\max_j \frac{\left|\bar{X}_{kj} - X_{kj}\right|}{\max_k \left|\bar{X}_{kj}\right|}.$$

The estimated component-wise backward error bound is the smallest change in any element of A or of B that would render the solution $\bar{X}$ exact. The reciprocal condition number is a measure of the closeness to singularity of the system of equations and was defined in equation (5.1.1). A good and stable model will be far from singular (to machine precision) at every step.

The reciprocal pivot growth factor is the ratio of uniform norms $\| \mathfrak{A} \|_\infty / \|U\|_\infty$, where $\mathfrak{A} = PLU$, P is the permutation matrix, L the lower triangular matrix and U the upper triangular matrix of the decomposition. If $\| \mathfrak{A} \|_\infty / \|U\|_\infty \ll 1$ the stability of the LU decomposition is poor and the computation should be aborted. The code (SKITTER) written for the modelling of the Planck multi mode horns sets a high threshold on both the reciprocal condition number and the reciprocal pivot growth factor to catch any instability and ensure high numerical precision. For standard corrugated horns no instability has ever been detected in models run with SKITTER, but it is a simple matter to alter corrugated horn geometry files to give a mechanically realistic horn that resulted in scattering operators close to singularity.

Since the code was written as a research tool (rather than for commercial applications) it was written to abort on encountering any hint of numerical instability. Mistakes in the horn geometry file have been caught by these stability checks, the code reporting the junction number where the error occurred, step size, and so on. In horn profile optimisation code, where the step size might want to go to zero, it would be more useful to detect the instability and then to use the procedures in section 5.2 to recalculate the operators P or Π, and then repeat the scattering calculation at the step.

PLU decomposition only needs to be done once and is a step in the process of finding an inverse matrix. Forward and backward substitution on an $N \times N$ array are processes of order $2N^2/3$ and LU decomposition is of order $N^3/3$; thus, the cost of computing H is of order n_E^3. With the PLU factorisation already available from the calculation of H, computing $A = X_{11}^{-1} Y_{11}$ is of order $2n_E^3/3$. It will be used twice more at a cost of $2n_E^3/3$. The total cost is then of order $3n_E^3$ whereas computing the inverse and using it three times would be a cost of order $5n_E^3$. (See reference [57] for a discussion of the complexity of these operations.)

The solutions A and H give $(X_{22} - X_{21}H)\,\mathfrak{S}_{21} = Y_{21} - X_{21}A$ with $\mathfrak{S}_{21}$ to be solved for. Let $T = (X_{22} - X_{21}H)$. Again use PLU decomposition of T and solve for $\mathfrak{S}_{21}$. This decomposition of T is of order $2n_M^3/3$ and the result will be used twice.

From the equation $X_{11}\,\mathfrak{S}_{11} = Y_{11} - X_{12}\,\mathfrak{S}_{21}$, use the existing PLU decomposition of X_{11} and values of X_{12} and $\mathfrak{S}_{21}$ to solve for $\mathfrak{S}_{11}$.

Equations (5.4.3-B and D) give the simultaneous equations to solve for $\mathfrak{S}_{22}$:

$$\begin{aligned}
X_{21}\,\mathfrak{S}_{12} + (X_{21}X_{11}^{-1}X_{12})\,\mathfrak{S}_{22} &= X_{21}X_{11}^{-1}Y_{12} \\
X_{21}\,\mathfrak{S}_{12} + \qquad\qquad X_{22}\,\mathfrak{S}_{22} &= Y_{22} \\
\Longrightarrow\ (X_{22} - X_{21}X_{11}^{-1}X_{12})\,\mathfrak{S}_{22} &= Y_{22} - X_{21}(X_{11}^{-1}Y_{12}) \\
\Longrightarrow\ T\,\mathfrak{S}_{22} &= Y_{22} - X_{21}(X_{11}^{-1}Y_{12}).
\end{aligned}$$

Again, the existing PLU decomposition of X_{11} is used to solve for B in $X_{11}B = Y_{12}$ giving $T\,\mathfrak{S}_{22} = Y_{22} - X_{21}B$ which is solved for $\mathfrak{S}_{22}$ using the existing PLU decomposition of T. Finally, calculate $\mathfrak{S}_{12} = B - H\,\mathfrak{S}_{22} = -H(I + \mathfrak{S}_{22})$ and the complete solution to S_{11} has been obtained. No operations of order greater than $\max\{n_E^3, n_M^3\}$ have been performed, and those operations are unavoidable matrix multiplications.

With S_{11} computed the remainder of the components of the full scattering operator S are computed in the following strict order:

$$\mathfrak{X} = I + S_{11}, \quad S_{21} = G\,\mathfrak{X}, \quad S_{12} = \mathfrak{X}\,K, \quad S_{22} = GS_{12} - I. \tag{5.4.4}$$

The S-matrix then has the form given in equation (3.3.2). In equation (5.4.4) the $G\,\mathfrak{X}$ and GS_{12} are simple multiplication of a real matrix times a complex matrix and requires no special coding; the product $\mathfrak{X}\,K$ can be carried out by a procedure similar to that followed in forming GK in §5.4.1 using the set multiplication tables derived there. Note that if the impedance ratio $Z = Z_a/Z_b$ is a complex scalar, then once S_{11} has been calculated it has to be multiplied by Z and the identity matrices in equation (5.4.4) have to be replaced by ZI.

The relative complexity of the methods for calculating the S-matrix at a junction described in this section and calculation directly from equations (3.3.2) are, counting order N^3 terms only: $8(n_E + n_M)^3$ from equations (3.3.2) provided no unnecessary calculations are made, compared with $20n_E^3/3 + 13n_M^3/3 + 8n_E^2 n_M + 6n_E n_M^2$ for the method above. Equating one clock cycle to one operation, the saving is approximately $4n_E^3/3 + 12n_E^2 n_M + 18n_E n_M^2 + 11n_M^3/3$ cycles. Thus, writing $\alpha = n_M/n_E$, the relative speed of identically compiled code to calculate the S-matrix at a junction would be a factor in the range $(1.846, 2.56]$ for $\alpha \in [1, \infty)$, with greatest gains as $\alpha \to 1_+$.

5.4.3 Computation of the scattering product

Equations (3.2.13-A to D) define the scattering product between two S-matrices. Once again the problem is to compute the product with the minimum of operations, but this time all matrices are complex. Write equations (3.2.13) as $C = A \odot B$. These equations are simply matrix products and sums with the exceptions of the two terms $(I - A_{22}B_{11})^{-1}$ and $(I - B_{11}A_{22})^{-1}$, so that it is in the calculation of those two terms that the efficiency can be obtained.

Observe that $(I - B_{11}A_{22})^{-1}B_{11} = B_{11}(I - A_{22}B_{11})^{-1}$ so that the system of equations can be rewritten as

$$C_{11} = A_{11} + (A_{12}B_{11})(I - A_{22}B_{11})^{-1}A_{21} \tag{5.4.5-A}$$

$$C_{12} = (A_{12}B_{11})(I - A_{22}B_{11})^{-1}(B_{11}^{-1}B_{12}) \tag{5.4.5-B}$$

$$C_{21} = B_{21}(I - A_{22}B_{11})^{-1}A_{21} \tag{5.4.5-C}$$

$$C_{22} = B_{22} + B_{21}(I - A_{22}B_{11})^{-1}(A_{22}B_{12}) \tag{5.4.5-D}$$

Clearly the products involving $(I - A_{22}B_{11})^{-1}$ must be solved, but the methods of PLU factorisation from subsection 5.4.2 can be used. The only question mark is over the efficiency of solving the term $B_{11}^{-1}B_{12}$.

At any step in the modelling process the array B will be the most recent S-matrix to have been calculated prior to the scattering product. The calculation of $D = B_{11}^{-1}B_{12}$ is to be done by solving $B_{11}D = B_{12}$, but that is has the form

$$\begin{aligned} B_{11}D = (I + KG)^{-1}(I - KG)D = B_{12} = (I + B_{11})K = 2(I + KG)^{-1}K \\ \implies (I - KG)D = 2K. \end{aligned} \tag{5.4.6}$$

From §5.4.2, the components of $Y = (I - KG)$ and of K are already in storage from the calculation of B_{11}, so the problem is to solve for D in $YD = 2K$. The problem here is smaller than the problem of solving for B_{11} because component $K_{ME} = 0$. It is also a smaller problem than solving $(I - B_{11}A_{22})^{-1}B_{12}$ both for the same reason and because of the saving of the order $(n_E + n_M)^3$ matrix multiplication $B_{11}A_{22}$. The problem is then to use the methods of §5.4.2 to solve for D_{21}, D_{11}, D_{22} and D_{12}, strictly in that order, from the set of equations

$$\begin{aligned} (Y_{22} - Y_{21}Y_{11}^{-1}Y_{12})D_{21} &= -2Y_{21}Y_{11}^{-1}K_{EE} \\ D_{11} &= 2Y_{11}^{-1}K_{EE} - Y_{12}D_{21} \\ (Y_{22} - Y_{21}Y_{11}^{-1}Y_{12})D_{22} &= 2(K_{MM} - Y_{21}Y_{11}^{-1}K_{EM}) \\ D_{12} &= 2Y_{11}^{-1}K_{EM} - Y_{12}D_{22}. \end{aligned} \tag{5.4.7}$$

Finally, the solution to equations (5.4.5) follows essentially the same procedure as given in 5.4.2.

In order to avoid excessive complexity of notation write

$$A_{22} = \begin{bmatrix} A_{EE} & A_{EM} \\ A_{ME} & A_{MM} \end{bmatrix}, \quad B_{11} = \begin{bmatrix} B_{EE} & B_{EM} \\ B_{ME} & B_{MM} \end{bmatrix}, \quad A_{21} = \begin{bmatrix} \mathfrak{A}_{11} & \mathfrak{A}_{12} \\ \mathfrak{A}_{21} & \mathfrak{A}_{22} \end{bmatrix}$$

and define $X = (I - A_{22}B_{11})$ so that

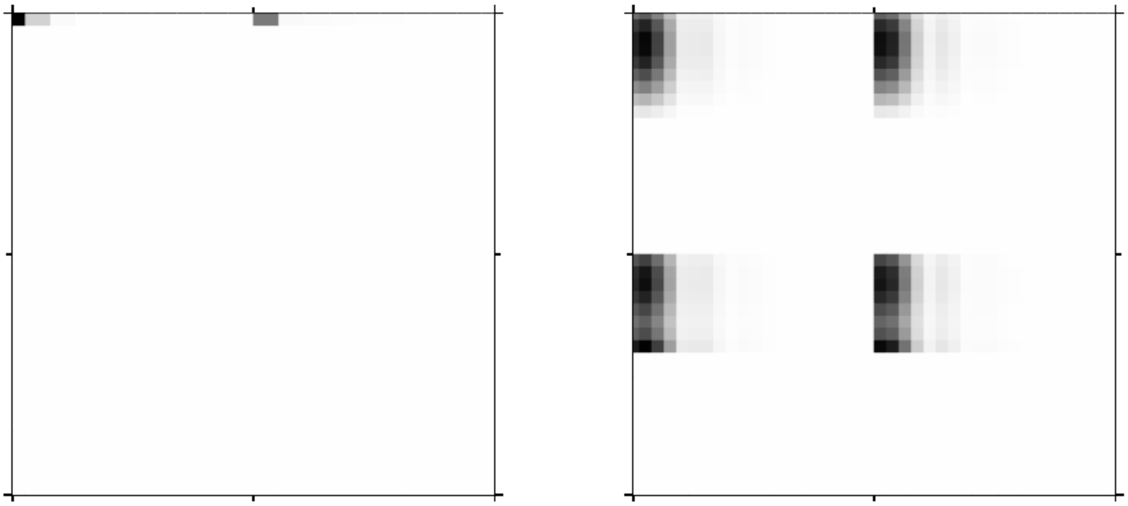
$$\begin{aligned} X_{11} &= I - A_{EE}B_{EE} + A_{EM}B_{ME} & X_{12} &= A_{EE}B_{EM} + A_{EM}B_{MM} \\ X_{21} &= A_{ME}B_{EE} + A_{MM}B_{ME} & X_{22} &= I - A_{ME}B_{EM} + A_{MM}B_{MM}. \end{aligned}$$

Equation (5.4.5-A) then becomes $C_{11} = A_{11} + A_{12}B_{11}X^{-1}A_{21}$, the solution of which requires that first the solution, E say, to $XE = A_{21}$ be found. To do so requires solving $(X_{22} - X_{21}X_{11}^{-1}X_{12})E_{22} = X_{22}\,\mathfrak{A}_{22} - X_{21}\,\mathfrak{A}_{12}$ which is done by first finding the PLU factorisation of X_{11}, then solving for F in $X_{11}F = X_{12}$, and thereafter finding the PLU factorisation of $(X_{22} - X_{21}F)$. From E_{22} comes $E_{12} = \mathfrak{A}_{12} - FE_{22}$. In like manner solve for J in $X_{22}J = X_{21}$ and then $(X_{11} - X_{12}J)E_{11} = X_{11}\,\mathfrak{A}_{11} - X_{12}\,\mathfrak{A}_{21}$ followed by $E_{21} = \mathfrak{A}_{21} - JE_{11}$. Once that has been done, with D from equations (5.4.7) the solution to all of equations (5.4.5-A to D) is simply a matter of matrix multiplication and addition, using the PLU factorisations of $(X_{22} - X_{21}X_{11}^{-1}X_{12})$ and of $(X_{11} - X_{12}J) = (X_{11} - X_{12}X_{22}^{-1}X_{21})$.

The relative complexity of calculating the scattering product by an efficient but straightforward implementation of equations (3.2.13) compared with using the above method is approximately $14(n_E^3 + n_M^3) + 42n_En_M(n_E + n_M)$ compared with

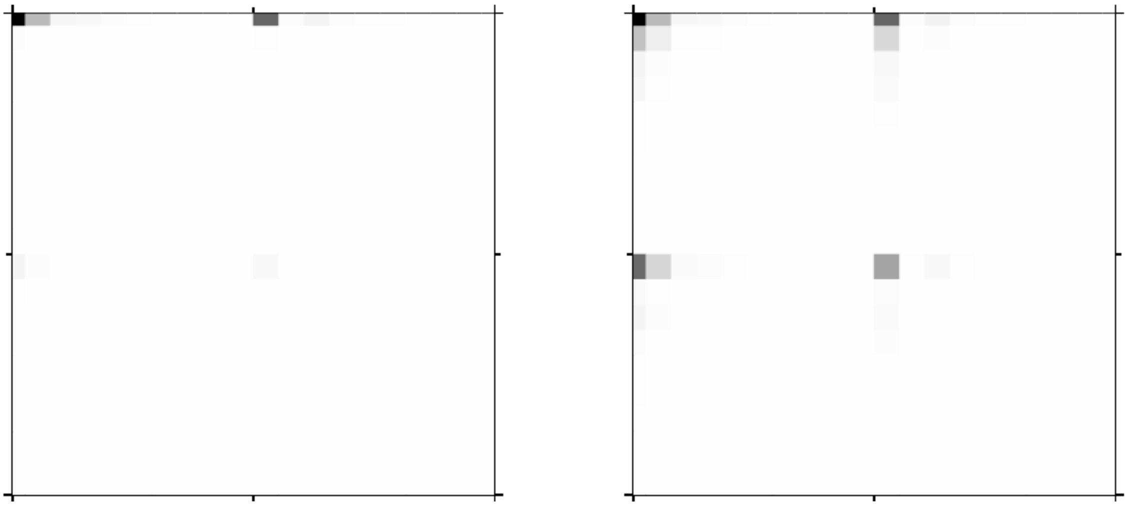
$$9(n_E^3 + n_M^3) + 26n_En_M(n_E + n_M) + 2n_M^2(n_E + n_M).$$

In addition to reductions in computational time, the methods outlined above make smaller demands on cache and stack size and can be expected to lead to higher precision as a result both of the reduction in floating point operations and less rounding.



(a) C*l*OVER 150 GHz.

(b) PLANCK 545 GHz full pixel.



(c) PLANCK 100 GHz full pixel.

(d) PLANCK 100 GHz Back-Back.

Figure 5.1: The scattering amplitudes in the S_{21} matrices for (a) the C*l*OVER 150 GHz horn, (b) the azimuthal order 1 field of the Planck 545 GHz multi-mode horn, (c) the full horn assembly of the Planck 100 GHz single-mode horn, and (d) the back-to-back section of the the horn in (c). Each grey square is the amplitude of an S_{21} element. These plots illustrate both the sparsity of the operator arrays and the modal purity of the C*l*OVER ultra-Gaussian horn and the full pixel assembly of a Planck single-mode horn. Compare (d) with (c): These plots also indicate that the accuracy of the field modelling in Planck would be compromised by modelling the back-to-back sections of the horns only, no matter how many modes were included in the models. That is particularly true for the multi-mode horns in which the full pixel has a more complex mode structure than the back-to-back section.

Chapter 6

Representation of fields: Schmidt triples

The problem at hand is to find the smallest set of data that completely encapsulates the scattering properties of the instrument. This set will turn out to be a set of vector-scalar-vector triples, the cardinality of the set being the rank of the scattering operator. The set is simple to calculate from the scattering operator and its application to system analysis and performance prediction leads to great efficiency when there is more than one mode.

For a single-mode system, at any one frequency, the field radiated from the aperture of the horn is represented by a single vector in the Hilbert space of aperture fields. In the mode-matching formalism that one vector is the image of the input vector under the action of the S_{21} operator. (Strictly speaking, it is essentially the image under S_{21} of the entire basis for the space of fields feeding into the horn from the bolometer cavity.) For a multi-mode system the radiated field is a set of fields. Each of those fields is, once again, simply a vector in the aperture Hilbert space. However, the representation of the field in terms of the columns of the matrix representation of S_{21} is not usually an efficient representation in the sense that the underlying space spanned by the fields is smaller in dimension than the number of fields in the scattering model – the columns of S_{21} form a non-minimal spanning set for the fields at the horn aperture. For example, in the Planck 857 GHz horns the structure of the horns leads to a requirement for large numbers of waveguide modes to give a stable model of the scattering process (stable in the sense that adding more modes to each azimuthal order in the model does not significantly alter the radiated field). The model requires in excess of sixty radial modes in each of three to five azimuthal orders, the precise numbers being frequency dependent, but the fields radiated from the horn do not span a dimension sixty dimensional space at any frequency over the band: the dimension of the vector space spanned by the field is both considerably smaller and frequency dependent. What is required is an optimal representation of the radiated fields: optimal in the sense of a minimal set of vectors spanning the subspace in which the field lies, and the vector coordinates of the field in that space. What will be found is an orthonormal basis for the smallest subspace containing the field and an expansion of the aperture fields in that basis, the image of the basis under the S_{21} operator and the contraction factors for each vector.

At each azimuthal order the model requires some N radial modes to represent the scattering in the waveguide at a particular frequency. We can always use the same number of radial modes at every scattering junction provided the number is sufficiently large, so the model of the space of radiated fields at any one azimuthal order has dimension N. The naïve way to model the radiated field pattern is to take these N fields at each azimuthal order and propagate them through whatever system follows the horn: the Planck telescope, in this instance. The answer would be correct and is a useful check on the method to be described below, but the process is grossly inefficient. Rather than propagate N fields in M azimuthal orders it would be better to find a new basis of fields that is of minimal dimension for the frequency and azimuthal order. For the Planck 857 GHz horn the scattering model was of the order 64×3 to 64×5 dimensional, depending upon the frequency, but these fields were representable in spaces that varied from ten to twenty dimensions as a function of frequency – ten dimensional at the low frequency end of the band, rising to twenty dimensional at the top end. The next section describes the method used to find these new representations of the fields – methods taken from functional analysis and linear algebra. Here the application is only to the relation between the input field and radiated field, but it applies equally to the fields at every stage in the scattering or propagation. The method necessarily has a more mathematical flavour than the rest of the thesis; however it is presented at a fairly non-technical level and basic mathematical terms are used without definition or explanation.

6.1 Representation of fields by Schmidt pairs

Let $S : \mathcal{H}_1 \longrightarrow \mathcal{H}_2$ be a compact linear operator between separable Hilbert spaces and $\{\boldsymbol{e}_n\}$, $n \in \mathbb{N}\cup\{0\}$, be a complete orthonormal sequence in the orthogonal complement of the kernel, $\mathrm{Ker}(S)^{\perp}$. Then every element $x \in \mathcal{H}_1$ can be written in the form $\boldsymbol{x} = \sum x^n \, \boldsymbol{e}_n + \boldsymbol{x}'$ for some $\boldsymbol{x}' \in \mathrm{Ker}(S)$, and the scalars x^n will generally be complex. Now suppose that the e_n are eigen-vectors of the Hermitian (self adjoint) bounded operator $S^{\dagger}S \in \mathscr{B}(\mathcal{H}_1)$, then for all $n \in \mathbb{N}\cup\{0\}$ there exist eigenvalues $\lambda_n \in \mathbb{R}_{\geq 0}$ and an eigenfunction expansion:

$$S^{\dagger}S\,\boldsymbol{x} = \sum S^{\dagger}\left(S\left(x^n\,\boldsymbol{e}_n + \boldsymbol{x}'\right)\right) = \sum x^n S^{\dagger}S\,\boldsymbol{e}_n = \sum x^n \lambda_n\,\boldsymbol{e}_n \tag{6.1.1}$$

in which the last sum is finite if the nullity of S, $\dim \mathrm{Ker}(S)$, is finite.

Let $s_n \overset{\text{def}}{=} \sqrt{\lambda_n}$ and for all $s_n > 0$ define $\boldsymbol{f}_n = s_n^{-1} S\,\boldsymbol{e}_n \in \mathcal{H}_2$. Any pair $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{H}_1 \times \mathcal{H}_2$ for which $S\,\boldsymbol{x} = s\,\boldsymbol{y}$ for some $s \in \mathbb{R}$ is called a *Schmidt pair.* By construction the pairs $(\boldsymbol{e}_n, \boldsymbol{f}_n)$ are Schmidt pairs for all n with $\boldsymbol{e}_n \in \mathrm{Ker}(S)^{\perp}$, and for all $\boldsymbol{x} = \sum x^n\,\boldsymbol{e}_n + \boldsymbol{x}'$

$$S\,\boldsymbol{x} = \sum s_n x^n\,\boldsymbol{f}_n\,. \tag{6.1.2}$$

The set $\{\boldsymbol{f}_j = s_j^{-1} S\,\boldsymbol{e}_j\}$ is an orthonormal basis for the image of the orthogonal complement of its kernel since, by definition of the adjoint

$$\left\langle \boldsymbol{f}_i \,|\, \boldsymbol{f}_j \right\rangle_{\mathcal{H}_2} = s_i^{-1} s_j^{-1} \left\langle S\,\boldsymbol{e}_i \,|S\,\boldsymbol{e}_j \right\rangle_{\mathcal{H}_2} = s_i^{-1} \left\langle \boldsymbol{e}_i \,|S^{\dagger}(s_j^{-1} S\,\boldsymbol{e}_j) \right\rangle_{\mathcal{H}_1} = s_i^{-1} s_j \left\langle \boldsymbol{e}_i \,|\, \boldsymbol{e}_j \right\rangle_{\mathcal{H}_1} = \delta^i_j,$$

If the domain and co-domain of a scattering operator S have orthonormal bases $\{\boldsymbol{e}_i\}$ and $\{\boldsymbol{f}_j\}$ respectively, then, since S is linear and $\mathcal{H}_1$ and $\mathcal{H}_2$ have inner products, there is a matrix representation of S which is the array $[s_{ij}]_{i,j=1}^{\infty}$ in which the s_{ij} are the generalised Fourier coefficients $s_{ij} = \left\langle \boldsymbol{f}_j \,|S\,\boldsymbol{e}_i \right\rangle_{\mathcal{H}_2}$. Clearly this is the case here: write $\mathcal{H}_1 = \mathscr{S}_1 \oplus \mathrm{Ker}(S)$ and $\mathscr{S}_2 = S(\mathscr{S}_1) \subset \mathcal{H}_2$, then the set of all Schmidt pairs $\{(\boldsymbol{e}_n, \boldsymbol{f}_n)\}_{n=1}^{N}$ is an orthonormal basis for $\mathscr{S}_1 \oplus \mathscr{S}_2$, and with respect to these orthonormal bases the matrix representation of S is diagonalised. Therefore, provided the scattering operators satisfy the required conditions, finding the Schmidt pair bases and the associated Schmidt numbers, s, for the operator will immediately give an expansion of the field in the radiating aperture in terms of an orthonormal basis and the magnitude of the contribution of the basis fields, the Schmidt modes, to the total field will be given by the Schmidt numbers. Henceforth the Schmidt numbers will be termed the s-numbers.

From now on the concept of Schmidt pair for an operator S will be extended to a triple, the set of which will be referred to as a Schmidt triple: $\{(\boldsymbol{e}_n, s_n, \boldsymbol{f}_n)\}_{n=1}^{N}$. This set encapsulates *all* information about the system described by S – for any system, but for a horn in particular, the Schmidt triple will be the complete and minimal encapsulation of its scattering properties that was sought.

If T_1 is any compact linear operator associated with a partial system and T_2 is any such operator for the next part of the system with $\mathrm{Dom}(T_2) = \mathrm{Codom}(T_1)$ (for example the transmission operator for a horn feeding a telescope and the transmission operator for the telescope) and if T_1 and T_2 have Schmidt triples $\{(\boldsymbol{e}_n, s_n, \boldsymbol{f}_n)\}_{n=1}^{N}$ and $\{(\boldsymbol{g}_k, t_k, \boldsymbol{h}_k)\}_{k=1}^{K}$, it is not necessarily true either that $\boldsymbol{g}_k = \boldsymbol{f}_n$ for any pair of indices (n, k), or that $K = N$, though clearly span $\langle \boldsymbol{f}_1, \ldots, \boldsymbol{f}_N \rangle$ and span $\langle \boldsymbol{g}_1, \ldots, \boldsymbol{g}_K \rangle$ are subspaces of the same Hilbert space and must have non-trivial intersection if there is to be any throughput. Rather, the operator product acts on the $\boldsymbol{e}_n$ by

$$T_2 T_1 \, \boldsymbol{e}_n = s_n T_2(\boldsymbol{f}_n) = \sum_{k=1}^{K} s_n t_k \, \langle \boldsymbol{g}_k \mid \boldsymbol{f}_n \rangle \, \boldsymbol{h}_k \, . \tag{6.1.3}$$

Therefore it cannot be expected that, even when the product operator has a physical meaning (such as is the case for the product of two total transmission operators as in equation (3.7.3)), a Schmidt field description of the system will be achieved from Schmidt field descriptions of the components. That would only be the case if either $\{\boldsymbol{f}_n\} \subseteq \{\boldsymbol{g}_k\}$ or $\{\boldsymbol{g}_k\} \subseteq \{\boldsymbol{f}_n\}$ as set inclusions, not merely as subspace intersections.

If the Schmidt triple description of a system is to be valid the operators must be shown to be of the required type: compact linear operators between separable Hilbert spaces. For any such operators that are non-hermitian and associated with some physical transfer process the analysis will apply, under the appropriate restriction to the orthogonal complement of its kernel. In an optical system the information transfer must be finite because the system has finite resolution. Physically this means that the system operator has a finite dimensional orthogonal complement of the kernel. Then $\mathcal{H} = \mathscr{S} \oplus \mathrm{Ker}(S)$ with $\mathscr{S} \subset \mathcal{H}$ closed and linear and there exists an isomorphism $\mathscr{S} \cong \mathbb{C}^N$ for some finite N. Since $\mathbb{C}^N$ is separable and complete, so is $\mathscr{S}$, and that is the space to which the theory applies, not to the whole of $\mathcal{H}$.

To argue for compactness use the fact that the operator describes a linear (by assumption) physical process and since the input has bounded energy, $\| \boldsymbol{x} \|_2^2 < \infty$, then $\|S(\boldsymbol{x})\|_2 < \infty$. In particular, there is a complete orthonormal sequence in $\mathscr{S}$ that is the sequence of eigenvectors of $S^\dagger S$, and $\sum \|S(\boldsymbol{e}_n)\|_2^2 \leq \sum \| \boldsymbol{e}_n \|_2^2 = \mathrm{rank}(S)$. Therefore S is Hilbert-Schmidt and therefore compact *a fortiori.* Mathematical details can be found in any text on functional analysis; see [22], [29], [58], [74]. Rather than this function-analytic argument for compactness a heuristic, physical cum mathematical, argument for the plausibility of compactness in applications to waveguides is the following: Every finite rank operator is compact since, if the range of S is R, then R inherits the norm from $\mathcal{H}_2$ and has the metric determined by this norm. It is then a finite dimensional metric space and closed bounded sets in a finite dimensional normed space are compact. A sequence, $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$, of vectors representing physical fields in $\mathcal{H}_1$ must be bounded, so its image $\{S\,\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ is bounded in R and hence has compact closure and so has a cluster point, $\boldsymbol{y}$ say, in R, but R is a metric space and so $S\,\boldsymbol{x}_n \to \boldsymbol{y}$. However a horn or waveguide operator is *essentially* of finite rank (meaning that there is a finite dimensional subspace onto which the fields can be projected without significant loss of information or change in beam profile) since propagation of the field through the guide results in attenuation of the field for all modes above cut-off. Thus, for a physical guide, given an input field $\boldsymbol{x} = \sum a_k\,\boldsymbol{x}_k$ and any $\varepsilon \in \mathbb{R}_{>0}$ (a lower bound on measurable power, say) there is some mode index, K_ε, such that the residual power in the 'tail' of the mode expansion is less than ε; that is

$$0 = \| \boldsymbol{x} - \sum_{k=1}^{\infty} a_k \boldsymbol{x}_k \|_2 \leq \| \boldsymbol{x} - \sum_{k=1}^{K_\varepsilon} a_k \boldsymbol{x}_k \|_2 + \| \sum_{k=K_\varepsilon+1}^{\infty} a_k \boldsymbol{x}_k \|_2 < \| \boldsymbol{x} - \sum_{k=1}^{K_\varepsilon} a_k \boldsymbol{x}_k \|_2 + \varepsilon.$$

Then for practical purposes the model of the propagated field can be take to be $\boldsymbol{x} \sim \boldsymbol{x}_\varepsilon \overset{\text{def}}{=} \sum_{k=1}^{K_\varepsilon} a_k \boldsymbol{x}_k$ with image $S\boldsymbol{x}_\varepsilon = \sum_{k=1}^{K_\varepsilon} a_k S\boldsymbol{x}_k \in \operatorname{span}\{S\boldsymbol{x}_1, \ldots, S\boldsymbol{x}_{K_\varepsilon}\}$, and the model can be taken to be an operator with both domain and range of dimension K_ε. The model of the operator is then a matrix in $\mathrm{M}(K_\varepsilon; \mathbb{C})$ and the model function spaces are just $\mathbb{C}^{K_\varepsilon}$. If this was not the case, then the entire processes of modelling horns by the scattering matrices would be invalid. In the next section explicit examples from the Planck multi-mode horns will be given showing that the horn models are essentially of rank much smaller than the dimension of the waveguide mode space from which the Schmidt fields are derived, the rank being frequency dependent.

Note that the truncation of the partial series at the K^{th} term does not imply the truncation of the waveguide mode expansion of the vectors $\boldsymbol{e}_j$ or $\boldsymbol{f}_j$. These remain l^2 sequences, $\boldsymbol{f}_j = (f_1^{(j)}, f_2^{(j)}, \ldots, f_n^{(j)}, \ldots)$ with $\| \boldsymbol{f}_j \|_2 < \infty$ and entries $f_n^{(j)} \to 0$ as $n \to \infty$; likewise for all $\boldsymbol{e}_j$.

The relationship between the Schmidt fields for the finite dimensional model and singular value decomposition is simple and leads to one method for calculating the Schmidt triples: the singular value decomposition of a matrix gives $S = UDV^\dagger$ where V is unitary in the domain, U is unitary in the co-domain and D is diagonal. The rank of D is the rank of S and the columns of U corresponding to the non-zero elements of D form a unitary basis for the co-domain, while the columns of V form a basis for $\mathscr{S} = \operatorname{Ker}(S)^\perp$ which spans the observable fields. By unitarity of U there is a unitary equivalence $S^\dagger S = VD^2V^\dagger$ between $S^\dagger S$ and D^2 and if d_n is the n^{th} non-zero column of D, then $S^\dagger S\,\boldsymbol{e}_n = VD^2(0 \cdots 0\, 1_n\, 0 \cdots 0)^T = V(0 \cdots 0\, d_n^2\, 0 \cdots 0)^T = d_n^2\, \boldsymbol{e}_n$ and the required Schmidt triple is $(\boldsymbol{e}_n, s_n, \boldsymbol{f}_n)$, where $\boldsymbol{f}_n$ is the n^{th} column of U and $s_n = d_n = D_{nn} > 0$.

When the system Schmidt triples $\{(\boldsymbol{e}_n, s_n, \boldsymbol{f}_n)\}_{n=1}^{N}$ have been found, given any set of input fields, $\{\boldsymbol{g}_k\}_{k=1}^{K}$, the output of the system is given by the set of fields

$$\left\{ s_n \, \boldsymbol{f}_n \sum_{n=1}^{N} \langle \boldsymbol{g}_k \,|\, \boldsymbol{e}_n \rangle \right\}_{k=1}^{K} . \tag{6.1.4}$$

When modelling the aperture fields of a horn it is primarily the non-zero columns of the U derived from the S_{21} operator that is of interest; vectors that span the co-domain orthogonal to the image of the kernel. The Schmidt triples can be derived from either the S_{21} or S_{12} operators, and if they do not give the same fields, then somewhere in the modelling process the calculation has either been conducted with poor numerical accuracy, or incorrectly. That has been a useful check on the results of the Planck multi-mode modelling.

The Schmidt field expansion of the scattering operators was developed for the multi-mode systems, but is equally applicable to single-mode systems with a single triple characterising the system. The advantage of applying the Schmidt field transformations to the single-mode S_{21} operator is that it finds a hybrid field, expanded in the transverse field of the model, spanned by a single vector in both cavity and aperture, fields that are truly characteristic of the system within the limits imposed by the finite dimensional model. For a horn with very high spectral purity like the C*l*OVER horns the unitary transformation converts the columns of the S_{21} operator into a single aperture Schmidt field vector that is almost identical to the first column of the matrix, once s-number scaling is taken into account. For the Planck 100 GHz single-mode horns the spectral purity of the operator is not as great; consequently, more information is contained in the S_{21} matrix columns other than the first, and the unitary transformation to the Schmidt field recovers this information about field structure that would be lost if the simple procedure of taking the first column of the operator matrix as the field vector was followed. Propagating all column vector fields through the optical system to derive the far field beam would also recover the same information, but at unnecessary computational cost since the other columns contain little information. The models of the C*l*OVER 150 GHz and the Planck 100 GHz horns discussed in section 2.3 were re-run with sixtyfour mode models for a direct comparison of the S_{21} operator matrices and the Schmidt fields. The plots in figure 5.1, page 141, illustrate the point about spectral purity, but detailed inspection of the 64-dimensional complex Schmidt vectors and matrix columns is required to appreciate the true differences and these are discussed in the next section.

The idea of applying the concept of the Schmidt pair to optical systems seems to have been first put forward by Withington in [70] in the context of signal detected when imaging with phased arrays, following on from a series of papers [71][72][73]. Withington refers to the input fields as 'eigen-fields'. The term 'Schmidt field' is used here instead of 'eigen-field' for the express reason that eigen means 'the same', but the point of the Schmidt fields is that the fields are changed in the process of propagation, and the Schmidt field pairs for any system are 'optimal' basis for $\mathscr{S}_1 \oplus \mathscr{S}_2 \subset \mathcal{H}_1 \oplus \mathcal{H}_2$ in the sense that the fields propagated by the S_{21} operator can be successively approximated by partial sums

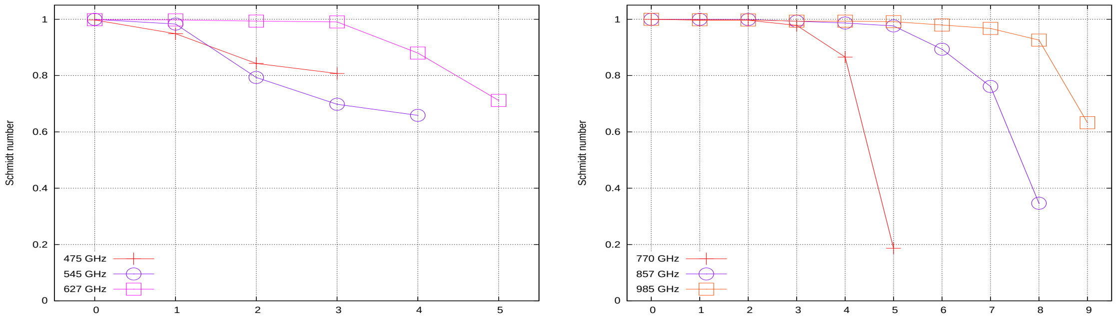
$$S_{21}(\boldsymbol{x}) = S_{21}\left(\sum_{n=1}^{N} x^n\, \boldsymbol{e}_n\right) = \sum_{n=1}^{K} s_n x^n\, \boldsymbol{f}_n + \boldsymbol{\varepsilon}_K, \qquad K \leq N \tag{6.1.5}$$

in a basis for which the 2-norm of the remainder vector, $\|\,\boldsymbol{\varepsilon}_K\,\|_2$, is smaller than for any other K-term approximation in any other basis, the ordering being the natural ordering of the s-numbers: $s_1 \geq s_2 \geq \cdots \geq s_K \geq \cdots \geq s_N > 0$. Here it is stressed that the concept is applicable to the fields propagated through an optical system of any kind (indeed to any physical process described by a compact linear operator between separable Hilbert spaces) and to scattering in a corrugated waveguide system in particular. Although the above presentation has skirted around analytical details, a careful analysis shows that the functional analytic theory is applicable to the situation.

6.2 ClOVER and Planck horn Schmidt triples

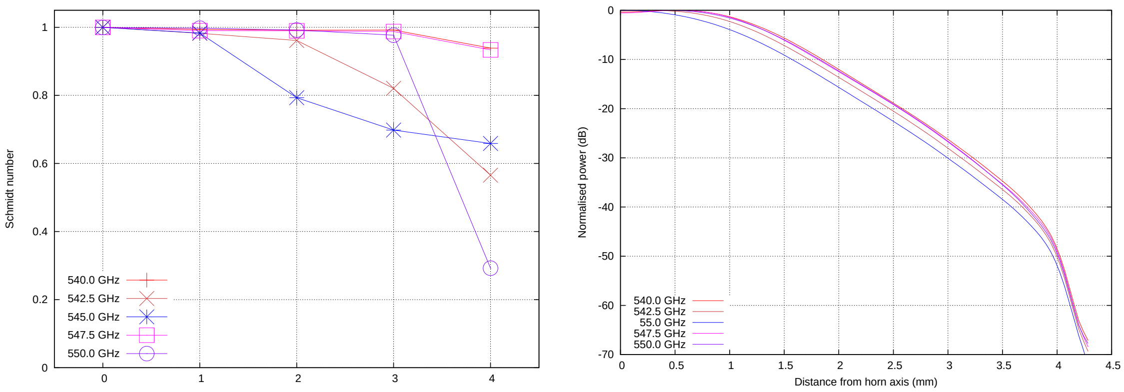
The S_{21} operator of the waveguide system is a function of frequency. Therefore, the characteristic triples of the system, $\{(\boldsymbol{\varphi}_n, s_n, \boldsymbol{\psi}_n)\}_{n=1}^{N}$, derived from the operator are also frequency dependent. Thus we should write $\{(\boldsymbol{\varphi}_n(\nu), s_n(\nu), \boldsymbol{\psi}_n(\nu))\}_{n=1}^{N(\nu)}$ to be precise, and that is to be understood hereafter. That this is the case over a frequency interval spanning a cut on is obvious because the dimension of the subspace spanned by the Schmidt vectors changes. Away from a cut on region the space remains of a constant dimension, but because a waveguide is a resonant system, the subspace itself, or the triples within the subspace, can vary. A useful mental picture is of a finite dimensional hyperplane wobbling slightly in the space of fields and the $\boldsymbol{\varphi}_n$, $\boldsymbol{\psi}_n$ and the s_n undergoing a Brownian motion type dance around the mean triple over the frequency interval. Call this mean point the local stability point for the frequency interval. When the frequency is increased to above, or crosses, an adjacent cut on frequency, the system $\{(\boldsymbol{\varphi}_n, s_n, \boldsymbol{\psi}_n)\}_{n=1}^{N}$ rapidly becomes a system $\{(\boldsymbol{\varphi}'_n, s'_n, \boldsymbol{\psi}'_n)\}_{n=1}^{N'}$ with $N' \geq N+1$ and settles down to wobble around a new local stability point. As seen with transitory cut-in of the azimuthal order four modes in the Planck 545 GHz back-to-back horn, this transition may be chaotic (see the order 4 curve in figure 2.15 (a), page 33).

Figure 6.3 shows the absolute values of the complex coordinates of the Schmidt vectors in the complex mode space $\mathscr{E} \oplus \mathscr{M}$ of the ClOVER 150 GHz horn aperture at 5 GHz intervals from 130 GHz to 180 GHz. The mode space is infinite dimensional, though in the numerical model giving these results the model space was $\mathbb{C}^{32} \oplus \mathbb{C}^{32}$. All Schmidt vectors are unit vectors by construction, and so the vector is a point on the unit sphere of fields with $\| \boldsymbol{\psi} \|_2 = 1$. The Schmidt field itself lies on the complex line through that point at a distance s that is the s-number for the field.

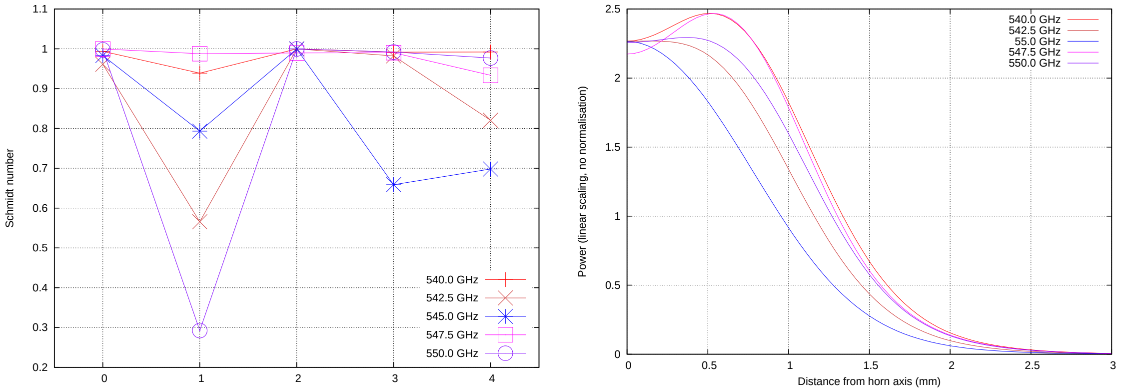


(a) Planck 545 GHz multi-mode horn. (b) Planck 857 GHz multi-mode horn.

Figure 6.1: The s-numbers for the aperture fields at three spot frequencies at the bottom, centre and top ends of the operating bands of (a) the Planck 545 GHz and (b) the 857 GHz horn assemblies. In both plots only those s-numbers $s_n > 10^{-4}$ are plotted. The TEM model space had dimension 54 for the 545 GHz horn and 64 for the 857 GHz, but it is seen from these plots that the effective dimension of the field space varies across the band and is always smaller than 18 (counting both polarisations).

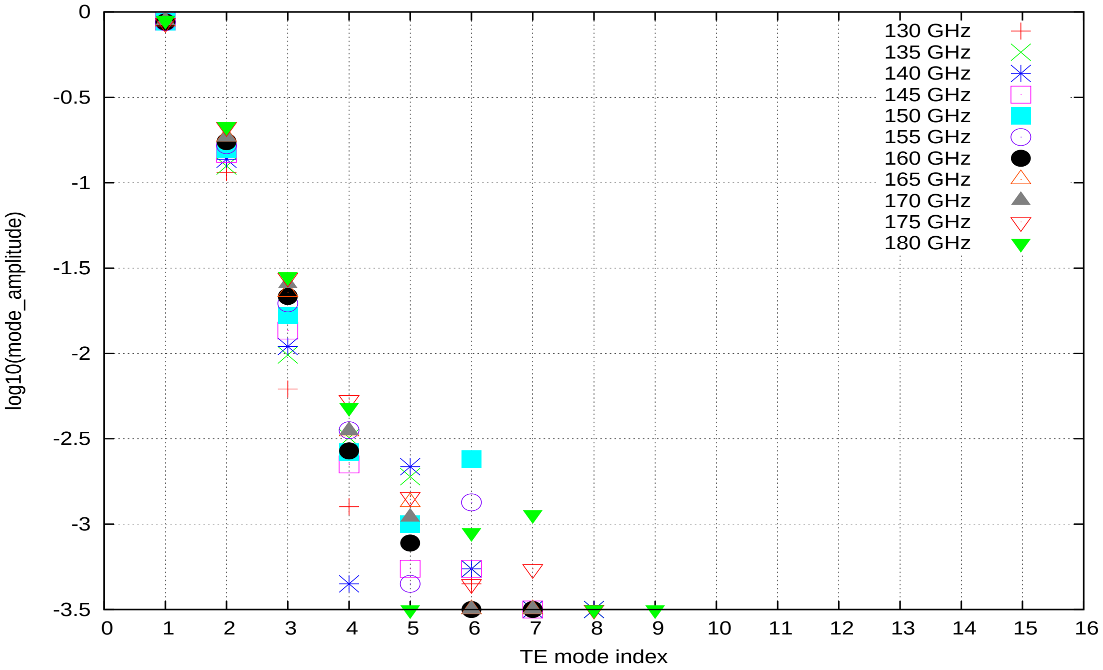


(a) Planck 545 GHz multi-mode horn. (b) Planck 545 GHz multi-mode horn.

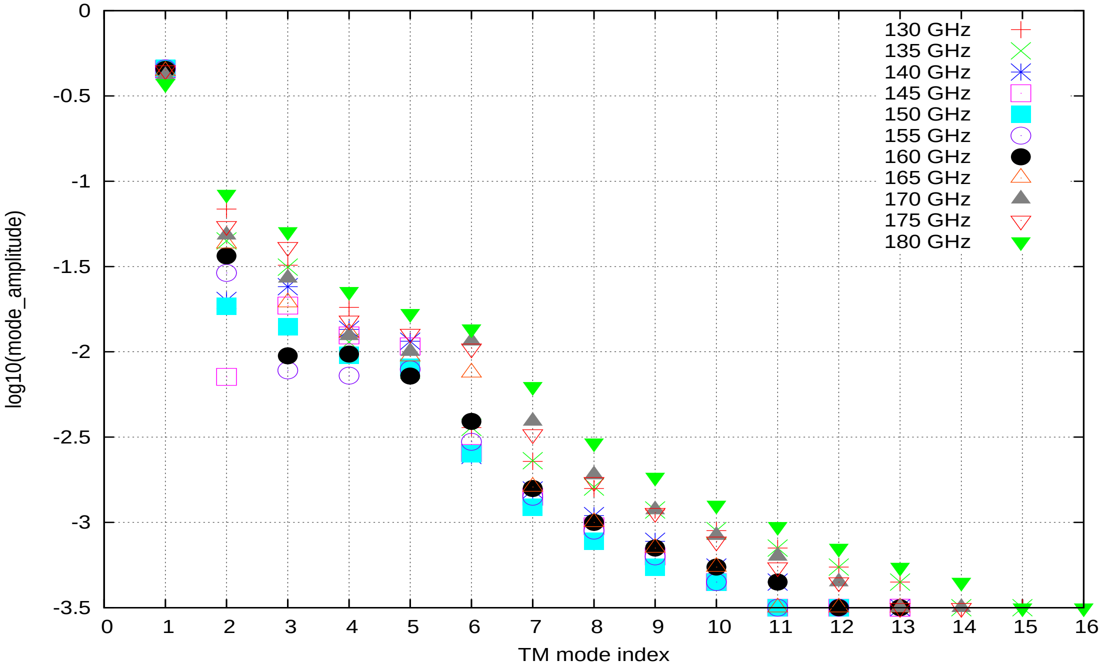


(c) Planck 545 GHz multi-mode horn. (d) Planck 545 GHz multi-mode horn.

Figure 6.2: Graph (a) shows the changing s-numbers in order of decreasing field power contribution and (b) shows the corresponding (normalised) aperture field cross-sections. All graphs cover the 540 to 550 GHz band in 2.5 GHz steps centred on 545 GHz. This is the chaotic cut-on region of figure 2.14 just above the cut on of an additional azimuthal order 3 field component and below the cut on of an additional order 1 component. The narrowness of the 545 GHz aperture field is primarily due to suppression of azimuthal orders 2 and 3. Figure (c) shows the contributions from the different azimuthal orders at each frequency in the following order on x-axis ordinate 0 through to 4: [01], [02], [11], [21] and [31]. Figure (d) is a linear version of (b) and clearly shows the changing beam shapes that result from the Schmidt mode contributions in (c). The colours in all plots correspond to the same frequencies.

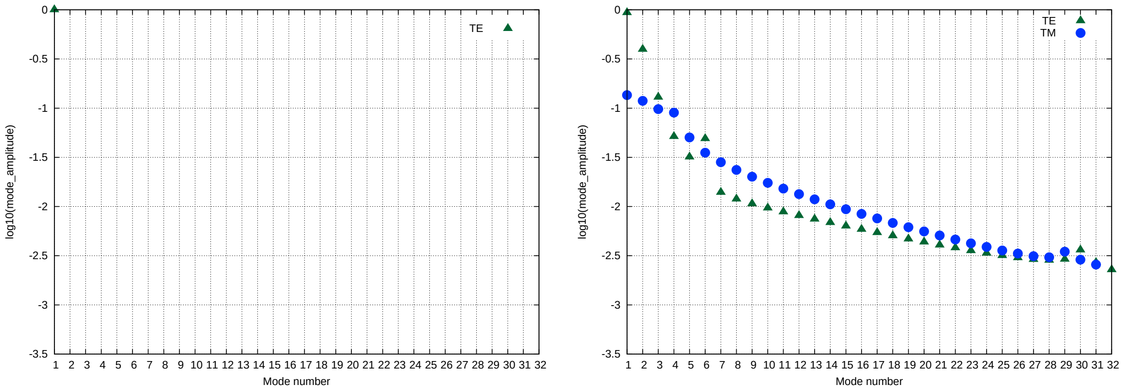


(a) ClOVER 150 GHz horn aperture field TE mode contributions.

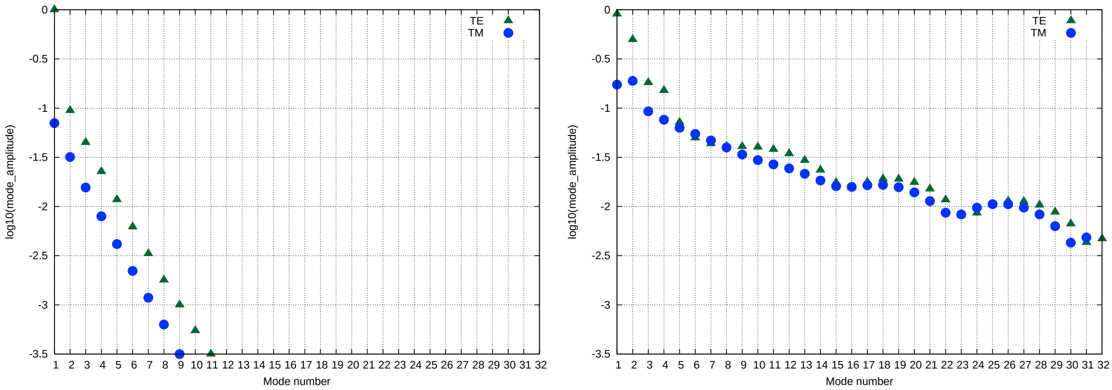


(b) ClOVER 150 GHz horn aperture field TM mode contributions.

Figure 6.3: The amplitudes of (a) the TE, and (b) the TM mode contributions to the aperture field Schmidt vectors of the ClOVER 150 GHz horn at eleven frequencies across the band. The mode ordering is the natural ordering by Bessel function root. These are the absolute values of the coordinates of the Schmidt field in the mode space $\mathscr{E} \oplus \mathscr{M}$ of the horn aperture. It is seen that at all frequencies the dominant mode is the fundamental TE mode followed by the fundamental TM mode. All models have 32 TE and 32 TM modes, but the amplitudes of the higher modes are below 5×10^{-4} and are not plotted.



(a) ClOVER 150 GHz horn waveguide feed. (b) ClOVER 150 GHz horn aperture.

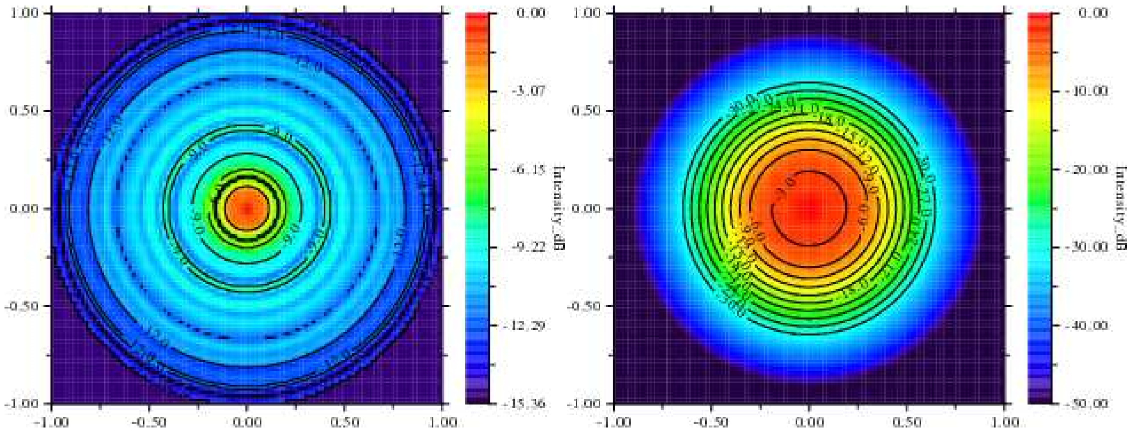


(c) Planck 143 GHz horn waveguide feed. (d) Planck 143 GHz horn aperture.

Figure 6.4: The amplitudes of the electric and the magnetic mode contributions to the waveguide field and aperture field Schmidt vectors of the ClOVER 150 GHz and the radiating section of the Planck 143 GHz horns at their band centres. Modes with amplitudes below 5×10^{-4} and are not plotted. In (b) the modes construct the hybrid $HE_{11} + HE_{12}$ field. The y-scales are $\log_{10}$(amplitude).

For a multi-mode system the picture is similar. As an example take the Planck 857 GHz horn at 995 GHz (the top end of the band). There, in each polarisation, there are eleven s-numbers greater than 10^{-4} (contributing a power fraction less than $10^{-8}/11$ to the beam; see figure 6.6 below). These are the calculated s-numbers for both the S_{21} and S_{12} operators. To each of the s-numbers there is associated a Schmidt vector in both the cavity and the aperture. Each set in its respective domain comprises mutually orthogonal vectors on the unit sphere, the set determining the hyperplane (a copy of $\mathbb{C}^{11}$) in which the field exists with coordinates given by the s-numbers. The model space here was isomorphic to $\mathbb{C}^{320}$ (five azimuthal orders of 64 modes) but the space containing the Schmidt fields used in modelling the beam is $\mathbb{C}^{10}$ since the [32]-mode contains too little power to warrant its inclusion in the beam pattern model; thus, the nullity of the horn operator in this model was 310 and a considerable computational time saving in calculating the far field beam patterns is to be expected. The matter of computation efficiency is discussed on page 158. The transverse waveguide mode contributions to each of these eleven Schmidt vectors are plotted in figures 6.7 on page 156 and 6.8 on page 157. Each field in each set is the image of the equivalent field (with the same label and colour) in the other set. Thus, if we take as an example the [21]-mode in each set (the black dots in the plots) and the corresponding s-number from figure 6.6, we have a Schmidt triple for the horn at this frequency. Since the cavity mode space and the aperture mode space are spanned by the same transverse modes (up to radial normalisation) it is immediately clear from the plots 6.7 and 6.8 that the Schmidt fields in a pair are radically different, so giving rise to the radically different fields in the cavity and aperture illustrated in figure 6.5.

In the examples above it is seen that the image space of the horn operator is of complex dimension one for a single-mode horn and of a small, frequency dependent, number of dimensions for the Planck multi-mode horns. Any optical system has finite resolution and that can be understood as a finite basis for the range of the system operator. For a telescope such as Planck or Herschel, the modes in the bolometer cavity are mapped through the optical system onto the sky. The telescope transforms the horn aperture distribution into a telescope aperture distribution. That aperture distribution can be mapped to the far field in a number of ways, but most obviously by Fourier transform or spherical wave expansion. Whatever method is chosen the basis set required for the expansion will be finite (if they were not, then the basis set would span the entire space and be complete, so giving unlimited resolution). In the particular case of the Fourier spectrum the basis set used would be the Hermite functions; in any case a discrete spectrum for both domain and co-domain is required. The system scattering operator will then have finite range simply because the instrument resolution is finite. Then compactness of the operator is immediate.



(a) Cavity field power pattern.

(b) Aperture field power pattern.

Figure 6.5: Power patterns of the Planck 857 GHz horn (a) cavity field and (b) aperture field, both at 995 GHz. These two fields patterns are constructed form the Schmidt triples for the S_{21} operator with amplitudes as plotted is figures 6.7 and 6.8, pages 156 and 157 respectively. Each component of the aperture field is a $\boldsymbol{\psi}_n$ corresponding to the component $\boldsymbol{\varphi}_n$ of the cavity field. For plotting the waveguide radius has been normalised to 1 in both plots. In (a) the waveguide radius at the corrugation depth to cavity radius ratio is 0.211 and corresponds to the centre reddish disc.

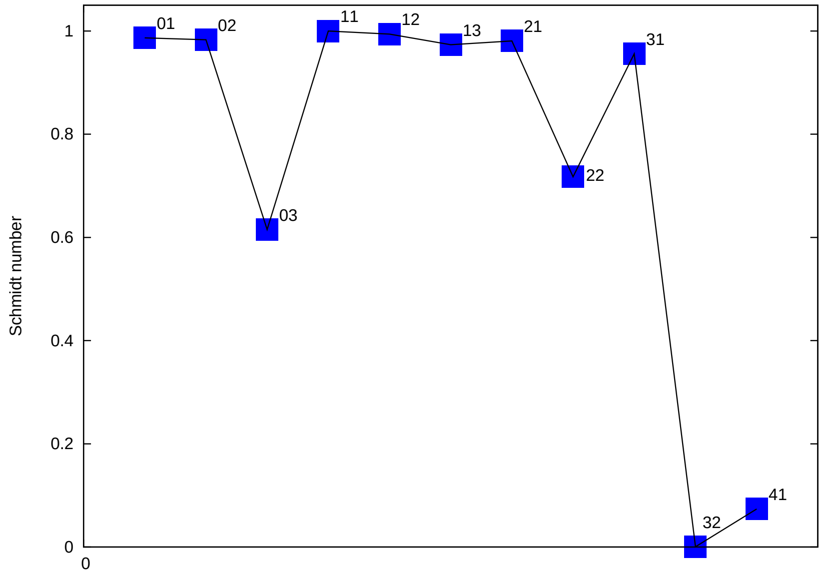
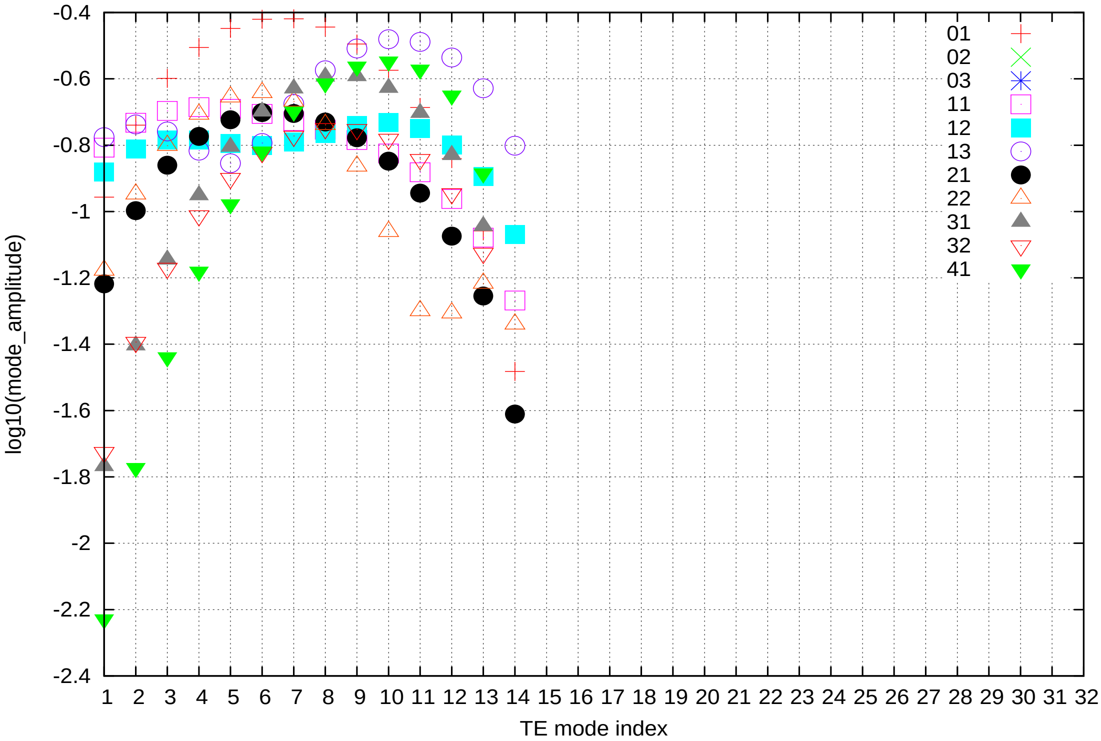
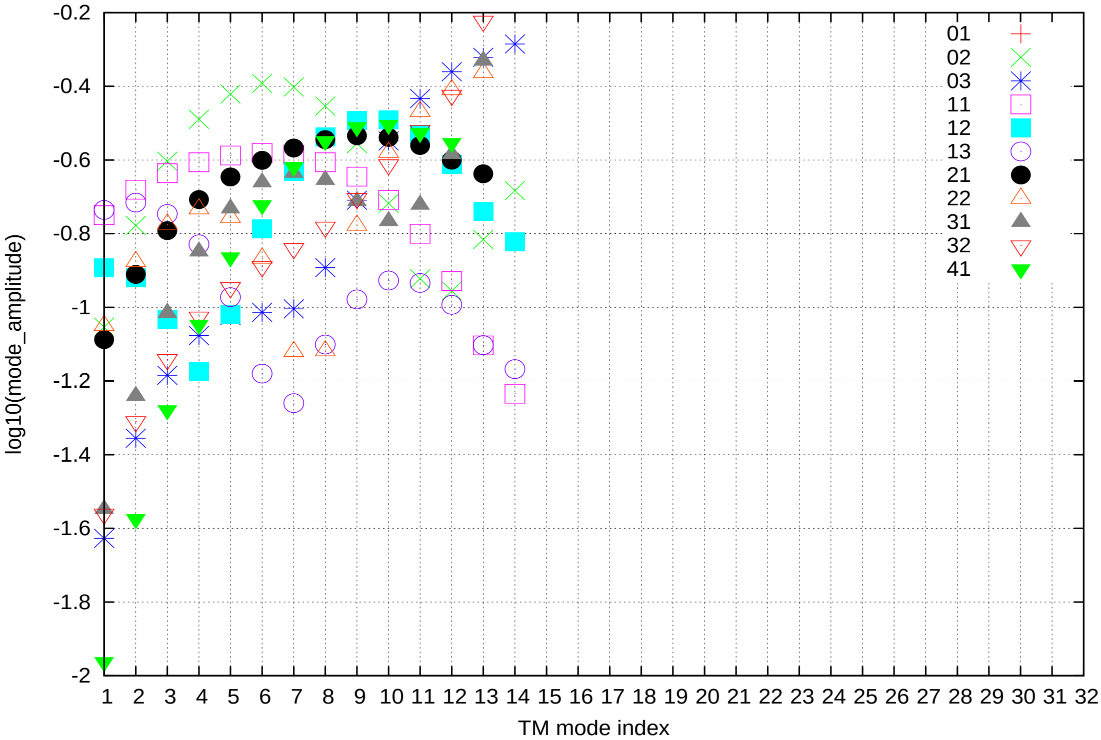


Figure 6.6: The s-numbers for the Planck 857 GHz horn operating at 995 GHz. The [01]-mode is purely electric, the [02] and [03] modes are purely magnetic; all others are hybrid. These are the s-numbers for the eleven fields containing at least 10^{-9} of the total field power encoded in both the S_{21} and S_{12} operators for azimuthal orders 0, 1, 2, 3 and 4. Although the space is 11 dimensional there is so little power in the [32] mode that it would not be included in a model and the state space can be considered to be 10 dimensional.

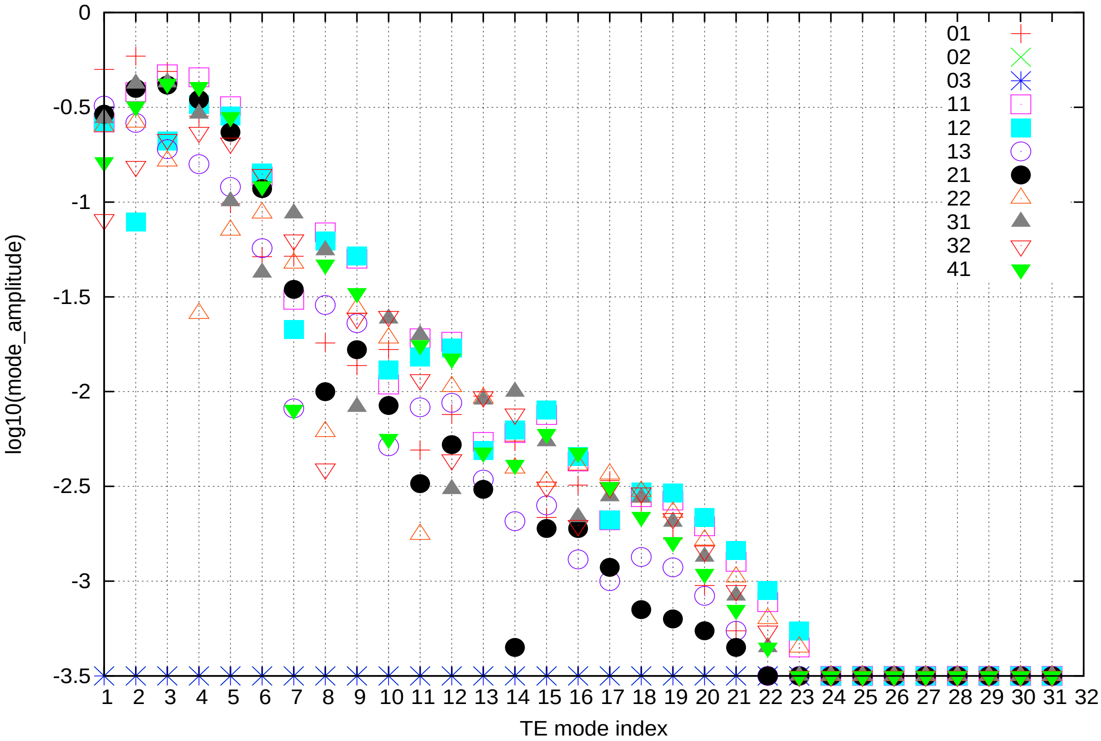


(a) Planck 857 GHz horn at 995 GHz, TE mode contribution to the $\boldsymbol{\varphi}_n$.

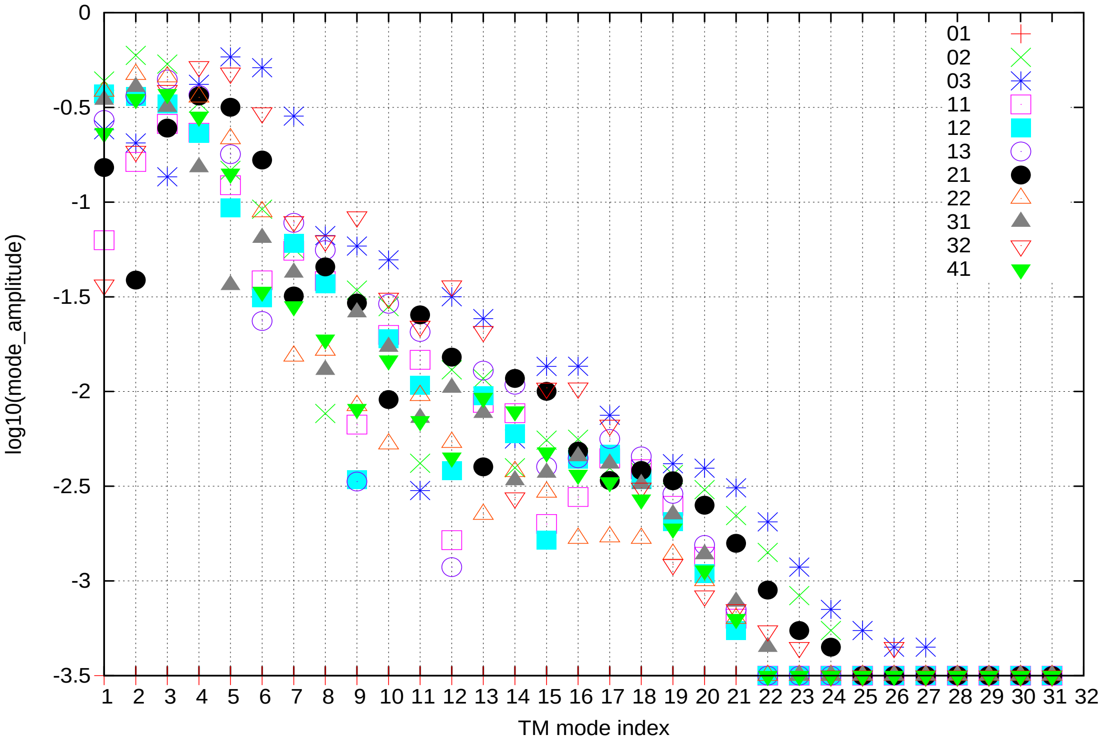


(b) Planck GHz horn at 995 GHz, TM mode contribution to the $\boldsymbol{\varphi}_n$.

Figure 6.7: The amplitudes of (a) the TE, and (b) the TM mode contributions to cavity field (the $\boldsymbol{\varphi}_n$) of the eleven Schmidt vectors of the five Planck 857 GHz horn S_{12} operators at 995 GHz. In (a) the [02] mode and [03] mode are listed, but they are zero because they are pure magnetic modes; likewise for the purely electric [01] mode in (b). The images of these fields, $\boldsymbol{\psi}_n = s_n^{-1} S\, \boldsymbol{\varphi}_n$, are plotted in figure 6.8 below.



(a) Planck 857 GHz horn at 995 GHz, TE mode contribution to the $\boldsymbol{\psi}_n$.



(b) Planck GHz horn at 995 GHz, TM mode contribution to the $\boldsymbol{\psi}_n$.

Figure 6.8: The amplitudes of (a) the TE, and (b) the TM mode contributions to the aperture field ($\boldsymbol{\psi}_n$) of the eleven Schmidt vectors of the five Planck 857 GHz horn S_{21} operators at 995 GHz. These vectors are the S_{21}-images of the vectors in figure 6.7 above; the colour coding is the same in both plots. In figure 6.5-(b) the field generated by these vectors is plotted.

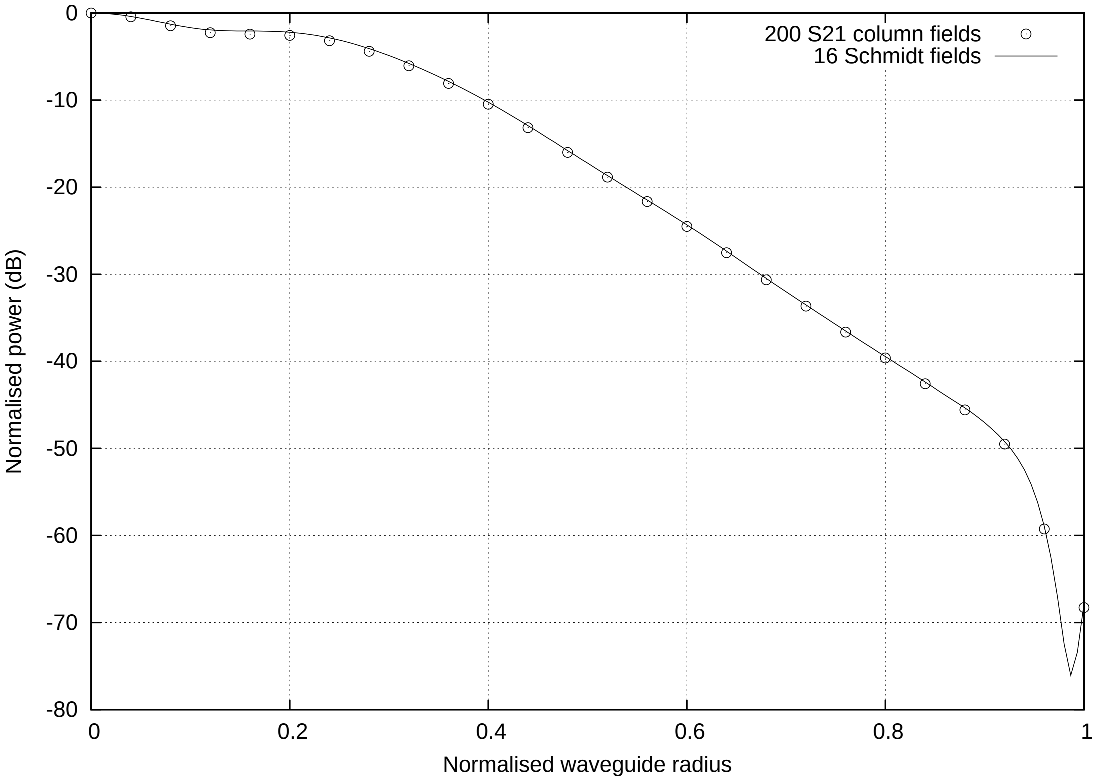


Figure 6.9: Planck 857 GHz horn aperture field total power pattern cross-section at band centre: (a) The field reconstructed from the 200 fields encoded in 100 columns from the five S_{21} matrices. (b) The same field reconstructed from 16 Schmidt fields derived from the same five S_{21} matrices. For both reconstructions the threshold for inclusion of a component field was set at $\| \boldsymbol{x} \|_2 \geq 10^{-4}$. For the same horn at 995 GHz the S_{21} column reconstruction requires 210 fields and the Schmidt field reconstruction requires 20 fields.

As an illustration of the accuracy and efficiency of the Schmidt field reconstruction, consider the Planck 857 GHz horn at the band centre. Figure 6.9 shows a decibel plot of the cross-section through the total power pattern of the aperture field at the band centre (both polarisations). Setting the Euclidean norm of the component fields at 10^{-4} as an inclusion threshold, the aperture field model uses 200 fields derived from the columns of the S_{21} matrix (100 per polarisation) while the Schmidt field reconstruction requires 16 fields (eight per polarisation). Given that reconstruction of the far field beam of the Planck telescope from the aperture fields is approximately linearly dependent on the number of fields, the broad band modelling of the telescope beams from the sixteen fields of the Schmidt field reconstruction was approximately 12×(number of frequencies) times faster than it would have been with the two hundred S_{21} column fields. The broad band reconstruction for the 857 GHz horn was done with 54 frequencies and took 576 computer hours using the Schmidt fields. Thus, with the straightforward reconstruction from the S_{21} matrix fields the same computation would have taken some 6219 hours – roughly eight and a half months per horn. For the 545 GHz horn the time saving factor was approximately a factor of seven. For a single mode horn the computation time is the same in both cases. Thus, while the approach is applicable in all cases, the gain depends upon the dimension of the hyper-plane spanned by the aperture fields in the Hilbert space of aperture fields.

6.3 Measurement of power

The beam pattern can be measured by scanning with a point-like source and correcting for the beam pattern of the source. If the geometry of the horn is known sufficiently precisely then the Schmidt triples can be assumed known and the response of the bolometer to the individual cavity Schmidt fields deduced from a comparison of the measured and modelled beams (up to a common scale factor). For a single-mode horn there is no difficulty because there is just one Schmidt mode with unit amplitude and the measurement gives a direct measurement of the Schmidt field in the aperture and, if the input power density is known, of the unknown bolometer response. For a multi-mode system like the Planck horns the argument becomes circular: it is not known precisely what has been built, so the Schmidt triples are not known precisely; the beam pattern can be measured and the probe pattern deconvolved, but what is being measured is the total power in a field in a multi-dimensional space plus the bolometer response to the unknown Schmidt fields. The modelled Schmidt field and numbers could be compared with the measurements, but only if the bolometer response to individual waveguide modes was know and it was also known that there was no correlation between the power measured simultaneously in pairs of modes regardless of their relative phase and amplitudes. It follows that, whereas for a single-mode horn the measurement of the beam pattern gives the Schmidt triple of the system when all experimental artifacts have been accounted for, with a multi-mode system no such assurance can be given. All that can be given is a probability amplitude for the Schmidt triples: a probability distribution on the space of all possible Schmidt triples for the system is required, a distribution which would be essentially zero outside a small neighbourhood of the designed Schmidt triples for an ideally designed and manufactured system.

A geometrical picture of this situation is helpful. If there are azimuthal orders 0 to n and the i^{th} order has K_i radial modes, then the model space is $\mathbb{C}^{K_0} \times \cdots \times \mathbb{C}^{K_n}$. If the number of Schmidt fields in the i^{th} azimuthal order is $N_i \geq 1$, then those fields lie on the sphere of unit modulus vectors, $S^{2K_i-1} \subset \mathbb{C}^{K_i}$, and define a plane that cuts the sphere in a sphere of dimension $N_i - 1$. Because the fields of different azimuthal orders are independent, these n spheres are to be treated independently. The group that acts on $\mathbb{C}^{K_i}$ preserving the S^{2K_i-1} sphere and orientation is the special unitary group $SU(K_i)$. A small perturbation to the sphere is described by an element of $SU(K_i)$ in a small neighbourhood of the identity; consequently it is useful to think of a perturbation as an element of the tangent space to the group at the identity. That tangent space is the Lie algebra, $\mathfrak{su}_{K_i}$, of traceless skew-hermitian matrices which acts on $\mathbb{C}^{K_i}$ via the exponential mapping, $\exp : \mathfrak{su}_n \to SU(n)$, $u(\boldsymbol{\psi}) \overset{\text{def}}{=} e^u\, \boldsymbol{\psi}$, [59][61]. In general this will not preserve the sphere $S^{N_i-1} \subset S^{K_i-1}$, but it will map Schmidt fields to possible Schmidt fields.

The small perturbations of the complete set of Schmidt fields is therefore an element of $\mathfrak{su}_{K_0} \times \cdots \times \mathfrak{su}_{K_n}$ with each coordinate element acting independently on each coordinate in $\mathbb{C}^{K_0} \times \cdots \times \mathbb{C}^{K_n}$. Since the Lie algebras are real vector spaces the perturbations can be given real coordinates. (Only the linear structure and mapping into $SU(n)$ will be used here; the Lie algebraic structure is not needed.) Thus, since precise knowledge of what has been built is not possible, but it is reasonable to assume that what has been built is close to what has been designed, the actual Schmidt fields at any given frequency can be thought of as a perturbation of the designed fields: if in the i^{th} azimuthal order the design was for Schmidt fields $\boldsymbol{\psi}_1^{(i)}, \ldots, \boldsymbol{\psi}_{N_i}^{(i)}$, then the actual fields will be $e^{u_i}\, \boldsymbol{\psi}_1^{(i)}, \ldots, e^{u_i}\, \boldsymbol{\psi}_{N_i}^{(i)}$ for some $u_i \in \mathfrak{su}_{K_i}$ close to the zero matrix. The *possible* horns that would be produced from a given design would be described by a distribution on each of the n Lie algebras, each distribution centred on the zero matrix (assuming no bias).

As a real vector space $\mathfrak{su}_n \cong \mathbb{R}^{n^2-1}$ so that there are n^2-1 linearly independent infinitesimal perturbations to the Schmidt vectors in $\mathbb{C}^n$. Fix the isomorphism and let A_k be the $\mathfrak{su}_n$ image of the k^{th} unit vector in $\mathbb{R}^{n^2-1}$; then the n^2-1 paths

$$\gamma_{A_k} : \mathbb{R} \to SU(n), \quad \text{defined by} \quad \gamma_{A_k}(t) = e^{tA_k}$$

are paths through, and mutually orthogonal at, the identity that describe (locally) independent perturbations of the unit sphere on which the Schmidt fields reside. Assume that, if a vast number of nominally identical copies of a Planck multimode horn (or any other horn) were made and the Schmidt fields measured in some way, and they were found to follow some distribution. For the sake of an example assume they are normally distributed about the design values. Then the distribution can be modelled as a normal distribution of perturbations on the Lie algebras: In each of the $n+1$ azimuthal orders choose an orthonormal basis $\{A_1, \ldots, A_{K_i}\}$ for $\mathfrak{su}_{K_i}$; fix normally distributed t_j for $j = 1, \ldots, K_i$ so that each has distribution $p_i(t_j) = e^{-(t_j-\mu_j)^2/2\sigma_j^2}/2\pi\sigma_j$ which can all have independent variances and zero mean. Then the perturbation of the Schmidt fields in this azimuthal order are given by

$$\boldsymbol{\psi} \to e^{t_1 A_1} \cdots e^{t_{K_i} A_{K_i}} \, \boldsymbol{\psi} = e^{t_1 A_1 + \cdots + t_{K_i} A_{K_i}} \, \boldsymbol{\psi} \, . \tag{6.3.1}$$

This equation is useful for the actual modelling of scattering operators from the Schmidt field descriptions of the design as will be discussed in section 6.4.

It is important to appreciate that the perturbation works both at the level of the Schmidt fields and at the level of the S matrices because the perturbation matrix is unitary: $(AS)^\dagger(AS) = S^\dagger S$, and $s_n A \, \boldsymbol{\psi}_n = A(s_n \, \boldsymbol{\psi}_n) = A(S \, \boldsymbol{\varphi}_n) = (AS) \, \boldsymbol{\varphi}_n$. Consequently it does not matter whether the perturbation is applied before or after calculation of the Schmidt fields. From the perspective of computational efficiency it is better to derive the fields and then to apply perturbations.

The next unknown to consider is the distribution of the Schmidt numbers. There are $\sum_{i=0}^{n} N_i$ Schmidt fields each with Schmidt number $s_{k_i} \in [0,1]$. The design is for some $\{s_{k_i} : \ i = 0, \dots, n; \ k_i = 1, \dots, N_i\}$ but for each k_i there is a distribution τ_{k_i} on $[0, 1/s_{k_i}]$ and the possible values are $s'_{k_i} = \tau_{k_i} s_{k_i} \in [0, 1]$.

Finally, roll all other unknowns (such as exact filter transmission and readout noise) into a single unknown of bolometer response. An ideal bolometer would absorb all incident radiation and detect unit power from all possible Schmidt fields. In practise the bolometer will not be perfect, so that given input $\boldsymbol{\psi}$ the measured power will be an efficiency $B(\boldsymbol{\psi})^2 \leq 1$. The efficiency may vary between azimuthal orders, so there are mappings $B_i : S^{2K_i - 1} \to [0, 1]$.

The complete picture becomes the following: The model space is $\mathbb{C}^{K_0} \times \dots \times \mathbb{C}^{K_n}$ and the design is for Schmidt field and s-number pairs $\{(\boldsymbol{\psi}_{k_i}, s_{k_i}) : \ i = 0, \dots, n; \ k_i = 1, \dots, N_i\}$. The actual set is $\{(e^{u_i}\, \boldsymbol{\psi}_{k_i}, \tau_{k_i} s_{k_i}) : \ i = 0, \dots, n; \ k_i = 1, \dots, N_i\}$ for some $\{u_i \in \mathfrak{su}_{K_i} \ : \ i = 1, \dots, n\}$. However the B_i scale the $\tau_{k_i} s_{k_i}$ and the *total observed* power in a single polarisation with all modes excited is

$$P_{TO} = \sum_{i=0}^{n} \sum_{k_i=1}^{N_i} \| B_i(e^{u_i}\, \boldsymbol{\psi}_{k_i}) \tau_{k_i} s_{k_i} e^{u_i}\, \boldsymbol{\psi}_{k_i} \|_2^2 = \sum_{i=0}^{n} \sum_{k_i=1}^{N_i} B_i(e^{u_i}\, \boldsymbol{\psi}_{k_i})^2 \tau_{k_i}^2 s_{k_i}^2. \qquad (6.3.2)$$

Assume that, in the beam pattern measurement setup, the horn aperture is centred at the origin of the coordinate system and the horn radiates in the positive z direction. The source probe scans the aperture from a plane $\{(x, y, z_0) : \ x, \ y \in \mathbb{R}\}$ at a distance z_0 from, and parallel to, the aperture. If the probe has aperture A in which the fields (after accounting for the edge currents and impedance step to free space) are $\boldsymbol{E}$ and $\boldsymbol{H}$, then the electric field at a point $\boldsymbol{x} = (x, y, 0)$ in the horn aperture, due to the probe centred at $\boldsymbol{x}' = (x', y', z_0)$, is given in the physical optics approximation (see [18], [49]) by

$$\begin{aligned} \boldsymbol{E}(\boldsymbol{x}, \boldsymbol{x}') = \frac{k_0^2 Z_0}{4\pi} \int_A \Bigg[& \boldsymbol{J}_E \left(-\frac{j}{k_0 R} - \frac{1}{k_0^2 R^2} + \frac{j}{k_0^3 R^3} \right) \\ & + (\boldsymbol{J}_E \cdot \hat{\boldsymbol{r}}) \hat{\boldsymbol{r}} \left(\frac{j}{k_0 R} + \frac{3}{k_0^2 R^2} - \frac{3j}{k_0^3 R^3} \right) \Bigg] e^{-jk_0 R} \, \mathrm{d}\boldsymbol{a} \qquad (6.3.3) \\ & - \frac{k_0^2}{4\pi} \int_A (\boldsymbol{J}_M \times \hat{\boldsymbol{r}}) \frac{1 + jk_0 R}{k_0^2 R^2} e^{-jk_0 R} \, \mathrm{d}\boldsymbol{a}, \end{aligned}$$

where $\boldsymbol{J}_E \overset{\text{def}}{=} -2\hat{\boldsymbol{z}} \times \boldsymbol{H}$ and $\boldsymbol{J}_M \overset{\text{def}}{=} -2\hat{\boldsymbol{z}} \times \boldsymbol{E}$, the integral is over the probe aperture with coordinates $\boldsymbol{x}''$ centred on $\boldsymbol{x}'$, $\boldsymbol{r} = \boldsymbol{x} - \boldsymbol{x}''$ and $R = |\,\boldsymbol{r}\,|$. What the horn actually measures is given by the sum of the inner products of the field illuminating the aperture, $\boldsymbol{E}(\boldsymbol{x}')$, with the Schmidt fields of the horn, scaled as in equation 6.3.2:

$$P_O(\boldsymbol{x}') = \sum_{i=0}^{n} \sum_{k_i=1}^{N_i} B_i(e^{u_i}\,\boldsymbol{\psi}_{k_i})^2 \tau_{k_i}^2 s_{k_i}^2 \left\langle \boldsymbol{E}(\boldsymbol{x}') | e^{u_i}\,\boldsymbol{\psi}_{k_i} \right\rangle^2 . \tag{6.3.4}$$

This is an idealised single frequency mapping whereas every term on the right hand side of equation (6.3.4) is a function of frequency. The complete observed beam pattern is the set $\{\tilde{P}_O(\boldsymbol{x}') : \boldsymbol{x}' = (x, y, z_0)\}$ where $\tilde{P}_O$ is the integral over the measurement frequency band of the P_O. In a real measurement there is a finite set of values measured at a finite set of points $\boldsymbol{x}'_j$, $j = 1, \ldots, J$. Suppose that the measurements are known to within $\pm\delta$, which will probably be a function of frequency and of measured power, but for the sake of simplicity assume that it is constant. Then if two horn assemblies are measured with the same setup giving powers $\tilde{P}_O$ and $\tilde{P}'_O$, and

$$\left| \tilde{P}_O(\boldsymbol{x}'_j) - \tilde{P}'_O(\boldsymbol{x}'_j) \right| \leq \delta \quad \text{for all} \quad j = 1, \ldots, J,$$

it would not be possible to distinguish the horns over this frequency band with this measurement setup. It does not mean that the horns are the same, only that sums in equation (6.3.4) integrated across the measurement band for both horns are the same to within error. Nor does it mean that the horns will be indistinguishable over some other frequency band and, most importantly, even if there is agreement to within measurement error as band averages it does not follow that the Schmidt fields, s-numbers or bolometer responses are the same, only that if the number of modes is small there is a reasonable chance that they are close. The situation is in stark contrast with a single-mode system where it is certain that the Schmidt field structure is the same to within experimental error if there exists a scalar function of frequency, $h(\nu)$ say, such that at every measurement datum $\boldsymbol{x}'$,

$$\begin{aligned} \int_{\nu_0}^{\nu_1} \Big| h(\nu) B(e^{u(\nu)}\,\boldsymbol{\psi}(\nu))^2 \tau(\nu)^2 s(\nu)^2 \left\langle \boldsymbol{E}(\boldsymbol{x}';\nu) | e^{u(\nu)}\,\boldsymbol{\psi}(\nu) \right\rangle^2 \\ - B(e^{u'(\nu)}\,\boldsymbol{\psi}(\nu))^2 \tau'(\nu)^2 s'(\nu)^2 \left\langle \boldsymbol{E}(\boldsymbol{x}';\nu) | e^{u'(\nu)}\,\boldsymbol{\psi}(\nu) \right\rangle^2 \Big| \, \mathrm{d}\nu \leq \delta \end{aligned} \tag{6.3.5}$$

over all measurement bands. Then the mode structure is the same to within experimental error, $e^{u'}\,\boldsymbol{\psi} = e^{u}\,\boldsymbol{\psi}$, even if the unknown $B(e^{u(\nu)}\,\boldsymbol{\psi}(\nu))\tau(\nu)s(\nu)$ products are not identical, because they simply scale the overall pattern.

The mappings B_i and τ_{k_i} and the Lie algebra elements $u_i \in \mathfrak{su}_{K_i}$ are all unknown and, to complicate matters, are themselves functions of frequency that might be smooth, but might not even be continuous if there are resonances as in the Planck multi-mode horn assemblies. Both the models reported in section 2.4 (simulations of idealised instantaneous measurements at pure frequencies) and the measurements reported in the same section, indicate that the behaviour is quite chaotic. In any case neither the measurement process nor the models can truly distinguish discontinuous, continuous or differentiable behaviour, but the measurement process can be modelled as an approximation to a Riemann-Stejltz integration, and that seems to be the most appropriate integration theory.

There are three sets of distributions in total: (a) the distributions on the Lie algebras of $\mathfrak{su}_{K_i}$ for each azimuthal order $i = 0, \ldots, n$, (b) the distribution for the perturbations of the s-numbers, and (c) the distributions describing the uncertainties rolled into in the bolometer response. The distributions on the Lie algebras are determined by $\sum_{i=0}^{n}(K_i^2 - 1)$ random real numbers following the chosen distribution. The $\sum_{i=0}^{n} N_i$ s-numbers s_i each have perturbation $\tau_i \in [0, 1/s_i]$, and to each of these there is an unknown bolometer response which can be modelled as a distribution on $[0, 1]$.

6.4 Perturbing the field models

Given a completed horn design and an assumed bolometer response to the Schmidt fields of the design, it would aid performance prediction if the model was perturbed in a realistic way. The tolerance models of section 2.4 were generated by running multiple horn geometries; if the design is reasonably stable to manufacturing tolerances, then it would be much more efficient to perturb the finished design. (In the case of the Planck 857 GHz horn models, direct modelling of a perturbed geometry file takes about ten minutes for the five azimuthal orders at a single frequency whereas generating a perturbation of the Schmidt structure or five S_{21} matrices takes of the order one second.) Because very few multi-mode horns have been built and the mode structure of those horns has proven difficult to measure, essentially nothing can be said about the probability distributions of the Schmidt triples, though it is clear from the power measurements plotted in figure 2.17 on page 35 and from the equivalent measurements for the 857 GHz horns, that there is considerable variation in performance between nominally identical horns. Here the aim is to show that the design can be perturbed in a simple way so that the performance of many physically realistic horns, close to the design, can be assessed. To produce a perturbed model the perturbations have to be chosen randomly and that requires a choice of probability distribution. In principle any distribution can be used; here the Gaussian with mean and variance chosen to bias the perturbation towards the design will be used, though purely for illustrative purposes.

Assume that nothing is known about the sensitivity of the design to the perturbation of individual modes. As in the previous section let there be K_i waveguide modes in the model for the i^{th} azimuthal order. Then the appropriate Lie algebra for generating the perturbations is $\mathfrak{su}_{K_i}$ and $\dim_{\mathbb{R}}(\mathfrak{su}_{K_i}) = K_i^2 - 1$. Fix a probability distribution and generate $K_i^2 - 1$ random real numbers over a small interval centred on zero: $[-1, 1]$ for example. Denote the resulting sequence $\{r_1, r_2, \ldots, r_{K_i^2-1}\}$. From these numbers construct an element of $u \in \mathfrak{su}_{K_i}$ as follows: For $k = 1, \ldots, K_i - 1$ set $u_{kk} = ir_k$, and $u_{K_iK_i} = -i\sum_{k=1}^{K_i-1} r_k$ so that $\mathrm{Tr}(u) = 0$. Then set $n = K_i - 2$, choose any $t \in [-1, 1]$ as the distance along the path from the zero matrix, and proceed to fill the remainder of the upper triangle of u:

$$\text{For } j = 1, \ldots, K_i - 1 \ \ \{\text{For } k = 2, \ldots, K_i \ \ \{n = n + 2, \ \ u_{jk} = (r_n + ir_{n+1})\}\}.$$

Since matrix is hermitian it is not necessary to use the sub-diagonal elements to form sums or products if one of the standard packed storage schemes is used, for example the BLAS, [28]. Since the elements of u are small, the exponential of u will be well approximated by a finite series expansion that can be terminated when $\|t^n u^n / n!\|_F \leq \varepsilon$ for some chosen ε greater than the machine epsilon. Then $\gamma_u(t) \simeq I + tu + t^2u^2/2 + t^3u^3/3! + \cdots + t^{n-1}u^{n-1}/(n-1)!$ is the generated $SU(K_i)$ perturbation matrix. As an example, a random element of $SU(64)$ that would be suitable for perturbing a typical model S_{21} for the Planck multi-mode horns can be generated in

~ 0.05 seconds with all columns satisfying $|\,\boldsymbol{x}_i \cdot \boldsymbol{x}_j^\dagger\,| = \delta_{ij} \pm 10^{-17}$, the magnitude of the error depending upon the path length factor t and the details of the algorithm and compiling. A simple test of the output is to calculate either one of the Frobenius norm $\|\gamma_u(t)^\dagger\gamma_u(t)\|_F$, the uniform norm $\|\gamma_u(t)^\dagger\gamma_u(t)\|_\infty$ or $\|\gamma_u(t)^\dagger\gamma_u(t) - I\|_{\max}$, which should return the values K_i, 1 and 0 respectively to within $\sim 10^{-16}$.

An important point reflecting the physics of the model has been glossed over. With reference to equation (6.3.1) on page 161, infinitesimal rotations are determined by displacements in the Lie algebra tangent to the direction of individual modes. The effect of the resulting rotation is to transfer power out of one mode into other modes. An arbitrarily generated rotation could model the transfer of power out a propagating mode and into evanescent modes, and that could lead to unrealistic models. The perturbation model must be set up so that the dominant rotation has as axis the hyperplane determined by the evanescent modes in each of the TE and TM subspaces. That rotation is to be combined with a small random rotation allowing for the weak intermixing of any mode. Computationally this is simple: if, in the model, there are n_e TE modes of which the first k_e are propagating (in the aperture space) and n_m TM modes of which the first k_m are propagating, then the dominant perturbation is generated by an $A \in \mathfrak{su}_{n_e+n_m}$ matrix which is zero everywhere except in the upper left $k_e \times k_e$ block and the $k_m \times k_m$ block with upper left corner at A_{n_e+1,n_e+1}, each non-zero sub-block being traceless as required of an element in the Lie algebra. The exponential of such a matrix will leave the evanescent modes unchanged. (Clearly any other linear embedding $\mathfrak{su}_{k_e+k_m} \hookrightarrow \mathfrak{su}_{n_e+n_m}$ which preserves the evanescent modes pointwise will do as well.) To that is added a random element of $\mathfrak{su}_{n_e+n_m}$ that is much closer to the zero matrix than the dominant rotation. The exponential of the sum will be a perturbation that leaves the total power in the evanescent field little changed, as is required for a physically realistic model, while allowing the redistribution of power between propagating modes to dominate the perturbation.

The other two perturbations required are the τ_i perturbations of the s_i and the function B that incorporates the frequency dependent bolometer response, filter transmission and other 'losses' in the measurement of the Schmidt fields. The range of τ_i is $[0, 1/s_i]$ and to be realistic needs to be centred at one. A reasonable model would be a Gaussian with small variance and mean 1, normalised to a maximum value of one: $\tau_i(x) = \exp[-(x-1)^2/2\sigma_i^2]$. To choose the sample point, x, use any probability distribution on $[0, 1/s_i]$; here the Gaussian $p(\xi) = \exp[-(\xi-1)^2/2\sigma^2]/2\pi\sigma$, with $\sigma \sim 0.5$ will be used to generate the sample point x: generate uniformly distributed random samples $\xi \in [0, 1/s_i]$ and $\eta \in [0, 1/2\pi\sigma]$; if $\eta \leq p(\xi)$, then $\tau_i(\xi)$ is used as the perturbation for s_i, otherwise generate new (ξ, η) pairs until a validly distributed ξ is found.

The unitary transformations that transform the design S_{21} matrix into the S_{12} matrix for a tolerance model can be derived form the two matrices in the following way. Having calculated the Schmidt vectors for the designed and tolerance systems the result is two sets of unimodular vectors. Each set can be arranged into a square matrix preserving the column ordering between the two models. Since the columns are all unimodular and orthogonal, the result is unitary. Denote the two matrices S and S' respectively, then the problem is to solve for A in $AS = S'$ using standard computational linear algebraic methods. Necessarily A will be special unitary and the associated Lie algebra element $u \in \mathfrak{su}_N$ such that $e^u = A \in SU(N)$ can be found by taking the logarithm base e as a power series, $\log_e(I + A) = \sum_{n=1}^{\infty}(-1)^N A^n/n$, testing the trace of the partial sums for convergence in $\mathfrak{su}_N$. By running a few tolerance models, then constructing the deformation operators in $\mathfrak{su}_N$, the magnitude and phase of the off diagonal elements can be used to set realistic bounds on distributions used to generate the random perturbations of the design, and thus generate many tolerance models with extreme rapidity.

6.5 A hypothetical calibration scheme

The following scheme was devised with the intention that it be tested with multimode horn measurements to be made at Manchester in collaboration with Maynooth. Since the measurements are to proceed the scheme is presented lest it be useful to those involved in the measurement program, but has not been tested and improved. The scheme is first presented under the assumption that Schmidt mode filters can be manufactured for preliminary calibration; finally the method is revised under the assumption that bolometer response to individual Schmidt modes cannot be measured directly an so no preliminary calibration can be made.

The problem of calibration is to determine the Schmidt vectors and numbers and the efficiency of the bolometer response. The usual way to do this would be to scan the beam, but as pointed out above the beam scan gives limited information except in the case of a single-mode system. What is put forward here is a hypothetical scheme for a detailed calibration over a narrow frequency band starting from the essential assumption that the manufactured system is sufficiently close to the design that accurate estimates of both the left and right Schmidt fields, the $\boldsymbol{\psi}_n$ and $\boldsymbol{\varphi}_n$ respectively, have been made. In what follows problems of implementation such as the likelihood of exciting standing waves due to the inclusion of mode filters, are ignored.

Knowing the cavity Schmidt fields of the design, $\{\boldsymbol{\varphi}_n\}_{n=1}^{N}$, build a 'cavity' that consists simply of the bolometer with its back short terminating a section of waveguide of the correct radius. Design an illumination system that will give a known power and phase distribution over the open end of the waveguide; for the sake of argument let it have a locally planar phase front and Gaussian power distribution. A single Schmidt field, $\boldsymbol{\varphi}_n$, is a known (though possibly complicated) amplitude and phase distribution. Therefore, at least in principle, it would be possible to design a phase and amplitude mask that, placed within the opening of the waveguide, would admit only field $\boldsymbol{\varphi}_n$. By inserting neutral density filters, or otherwise adjusting the input power, the bolometer response $B(\boldsymbol{\varphi}_n)$ could then be measured directly for this field.

Having measured the $\{B(\boldsymbol{\varphi}_n)\}$ the horn is then assembled and a second set of filters prepared for the $\boldsymbol{\psi}_n$. If the system was a perfect reproduction of the design, then the measured power would be $B(\boldsymbol{\varphi}_n)^2\tau_n^2 s_n^2$ so that, with the bolometer already calibrated, the product $\tau_n s_n$, and thus τ_n, would be known. From these measurements the beam pattern for the system with the assumed Schmidt fields can be calculated from the individual fields and equation (6.3.3) with the rôles of probe and horn reversed. The predicted beam pattern would then be compared with the measured beam pattern at the frequencies used for calibration. If the agreement was good to within experimental uncertainty at a number of frequencies, then the hypothesis that the Schmidt mode structure, $\{\boldsymbol{\varphi}_n, s_n, \boldsymbol{\psi}_n\}_{n=1}^{N}$, of the horn was known would have strong support.

To make assessments of the usefulness of the beam pattern measurement, and to talk about the information content of the beam measurement, a random variable and probability distribution functions are required. The domain of the random variables will be $\Omega = S^{2N-1} \times [0,1]$, a point $\omega = (\boldsymbol{\psi}, \tilde{s}) \in \Omega$ is a possible Schmidt field and a real number $\tilde{s} = B(\boldsymbol{\psi})\tau s$. In conventional statistics and information theory a random variable takes a value either in a discrete set of numbers or in a number field; here the 'value' is a beam pattern. Nevertheless, the measurement plane is divided into sample points indexed by probe position, and these can be given any ordering to give a set $X = \{\boldsymbol{x}_L : L = 1, \ldots, L\}$. At each sample point equation (6.3.4) gives a real number, and each $\omega \in \Omega$ takes as value the vector $(P_O(\boldsymbol{x}_1), \ldots, P_O(\boldsymbol{x}_L)) \in \mathbb{R}^L$ of values of the power in a beam due to $\omega = (\boldsymbol{\psi}, \tilde{s})$. Since the outcome of the measurement is dependent upon the choice of set X it seems natural to indicate this by writing $\omega_{\boldsymbol{\psi},s}(X) = (P_O(\boldsymbol{x}_1), \ldots, P_O(\boldsymbol{x}_L))$ and interpret $\omega_{\boldsymbol{\psi},s}(\,\cdot\,) : \mathbb{R}^2 \to \mathbb{R}$ in the continuum limit as the measured power function.

Equation (6.3.3) on page 162 with the horn aperture as the source gives, for a Schmidt field over the aperture, the electric field strength at $\boldsymbol{x}$ in the measurement plane. Integrated over the probe aperture and convolved with the probe field gives a prediction of the measurement, and taking the modulus squared gives the power up to normalisation. When the performance of the horn is modelled it is the individual $\omega_{\boldsymbol{\psi},s}(X)$ over the set of the Schmidt triplets that is modelled and the model beam power pattern is the sum of all the vectors. If the aperture filters for the Schmidt modes were constructed to be illuminated by the coherent sum of all fields due to the probe at its complete set of measurement loci, X, (a partial Huygens wavelet type construction) then the Schmidt field beams could be individually measured and compared with the model. Here it will be assumed that this cannot be done. What is actually measured is just the total power at each $\boldsymbol{x} \in X$.

Let $\bar{\omega} = (\omega_{\boldsymbol{\psi}_1, s_1}, \ldots, \omega_{\boldsymbol{\psi}_N, s_N})$ be a vector in Ω^N determined by N orthogonal $\boldsymbol{\psi}_n$ and their supposed s-numbers, and let $Y = (y_1, \ldots, y_L) \in \mathbb{R}^L$, then define a distribution vector of distribution function values by

$$F_{\bar{\omega},X}(Y) = \big(F_{\bar{\omega},\boldsymbol{x}_1}(y_1), \ldots, F_{\bar{\omega},\boldsymbol{x}_L}(y_L)\big) \tag{6.5.1}$$

where for each $l = 1, \ldots, L$, the coordinate functions are $F_{\bar{\omega},\boldsymbol{x}_l}(y_l) = P\{\bar{\omega}(\boldsymbol{x}_l) \leq y_l\}$, with each $\bar{\omega}(\boldsymbol{x}_l) = P_O(\boldsymbol{x}_l)$ being the total power observed from sampling position $\boldsymbol{x}_l$, as in equation (6.3.4). In the terminology of the previous two sections there is, for each azimuthal order i, a set of N_i Schmidt vectors in $S^{2K_i - 1}$, a sample space $\Omega_i = S^{2K_i-1} \times [0, 1]$ with random variables $\bar{\omega}_i = (\omega_{\boldsymbol{\psi}_1^{(i)}, s_1^{(i)}}, \ldots, \omega_{\boldsymbol{\psi}_{N_i}^{(i)}, s_{N_i}^{(i)}})$, and distribution vector $F_{\bar{\omega}_i,X}(Y)$. From the analytical and the physical perspective the $\bar{\omega}_i$ are completely separate, but from the measurement perspective they are inseparable since the vector $\bar{\omega}_X(Y) = \sum_i \bar{\omega}_{i,X}(Y)$ is the measured total power at the sample points. That is the essence of the problem; if the interest lies in knowing only the total beam pattern, then measuring $\bar{\omega}(X)$ is all that is required, but if the interest is the structure of the beam for comparison with the design, then in an ideal world the individual $(\boldsymbol{\psi}_l^{(i)}, s_l^{(i)})$ pairs need to be known for all azimuthal orders. However the individual $(\boldsymbol{\psi}_l^{(i)}, s_l^{(i)})$ cannot be measured, and all that can be asked is what the probability is that the observed $\bar{\omega}(X)$ is compatible with the assumed $(\boldsymbol{\psi}_l^{(i)}, s_l^{(i)})$ given the calibrations $(B(\boldsymbol{\psi}_l^{(i)}), \tau_l^{(i)})$, everything being averaged over each measurement frequency band separately. Thus, what is sought is a conditional probability density $p(\bar{\omega}_0 \cdots \bar{\omega}_N | \bar{\omega})$ for parameter estimation. That is rather daunting, particularly if the total number of Schmidt modes is large.

What is proposed here is that the perturbation method of section 6.4 be used in a stochastic sampling of the space of $[0, 1]$-valued functions over the Schmidt fields, namely $\Omega = \Omega_0 \times \cdots \times \Omega_N$ for an $N + 1$ azimuthal order system. This is perfectly feasible since, on a standard PC, it takes only about *1/15*th of a second to compute a perturbation of the S_{21} matrix and to derive the Schmidt vectors and numbers and to calculate the aperture fields from them. The slowest part of the process is to compute the beam pattern from the aperture fields. Nevertheless, if the procedure followed was to prepare the beam measurement set-up, and then to start the stochastic simulations, given the time that it takes to make a measurement of the beam pattern it would be reasonable to expect roughly one completed simulation per minute. (Here it is assumed that the model is coded efficiently and compiled properly). The final step in the procedure is to make a comparison between the beam measurements and the output of the models and to weight the outcome: If the beam power measurement at datum $\boldsymbol{x}$ is $m(\boldsymbol{x})$ and the j^{th} model gives $m_j(\boldsymbol{x})$, the pointwise error is $|m(\boldsymbol{x}) - m_j(\boldsymbol{x})|$ and the total distance between measurement and model measured at $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L\}$ is given by $d(m, m_j) = \sum_j \sum_l |m(\boldsymbol{x}_l) - m_j(\boldsymbol{x}_l)|$ if the central part of the main beam is to dominate the comparison, or by $d(m, m_j) = \sum_j \sum_l |1 - m_j(\boldsymbol{x}_l)/m(\boldsymbol{x}_l)|$ if all data are to be treated equally (though beware of noise in low power measurements with such a metric). The model at the minimum distance is the best 'fit' given the calibration assumptions and results. The outcome of the stochastic model would indicate how sensitive the system was to perturbation. If the system shows a clear localisation of the results, then the probability that the best fit model is a reliable indicator of what has been built is high, but if there is no clear localisation, or two or more local minima, then a further stochastic model with restricted feasible domain would have to be run. Note that two models are close if the Euclidean distance between their perturbation vectors in the real vector space $\mathfrak{su}_{K_0} \times \cdots \times \mathfrak{su}_{K_N}$ is small and the results are localised if there is a small open neighbourhood in $\mathfrak{su}_{K_0} \times \cdots \times \mathfrak{su}_{K_N}$ over which the $d(m, m_j)$ are all small. It is the hyper-volume and shape of this open neighbourhood that indicates localisation. Figure 6.10 below summarises the process.

INPUT FROM SYSTEM DESIGN AND ASSUMPTIONS:

1. Design S_{21} matrix or Schmidt triples.
2. Assumed distribution function for
   (a) Schmidt field perturbations: $\mathfrak{su}_{K_0} \times \cdots \times \mathfrak{su}_{K_N}$ coordinate distributions.
   (b) Generalised s-number perturbations $B(\,\cdot\,)\tau$.

STOCHASTIC MODEL PREPARATION:

For model index $j = 1, \dots, M$:

Stochastic perturbation generator:
Produces $u_i \in \mathfrak{su}_{K_i}$ and τ_{ik} for azimuthal orders $i = 0, \dots, N$. ⟶ Create j^{th} model file: Store u_i and τ_{ik}.

Perturbed model:
All Schmidt triples $(\boldsymbol{\psi}_n, s_n, \boldsymbol{\varphi}_n)$ for for the j^{th} perturbed model. ⟶ Copy to j^{th} model file.

Aperture fields:
For all $(\boldsymbol{\psi}_n, s_n)$.

Simulated beam power pattern:
m_j: incoherent sum of all aperture fields. ⟶ Copy to j^{th} model file.

POST PROCESSING:

Input:	**Output:**
1. Measured beam pattern, m. 2. All simulated beam patterns, m_j.	Metrics $d(m, m_j)$ and $\mathfrak{su}_{K_0} \times \cdots \times \mathfrak{su}_{K_N}$ volume of feasible perturbations.

Figure 6.10: Scheme for the stochastic search for the probable Schmidt structure of a multi-mode horn. The final output is a hyper-volume in $\mathfrak{su}_{K_0} \times \cdots \times \mathfrak{su}_{K_N}$, the elements of which generate feasible perturbations of the design that give modelled beam patterns compatible with the measured beam.

What is sought is the conditional probability $p(\bar{\omega}_0 \cdots \bar{\omega}_N | \bar{\omega})$, but we always have a finite set of measurements and models and therefore cannot define a true density. However, given the output of the post processing of the models, define the normalisation of the measurement and model values to be $\bar{m}(\boldsymbol{x}) = m(\boldsymbol{x}) / \max_i\{m(\boldsymbol{x}_i)\}$ so that the peak power is one in both maps, then let

$$P(\bar{\omega}_0 \cdots \bar{\omega}_N | \bar{\omega}) = 1 - \frac{\sum_j \sum_{l=1}^{L} |\bar{m}(\boldsymbol{x}_l) - \bar{m}_j(\boldsymbol{x}_l)|}{\sum_j \sum_{l=1}^{L} |\bar{m}(\boldsymbol{x}_l) + \bar{m}_j(\boldsymbol{x}_l)|}, \tag{6.5.2}$$

where the sum over j is understood to be over the set of all Schmidt fields. With this definition $0 \leq P(\bar{\omega}_0 \cdots \bar{\omega}_N | \bar{\omega}) \leq 1$ and it takes the value zero only when the model propagates no power to the far field and the value one when the model agrees perfectly with the measurement. This is therefore a true conditional probability that the Schmidt field structure is in agreement with measurement to within the limits of modelling and measurement accuracy. Since evanescent fields influence the modelled beam and measurement only very weakly, the uncertainty in the evanescent field structure will always be relatively high for an imperfect model; however, the Schmidt fields have unit power and $B(\boldsymbol{\psi})\tau s \leq 1$ so that the modelled beam cannot both match the non-evanescent structure of the measured beam and have unrealistic power in the evanescent component. Thus $P(\bar{\omega}_0 \cdots \bar{\omega}_N|\bar{\omega}) \simeq 1$ will ensure a physically realistic field structure.

In the above discussion a critical point has been raised that is not explicitly written in the expression for $P(\bar{\omega}_0 \cdots \bar{\omega}_N|\bar{\omega})$: the expression assumes a particular Schmidt triple structure and requires a preliminary precise mode response calibration subject to that assumed structure to provide the estimated $B(\boldsymbol{\psi})\tau$ values. Thus, the probability is really a conditional probability subject to the probability that the calibration values are correct. The choice of calibration modes and the accuracy of the calibration values will bias the result and the probability is really the probability subject to that calibration. If the probability distribution indicates a Schmidt mode structure that is very close to the calibration mode structure, then the result can be taken to be reliable. Otherwise the entire process of beam pattern modelling ought to redone following a recalibration using the most probable Schmidt mode structure.

Possibly the experimenter is not going to have the time or the finances to do any calibration at all, or the construction of the mode filters might prove impossible. In that case the same process as above can be repeated assuming the values $B(\boldsymbol{\psi})\tau = 1$ for all Schmidt fields. In that case the output of the stochastic modelling is a new model S_{21} operator and Schmidt triples, but for the Schmidt field-number pair $(\boldsymbol{\psi}, s)$ returned by the analysis the number s is really $\tilde{s} = B(\boldsymbol{\psi})\tau s$ for unknown $B(\boldsymbol{\psi})\tau$. The stochastic analysis ought then to be repeated with the new Schmidt structure as the assumed value and a new optimal perturbation found. If this new perturbation is represented by an $SU(K_0) \times \cdots \times SU(K_N)$ matrix $A_0 \times \cdots \times A_N$ with each A_n having $\|A_n - I\|_F \simeq 0$, (or equivalently for the Lie algebra element u_n with $A_n = e^{u_n}$ has $\|u_n\|_{\max} \simeq \|u_n\|_F \simeq 0$), then the result of the preliminary analysis is reliable. However, even when a reliable result has been found, without the mode calibration no true value for the s-number can be deduced because the bolometer response to the individual Schmidt modes is not known. Even if the returned s-number equals the assumed s-number it cannot be known if the result occurred because the bolometer response to the mode is perfect, or because the Schmidt number was higher than predicted and the bolometer response sub-optimal, except when $s = \tilde{s} = 1$.

Chapter 7

Planck reflector surface fitting

The reflectors of the Planck telescope, when at L2, were expected to cool to around 50 Kelvin. The 50 K predictions for the shape of the Planck secondary reflector flight model (SRFM) and the primary reflector flight model (PRFM), obtained by linear regression analysis on the best fit surfaces to the room temperature coordinate measuring machine (CMM) and cryogenic videogrammetry data, are presented. The work was carried out at the request of ESTEC in 2008 as part of the programme to obtain the best possible pre-launch knowledge of the optical properties of the telescope. The work divided into two stages: firstly, conic surface fitting to the measurement data using the non-linear method of orthogonal distance regression; secondly, linear regression analysis to attempt to predict the shapes and positions to which the reflectors would have contracted when in a steady state at 50 K – the nominal in-flight temperature. The results were subsequently used as input to the pre-launch multi-mode beam pattern prediction (see 2.4) and then the post-launch reverse engineering of the telescope – work described in chapter 7. The work presented here has been extracted form technical reports [52], [51] and [50] that were presented to ESTEC in 2008. All cryogenic measurements were made in the test facilities at CSL, Belgium, and provided for the purposes of analysis by ESTEC.

This work was entirely driven by the engineering requirements of the Planck project, and the presentation of the results, as tables of derived data, reflects the needs of the engineers. The tables give a summary of the fitting results to all data sets, basic statistics for the form error measured along the normal to the best fit surface, and all surface parameters derived from the fitting of an ellipsoid to the data sets by orthogonal distance regression. In addition, for the optical model parameters (the semi-major and semi-minor axis lengths, the ellipsoid centre displacement from the aperture coordinate system, and the rotations about the aperture coordinate x-axis and y-axis) the 95% confidence intervals and estimated standard deviations for the fit of each parameter is listed. All results are presented in the tables for clarity and ease of comparison.

The contract with ESTEC required the fitting of ellipsoidal surface models to the reflectors and the provision of the best fit model data and residual surface form error maps to the Planck engineering team, and only the fits to ellipsoidal models is presented here. However, to investigate the surface distortion a spheroidal surface model was used. Fitting a spheroid uses a three-axis model rather than the two axis model of an ellipsoid (as well as surface rotations and displacements) and gives a third rotation angle as well as axis length. The result was a model surface fit with lower residual errors, and the different changes in the lengths of the two semi-minor axes could be seen very clearly in the fits to the measurement data both at different temperatures and at the same temperature at different stages in the cryogenic cycle – clear evidence for thermally induced distortion. This was found to be more pronounced in the PRFM than the SRFM. The greater distortion may have been due simply to the larger reflector not being in as good a thermal equilibrium as the smaller reflector, or may have resulted from the greater distortion of the reflector with the greater radius of curvature; most likely both.

During the linear regression analysis the SRFM returned a good to high confidence measure for the linear model; for the PRFM the confidence measure was not as good. This reflects two things about the differences in the data available for each reflector. Firstly, the PRFM showed greater distortion during the cryogenic cycling than did the SRFM. Secondly, the SRFM surface data was noisier than the PRFM data, and also there was less of it. As a result, the confidence intervals for the SRFM were relatively large; for the PRFM fitting the confidence intervals were very small which reflected a very high estimate of the reliability of the orthogonal distance regression fit to the surface. With the larger confidence intervals it is easier to get a linear fit to the data (see the graphs 7.1 and 7.4, both of which show the 95% confidence intervals for the fitting to each data set).

The contract deliverables were GRASP models of the Planck telescope optics at the estimated in-flight temperature with surface form error maps of two kinds: large scale maps and detailed surface quilting maps. These were delivered to ESTEC-TEC MMO and to industrial contractors to ESTEC involved in the design and building telescope (ASTRIUM, ThalesAlenia Space and TICRA). In the following sections a general description of the analysis methods that were adopted is given, but no commercially sensitive information is included.

7.1 Surface fitting by orthogonal distance regression

The problem of fitting a surface to the measurement data is a nonlinear problem with unknown errors on the measurement data. The model used for the fitting was required to be an ellipsoidal surface, though a general spheroidal surface model was also used. Two techniques were combined: orthogonal distance regression as described in [9] and [10] for the model fitting and error estimation, and methods taken from non-sequential ray tracing to establish the nearest point on the trial model surface to the measurement data.

The problem at hand is to fit a spheroidal surface to a set of measurement data. The data comprised either sets of coordinates of points on the reflector surface obtained by a coordinate measuring machine at 293 K, or sets comprising centroid positions of markers on the surfaces of the Planck reflectors that were measured by videogrammetry at a number of temperatures on the cryogenic cool-down and warm-up cycle between 293 K and 95 K. The markers were bonded to the surface of the reflectors and during thermal cycle some of the markers gradually became detached from the surface, and sometimes fell off. There is an intrinsic and unknown measurement error for each centroid, and for those markers that became even partially detached over the sequence of measurements that error is not constant. Moreover, unless a marker became fully detached at some point in the cycle, there is no way to determine whether or not any individual marker underwent a shift in position on the surface, or if the surface has deformed. The problem is to fit the model to the data and to find the maximum likelihood estimator: pointwise weights reflecting the reliability of the data.

Let (x, y, z) denote the measured centroid of a marker in the coordinate system of the measurement apparatus. The coordinate system was established by measuring the centroids of a collection of pin-balls attached to the reflector within the cryostat at each temperature and, for the present purposes, can be taken as given. The centroids are then measured relative to these coordinate systems.

Ellipsoid fitting requires a seven parameter model: the semi-major and semi-minor axis lengths, the rotations about the two semi-minor axes (pitch and yaw) and three displacements of the entire surface. Spheroid fitting requires nine parameters: the seven for the ellipsoid plus the second semi-minor axis length and rotation about the semi-major axis. In the spheroidal models the rotation about the major axis proved to be redundant, reflecting no detectable roll about that axis. Only the seven parameter model will be discussed as the nine parameter model is derived in the same way.

Let there be n data; then the input to the model is the preliminary guess for the best fit surface parameters $\boldsymbol{\beta}$ and the measurement data array x:

$$\boldsymbol{\beta} = (\beta_1, \cdots, \beta_7), \quad x = \{\boldsymbol{x}_i = (x_i, y_i, z_i),\ i = 1, \cdots, n\}.$$

Associated with each $\boldsymbol{x}_i \in x$ there is a vector of unknown errors, $\boldsymbol{\xi}_i$, arising from both measurement and surface error, and weight, w_i:

$$\xi = \{\boldsymbol{\xi}_i = (\xi_x, \xi_y, \xi_z),\ i = 1, \cdots, n\}, \quad w = \{w_i \in \mathbb{R}_{\geq 0},\ i = 1, \cdots, n\}.$$

The weights describe the estimated confidence in the accuracy of the centroids and can be adjusted in the light of successive fits and, for markers that came detached from the surfaces, given decreasing values between the initial, room temperature measurements and the last measurement set at which the marker was still attached (or zero if preferred).

Let $A \in SO(3;\mathbb{R})$ denote the matrix for rotation about semi-minor x-axis (parameter β_3) followed by rotation about semi-minor y-axis (parameter β_4). Let $\boldsymbol{d} = (\beta_5, \beta_6, \beta_7)$ be the translation vector of the ellipsoid centre, and write

$$\bar{\boldsymbol{x}}_i = (\bar{x}_i, \bar{y}_i, \bar{z}_i) = A(\boldsymbol{x}_i + \boldsymbol{\xi}_i + \boldsymbol{d}) \tag{7.1.1}$$

for the position of the corrected measurements. Once the correct model parameter, $\boldsymbol{\beta}$, has been found, the set $\{\bar{\boldsymbol{x}}_i\}$ will be the best fit ellipsoid to the data in standard form. For the ellipsoid the problem is to find:

$$\min_{\boldsymbol{\beta},\boldsymbol{\xi}} \sum_{i=1}^{n} w_i |\,\boldsymbol{\xi}_i\,|^2 \quad \text{subject to} \quad P(\bar{\boldsymbol{x}}_i) = \frac{\bar{x}_i^2 + \bar{y}_i^2}{\beta_1} + \frac{\bar{z}_i^2}{\beta_2} = 0 \quad \forall\, i, \tag{7.1.2}$$

where parameters β_1 and β_2 are the semi-minor and semi-major axis lengths, respectively. The errors are to be the orthogonal distance from the transformed measurement coordinate to the trial surface: for each i, $\boldsymbol{\xi}_i = \boldsymbol{x}_i - \boldsymbol{s}_i$ where $\boldsymbol{s}_i$ is that unique point on the trial surface minimising $|\,\boldsymbol{\xi}_i\,|$. With this definition of the error the closest model surface to the measured data will be found in the proper sense of metric space theory, subject only to the limitations of having a finite data set.

The first part of the problem is to find, for each trial $\boldsymbol{\beta}$, the associated error vector set $\{\boldsymbol{\xi}_i\}_{i=1}^n$. To do this a method used in non-sequential ray tracing was used: Given any spheroid and a ray intercept with that spheroid, to find the reflected ray it is necessary to find the normal to the surface at that intercept. However since any conformal deformation of the surface preserves the angles between lines at any point, the problem can be reduced to the trivial problem of finding the normal to the unit sphere. This is the way reflection and scattering from spheroidal surfaces is calculate in the non-sequential ray-tracing tools used for both stray light analysis in optical systems and in animation [30]. In the current problem the method is to apply the conformal mapping (of the trial ellipsoid onto the unit sphere) to the measured data: $\boldsymbol{x}_i \mapsto f_{\boldsymbol{\beta}}(\boldsymbol{x}_1)$, say. Each datum determines a radial vector which intercepts the sphere: $\sigma_i = f_{\boldsymbol{\beta}}(\boldsymbol{x}_1)/|f_{\boldsymbol{\beta}}(\boldsymbol{x}_i)|$ giving the required error estimate as $\boldsymbol{\xi}_i = \boldsymbol{x}_i - f_{\boldsymbol{\beta}}^{-1}(\sigma_i)$.

The orthogonal distance regression algorithm used was that described in [9] which returns standard deviation and 95% confidence interval estimates for each component of the model parameter, $\boldsymbol{\beta}$. It was programmed in standard conforming FORTRAN 95, and all data presented here is the output of the programs written for the purposes of fulfilling the requirements of the ESTEC contract. In the following three sections the tables of results are presented, followed by discussion of the evidence for reflector distortion due to residual thermal stress in the cryogenically cooled reflectors.

7.2 Results: fitting the SRFM measurement data

The measurement data for the SRFM comprised one CMM data set of 8949 surface coordinates, taken at 293 K, and eleven data sets taken by videogrammetry within the cryostat at CSL during the thermal cycling of the reflector over the temperature range range 293 K to 95 K, each set comprising approximately 2860 points. In all of the following tables the data is listed in the order of the data sets, starting with the CMM data and followed by the videogrammetry data in the temperature cycle order as follows:

293 K	293 K	293 K	140 K	110 K	95 K
CMM_SRFM	SRFM_M01	SRFM_M02	SRFM_M03	SRFM_M04	SRFM_M05
140 K	95 K	140 K	200 K	293 K	293 K
SRFM_M06	SRFM_M07	SRFM_M08	SRFM_M09	SRFM_M10	SRFM_M11

In all of the following tables the parameters are: **A** – semi-minor axis length (β_1); **B** – semi-major axis length (β_2); α – rotation about the x-axis (β_3); β – rotation about the y-axis (β_4); and **dX**, **dY** and **dZ** are longitudinal displacements along the x, y, and z axes ($\beta_5, \beta_6, \beta_7$); **Int FD** is the inter-focal distance and **VF dist** is the vertex to focus distance of the ellipsoid. All lengths are in millimetres, except for the residual surface errors given in table 7.4 which are in microns; all rotations are given in degrees. Displacement and rotation are measured relative to the coordinate system determined by the three point mount which is nominally co-centred with, and parallel to, the aperture coordinate system, with the semi-major axis aligned to the coordinate system z axis. The optical parameters for the surface – the radius of curvature and conic constant, or equivalently the inter-focal distance and vertex distance – are derived form the fit parameters A and B defined above as follows:

$$k = \left(\frac{A}{B}\right)^2 - 1, \qquad R = A^2/B,$$

$$\text{IF dist} = 2\sqrt{B^2 - A^2}, \qquad \text{VF dist} = B\left(1 - \sqrt{-k}\right).$$

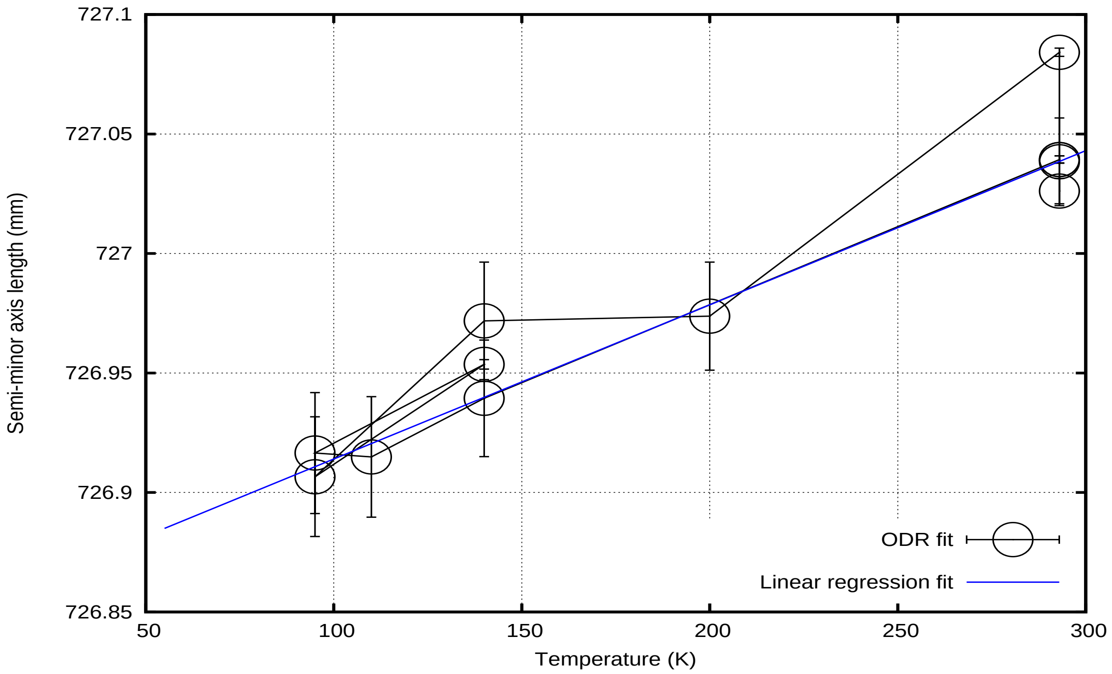
T (K)	A	σ_A	B	σ_B	α	σ_α	β
293	727.0261	0.00271	820.7323	0.0065	-0.0000002	0.000320	1.1695044
293	727.0393	0.00077	820.7485	0.0019	0.0147302	0.000120	1.0285662
293	727.0393	0.00077	820.7485	0.0019	0.0147302	0.000120	1.0285662
140	726.9394	0.01243	820.5658	0.0302	0.0279317	0.001927	1.0697632
110	726.9149	0.01285	820.5156	0.0312	0.0289275	0.001993	1.0799602
95	726.9165	0.01288	820.5226	0.0312	0.0308254	0.001997	1.0763719
140	726.9536	0.00100	820.5678	0.0025	0.0298980	0.000157	1.0705069
95	726.9066	0.01279	820.5093	0.0310	0.0282322	0.001984	1.0788939
140	726.9718	0.01253	820.6074	0.0304	0.0260809	0.001941	1.0715667
200	726.9738	0.01153	820.6312	0.0280	0.0299922	0.001786	1.0545999
293	727.0842	0.00085	820.8199	0.0021	0.0139935	0.000133	1.0171727
293	727.0384	0.00934	820.7739	0.0227	0.0183574	0.001444	1.0232384

T (K)	σ_β	dX	σ_{dX}	dY	σ_{dY}	dZ	σ_{dZ}
293	0.000959	330.305	0.005024	-0.000	0.0026005	432.778	0.0119650
293	0.000277	334.689	0.001514	-0.219	0.0009901	443.698	0.0034819
293	0.000277	334.689	0.001514	-0.219	0.0009901	443.698	0.0034819
140	0.004412	334.899	0.023970	-0.327	0.0158200	443.255	0.0556710
110	0.004562	334.952	0.024780	-0.335	0.0163580	443.142	0.0575490
95	0.004572	334.930	0.024837	-0.351	0.0163950	443.173	0.0576810
140	0.000361	334.914	0.001970	-0.343	0.0012893	443.245	0.0045290
95	0.004541	334.940	0.024666	-0.330	0.0162830	443.144	0.0572850
140	0.004446	334.930	0.024152	-0.313	0.0159400	443.277	0.0560960
200	0.004089	334.830	0.022220	-0.311	0.0146600	443.412	0.0516090
293	0.000306	334.640	0.001674	-0.213	0.0010944	443.834	0.0038496
293	0.003308	334.651	0.017985	-0.248	0.0118570	443.761	0.0417780

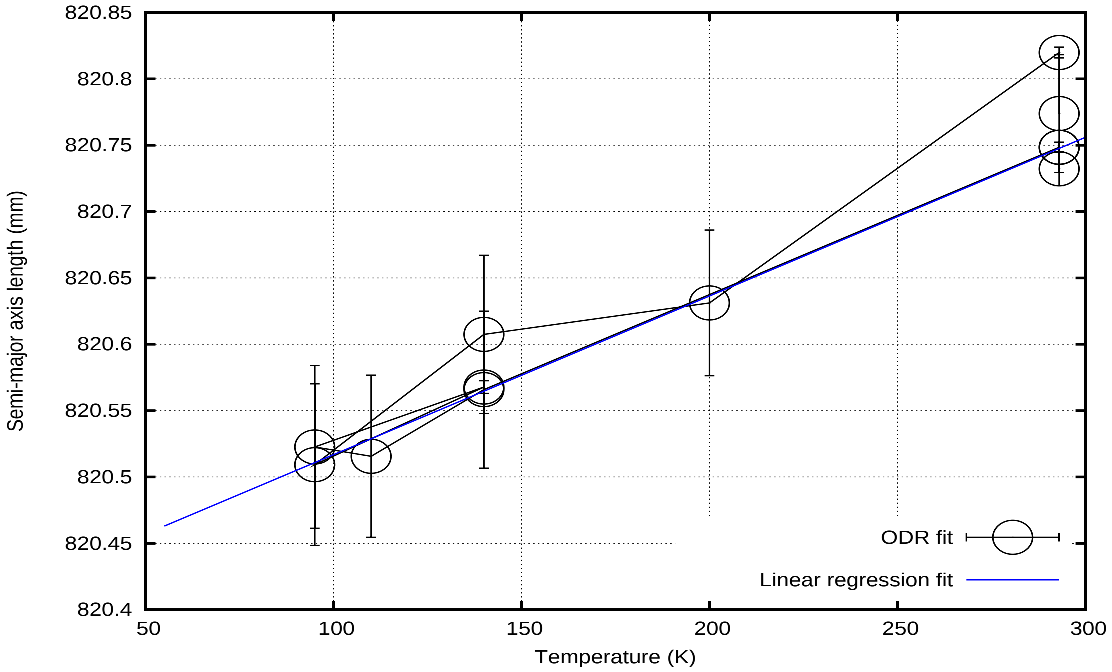
Table 7.1: The results obtained from fitting the seven parameter ellipsoid model surface by orthogonal distance regression to the 12 data sets for the SRFM.

Semi-minor axis:	726.885	mm	(dX, dY, dZ):	(335.005,-0.376,443.048)
Semi-major axis:	820.463	mm		
Radius of curvature:	643.977	mm	α:	0.034°
Conic constant:	-0.215104		β:	41.091°
Inter-focal dist:	761.048		Vertex-focus dist:	439.937

Table 7.2: 55K estimates for the SRFM derived by linear regression on the results obtained by ODR fitting to the CMM and videogrammetry data files

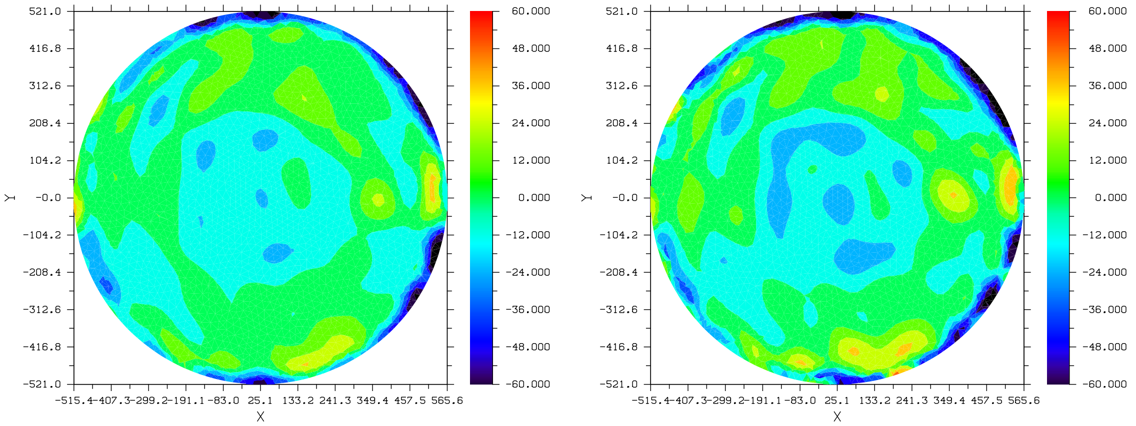


(a) Fitted semi-minor axis as a function of temperature. The ratio of the 95% confidence interval width to the semi-minor axis length is ±3.5 × 10^{-5} for the videogrammetry data.

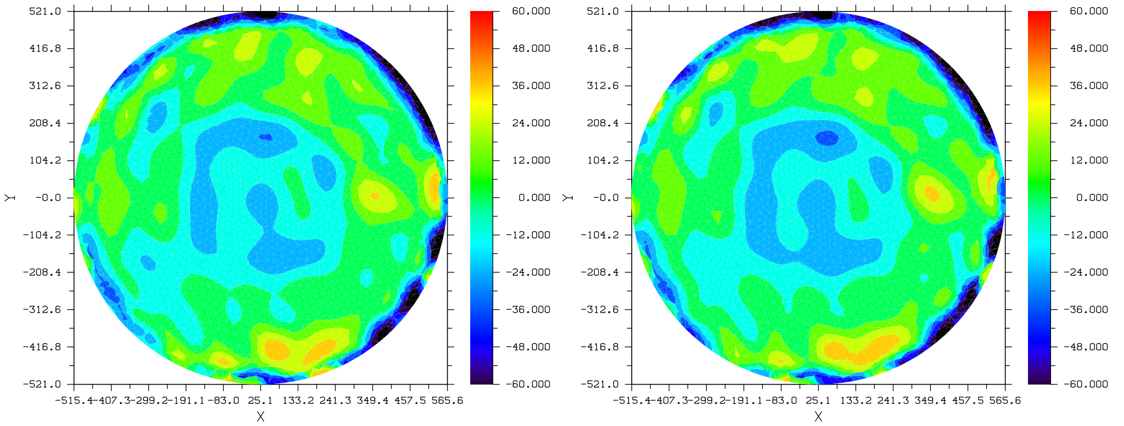


(b) Fitted semi-major axis as a function of temperature. The ratio of the 95% fit confidence interval width to the semi-major axis length is ±6 × 10^{-5} for the videogrammetry data.

Figure 7.1: SRFM semi-axes and linear regression fit. To derive the fit point at 55K the outlier points at 140K and 293K under vacuum were removed from the data. Including them changes both axis lengths by less than 20 μm (approximately 2.5 × 10^{-3} % in each case) but lowers the probability estimate of correctness of fit the the semi-minor axis length below the usually accepted minimum.



(a) CMM_SRFM at 293K (b) SRFM_M09 at 200K



(c) SRFM_M06 at 140K (d) SRFM_M05 at 95K

Figure 7.2: The changing surface form sagitta error (in aperture coordinates) as the SRFM cools. These maps are the residual when the best fit ellipsoid has been subtracted from the surface data and are measured along the ellipsoid normal. All maps are on a common scale from $+60\,\mu$m to $-60\,\mu$m.

T(K)	A	A −	A +	B	B −	B +	Data Set
293	727.0261	727.0208	727.0314	820.7323	820.7196	820.7450	CMM_SRFM
293	727.0393	727.0378	727.0408	820.7485	820.7448	820.7522	SRFM_M01
293	727.0393	727.0378	727.0408	820.7485	820.7448	820.7522	SRFM_M02
140	726.9394	726.9150	726.9638	820.5658	820.5066	820.6250	SRFM_M03
110	726.9149	726.8897	726.9401	820.5156	820.4545	820.5767	SRFM_M04
95	726.9165	726.8912	726.9418	820.5226	820.4613	820.5839	SRFM_M05
140	726.9536	726.9516	726.9556	820.5678	820.5630	820.5726	SRFM_M06
95	726.9066	726.8816	726.9317	820.5093	820.4484	820.5702	SRFM_M07
140	726.9718	726.9472	726.9964	820.6074	820.5478	820.6671	SRFM_M08
200	726.9738	726.9512	726.9964	820.6312	820.5763	820.6861	SRFM_M09
293	727.0842	727.0825	727.0859	820.8199	820.8158	820.8240	SRFM_M10
293	727.0384	727.0200	727.0567	820.7739	820.7294	820.8183	SRFM_M11

Table 7.3: SRFM semi-minor and semi-major axis lengths and 95% confidence intervals obtained by orthogonal distance regression. Results are shown in red should be compared with other data taken at the same temperature.

All of the fitting results indicate that the SRFM was not in a state of thermal equilibrium at the times in the thermal cycle at which the videogrammetry measurements were made. With reference to table 7.3 above, the first row is the room temperature fit to the CMM data. This data is intrinsically more accurate than the videogrammetry data and the data set is more than three times the size. Furthermore, the data are taken over the entire surface, right out to the edge of the reflector where the surface form error is greatest. Effectively, the CMM data is the most reliable data set, and the videogrammetry data is available only over a smaller portion of the reflector over which the surface shape is closest to an ellipsoid. That is why the CMM fit returns an ellipsoid with semi-axes of the order $15\,\mu$m shorter than the first pair of videogrammetry results at the same temperature (the extreme edges of the reflectors curl up very slightly).

The second and third rows in table 7.3 show repeated videogrammetry measurements taken before the surface thermal cycle. As would be hoped, they return identical fits to within error. The reflector was then cooled to 95 K. The two measurements at 95 K are different, but both are within the error bounds of the other and the 95% confidence interval for the two semi-axis lengths are virtually identical. Looking at the three measurements made at 140 K shows that the second time the reflector returned to 140 K it showed less distortion than in the first and third measurements, and both those measurements returned axis lengths outside the 95% confidence interval for the second measurement. This is seen clearly in the positions of the three circles at 140 K in figure 7.1 where the first measurement is clearly anomalous due, presumably, to the reflector not being at a uniform temperature when the measurements were made.

Finally, the last two rows show the fits when the system has nominally returned to its original temperature, but clearly the reflector itself was in a state of thermal stress when data set SRFM_M10 was taken. When measurement set SRFM_M11 was taken the reflector had returned to its original shape to within the 95% confidence interval.

With the ellipsoid fitted to the data, the residual fit errors are the orthogonal distances from the surface of the ellipsoid to the measurement datum. These distances provide a form error map, the statistics for which are presented in table 7.4. From these maps an estimate for the 50 K form error was derived by the methods presented in section 7.5 below.

T	Surface shape parameters (mm)				Surface form error (μm)			
(K)	Rad	Conic	Int FD	VF dist	Stdev	Min	Max	P-V
293	644.0188	-0.2153	761.668	439.898	11.6344	-38.4107	44.2549	82.67
293	644.0294	-0.2153	761.687	439.905	8.9235	-29.9294	35.1534	65.08
293	644.0294	-0.2153	761.687	439.905	18.7312	-147.8871	52.3168	200.20
140	643.9957	-0.2152	761.282	439.925	8.9235	-29.9294	35.1534	65.08
110	643.9918	-0.2151	761.158	439.936	18.7312	-147.8871	52.3168	200.20
95	643.9891	-0.2151	761.183	439.931	11.7547	-43.3531	46.2428	89.60
140	644.0193	-0.2152	761.236	439.950	24.8199	-153.5048	74.7819	228.29
95	643.9821	-0.2151	761.163	439.928	12.1530	-40.4716	50.5232	90.99
140	644.0205	-0.2152	761.337	439.939	25.7199	-169.2620	76.2625	245.52
200	644.0053	-0.2152	761.432	439.915	12.1803	-41.7040	45.3714	87.08
293	644.0529	-0.2154	761.824	439.908	25.8721	-179.0477	78.9238	257.97
293	644.0078	-0.2154	761.800	439.874	24.8788	-161.4504	83.8810	245.33

Table 7.4: Best fit radius of curvature, conic constant, inter-focal distance and vertex-focus distance derived for the SRFM by orthogonal distance regression fitting to the CMM and videogrammetry data files. With each set of results is given basic statistics for the residual fitting errors.

T (K)	Rad	Rad +	Rad −	Conic	Conic -	Conic +	Data Set
293	644.0188	644.0382	643.9994	-0.2153	-0.2153	-0.2153	CMM_SRFM
293	644.0294	644.0350	644.0238	-0.2153	-0.2153	-0.2153	SRFM_M01
293	644.0294	644.0350	644.0238	-0.2153	-0.2153	-0.2153	SRFM_M02
140	643.9957	644.0853	643.9060	-0.2152	-0.2150	-0.2153	SRFM_M03
110	643.9918	644.0845	643.8992	-0.2151	-0.2150	-0.2153	SRFM_M04
95	643.9891	644.0820	643.8962	-0.2151	-0.2150	-0.2153	SRFM_M05
140	644.0193	644.0266	644.0121	-0.2152	-0.2151	-0.2152	SRFM_M06
95	643.9821	644.0743	643.8899	-0.2151	-0.2150	-0.2153	SRFM_M07
140	644.0205	644.1109	643.9302	-0.2152	-0.2150	-0.2154	SRFM_M08
200	644.0053	644.0885	643.9222	-0.2152	-0.2151	-0.2154	SRFM_M09
293	644.0529	644.0591	644.0467	-0.2154	-0.2153	-0.2154	SRFM_M10
293	644.0078	644.0752	643.9405	-0.2154	-0.2152	-0.2155	SRFM_M11

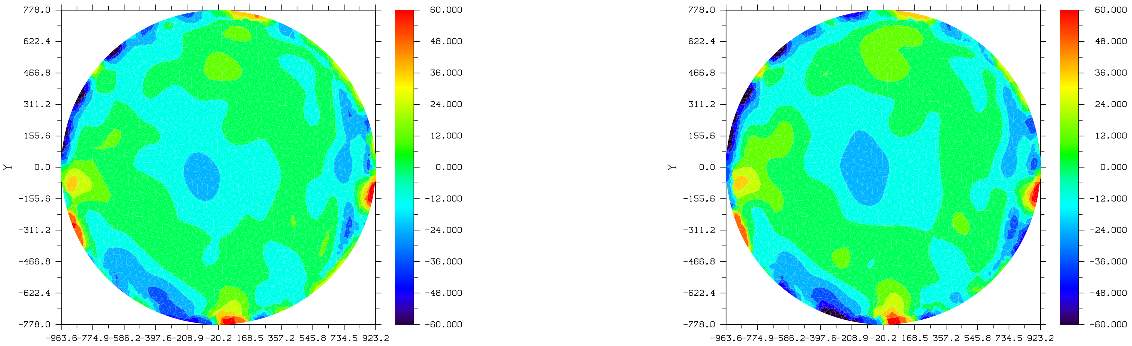
Table 7.5: Orthogonal distance regression fits to the SRFM radius of curvature and conic, with 95% confidence intervals.

7.3 Results: fitting the PRFM measurement data

The procedures described above for the fitting of the SRFM were followed exactly for fitting the PRFM. All data for the PRFM is tabulated in exactly the same way as for the SRFM and all nomenclature is as before. The same general comments about reflector distortion apply as before, and will not be repeated. The only additional comment is that the fitting of a spheroid showed that the two semi-minor axes differed slightly and the general level of surface distortion was higher than for the SRFM. That would be expected because the PRFM is relatively flat, and being of the same thickness would therefore be less rigid.

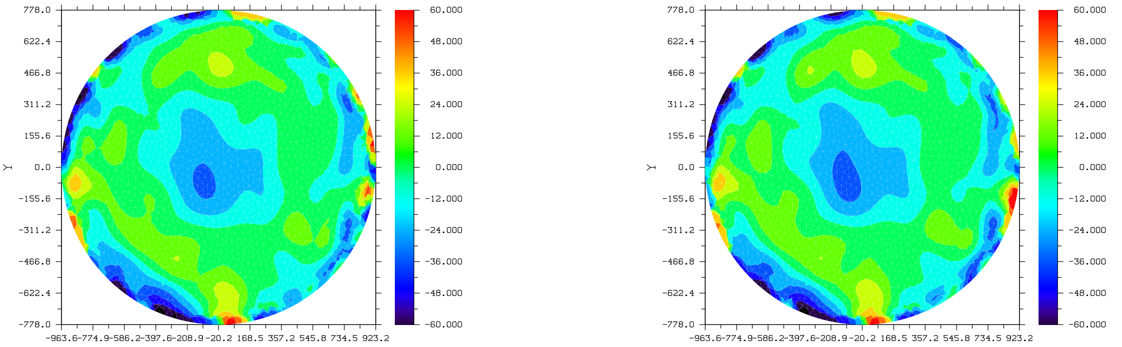
In all of the following tables the data is listed in the order of the data sets, starting with the CMM data and followed by the videogrammetry data in the temperature cycle order as follows:

293 K	293 K	293 K	140 K	110 K	95 K
CMM_PRFM	PRFM_M01	PRFM_M02	PRFM_M03	PRFM_M04	PRFM_M05
140 K	95 K	140 K	200 K	293 K	293 K
PRFM_M06	PRFM_M07	PRFM_M08	PRFM_M09	PRFM_M10	PRFM_M11



(a) CMM_PRFM at 293K

(b) PRFM_M09 at 200K



(c) PRFM_M06 at 140K

(d) PRFM_M05 at 95K

Figure 7.3: The changing surface form sagitta error (in aperture coordinates) as the PRFM cools. All maps are on a common scale of $+60\,\mu$m to $-60\,\mu$m.

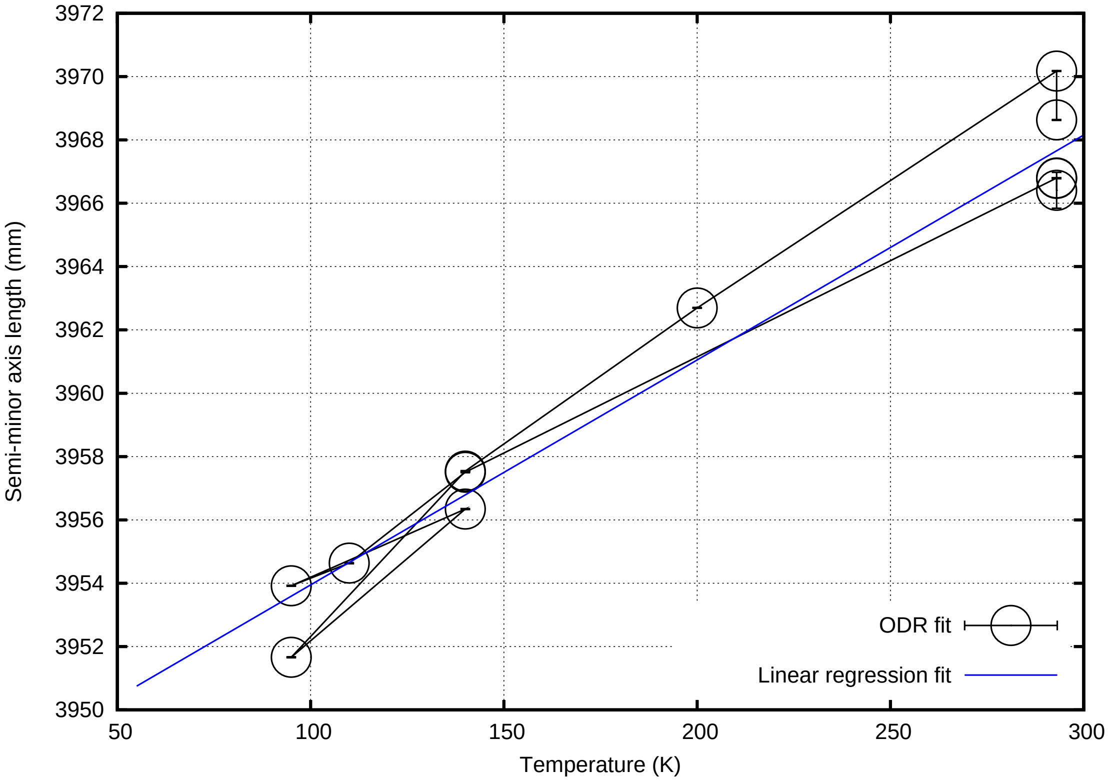
Table 7.6: The results obtained from fitting the seven parameter ellipsoid model surface by ODR to the 12 data sets for the PRFM.

T (K)	A	σ_A	B	σ_B	α	σ_α	β
293	3966.409	0.2891	10923.553	1.6013	-0.0000	0.0001	-37.2041
293	3966.789	0.0062	10926.226	0.0346	0.3692	0.0000	-37.2224
293	3966.797	0.0062	10926.249	0.0345	0.3691	0.0000	-37.2223
140	3957.509	0.0077	10879.348	0.0426	0.3630	0.0000	-37.2467
110	3954.634	0.0079	10864.014	0.0436	0.3627	0.0000	-37.2540
95	3953.920	0.0080	10860.846	0.0442	0.3616	0.0000	-37.2558
140	3956.341	0.0077	10872.837	0.0425	0.3627	0.0000	-37.2488
95	3951.662	0.0080	10849.105	0.0440	0.3612	0.0000	-37.2589
140	3957.542	0.0076	10878.482	0.0420	0.3630	0.0000	-37.2473
200	3962.696	0.0070	10905.552	0.0386	0.3636	0.0000	-37.2364
293	3970.175	0.0067	10943.809	0.0369	0.3685	0.0000	-37.2167
293	3968.632	0.0063	10935.749	0.0351	0.3680	0.0000	-37.2190

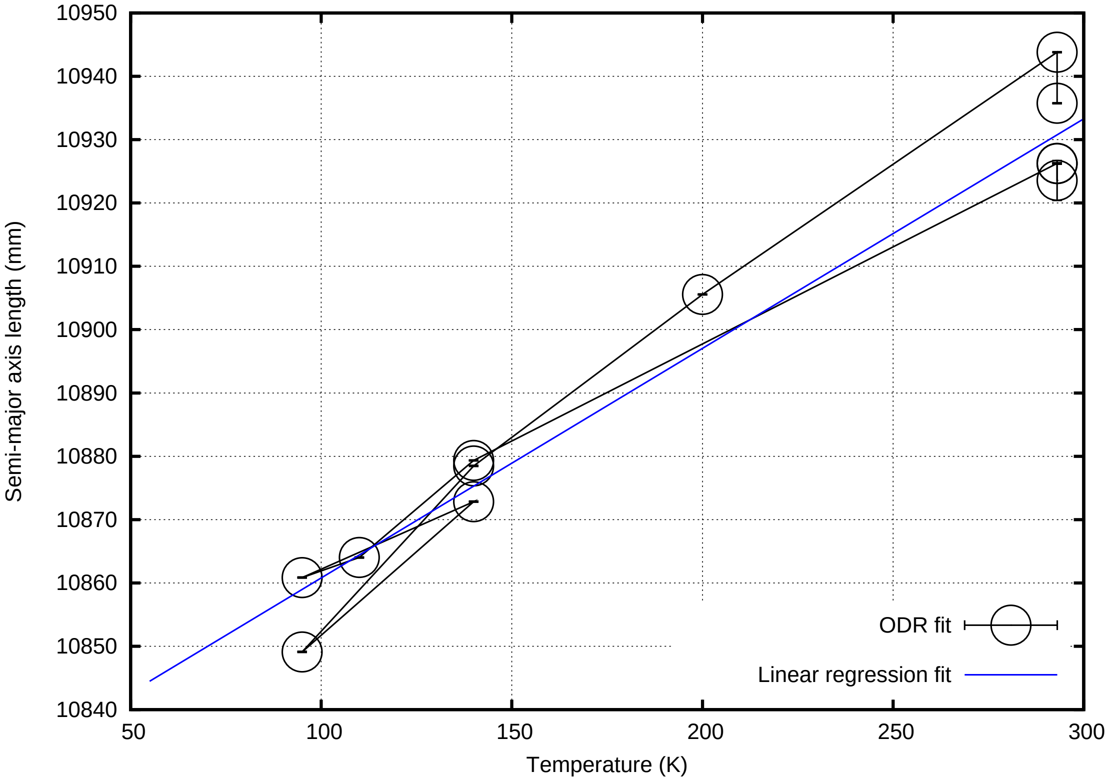
T (K)	σ_β	dX	σ_{dX}	dY	σ_{dY}	dZ	d σ_{dZ}
293	0.0005	-1050.7988	0.0108	0.0000	0.0038	10311.875	1.6115
293	0.0000	-1052.3921	0.0002	-16.3600	0.0001	10320.353	0.0348
293	0.0000	-1052.3917	0.0002	-16.3578	0.0001	10320.377	0.0347
140	0.0000	-1052.6020	0.0003	-16.1729	0.0001	10273.050	0.0429
110	0.0000	-1052.7330	0.0003	-16.1657	0.0001	10257.561	0.0439
95	0.0000	-1052.7083	0.0003	-16.1325	0.0001	10254.380	0.0444
140	0.0000	-1052.6529	0.0003	-16.1667	0.0001	10266.450	0.0427
95	0.0000	-1052.6936	0.0003	-16.1189	0.0001	10242.620	0.0443
140	0.0000	-1052.7182	0.0003	-16.1779	0.0001	10272.127	0.0423
200	0.0000	-1052.5934	0.0003	-16.1948	0.0001	10299.403	0.0388
293	0.0000	-1052.3819	0.0003	-16.3417	0.0001	10338.017	0.0371
293	0.0000	-1052.3846	0.0002	-16.3281	0.0001	10329.927	0.0353

Table 7.7: 55 K estimates for the PRFM derived by linear regression on the results obtained by ODR fitting to the CMM and videogrammetry data files

Semi-minor axis:	3950.749	mm	dX:	-1052.799	mm
Semi-major axis:	10844.348	mm	dY:	-16.080	mm
			dZ:	10238.622	mm
Radius of curvature:	1439.314	mm	α:	0.35985°	
Conic constant:	-0.867275		β:	-37.263324°	
Inter-focal distance:	2019.154	mm			
Vertex-focus distance:	745.226	mm			



(a) Fitted semi-minor axis as a function of temperature.



(b) Fitted semi-major axis as a function of temperature.

Figure 7.4: PRFM semi-axes and linear regression fit. To derive the fit point at 55K the outlier point on warm up at 293K under vacuum were removed form the data.

T	Surface shape parameters (mm)				Surface form error (μm)			
(K)	Rad	Conic	Int FD	VF dist	Stdev	Min	Max	P-V
293	1440.227	-0.8682	20355.993	745.557	20.846	-87.959	129.668	217.63
293	1440.151	-0.8682	20361.434	745.509	13.339	-105.206	48.015	153.22
293	1440.154	-0.8682	20361.477	745.510	16.656	-77.657	127.442	205.10
140	1439.593	-0.8677	20268.041	745.327	13.283	-105.069	47.161	152.23
110	1439.535	-0.8675	20237.359	745.335	16.576	-77.922	127.276	205.20
95	1439.436	-0.8675	20231.113	745.289	16.710	-108.067	65.029	173.10
140	1439.609	-0.8676	20254.970	745.352	21.023	-80.727	130.969	211.70
95	1439.348	-0.8673	20207.667	745.272	17.058	-109.300	68.766	178.07
140	1439.735	-0.8677	20266.153	745.406	21.548	-87.605	132.486	220.09
200	1439.905	-0.8680	20320.246	745.429	17.223	-106.050	69.131	175.18
293	1440.293	-0.8684	20396.534	745.541	21.666	-85.837	128.549	214.39
293	1440.234	-0.8683	20380.439	745.530	16.515	-106.987	66.485	173.47

Table 7.8: Best fit radius of curvature, conic constant, inter-focal distance and vertex-focus distance for the PRFM derived by orthogonal distance regression fitting to the CMM and videogrammetry data files. With each set of results is given basic statistics for the residual fitting errors.

T (K)	A	A -	A +	B	B -	B +
293	3966.409	3965.842	3966.975	10923.553	10920.414	10926.692
293	3966.789	3966.777	3966.801	10926.226	10926.158	10926.294
293	3966.797	3966.785	3966.809	10926.249	10926.181	10926.316
140	3957.504	3957.489	3957.519	10879.348	10879.264	10879.431
110	3954.634	3954.618	3954.649	10864.014	10863.929	10864.100
95	3953.920	3953.905	3953.936	10860.846	10860.759	10860.932
140	3956.341	3956.326	3956.356	10872.837	10872.754	10872.920
95	3951.662	3951.647	3951.678	10849.105	10849.019	10849.192
140	3957.542	3957.527	3957.557	10878.482	10878.400	10878.565
200	3962.696	3962.683	3962.710	10905.552	10905.476	10905.627
293	3970.175	3970.162	3970.189	10943.809	10943.736	10943.881
293	3968.632	3968.620	3968.645	10935.749	10935.681	10935.818

Table 7.9: PRFM semi-minor and semi-major axis lengths and 95% confidence intervals.

Table 7.10: PRFM radius of curvature and conic 95% confidence intervals.

T (K)	Rad	Rad -	Rad +	Conic	Conic -	Conic +
293	1440.2271	1441.0527	1439.4021	-0.8682	-0.8680	-0.8683
293	1440.1509	1440.1687	1440.1330	-0.8682	-0.8682	-0.8682
293	1440.1538	1440.1716	1440.1360	-0.8682	-0.8682	-0.8682
140	1439.5933	1439.6153	1439.5713	-0.8677	-0.8677	-0.8677
110	1439.5348	1439.5574	1439.5122	-0.8675	-0.8675	-0.8675
95	1439.4355	1439.4583	1439.4126	-0.8675	-0.8675	-0.8675
140	1439.6090	1439.6310	1439.5870	-0.8676	-0.8676	-0.8676
95	1439.3476	1439.3704	1439.3247	-0.8673	-0.8673	-0.8673
140	1439.7354	1439.7571	1439.7137	-0.8677	-0.8676	-0.8677
200	1439.9052	1439.9251	1439.8853	-0.8680	-0.8680	-0.8680
293	1440.2932	1440.3122	1440.2742	-0.8684	-0.8684	-0.8684
293	1440.2343	1440.2523	1440.2162	-0.8683	-0.8683	-0.8683

7.4 Surface form error maps

At each temperature, the orthogonal distance regression fitting of the ellipsoid to the videogrammetry data gave, at each datum, a residual surface error vector $\boldsymbol{\varepsilon} = (\varepsilon_x, \varepsilon_y, \varepsilon_z)$ orthogonal to the best fit surface. The position of each datum was measured relative to the mounting coordinates, and having derived the ellipsoid parameters the coordinate transformation could be made to the ellipsoid vertex coordinate system and to the reflector aperture coordinate system. The plane of the aperture coordinate system is parallel to the plane of the mounting coordinate system for both reflectors. If P_a is the projection operator onto the aperture plane, then at each datum the scalar $s = \boldsymbol{\varepsilon} - P_a \boldsymbol{\varepsilon}$ is a surface form error vector at the datum orthogonal to the aperture. These maps can be used in the GRASP model to add form error to the perfect ellipsoid.

The problem was to derive a form error map for each of the two reflectors that extended to the reflector rims and would be a reasonable approximation of the true, but unknown, form errors at operating temperature. By the term ‘reasonable approximation' is meant the following: (a) the form error statistics of the extrapolated Zernike surface should follow the temperature dependent trend set by the videogrammetry data, particularly the RMS form error that is the most important measure for wavefront error, and (b) that the overall shape of the error surface should follow the observed temperature dependent trend (see figures 7.2 and 7.3).

The method adopted was to use the Zernike surface reconstruction methods developed for the study of the thermoelastic deformation of the Herschel telescope primary reflector, [49][48]: first obtain a very high fidelity Zernike polynomial fit to each sagitta error map at each temperature, then use linear regression analysis to extrapolate each of the Zernike coefficients individually to 50 K, and finally rebuild the surface sagitta error estimate map at 50 K using those extrapolated values. Technical details of the fitting and mathematical details are to be found in [49]. For this study the first 325 Zernike polynomials were used (in the ordering due to Zernike; see [11] chapter 9 and Appendix VII, and [49] chapter 6) which include all radial and azimuthal orders 0 to 24.

The difficulty with the method was that the Zernike coefficients, whilst reproducing the videogrammetry maps with very high precision, showed marked fluctuation even between videogrammetry maps taken at the same temperature. That reflected the fact that, at the scale of the surface form errors, the reflector surfaces showed marked changes with temperature and residual surface stress. The only reasonable test of the likelihood that the resulting surface was realistic was to look at the surface form error statistics and to compare those with the interferogram data (available only for the SRFM). The reconstructed map covered the entire surface whereas the interferogram did not include the edges where roll-off and distortion were greatest. Therefore the Zernike data should show slightly larger RMS form error and significantly larger peak-to-valley error. Furthermore, the RMS value should be close to the 95 K videogrammetry map. To make a good comparison with the 95 K map the full aperture form error map was built for the second of the two 95 K data sets (for both reflectors) using the Zernike reconstruction of the error maps. The 50 K maps and 95 K maps for each reflector were built on the same triangulation of the reflector aperture. The surface for error statistics for the 95 K surfaces presented in the following table are derived form those maps and should be compared with the statistics for data sets SRFM_M07 and PRFM_M07 in rows six and eight of tables 7.4, page 185, and 7.8 page 189, respectively. For the SRFM and PRFM the results were as follows:

	SRFM 50 K Zernike	SRFM at 95 K	Interferogram
Mean (μm) :	0.217	-0.593	0.000
RMS (μm) :	13.508	12.018	10.200
Maximum (μm) :	74.335	63.231	no data
Minimum (μm) :	-121.964	-112.013	no data
P-V (μm) :	196.299	175.244	114.000

	PRFM 50 K Zernike	PRFM at 95 K	Interferogram
Mean (μm) :	0.866	0.000	no data
RMS (μm) :	21.273	19.812	no data
Maximum (μm) :	314.740	246.178	no data
Minimum (μm) :	-164.079	-108.698	no data
P-V (μm) :	478.819	354.876	no data

Table 7.11: Estimates of the 50 K surface form error statistics for the SRFM and the PRFM compared with the statistics for the Zernike reconstruction of the 95 K surface and the interferometer measurement over the measurable section of surface.

From these tabulated data we get the only indication available that the RMS wavefront error statistics, the key measure of the reflected beam quality, will be reasonably reliable:

SRFM: 27.016 μm	PRFM: 42.546 μm	Total: 50.399 μm

Table 7.12: Estimated in-flight wavefront error contributions from the SRFM and PRFM, and the total wavefront error. These estimates assume uniform aperture illumination of the entire surface of both reflectors, and, for the total, statistical independence for the form error in the two reflectors.

For the highest frequency channel at 857 GHz, where surface form error is of greatest concern, this total RMS wavefront error corresponds to $\lambda/7$. Furthermore, this estimate will be pessimistic because the illumination of the apertures is Gaussian, not uniform. At mid band in the lowest HFI frequency channel, 100 GHz, these form errors give $\lambda/57$ RMS wavefront error. The generally accepted definition of diffraction limit is $\lambda/4$, and though there will be contributions to wavefront error from optical misalignment and the departure of the surfaces from the ideal shape, this analysis suggested that the contribution form form error would not be significant.

7.5 Quilted surface form error maps

The Planck reflectors were constructed form resin bonded carbon fibre in the form of a front and back skin separated by a honeycomb. As the structure cools the bonding of the skins to the honeycomb causes stress in the surface with the result that the honeycomb shows through the reflector surface as print-through, and the surface quilts. The quilting is not uniform and does not follow the common models of a simple cosine surface sagitta over each honeycomb cell. That was revealed by the series of cryogenic interferograms taken by Robert Daddato (ESTEC). The sequence of interferograms also showed that the quilting over individual cells could completely reverse the quilt sagitta sign over the cooling cycle. A full technical report on the interferometry set-up and surface recovery for both the SRFM and PRFM was given in [16]. The videogrammetry gave a general surface shape and form error map at 50 K, but with only one or two data per cell, it could not reveal the quilting detail. The final requirement of the surface modelling contract was to produce a surface form error map, to be included in the GRASP model of the telescope, that included the quilting as well as the general form error.

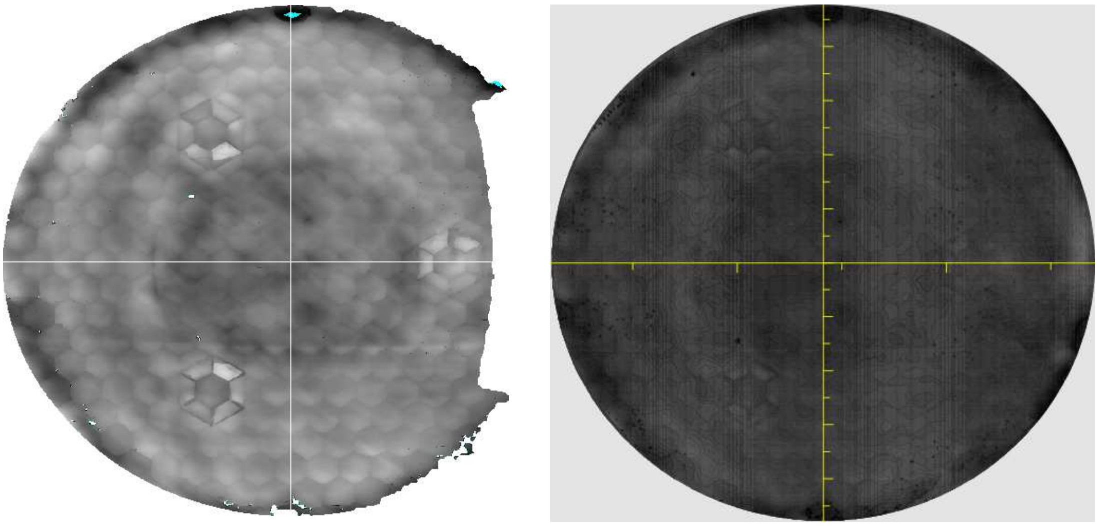
Reconstruction of the complete surface form error maps was made difficult by two circumstances: For the SRFM the interferogram was good but only covered part of the reflector (see figure 7.5); for the PRFM technical difficulties described in [16] meant that the maps were incomplete and unreliable. (Essentially the problem was that the reflector surface gradients at, and normal to, the cell boundaries where too great to permit phase unwrapping to derive quilt depth.) Therefore, for the SRFM the problem was how to use the interferogram and videogrammetry maps to extend the quilting map over the entire surface, while for the PRFM the problem was how, given the limited information available, to construct a quilted surface map that, although it could not be strictly correct, was at least a quilting on the correct scale in terms peak-to-valley and RMS error, and correctly positioned. The method adopted is described in the following two sections.

7.6 The SRFM surface quilting map

The best available surface error map at around operating temperature is derived form the interferometry map *09_SRFM_2_1_50K_30JUN05_1STIT_SFE_sag_Zvrf.itx*. This file was used to derive the quilting map on both surfaces, but in different ways. The interferometry map itself is of the residual surface error over as much of the reflector surface as it was possible to illuminate in the interferometry setup. It is in the aperture coordinate system, O-M2C, in which the honeycomb cell structure forms a regular hexagonal pattern. The sampling of the surface is on the scale 1 mm, which is too fine for practical use in the the GRASP model, but does provide great detail to use in building a map. As seen in figure 7.5-(a), the interferogram does not extend quite to the reflector rim and the part of the reflector surface closest to the parent ellipsoid's vertex is missing.

The hexagonal cell structure has a wall spacing of 600 mm. To extract the surface detail, upon each cell in the interferogram a disc of radius 600 mm was centred with an equilateral triangulation of 4 mm on a side. Thus, the triangulation of the disc mapped to a triangulation of the cell skin over the cell and overlapped the adjacent cells as far as the cell centre. The interferogram data that was supported on the disc was interpolated onto the vertices of the triangulation using Akima's method [3][4][49]. Triangulation vertices that fell beyond the bounds of the interferogram map were assigned the height value zero. From each of these discs a second copy was made with the best fit plane removed. These 'flattened' cell maps have essentially only the residual quilting structure (since the surface curvature has already been removed in taking the interferogram).

A smooth partition of unity was then constructed that took the value 1 over the central hexagon and fell to zero at the boundary of the disc. Thus, using the partition of unity, the entire quilted surface map could be reconstructed from the family of



(a) Interferogram of the SRFM at 50 K (b) Reconstruction of the SRFM at 50 K

Figure 7.5: The interferogram map of part of the SRFM at 55 K, and a plot of the reconstruction of the entire surface form error map using the videogrammetry measurements of the full surface. Note that the quilting depth is identical in both maps over the area covered by the interferogram, but the low resolution grey scale of the reconstruction plot does not show the detail of the interferogram. Data courtesy of Robert Daddato (ESTEC)

discs, and this reconstruction was identical to the original data set interpolated onto the vertices of the 4 mm side length triangulation.

The quilting map was extended to give the entire surface map as follows: On the section of the map where the interferometric map is complete the original discs were used so that on that part of the surface the interferometric map is recovered, but with the new and coarser sampling. On the missing data part, a random selection was made from the 'flattened' discs – one for each missing cell. To the sagitta of the points in each disc was added the sagitta of the 50 K surface form error model at the coordinates of the triangulation vertices. These discs were then pieced together with all of the other discs using the partition of unity. The resulting map is illustrated in figure 7.5-(b). It must be stressed that, over the area covered by the interferogram it recovers the interferogram map exactly, and it blends smoothly into the videogrammetry data over the remainder of the surface giving a quilting over the whole surface with a smooth transition across the patch boundary.

The output map was written to a text file in the GRASP SFC file format, [55], to give a residual surface error map that could be used in the Physical Optics modelling of the 50 K telescope for beam pattern prediction. Note that the O-M2C coordinate system has height axis +Z downwards (into the reflector) so that the local depression in the centre of the map is a high region on the reflector. In the GRASP model supplied the coordinate system is correctly oriented for interface control conformance and use with the deformation map.

7.7 The PRFM surface quilting map

The PRFM map is derived from (a) the 50 K Zernike surface model described in section 7.4, and (b) 'flattened' discs of the quilt model described above. The size of the PRFM means that to avoid excessive computation times (with no demonstrable gain in accuracy) it was sensible to build the model with coarser sampling than is used on the SRFM. A sample spacing of 6mm was used. Again, an equilateral triangular grid of points was used, and the residual SRFM quilting surface re-sampled at the new scale and the sample discs 'flattened'. (The sample spacing on the discs had to be chosen to be an integer divisor of the cell's inter-wall spacing so that in all models the quilt boundary was correctly reproduced.) All discs that overlap the edge of the SRFM were then rejected because they have atypical structure. Furthermore, the 21 patches that adjoin the mount points (the three obvious large hexagons in the interferogram 7.5) were separated out, leaving a subset of 174 'interior' cell patches.

The underlying quilt was then constructed as follows: For every cell in the PRFM that is not adjacent to a mount point, randomly select a disc from the set of 174 and, for each disc, randomly rotate it through $n\pi/3$ radians, where $n \in \{0, 1, \cdots, 5\}$. For the cells adjacent to a mount point select an appropriate cell according to its position and orientation. Patch all of the 745 cells (illustrated in figure 7.6) together using a smooth partition of unity. The result was a map of 225353 points representative of the type of quilting that was to be expected for the PRFM at operating temperature, without the general surface form error.

This quilting map, by construction, has mean height essentially zero, and no unrealistic local slopes. Because the depth of the quilts was unknown (due to the phase unwrapping difficulties) the reconstruction allowed for an optional depth scaling of the quilt. Having made the quilt map, the next step was to build the 50 K Zernike surface error map, as described in section 7.4, on the same set of (x,y)-coordinates as the points of the quilting map. Note that the Zernike map was derived from the videogrammetry data that had an average of only three sample points per cell; consequently there is virtually no cell deformation information in the Zernike map. The resulting large scale surface form error sagitta were then added to the fine scale quilting map to produce the final form error map for the PRFM. (In the technical report [50] a plot of the final map was given. It is not reproduced here because the quilt structure is almost indiscernible on an A4 plot).

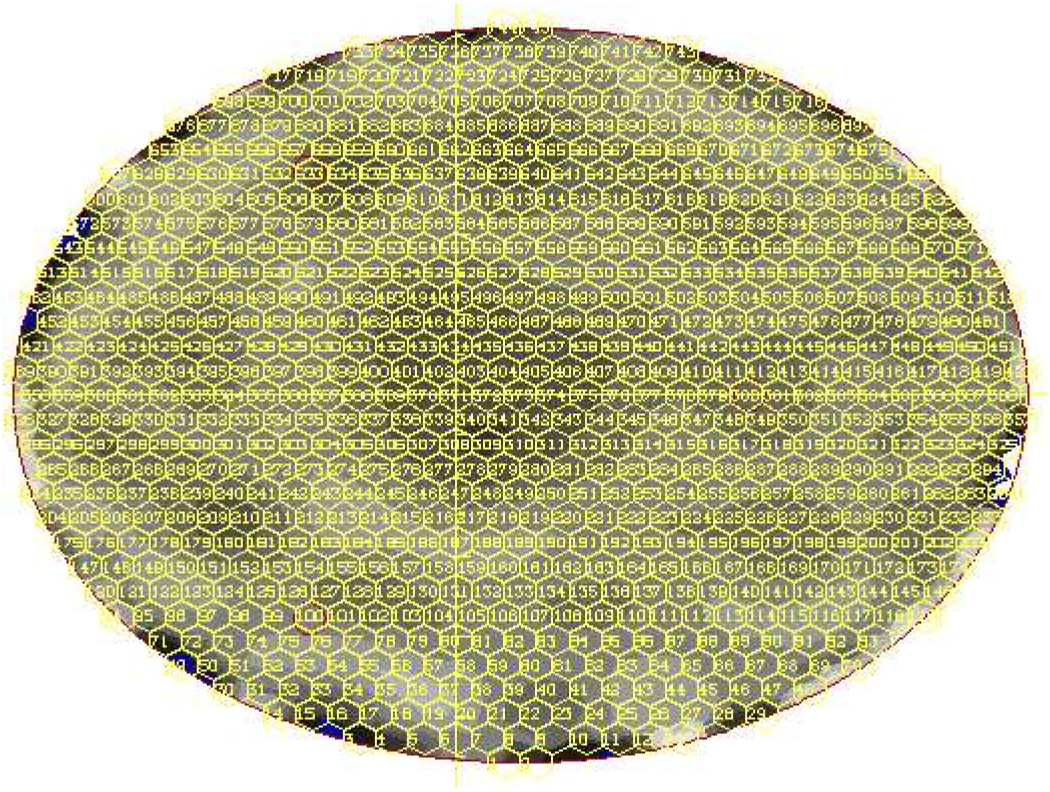


Figure 7.6: The internal cell structure of the PRFM superimposed on the Zernike polynomial construction of the surface form error map at 50 K. The plane of the map is the plane of the aperture coordinate system.

Chapter 8

The Planck telescope: reverse engineering

The material in this chapter has been extracted from the final report, [53], on the author's work on the reverse engineering of the Planck telescope; work that was conducted on behalf of the Planck Core Team and funded by PRODEX 90258-CN2 between January and September 2010. The aim of the work was to revise the pre-launch model of the Planck telescope (discussed in chapter 7) to build a model of the telescope that, so far as was possible, would return the same beam patterns, with the same pointing directions, that was being derived from the calibration measurements made on Mars and Jupiter and processed by the HFI beam analysis team. With that achieved the models could be used by the telescope builders and ESA to derive knowledge of the telescope optics as they were at L2, and to compare what had been built with what had been designed.

8.1 Methodology

A series of models of the telescope were built incorporating all thirty two polarised and twelve unpolarised single-mode beams of the HFI. All models were in conformance with (a) the predicted in-flight reflector shapes and surface form errors (see chapter 7), and (b) mechanical structure of the telescope as measured: [67] and [68] and the relative positions of the reflectors given in [66] and [20] and the knowledge of the telescope focal plane assembly reported in [21]. Preliminary analysis of the models was conducted at the centre frequency of each band; detailed analysis was conducted with broad-band models spanning the full spectrum of each pixel in sixteen or seventeen frequencies per band. (Since the study was conducted using only the single-mode channels there was no change in beam structure across the band and the large numbers of frequencies needed for multi-mode channels were not required.) With the aim being to produce a best estimate of the telescope configuration as it was at L2, the defining parameters of all models were constrained to lie within a feasible domain determined by realistic engineering considerations: measurement error analysis and finite element models of the telescope structure under in-flight conditions provided by industry, [63]. The parameters adjusted in the models where the conic constants and curvatures of both reflectors, their relative positions and tilts, the position of the focal plane assembly relative to the reflectors, and the rotation of the focal plane assembly. The analysis was concentrated upon matching both the HFI focal plane map and the beam shapes that had been obtained by the HFI team from the scans of Mars and Jupiter at the time of the contract.

The modelled beam power patterns for all single-mode horns were analysed to obtain centroids and best fit elliptical Gaussians, and those parameters compared with the tabulated measurements from the two issued data sets (Mars_DataV32, Mars_DataV41 and the equivalent sets for Jupiter) available at the time when the work was undertaken. Between the Mars and Jupiter sets there was essentially no difference, so only the data derived from observations of Mars is reported below. The result was a range of closely related models that were mechanically realistic and give good fits to the latest measurement data sets for all HFI pixels in the 100 GHz to 353 GHz range. Overall it appeared that the optics were within the measurement error estimates and that the only essential mechanical change was the displacement of the focal plane by 0.5 mm towards the secondary reflector. The results are described, qualified, and tabulated in what follows.

Good overall agreement was attained between the HFI focal plane distortion, individual beam centroids, and beast fit elliptical Gaussian beam size. The discrepancies between in-scan measurement and model beam centroids in the data set V32 was nearly twice as bad as for the data set V41, suggesting that the pixel timing parameters had improved, though it remains large for some pixels (see figure 8.3, page 211). In-scan, where the uncertainties in timing are significant, the global model/measurement centroid fit error had mean -0.0907″ and standard deviation 9.2624″, cross-scan, where the uncertainty in the timing has no effect, the mean and standard deviation are 0.0852″ and 3.3256″ respectively; see table 8.2 and figure 8.3.

One of the complications that the measurement data presented in the attempt to get ‘perfect’ agreement between model and measurement was the offsets between orthogonal polarisations ()referred to as A and B polarisations) in some pixels. Broad band physical optics modelling of the system indicated that a perfect horn would exhibit centroid offsets induced by the optics between orthogonally polarised beams at the sub-arc second level. There were cross-scan polarisation centroid offsets as large as 8.8″, and these appear to be real (present in both measurement sets). There is a reasonable physical explanation for how such an offset can arise in a single-mode pixel that is discussed in §8.3.1.

Since no measurement is exact there is no unique solution to the reverse engineering problem. Several similar models were build, run and analysed, all were mechanically feasible, and all showed very similar agreement with the measurements. That meant that the optical design was stable to small perturbations (was reasonably tolerant of manufacturing error). At the time of the contract, because many aspects of the pixel performance were not fully understood and no measurement error estimates were available, there was no way to discriminate between them and to decide which was the “best” model. As confidence in the beam centroid positions and shapes grows, the models could be returned to and tuned. With the then available V41 data set it appeared that the models could be improved very slightly by a very small tilt to the focal plane, of the order 1/100°, but the models were in good overall agreement with measurement. Sources of modelling error (where the term 'error' is used in the scientific sense of uncertainty) are discussed in sections 8.2.1 to 8.2.3 below.

8.2 The telescope model

The Planck telescope is a fairly conventional unobscured, off-axis, quasi-Gregorian, but used wide field with the focal plane position below optic axis of the telescope (intersecting the focal surface approximately at the top of the focal plane layout diagram shown in figure 8.1 on page 206). The nominal paraxial focal length of the system was 1600 mm and, prior to this study, the configuration of the telescope was described by the following position and axis orientation, measured relative the geometric centre of the focal plane assembly (see figure 8.1 on page 206) and the radius of curvature and conic constant:

Primary reflector vertex coordinates:	
origin	x: -69.2382 mm, y: -0.056 mm, z: -136.9882 mm
x_axis	x: 0.8534851114, y: -0.3542000047E-03, z: -0.521117107
y_axis	x: 0.3449000117E-03, y: 0.9999999339, z: -0.1148000039E-03
R	1439.314 mm
conic	-0.867275
Secondary reflector vertex coordinates:	
origin	x: 474.4054 mm, y: 0.0107 mm, z: 914.1936 mm
x_axis	x: 0.9313259733, y: -0.3635999896E-03, z: -0.3641864896
y_axis	x: 0.3010999815E-03, y: 0.9999999286, z: -0.228499986E-03
R	643.977 mm
conic	-0.215104
Line of sight:	
x_axis	x: 0.996194693807249, y: 0.0, z: -8.715579171967110E-002
y_axis	x: 0.0, y: 1.0, z: 0.0

The reverse engineering of the telescope is being approached with the working assumption that what was least certain about the telescope construction before launch was most likely to be incorrect in the pre-flight model. That meant that particular attention had to be given to the curvature and conic constant of the two reflectors since those four parameters had the dominant influence on the optical performance of the telescope. The remaining constructional parameters – those determined by the telescope support structure – were easier to measure and to model and were therefore known with much greater confidence, and had a weaker influence on the optical performance when adjusted within the limits of uncertainty. The reflector shapes, the focal plane tilt, the focal-plane to secondary reflector spacing, secondary to primary spacing, and the lateral displacement of the two reflectors were all adjusted in the investigation of the model. The surface form error maps, though included in the models, have only a minimal influence on the beam centroid and shape since, as shown in section 7.4, the form errors contribute less than $\lambda/7$ RMS wavefront error whereas beam distortion and aberration are significant in any two reflector optical telescope when used for wide field imaging, as is Planck.

There were four issues of prime concern in the reverse engineering of the telescope: (i) that the individual beam centroid coordinates predicted by the model should agree with with the coordinates derived from the planetary scans to within the limits on the accuracy of the measurements, and thus (ii) that the overall optical distortion maps from measurement and model should agree; (iii) that the modelled and measured beam widths should agree to within measurement error when analysed in the same way; and (iv) that the reconstructed telescope should agree in *every* respect to within the error bounds on the measured dimensions, component locations, tilts, reflector curvatures, and so on, of every component in the telescope as it was, in flight, at L2.

At the time when this work was undertaken there were three main obstacles to realising these four goals: (i) there were currently no estimates of the in-scan centroid measurement errors for each pixel; (ii) the coalignment of the star tracker and the telescope was not known to within $108''$, and (iii) the understanding of the bolometer response was in a state of flux; consequently the beam measurements were not fixed and certain. The first of these obstacles means that the in-scan beam centroids did not give as reliable an indication of the model's validity as the cross-scan measurements. However, since measurement error in the in-scan direction had no effect on the cross-scan centroid coordinate, the cross-scan coordinates of the beams give a reliable measure of agreement between optical distortion of model and telescope. The second obstacle had no influence on the optical validity of the model; it simply means that, once the mean of all centroid coordinates for model and measurements agree to within $108''$, nothing more can be known. The third issue was a serious handicap to progress.

In the search for agreement between the model and the measurements of the beam centroid locations derived from the scans of the planets (data sets Mars_DataV41, Mars_DataV32 and Jupiter_DataV32) the optical parameters of the models – the curvatures and conics, and the relative positions and tilt angles of the reflectors and the focal plane assembly – were progressively adjusted. After each modification the GRASP model was run, the far field beam pattern derived for all polarisations and pixels, and the results compared with the planet scan measurements. Throughout the modelling the values for curvatures and conics, and the secondary parameters of mirror offset, were constrained to vary within the error bounds on the measured values with the aim of attaining agreement between measurement and model for focal plane distortion and beam sizes. Component displacements within the ranges $450\,\mu$m – $500\,\mu$m for the focal plane assembly towards the SRFM, up to $100\,\mu$m for the SRFM displacement towards the focal plane, and up to $350\,\mu$m for the PRFM towards the SRFM were found to give improved conformance to measurement while remaining within the limits on the uncertainties on the relative positions of the mirrors derived form [20]. It would have been preferable to use the raw metrology data for the cold telescope structure to further constrain the models, but the data were not available during the contract. Note that there is only very minor variation in performance between models with (a) the focal plane displaced by $500\,\mu$m and no other changes, and (b) the complete set of changes; this reflects the configuration tolerance of the optical system. Once estimates of on-sky centroiding errors become available it might become possible to make a choice between models; currently it is not possible.

In the remainder of this chapter five models are referred to, labelled A through to E. Most tabulated data refers to model A, though it is debatable which of models A, B or E are in best overall agreement with the measurements. In these models both the primary reflector and the secondary reflector are within the estimated uncertainty in the mirror shape. The relative locations of the focal plane assembly and the two reflectors are also within the tolerances in position than can be derived from the ThalesAlenia Space documents [66] and [20].

Because of the complexity and computational time involved in the modelling of the multi-mode channels they were not used in the reverse engineering of the telescope. Furthermore, knowledge of the multi-mode pixel bolometer properties was less developed than for the single-mode pixels, and there were greater problems in centroiding planet scans. Consequently, since both the centroid coordinates and the beam sizes were less reliable than for the single-mode channels, it seemed sensible to ignore the multi-mode channels until such time as both measurement and model attained their final state.

8.2.1 Sources of model error: reflector shape and form error

The shapes of the two reflectors in the pre-flight model were derived both directly by linear extrapolation (regression) to operating temperature of the curvatures and conics derived from analysis of the videogrammetry measurements made at a range of temperatures down to 95 K, and by linear extrapolation of the semi axis lengths followed by derivation of the conic and curvature. For the SRFM there are also interferograms covering most of the surface. Interferometric measurements were made by Robert Daddato and CSL at a range of temperatures down to 45 K. From the analysis of the videogrammetry data it was clear that at no stage in either the cool down or warm up cycles were the reflectors at thermal equilibrium, and this lead to the uncertainties in the reflector shapes derived by extrapolation to the estimated operating temperatures that are given in this report.

Table 8.1 gives the pre-flight videogrammetry estimates of the radius of curvature and conic of both reflectors along with the interferometry estimates for the secondary reflector. The values are compared with the values used in the reverse engineering model A of the telescope.

Note that the reason for the slightly better agreement between the quoted videogrammetry estimates of the reflector shapes and this in-flight model compared with the interferometry estimates is that the interferometry measurements were made over a smaller area than the videogrammetry, and that biased the result (see Chapter 7 and figure 7.5). The effect of inclusion or removal, or even inversion, of the surface error maps on the modelled beam centroid coordinates is much smaller than that due to the indeterminacy, $\pm 20\,\mu$m, of the horn positions in the focal plane.

Primary reflector radius of curvature and conic				
	Pre-flight estimate	Error bound estimate	Model	Difference
R (mm)	1439.314	± 0.2	1439.42	0.106
k	-0.867275	$\pm 5.0 \times 10^{-5}$	-0.867266	9.0×10^{-6}
Secondary reflector radius of curvature and conic				
	Pre-flight estimate	Error bound estimates	Model	Difference
R (mm)	643.977	± 0.1	644.075	0.098
k	-0.215104	$\pm 5.0 \times 10^{-5}$	-0.215102	-2.0×10^{-6}
Secondary reflector interferometry estimates				
R (mm)	643.972	–	644.60	0.628
k	-0.215424	–	-0.21510	-3.24×10^{-4}

Table 8.1: Pre-flight radius of curvature and conic constant estimates for the PRFM and SRFM, the error bars on the estimates, the current values used in model A of the telescope, and the differences between pre-flight estimates and provisional model. In addition, the estimates derived from the interferometric measurements of the SRFM are given. Curvature and conic estimates are derived are part of the linear regression analysis output.

8.2.2 Sources of model error: focal plane metrology

The measurement of the relative positions and pointing directions for the horns in the HFI focal plane assembly were made at room temperature and documented in [21]. The documented uncertainty in the nominal relative phase centre position is $\pm 20\,\mu$m in the transverse direction. Axial position error was not recorded, but it is reasonable to assume that it is not critical for the single-mode pixels because they have a confocal distance of the order 11 mm or greater. There must be some change in the inter-focal spacing error upon cooling, but it is unknown. Finally the metrology was not accurate enough to determine individual pixel pointing, but the telescope images the phase centre onto the sky and the pointing error is of secondary importance. Simulations suggest that the pixel pointing error can be ignored in the model.

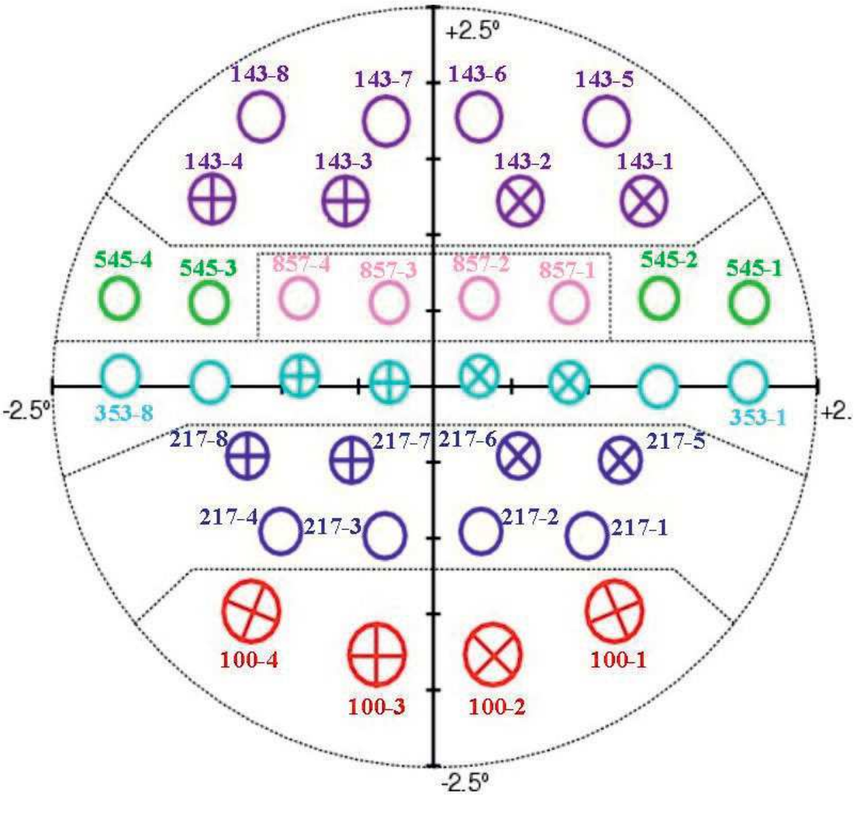


Figure 8.1: The layout of the HFI focal plane. The horns that are polarised are represented by cicles with crosses, the orientation of the two polarisations being indicated by the orientation of the arms of the cross. Horns that are unpolarised are represented by circles without croses.

No attempt was been made to use the freedom of the metrology tolerances to adjust the model focal plane to attain closer agreement between measurement though, with a design focal length of 1600 mm giving a plate scale of 129 arc seconds per millimetre, the $\pm 20\,\mu$m positional uncertainty for the pixels gives an angular uncertainty for the beam on the sky of $2.58''$. This matter is touched upon again in §8.3.

8.2.3 Sources of model error: pixel modelling and source spectrum

The models of the fields radiated by the pixels that were used in modelling the beam patterns assume a perfect black body response. In reality the response is biased by the bolometer and cavity design and by the spectral transmission of the filter stacks. No two filter stacks or bolometers are identical in their spectral response, so no two nominally identical pixels will give precisely the same beam. If the spectral response of each of the bolometers was known the modelled beams could be weighted accordingly. Furthermore, the models assumed that the measurements were made on a black-body.

The effect of these simplifications is an uncertainty in the centroid of the beam and in its FWHM. The GRASP9 models show a small spectral drift in the beam centroid of the order $1''$ as the frequency is scanned across the band of an individual pixel, and the overall spectral response and source spectrum of the pixel would modify the measured beam width. However Mars, with an angular extent of approximately $7''$ as seen from L2, is a reasonable approximation to a point-like black body, and if the measured 545 GHz and 857 GHz pixel spectral is fairly typical, the net effect on the beam width will be small. These effects are, however, unquantified.

8.3 Beam centroids: the focal plane map

Figure 8.2 illustrates the coincidence of the measured and modelled beam centroids for all single-mode horns in the HFI at the central frequency (model A). The plot illustrates the generally good overall agreement between the measurements and the model. The data set used for the plot is Mars_DatV41 for which the centroids and the model off-sets are tabulated in Table 8.5, page 225. The data set Mars_DatV32 and Jupiter_DatV32 has also been checked (the agreement between these two data sets is at the fraction of an arc second level). With the issue of data set V41 there was a marked improvement in the agreement between the model and the measured beam centroids in the in-scan direction due, not to changes in the model, but in the data reduction process. Very little changed in the cross-scan direction, but in the in-scan direction there was almost a halving of the peak-to-peak offset error.

Figure 8.3-(b) shows the difference between the measurements and the models for both polarisations of all pixels, data set V41. This plot is particularly informative because it shows that while the cross-scan focal plane map agrees to within $\pm 8''$ with a standard deviation of $3.33''$ (essentially unchanged from the V32 data set fit plotted in figure 8.3-(a)), the scatter in the in-scan measured beams centroids gives a peak-peak fitting error of $13.8''$ to $-19.1''$ with a standard deviation of $9.2624''$. The in-scan result for the V32 data set was $+42''$ to $-27''$ with a standard deviation of $18''$; see also table 8.2.

The error bars in Figure 8.3 are estimates of the total model errors: room temperature pixel-pixel focal plane metrology error $20\,\mu$m increased by cooling to $30\,\mu$m giving $2.2''$ on the sky, spectral drift contribution of $0.5''$, and modelled field centroiding tolerance $1''$. This last tolerance is subject to the choice in thresholding of the modelled far field. The plot shows error bars of $3.7''$ applied to both in-scan and cross-scan directions, but see the comments in section 8.5.

No centroiding errors in the measured beams were taken into account because no measurement error estimates were available. Furthermore, the measured (and probably real) pointing offsets between A and B polarisations in some pixels (tabulated and discussed in subsection 8.3.1) make the comparison between model and measurement below the $6''$ level difficult for some pixels. The statistics of the beam centroid differences are given in table 8.2 for all models, A, B, C, D and E.

The cross-scan difference are all less than $8.3''$. The in-scan differences are greater. The scatter in the in-scan offsets cannot be an optical phenomenon because of the complete the overlap of the beams on the primary, and extensive overlap of the secondary reflector, and so must be due to uncertainties in the timing. In-scan uncertainty does not influence the cross-scan offsets. Error bars are as described above.

Figure 8.4 illustrates the variation in the measured/modelled beam centroid offsets for four further mechanically realistic models. All models have the relative displacements of focal plane and the secondary and primary reflectors constrained to within $500\,\mu$m. The small differences in configuration show different offsets for individual beams, but the *overall* pattern is consistent across all models. (The measurement data set used for reference in *all* analysis was V41, the data taken from the table on Antoine Chamballu's wiki page of 15^{th} July 2010.) The same pattern is found in models that extend outside the range of mechanically plausibly models, such as were described in the Phase 1, 2 and 3 technical report presented to ESTEC prior to the Phase 4 report that forms the basis for the material in this chapter.

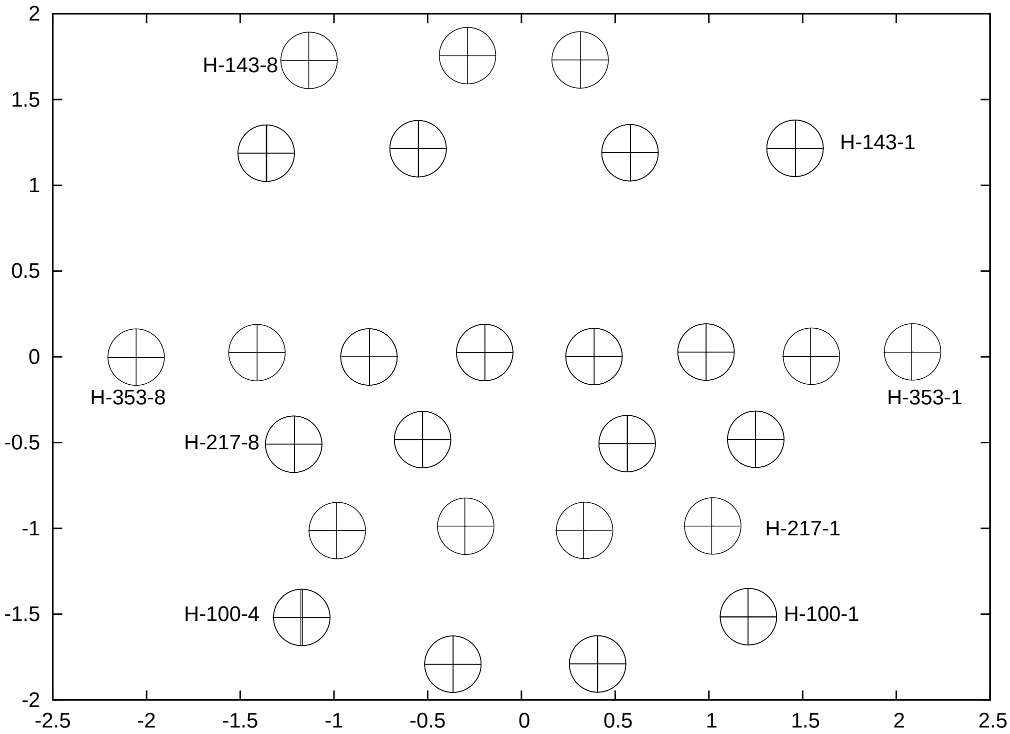
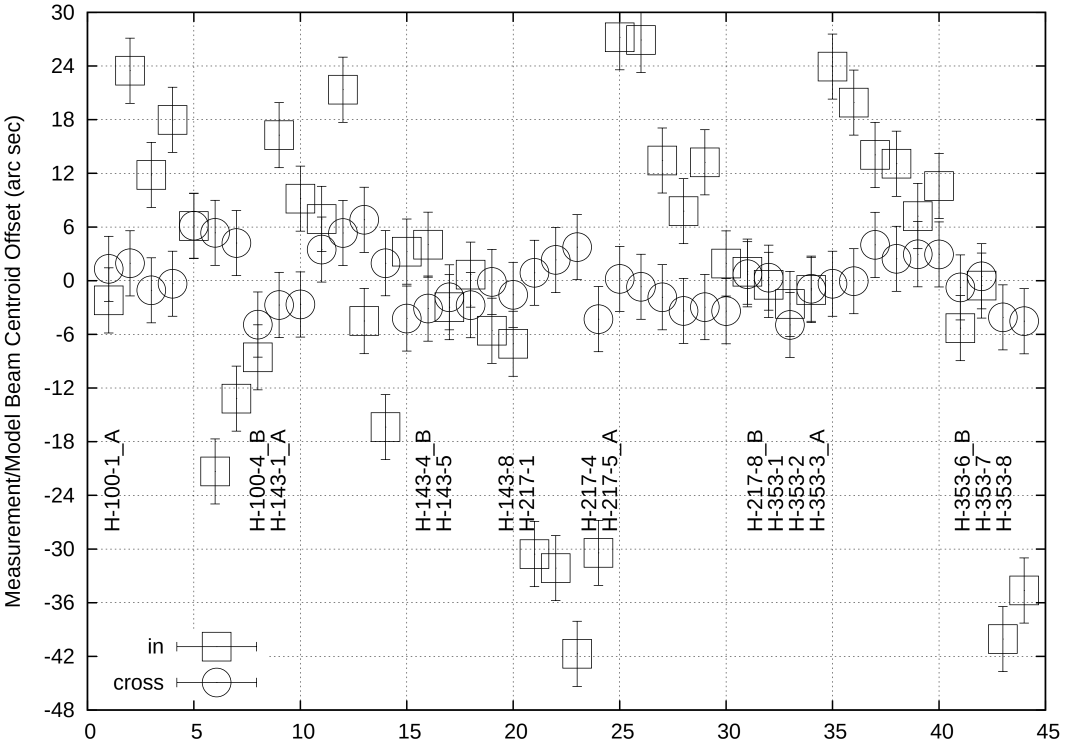


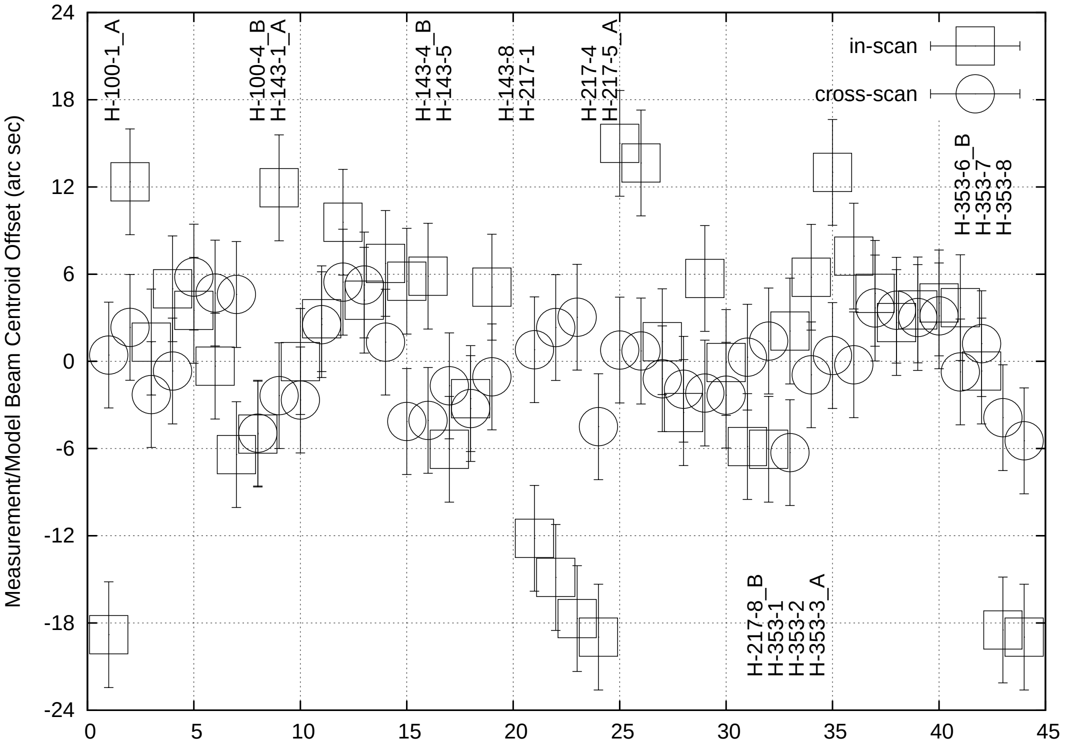
Figure 8.2: The HFI single-mode pixel focal plane map for Model A: coincidence of measured and modelled beam centroid locations. **Key:** + measured on Mars, data set Mars_DataV41, ○ modelled broad band. Axis scales are in degrees offset from the nominal LOS. For the 32 polarised beams the centroid of both polarisations are plotted, resulting in double crosses and circles. The 545 GHz and 857 GHz beams are not included in the analysis, and datum H-143-5 is not in the Mars_DataV41 table. Note that the orientation of the plus signs bares no relation to the polarisation angle, they are purely positional markers.

Model	Scan direction	Maximum	Minimum	Mean	Std. DataV.	Variance
A	in-scan	13.7954	-19.0594	-0.0907	9.2624	85.7912
	X-scan	6.9941	-8.2329	0.0852	3.3256	11.0597
B	In-scan:	15.8325	-21.0495	-0.1115	9.9208	98.4221
	X-scan:	7.2459	-9.6201	-0.3316	3.8292	14.6628
C	In-scan:	16.0917	-21.3015	-0.1605	9.6827	93.7546
	X-scan:	7.5414	-8.8822	-0.3271	3.7148	13.8000
D	In-scan:	15.5026	-20.3929	-0.1482	9.4922	90.1018
	X-scan:	7.5190	-9.1818	-0.3783	3.7695	14.2090
E	In-scan:	14.0551	-18.4469	-0.1399	9.0120	81.2155
	X-scan:	6.5805	-7.3433	0.0897	3.1857	10.1489

Table 8.2: Fitting statistics for the differences between the measured and modelled beam centroids for all 40 single-mode beams. Units: arc seconds. For model A the data are listed in Table 8.5 and plotted in figure 8.3 (b). For models B to E the in-scan and cross-scan centroid offsets are plotted in figures 8.4 (a) to (d), in the same order. All dimensions are arc seconds.

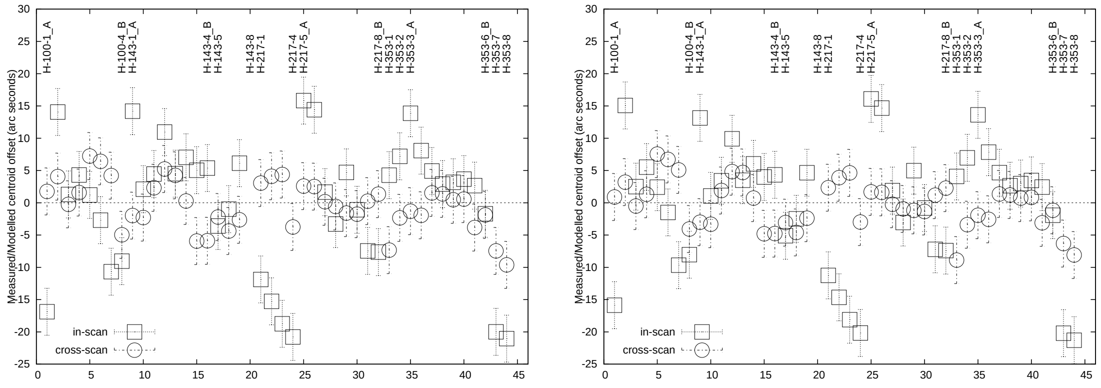


(a) Measurement data set Mars_DataV32, July 2010

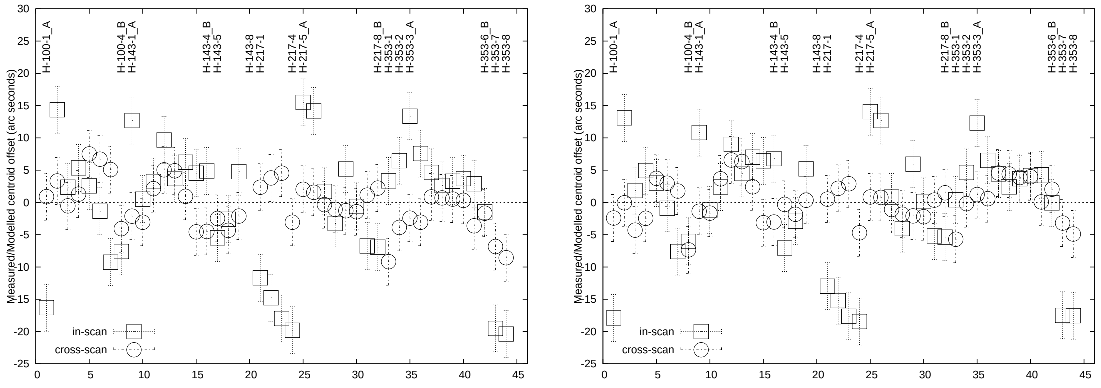


(b) Measurement data set Mars_DataV41, September 2010

Figure 8.3: Measured-modelled beam centroid differences in arc seconds, using measurement data from (a) V32 , and (b) V41. Model A. The with the issue of the V41 data set the differences between the modelled and measured in-scan offsets approximately halved; an improvement due to improvement in the instrument calibration.



(a) Model B (b) Model C



(c) Model D (d) Model E

Figure 8.4: Plots of the measured-modelled beam centroid offsets for four different telescope configurations. All models show the same overall pattern of fit between measured and modelled beam centroids. Model E shows marginally better fit statistics than the other models; see table 8.2.

8.3.1 Polarisation pointing offsets

Table 8.3 shows the measured in-scan and cross-scan pointing offsets between polarisations A and B in all thirty two polarised pixels, for data set V41. For the in-scan set there are large discursions from the sub arc second values that would be expected from the effects of the optics on the beam from a perfect pixel. These gross pointing offsets, $31''$ in the case of the 100 GHz pixel H-100-1, which has a nominal FWHM of $10'$, are probably largely the result of uncertainties in the time-line for the pixels, an hypothesis supported by the much greater offsets observed in data set V32 than in V41.

The measured cross-scan offsets are generally slightly larger than, but broadly in line with, those predicted by the models, showing offsets between measured beam

Table 8.3: Measured and modelled pointing offsets between the two orthogonal polarisations of the beam in all of the polarised horns, model A, broad band; units: arc seconds. The anomalous measured in-scan offsets can probably be attributed to timing error; the anomalous measured cross-scan offsets are more likely to be real.

Horn	in-scan		cross-scan	
	Measured	Modelled	Measured	Modelled
H-100-1	30.9149	0.0756	2.6591	0.3024
H-100-2	2.7081	0.3816	1.9725	0.1404
H-100-3	3.9593	0.0468	0.5257	0.2340
H-100-4	1.6829	0.1008	8.8182	0.2952
H-143-1	12.3692	0.3852	0.1800	0.1116
H-143-2	6.1458	0.4608	2.9821	0.0540
H-143-3	2.5072	0.2880	5.0643	1.1628
H-143-4	0.2377	0.1080	1.1466	1.2636
H-217-5	1.5756	0.2160	0.1581	0.2196
H-217-6	5.1366	0.2448	0.6197	0.1044
H-217-7	5.8283	0.0396	0.3409	0.2232
H-217-8	0.3016	0.1188	0.8543	0.2736
H-353-3	5.8826	0.1080	0.6581	0.0216
H-353-4	2.1304	0.1188	0.1916	0.0324
H-353-5	0.5062	0.0180	0.0781	0.0432
H-353-6	4.3611	0.0000	1.9030	0.0540

centroids for orthogonal polarisation in a single pixel of the order $1''$. However, pixel H-100-4 shows a measured $8.8''$ offset between orthogonal polarisations, and pixel H-143-3 a $5.1''$ offset. These may represent calibration errors, or they may be real offsets. They are fairly consistent between data sets V32 and V41.

Assuming that these offsets are real, an explanation needs to be given. If the two polarisations from a single horn are pointing in different directions the two polarised fields in the horn aperture must be asymmetrical and different. For that to be the case the radiating horn has to be supporting modes other than those that would be induced by the fundamental. Since the throat of the back-to-back horn acts as a mode filter and will not support higher order modes it means that the higher order modes that induce asymmetry must be induced in the radiating horn after the throat. That would occur if there is an obstruction in a corrugation or a damaged corrugation. In particular, if there is a trace of the mandrel upon which the horn was electroformed, the corrugation would not have axial symmetry and the result would be that the modes scattering out of the throat scatter into modes that would not be present in a horn with perfect axial symmetry. The remainder of the horn also acts as a weak mode filter, but a residual asymmetry in the aperture field can result.

Necessarily this is very speculative. It would not possible to prove that this is what was happening without retrieving the horns, but the fact that there is a measured pointing offset between two polarisation in a single horn places further limits on the agreement that can be expected between measurement and model. To put this in perspective, $8.8''$ corresponds to the centres of the A and B polarisations being offset by approximately 65 μm in a horn aperture of 14339 μm – a relative error of 0.45%. The presence of higher order modes in the radiated field would change the beam profile; but clearly the asymmetries are very weak, so the change in shape would be hard to detect.

8.4 Elliptical Gaussian fit FWHM for measured and modelled beams

Table 8.10 shows the results of elliptical Gaussian fitting to model A, the fitted FWHM for model B, and the measured FWHM for all beams from the data set Mars_DatV32 and Mars_DatV41, and the values obtained from these models are representative of those attained by all models. The exact results obtained by elliptical Gaussian fitting depend upon the algorithm used and the data thresholding. Typically, choosing a threshold of -10 dB returns a marginally narrower beam than a higher threshold (say -5 dB). The only truly valid comparison between measurements and models would be obtained by using the same algorithm and identical sampling and thresholding on the modelled fields as on the measurements. This was not possible because access to the data and to the data processing pipeline was restricted to the data processing team members.

For the fitting reported here two optimising power pattern fitting routines were written, one using the long established Simplex parameter space search method, the second using a linear quasi-Newton method; both searched the same parameter space (ellipse axis lengths, peak value, centroid coordinates and ellipse rotation angle) and returned almost identical results. The Simplex method is well known and an account can be found in [24] and code in [57], and will not be summarised here. The Newton methods make a quadratic approximation to the function being sought; a valid assumption in the case of a Gaussian power distribution. Thus the Gaussian power difference distribution, $G(\boldsymbol{x})$, is approximated by some quadratic $q(\boldsymbol{x})$ and, if $\boldsymbol{x}_*$ is the point in the parameter space at which the difference between the Gaussian and the data is a minimum, then at a point $\boldsymbol{x}^{(k)}$ near $\boldsymbol{x}_*$, the local quadratic approximation is $q(\boldsymbol{x}^{(k)} + \boldsymbol{\delta}) = G(\boldsymbol{x}^{(k)}) + \boldsymbol{g}^{(k)^T}\,\boldsymbol{\delta} + \frac{1}{2}\,\boldsymbol{\delta}^T\,G^{(k)}\,\boldsymbol{\delta}$, where $\boldsymbol{\delta} = \boldsymbol{x}_* - \boldsymbol{x}^{(k)}$, $\boldsymbol{g}^{(k)} = \nabla\,G(\boldsymbol{x}^{(k)})$, and $G^{(k)}$ is the Hessian of G at $\boldsymbol{x}^{(k)}$. Quasi-Newton methods make a positive definite approximation $H^{(k)} \approx G^{(k)}$ at each iteration, set a search direction $\boldsymbol{s}^{(k)} = -H^{(k)}\,\boldsymbol{g}^{(k)}$, find a new $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha^{(k)}\,\boldsymbol{s}^{(k)}$ for some scalar $\alpha^{(k)}$, and update the approximation to the Hessian to give $H^{(k+1)}$. Mathematical details can be found in [24].

Whatever method is used, the aim is to minimise the absolute difference in the measured or modelled power distribution and the test function. There are a couple of things that need to be considered. The first is that the 143 GHz beams are close to the optic axis of the telescope and will therefore exhibit little distortion or other aberration. The beams should therefore be found to be virtually circular, and indeed they are (see figure 8.8). The rotation angle is then redundant and the code needs to be able to handle the resulting degeneracy. More importantly, the beams are not all that Gaussian, particularly the 100 GHz beams that are furthest from the optic axis and will therefore show greatest distortion, and the outer 353 GHz beams that are both far from the axis and of relatively high frequency so that both beam distortion and aberration are expected to be high. It would have been more informative to find the Gauss-Laguerre or Ince-Hermite spectra, both of which are conformally related to the Hermite function analysis subsequently adopted for the multi-mode beam analysis, but elliptical Gaussian fitting was a contractual requirement.

	Model label					Measured	
Pixel	(A)	(B)	(C)	(D)	(E)	V32	V41
H-100-1A	9.45656	9.65344	9.47736	9.48893	9.45327	9.48183	10.0575
H-100-1B	9.64705	9.65204	9.49570	9.49198	9.46706	9.58726	10.1906
H-100-2A	9.61013	9.63286	9.47563	9.49176	9.44088	9.43678	9.70971
H-100-2B	9.61633	9.64123	9.47729	9.50464	9.44183	9.45047	9.38593
H-100-3A	9.60933	9.63181	9.46393	9.48926	9.43608	9.42714	10.0661
H-100-3B	9.62208	9.65092	9.48993	9.50685	9.44175	9.44621	10.0345
H-100-4A	9.63927	9.65464	9.48272	9.48872	9.46435	9.55189	10.2004
H-100-4B	9.65240	9.67337	9.50277	9.50600	9.46980	9.52794	9.91908
H-143-1A	7.07470	7.14870	7.10177	7.09117	6.98235	6.91896	6.96828
H-143-1B	7.01203	7.09407	7.05280	7.04226	6.93496	6.96804	6.95189
H-143-2A	7.01910	7.09322	7.04428	7.07071	6.95027	6.88849	7.01083
H-143-2B	6.99909	7.06695	7.01599	7.02965	6.92193	6.87046	6.98437
H-143-3A	6.97355	7.03827	6.97650	6.99459	6.87561	6.97519	7.08094
H-143-3B	7.00407	7.08094	7.01766	7.03783	6.90339	6.94679	6.86819
H-143-4A	7.03171	7.09430	7.06446	7.06146	6.95817	7.08521	7.01279
H-143-4B	7.04052	7.10760	7.05533	7.06940	6.94871	7.07287	7.08531
H-143-5	7.34458	7.46722	7.44043	7.45108	7.29816	7.18205	7.15966
H-143-6	7.28515	7.40490	7.36977	7.40065	7.23510	7.17162	7.12807
H-143-7	7.26726	7.38195	7.34621	7.37534	7.20941	7.17395	7.23309
H-143-8	7.33128	7.43874	7.40407	7.42635	7.27890	7.33856	–
H-217-1	4.69155	4.76403	4.71486	4.71079	4.65484	4.65464	4.67979
H-217-2	4.69778	4.77022	4.72640	4.74574	4.64157	4.74743	4.63756
H-217-3	4.69960	4.76945	4.72463	4.73798	4.63799	4.66260	4.57269
H-217-4	4.69204	4.75906	4.71252	4.71178	4.63119	4.60960	4.64027
H-217-5A	4.66891	4.73461	4.68728	4.65760	4.60462	4.73211	4.76929
H-217-5B	4.67061	4.74166	4.69238	4.68484	4.61412	4.74743	4.71350
H-217-6A	4.69044	4.76334	4.71798	4.72689	4.62754	4.66260	4.68057
H-217-6B	4.69304	4.76590	4.72473	4.72582	4.63667	4.63459	4.62484
H-217-7A	4.68687	4.75590	4.71366	4.71735	4.63556	4.62434	4.59254
H-217-7B	4.69307	4.76365	4.71867	4.72276	4.63770	4.66547	4.62591
H-217-8A	4.67147	4.72868	4.67704	4.66601	4.61394	4.69041	4.68530
H-217-8B	4.66977	4.72820	4.67802	4.67000	4.60746	4.74138	4.75164
H-353-1	4.44536	4.39336	4.37733	4.34558	4.40915	4.52602	4.53123
H-353-2	4.36215	4.30017	4.29591	4.29056	4.34476	4.43023	4.43329
H-353-3A	4.35483	4.29230	4.30645	4.31756	4.35891	4.42912	4.41162
H-353-3B	4.36495	4.29261	4.30791	4.31616	4.35878	4.42594	4.33692
H-353-4A	4.38212	4.31364	4.32339	4.34111	4.38041	4.39701	4.37221
H-353-4B	4.38280	4.31440	4.32371	4.33688	4.37553	4.39036	4.37819
H-353-5A	4.38533	4.38693	4.31985	4.33777	4.35929	4.31478	4.41046
H-353-5B	4.37914	4.31028	4.31763	4.33241	4.36648	4.39791	4.41973
H-353-6A	4.37978	4.30063	4.30398	4.31651	4.37487	4.46048	4.35391
H-353-6B	4.37672	4.30178	4.31634	4.32363	4.36488	4.42553	4.35162
H-353-7	4.39498	4.31506	4.30958	4.30116	4.36929	4.42893	4.40631
H-353-8	4.46431	4.53607	4.37213	4.34240	4.43055	4.40301	4.48774

Table 8.4: FWHM derived from elliptical Gaussian fitting to the modelled beam in the LOS coordinate system thresholding to -10 dB for models A, B, C, D and E and the measured FWHM from Mars_DatV32 and Mars_DataV41. (See figure 8.5.)

The general agreement is good to the level of approximately 6″ or better on the FWHM for most pixels. The measured unpolarised 143 GHz show an asymmetry in beam width across the plane of symmetry of the optics. That is unexpected and is not explicable in terms of the aberrations that a quasi-Gregorian system exhibits: if the beams close to the optic axis exhibited astigmatism (due to system misalignment) it would be even more pronounced in the rest of the pixels, but no such discrepancy between measurement and modelling is evident. Furthermore, the measurements of the 143 GHz beams are mostly 9″ to 16″ narrower than the more symmetrical modelled beams for model A, but note the spread of FWHM in both measurements and models tabulated in Table 8.4. This may be a data processing artifact, or may indicate differences in performance for the four pixels. Pixel models take no account of the source spectrum, nor of the pixel-by-pixel frequency dependent filter transmission or the spectral response of the cavities, neither of which are know.

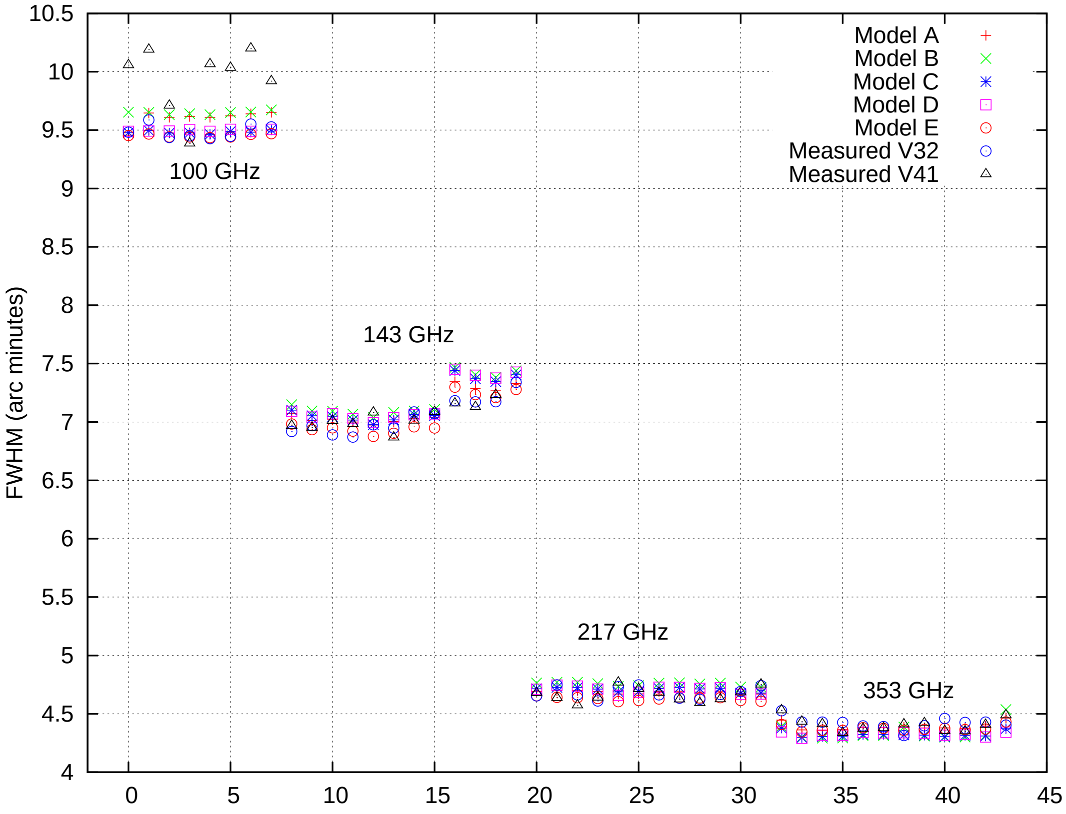
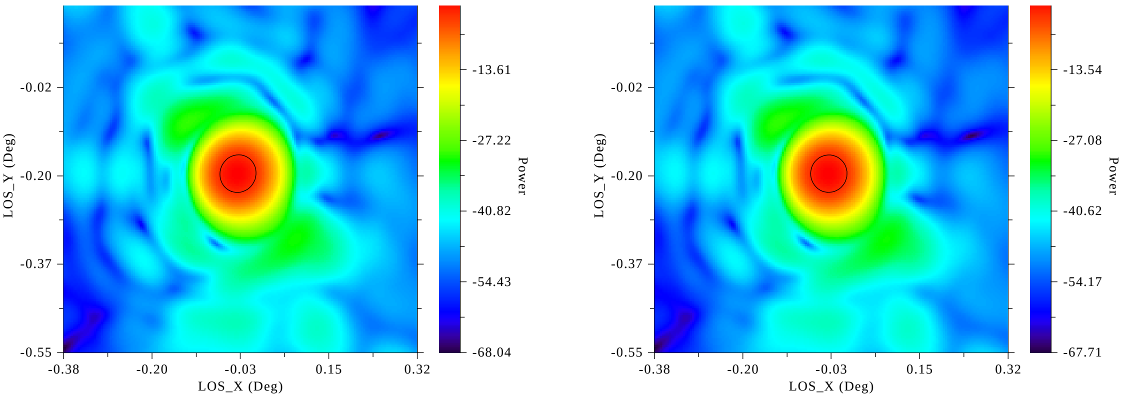
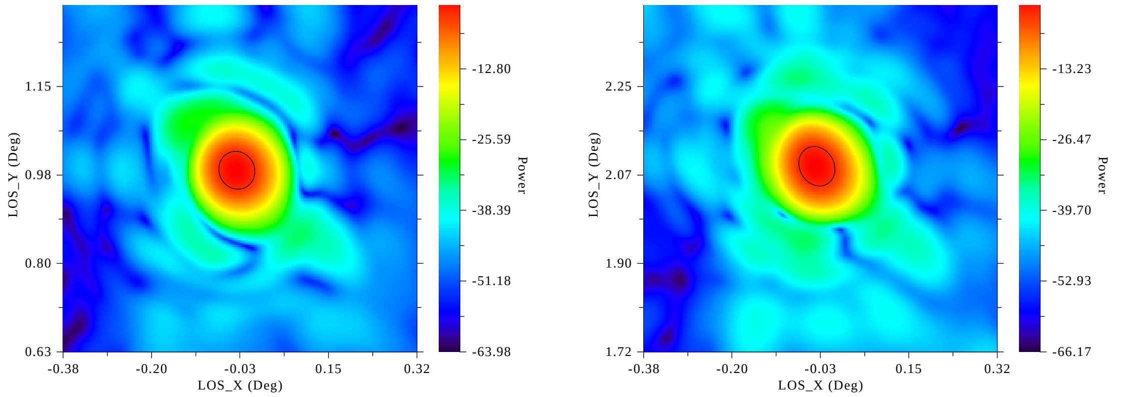


Figure 8.5: Plot of the FWHM for models A through to E and for data sets V32 and V41, as tabulated in Table 8.4. Note that the FWHM fitting to measurement set V41, 100 GHz horns appear anomalously large compared with all of the models and with the results of fitting to measurement set V32.

8.4.1 353 GHz beams



(a) H-353-4A (b) H-353-4B



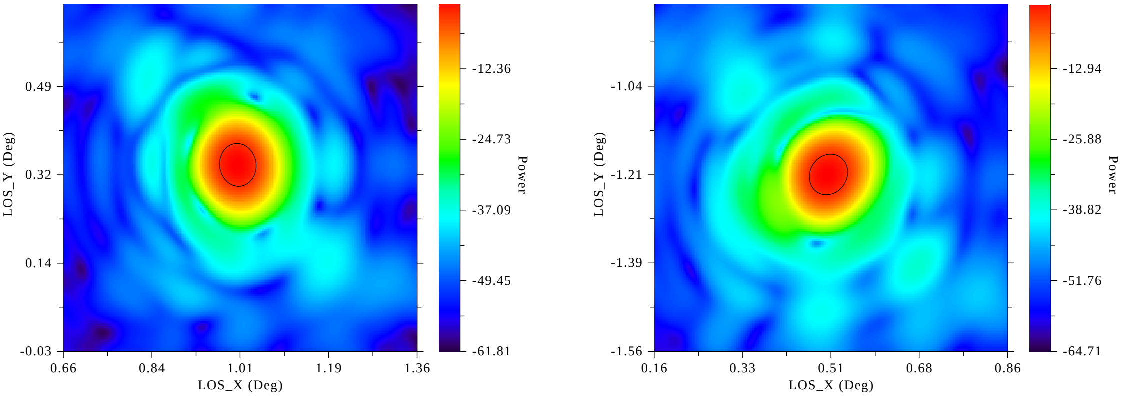
(c) H-353-6A (d) H-353-8

Figure 8.6: Normalised beam power maps (dB) for a subset of the 353 GHz pixel beams. The ellipses show the FWHM ellipses returned by elliptical Gaussian beam fitting as listed in Table 8.10. The beam models use sixteen equally spaced sample frequencies across the band.

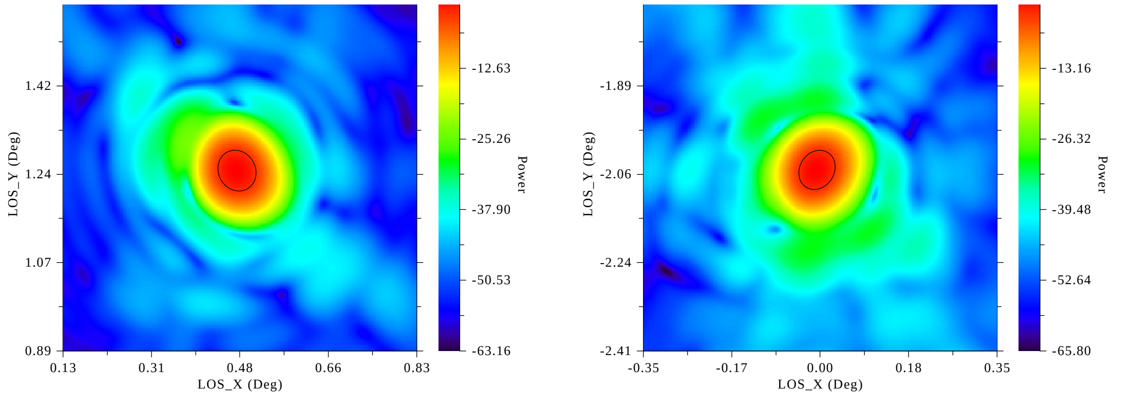
	In-scan (deg)			Cross-scan (deg)		
Pixel	Measured	Model	Diff (min)	Measured	Model	Diff (min)
H-353-1	-2.05511504	-2.054817	-0.01788255	-0.00311576	0.002814	-0.35578574
H-353-2	-1.40983655	-1.410721	0.05306723	0.02449356	0.029024	-0.27182654
H-353-3A	-0.80968414	-0.812701	0.18101185	-0.00001766	0.004245	-0.25575971
H-353-3B	-0.81131819	-0.812731	0.08476879	-0.00020046	0.004239	-0.26636763
H-353-4A	-0.19497610	-0.195808	0.04991426	0.02657979	0.030040	-0.20761287
H-353-4B	-0.19556789	-0.195841	0.01638689	0.02652657	0.030031	-0.21026588
H-353-5A	0.38806256	0.387454	0.03651374	0.00288485	0.006659	-0.22644926
H-353-5B	0.38820317	0.387459	0.04464994	0.00290654	0.006647	-0.22442753
H-353-6A	0.98577901	0.984992	0.04722035	0.02751943	0.032471	-0.29709426
H-353-6B	0.98456760	0.984992	-0.02546430	0.02804804	0.032456	-0.26447739
H-353-7	1.54258425	1.547846	-0.31570521	0.00238702	0.008365	-0.35867862
H-353-8	2.08207570	2.087346	-0.31621824	0.02738838	0.033946	-0.39345721
Difference statistics:						
In-scan:	Mean:	-0.01347810	Stdev:	0.14420351	Var:	0.02079465
Cross-scan:	Mean:	-0.27768355	Stdev:	0.05928328	Var:	0.00351451

Table 8.6: Table of in-scan and cross-scan beam centroids for the 353 GHz beams.

8.4.2 217GHz beams



(a) H-217-3 (b) H-217-5A



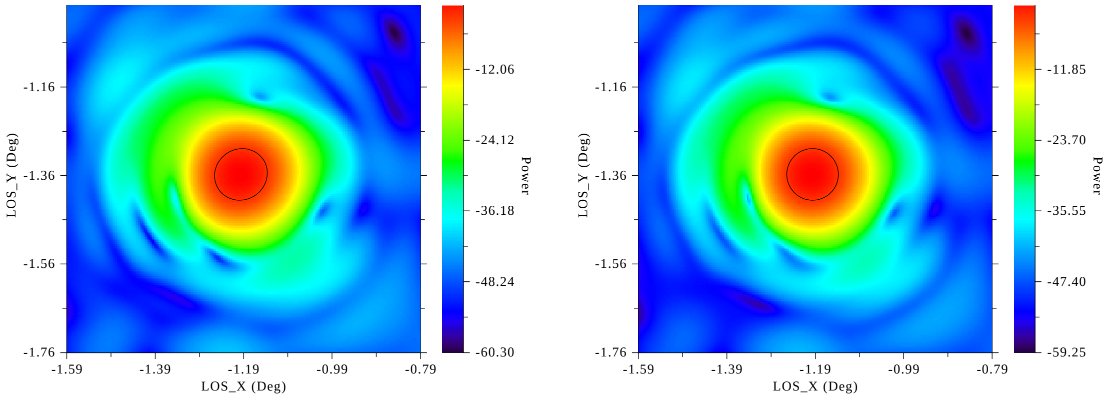
(c) H-217-8A (d) H-353-1

Figure 8.7: Normalised beam power maps (dB) for a subset of the 217 GHz pixels. The ellipses show the FWHM ellipses returned by elliptical Gaussian beam fitting as listed in Table 8.10. The beam models use sixteen equally spaced sample frequencies across the band.

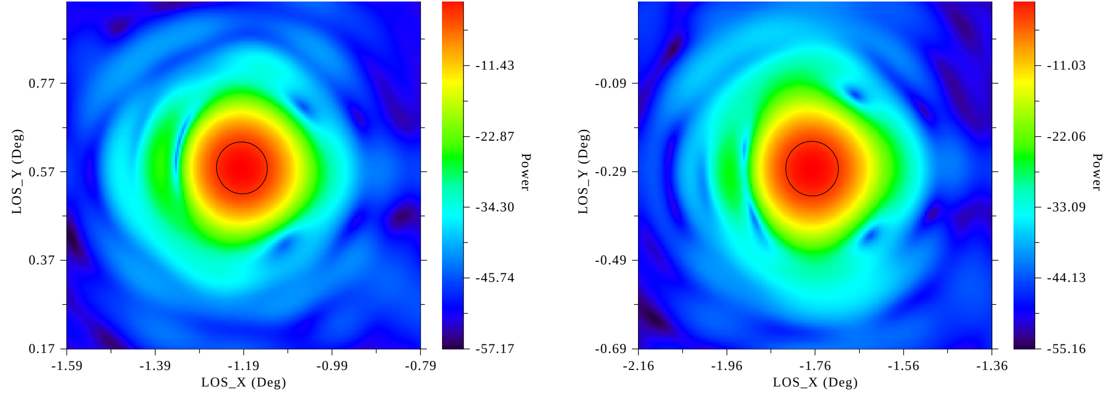
	In-scan (deg)			Cross-scan (deg)		
Pixel	Measured	Model	Diff (min)	Measured	Model	Diff (min)
H-217-1	-0.9862491	-0.982002	-0.2548265	-1.01309828	-1.008670	-0.26569673
H-217-2	-0.3016303	-0.296765	-0.2919156	-0.98726308	-0.983155	-0.24648473
H-217-3	0.3316057	0.337149	-0.3325994	-1.01150770	-1.007468	-0.24238204
H-217-4	1.0152104	1.020972	-0.3456991	-0.98724138	-0.980951	-0.37742254
H-217-5A	-1.2111753	-1.214540	0.2018806	-0.50907897	-0.504787	-0.25751828
H-217-5B	-1.2116130	-1.214600	0.1792212	-0.50903507	-0.504726	-0.25854392
H-217-6A	-0.5269182	-0.526616	-0.0181331	-0.48259759	-0.477626	-0.29829555
H-217-6B	-0.5283451	-0.526684	-0.0996631	-0.48276972	-0.477597	-0.31036319
H-217-7A	0.5657188	0.564611	0.0664708	-0.50647140	-0.501000	-0.32828416
H-217-7B	0.5640999	0.564600	-0.0300076	-0.50656608	-0.501062	-0.33024500
H-217-8A	1.2488630	1.250839	-0.1185618	-0.48078996	-0.475859	-0.29585779
H-217-8B	1.2487792	1.250806	-0.1216091	-0.48055267	-0.475935	-0.27705991
Difference statistics:						
In-scan:	Mean:	-0.09712027	Stdev:	0.17897748	Var:	0.03203294
Cross-scan:	Mean:	-0.29067949	Stdev:	0.03894866	Var:	0.00151700

Table 8.7: Table of in-scan and cross-scan beam centroids for the 217 GHz beams.

8.4.3 143GHz beams



(a) H-143-1A (b) H-143-1B



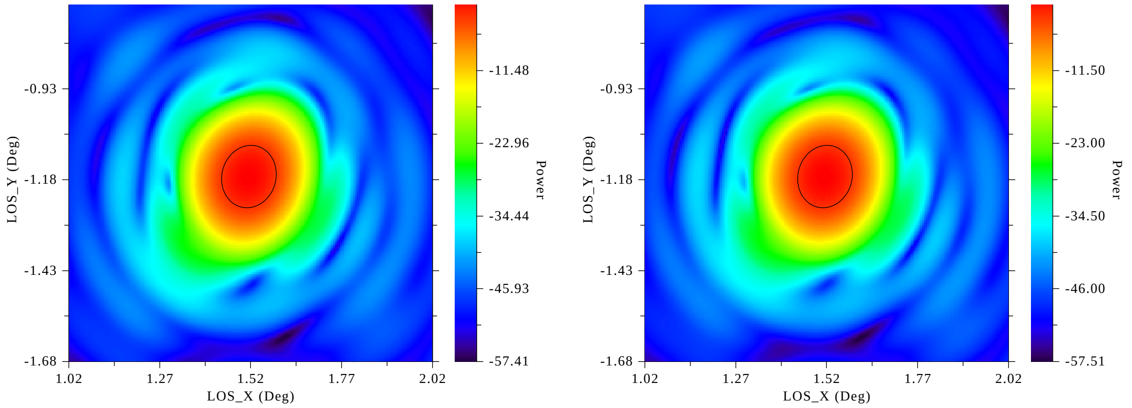
(c) H-143-3A (d) H-143-6

Figure 8.8: Normalised beam power maps (dB) for a subset of the 143 GHz pixels. The ellipses show the FWHM ellipses returned by elliptical Gaussian beam fitting as listed in Table 8.10. The beam models use seventeen equally spaced frequencies across the band.

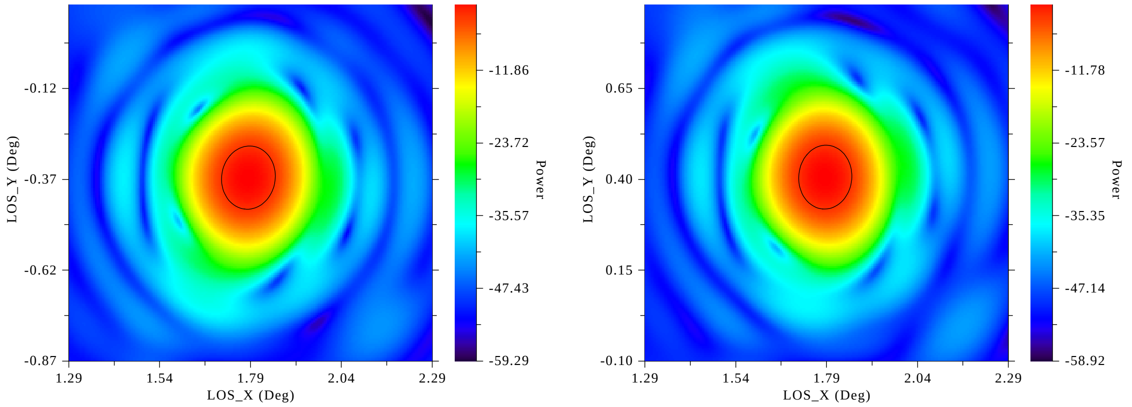
	In-scan (deg)			Cross-scan (deg)		
Pixel	Measured	Model	Diff (min)	Measured	Model	Diff (min)
H-143-1A	-1.35838761	-1.361215	0.16964341	1.18711421	1.191984	-0.29218766
H-143-1B	-1.36182350	-1.361322	-0.03009001	1.18706422	1.192015	-0.29704693
H-143-2A	-0.55013663	-0.550624	0.02924248	1.21449904	1.218144	-0.21869740
H-143-2B	-0.54842947	-0.550752	0.13935184	1.21532741	1.218159	-0.16989532
H-143-3A	0.58077395	0.579715	0.06353704	1.19186230	1.195002	-0.18838226
H-143-3B	0.58147039	0.579707	0.10580362	1.19045555	1.194679	-0.25340679
H-143-4A	1.46176415	1.460152	0.09672910	1.21450841	1.220466	-0.35745546
H-143-4B	1.46183019	1.460122	0.10249127	1.21418992	1.220115	-0.35550504
H-143-5	-1.13459442	-1.132583	-0.12068536	1.72834963	1.732967	-0.27704224
H-143-6	-0.28817926	-0.287318	-0.05167589	1.75534438	1.760570	-0.31353740
H-143-7	0.31506409	0.313679	0.08310534	1.73121013	1.735959	-0.28493197
Difference statistics:						
In-scan:	Mean:	0.04895440	Stdev:	0.08187090	Var:	0.00670284
Cross-scan:	Mean:	-0.25067404	Stdev:	0.09414494	Var:	0.00886327

Table 8.8: Table of in-scan and cross-scan beam centroids for the 143 GHz beams.

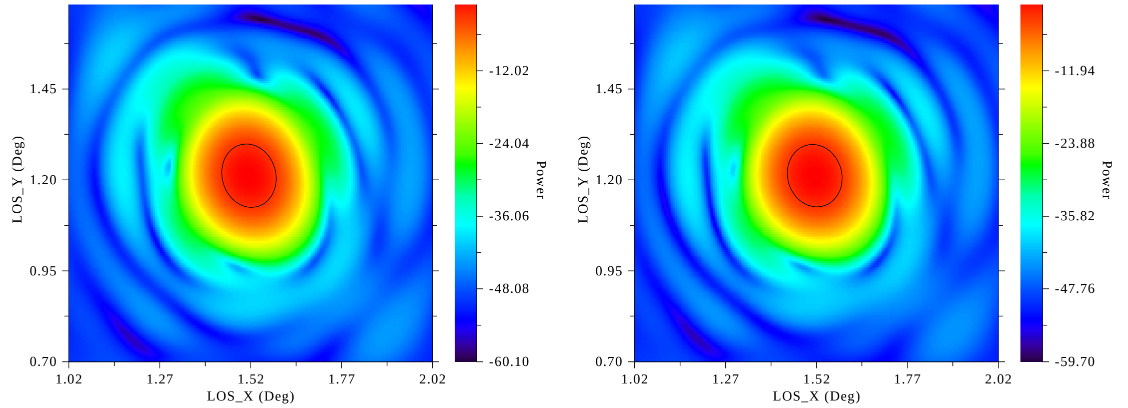
8.4.4 100GHz beams



(a) H-100-1A (b) H-100-1B



(c) H-100-2A (d) H-100-3A



(e) H-100-4A (f) H-100-4B

Figure 8.9: Normalised power maps (dB) for a subset of the 100 GHz beams. The ellipses show the FWHM ellipses returned by elliptical Gaussian beam fitting as listed in Table 8.10. The beam models use 16 equally spaced sample frequencies across the band for the 100 GHz horns. The beam LOS coordinate system in these plots and in Fig. 8.8 is the engineering LOS which differs from the measurement data table definition by the transformation $x \rightarrow y$, $y \rightarrow -x$, i.e. a 90° rotation.

Table 8.9: Tables showing the measured in-scan and cross-scan beam centroid coordinates from the Mars data table V41, the modelled centroid coordinates, and the differences. Note that the absolute pointing is not known to better than 1.8 arc minutes and the measurement error was unknown. Modelled beam centroiding depends upon the method used and data threshold imposed; the numbers tabulated impose no thresholding and are true centroids (centre of mass). Mean focal plane displacement for all beams: $-0.028'' \times -0.286''$.

	In-scan (deg)			Cross-scan (deg)		
Pixel	Measured	Model	Diff (min)	Measured	Model	Diff (min)
H-100-1A	-1.1770384	-1.171320	-0.3431057	-1.51934640	-1.514141	-0.31232377
H-100-1B	-1.1684509	-1.171341	0.1734033	-1.51860775	-1.514057	-0.27304507
H-100-2A	-0.3651289	-0.364837	-0.0175145	-1.79265167	-1.786759	-0.35356035
H-100-2B	-0.3643767	-0.364943	0.0339812	-1.79210377	-1.786720	-0.32302616
H-100-3A	0.4068292	0.406794	0.0021142	-1.78957164	-1.785752	-0.22917844
H-100-3B	0.4057294	0.406781	-0.0630939	-1.78971767	-1.785687	-0.24184042
H-100-4A	1.2085305	1.211391	-0.1716296	-1.51492380	-1.510394	-0.27178793
H-100-4B	1.2089980	1.211419	-0.1452607	-1.51737331	-1.510312	-0.42367832
Difference	statistics:					
In-scan:	Mean:	-0.06638822	Stdev:	0.14514444	Var:	0.02106691
Cross-scan:	Mean:	-0.30355506	Stdev:	0.05986102	Var:	0.00358334

8.5 Conclusions, and advancing the model

The contract to undertake the reverse engineering of the telescope was awarded, the work undertaken and the contract concluded before the beam processing had developed to the stage where the essential input – centroid location, centroid measurement error, beam FWHM and FWHM error, ellipse orientation and orientation error – was known with confidence. Steady progress to establish the input was being made while the reverse engineering was ongoing, but the work would have benefited from a delay of a couple of years. With the knowledge of the beams that exists today – shortly after the end of the satellites working life – much more could be said with certainty about the conformance of the model to the observation.

Referring to the cross-scan offsets in figure 8.3-(b), if a measurement centroiding error of as little as $2''$ exists, then the existing models are all, to within measurement and modelling error, in agreement in the cross-scan direction. In-scan, the measurement had not sufficiently developed to establish the true centroid coordinate due to uncertainties in the time-lines of individual bolometers. Further, the measurements of beam size and shapes was still in a state of flux, and some of the observations, such as the ellipticities of the unpolarised 143 GHz beams, were not optically plausible. Given that state of affairs, as much progress was made as the available information permitted, and the models are in good overall agreement with observation.

Assume that the work was resumed with the most up-to-date beam measurement input; then the in-flight telescope configuration could be reliably derived as discussed above with the following changes. (a) The model requires, as input from the data pipeline, proper estimates of the measurement errors in all five parameters used in the elliptical Gaussian fit. Without these there can be no real confidence in the level of agreement between measurement and model. Furthermore, there is no quantifiable limit on the required accuracy from the model. (b) The output from the model needs to be processed by exactly the same method as the measurement data. For elliptical Gaussian fitting it would give a better (completely unequivocal) comparison between measurement and model if the far field beam pattern was converted directly into the data format in which the on-sky measurements are fed into the data pipeline, and the pipeline itself was used to determine the fit between measurement and model. Since the GRASP files are text files, this presents no technical difficulty. (c) Measurements of polarisation: From the linearly polarised vector fields that the models derive, two sets of information are easily derived: (i) the total power coupling into (one polarisation of) the horn, and (ii) the co-polar angle of the field on the sky for that horn in that polarisation. (The polarisation angle has not been used, but it is a simple matter to derive from the linearly polarised electric vector field, in any coordinate system, if it is required.) Models and measurements of polarisation could then be compared. Finally, (d) the estimates of the bounds within which the focal plane, and primary and secondary reflectors have be moved needs to be confirmed as valid by reference to the measurements made by industry.

Further improvement could be made if the measured filter transmissions could be taken into account in the modelled beams. For that it would be necessary to know which filter was in which horn. Alternatively an average transmission could be derived from all filter measurements at each frequency. Furthermore, the spectrum of the sources should also be included in the weighting.

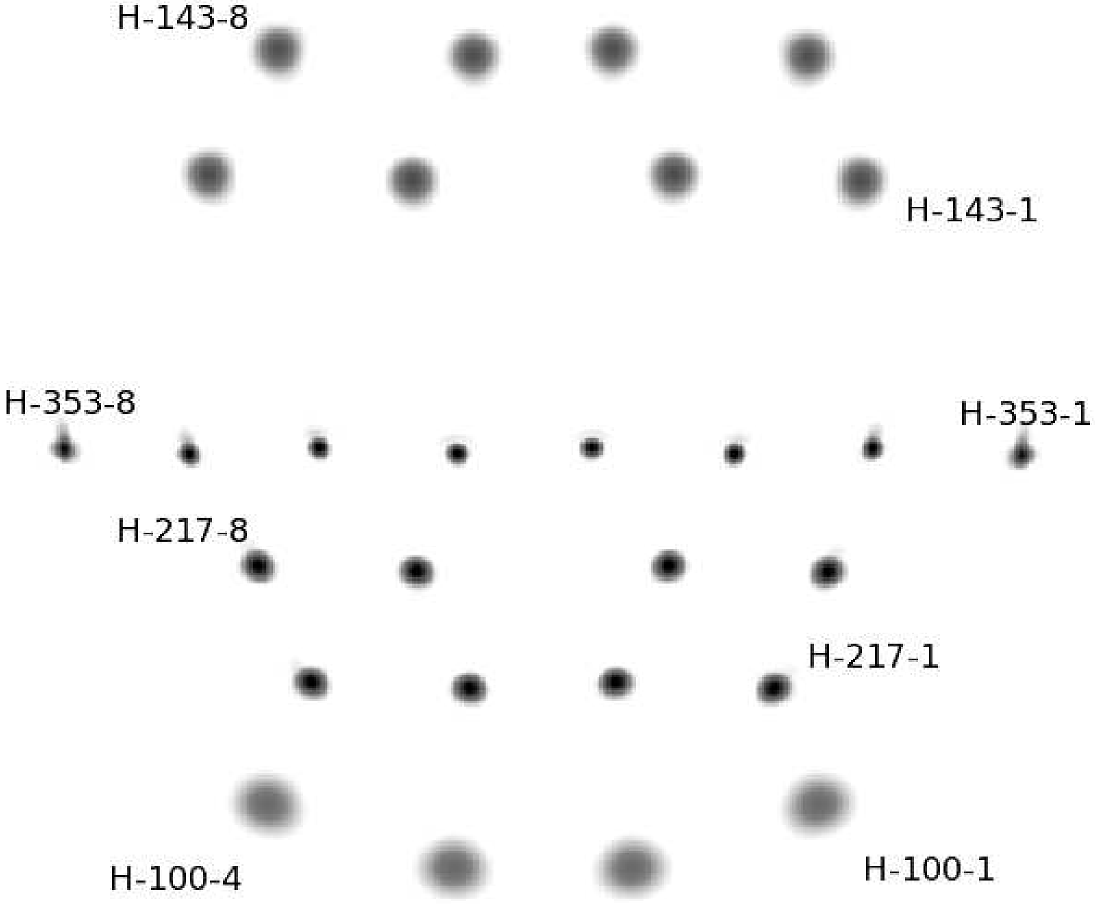


Figure 8.10: Single frequency (mid band) model of the location of the beams on the sky. The pixels are as follows: top row – unpolarised 143 GHz horns, second row – polarised 143 GHz horns; (missing row of the multi-mode 545 GHz and 857 GHz horns); third row – the eight 353 GHz horns; forth row – polarised 217 GHz horns; fifth row – unpolarised 217 GHz horns; at the bottom of the picture – the four polarised 100 GHz horns. The optic axis of the telescope intersects the focal plane above the two central unpolarised 143 GHz horns and, as the distance increases from the axis, the distortion in the beams, and the general level of aberration, increases in the manner typical of Gregorian-type telescopes. This is most apparent in the outermost 353 GHz beams, H-353-1 and H-353-8. The aberration in the four innermost 353 GHz beams are clearly seen to be dominated by coma, a fact not evident in the broad-band images in figure 8.6.

Table 8.5: Measured beam centroid positions in the LOS from the data set Mars_DataV41 and the offset of the *provisional* modelled centroid positions, model A. All units are degrees. See figure 8.2 and figure 8.3.

Pixel	in-scan	offset	cross-scan	offset
H-100-1A	-1.17703843	-0.00525105	-1.51934640	-0.00043100
H-100-1B	-1.16845094	0.00335743	-1.51860775	0.00022364
H-100-2A	-0.36512891	0.00017547	-1.79265167	-0.00111828
H-100-2B	-0.36437665	0.00103373	-1.79210377	-0.00060938
H-100-3A	0.40682924	0.00050261	-1.78957164	0.00095475
H-100-3B	0.40572944	-0.00058419	-1.78971767	0.00074372
H-100-4A	1.20853051	-0.00239312	-1.51492380	0.00024459
H-100-4B	1.20899799	-0.00195364	-1.51737331	-0.00228691
H-143-1A	-1.35838761	0.00329477	1.18711421	-0.00009540
H-143-1B	-1.36182350	-0.00003412	1.18706422	-0.00017639
H-143-2A	-0.55013663	0.00095475	1.21449904	0.00112943
H-143-2B	-0.54842947	0.00278991	1.21532741	0.00194280
H-143-3A	0.58077395	0.00152633	1.19186230	0.00163469
H-143-3B	0.58147039	0.00223077	1.19045555	0.00055095
H-143-4A	1.46176415	0.00207953	1.21450841	-0.00118320
H-143-4B	1.46183019	0.00217556	1.21418992	-0.00115069
H-143-5	-1.13459442	-0.00154405	1.72834963	0.00015702
H-143-6	-0.28817926	-0.00039389	1.75534438	-0.00045123
H-143-7	0.31506409	0.00185247	1.73121013	0.00002553
H-217-1	-0.98624911	-0.00377973	-1.01309828	0.00034611
H-217-2	-0.30163026	-0.00439788	-0.98726308	0.00066631
H-217-3	0.33160568	-0.00507595	-1.01150770	0.00073469
H-217-4	1.01521035	-0.00529427	-0.98724138	-0.00151598
H-217-5A	-1.21117532	0.00383205	-0.50907897	0.00048242
H-217-5B	-1.21161298	0.00345440	-0.50903507	0.00046533
H-217-6A	-0.52691822	0.00016516	-0.48259759	-0.00019720
H-217-6B	-0.52834505	-0.00119368	-0.48276972	-0.00039833
H-217-7A	0.56571884	0.00157521	-0.50647140	-0.00069701
H-217-7B	0.56409987	-0.00003275	-0.50656608	-0.00072969
H-217-8A	1.24886297	-0.00150865	-0.48078996	-0.00015657
H-217-8B	1.24877918	-0.00155944	-0.48055267	0.00015673
H-353-1	-2.05511504	0.00016933	-0.00311576	-0.00115537
H-353-2	-1.40983655	0.00135183	0.02449356	0.00024395
H-353-3A	-0.80968414	0.00348424	-0.00001766	0.00051173
H-353-3B	-0.81131819	0.00188019	-0.00020046	0.00033493
H-353-4A	-0.19497610	0.00129928	0.02657979	0.00131418
H-353-4B	-0.19556789	0.00074049	0.02652657	0.00126996
H-353-5A	0.38806256	0.00107594	0.00288485	0.00100024
H-353-5B	0.38820317	0.00121154	0.00290654	0.00103393
H-353-6A	0.98577901	0.00125438	0.02751943	-0.00017718
H-353-6B	0.98456760	0.00004297	0.02804804	0.00036644
H-353-7	1.54258425	-0.00479438	0.00238702	-0.00120359
H-353-8	2.08207570	-0.00480293	0.02738838	-0.00178323

	Model (A)			Modelled FWHM		Measured FWHM	
Pixel	Minor	Major	Eccen (ε)	(A)	(B)	(1)	(2)
H-100-1A	8.60981	10.38659	0.55935	9.45656	9.65344	9.48183	10.0575
H-100-1B	8.94101	10.40885	0.51201	9.64705	9.65204	9.58726	10.1906
H-100-2A	8.79809	10.49713	0.54545	9.61013	9.63286	9.43678	9.70971
H-100-2B	8.82506	10.47855	0.53916	9.61633	9.64123	9.45047	9.38593
H-100-3A	8.74614	10.55771	0.56012	9.60933	9.63181	9.42714	10.0661
H-100-3B	8.89560	10.40789	0.51913	9.62208	9.65092	9.44621	10.0345
H-100-4A	8.80328	10.55466	0.55167	9.63927	9.65464	9.55189	10.2004
H-100-4B	8.95783	10.40082	0.50816	9.65240	9.67337	9.52794	9.91908
H-143-1A	6.89149	7.26278	0.31564	7.07470	7.14870	6.91896	6.96828
H-143-1B	6.99530	7.02881	0.09754	7.01203	7.09407	6.96804	6.95189
H-143-2A	6.89198	7.14856	0.26551	7.01910	7.09322	6.88849	7.01083
H-143-2B	7.00019	6.99799	0.26551	6.99909	7.06695	6.87046	6.98437
H-143-3A	6.88718	7.06102	0.22053	6.97355	7.03827	6.97519	7.08094
H-143-3B	6.90941	7.10002	0.23016	7.00407	7.08094	6.94679	6.86819
H-143-4A	6.90241	7.16342	0.26748	7.03171	7.09430	7.08521	7.01279
H-143-4B	6.82742	7.26027	0.34012	7.04052	7.10760	7.07287	7.08531
H-143-5	7.22662	7.46446	0.25042	7.34458	7.46722	7.18205	7.15966
H-143-6	7.15640	7.41622	0.26238	7.28515	7.40490	7.17162	7.12807
H-143-7	7.15090	7.38551	0.25005	7.26726	7.38195	7.17395	7.23309
H-143-8	7.31427	7.34833	0.09617	7.33128	7.43874	7.33856	–
H-217-1	4.31100	5.10569	0.53579	4.69155	4.76403	4.65464	4.67979
H-217-2	4.35230	5.07068	0.51310	4.69778	4.77022	4.74743	4.63756
H-217-3	4.35224	5.07467	0.51425	4.69960	4.76945	4.66260	4.57269
H-217-4	4.32493	5.09032	0.52737	4.69204	4.75906	4.60960	4.64027
H-217-5A	4.35619	5.00407	0.49211	4.66891	4.73461	4.73211	4.76929
H-217-5B	4.40716	4.94980	0.45523	4.67061	4.74166	4.74743	4.7135
H-217-6A	4.43976	4.95527	0.44412	4.69044	4.76334	4.66260	4.68057
H-217-6B	4.46963	4.92762	0.42101	4.69304	4.76590	4.63459	4.62484
H-217-7A	4.42785	4.96104	0.45100	4.68687	4.75590	4.62434	4.59254
H-217-7B	4.47850	4.91792	0.41318	4.69307	4.76365	4.66547	4.62591
H-217-8A	4.37704	4.98571	0.47881	4.67147	4.72868	4.69041	4.6853
H-217-8B	4.40213	4.95368	0.45857	4.66977	4.72820	4.74138	4.75164
H-353-1	4.07349	4.85117	0.54306	4.44536	4.39336	4.52602	4.53123
H-353-2	4.10935	4.63051	0.46090	4.36215	4.30017	4.43023	4.43329
H-353-3A	4.18398	4.53266	0.38462	4.35483	4.29230	4.42912	4.41162
H-353-3B	4.22293	4.51174	0.35203	4.36495	4.29261	4.42594	4.33692
H-353-4A	4.28993	4.47629	0.28554	4.38212	4.31364	4.39701	4.37221
H-353-4B	4.31107	4.45572	0.25273	4.38280	4.31440	4.39036	4.37819
H-353-5A	4.27629	4.49715	0.30953	4.38533	4.38693	4.31478	4.41046
H-353-5B	4.30637	4.45314	0.25463	4.37914	4.31028	4.39791	4.41973
H-353-6A	4.19676	4.57078	0.39619	4.37978	4.30063	4.46048	4.35391
H-353-6B	4.21780	4.54163	0.37084	4.37672	4.30178	4.42553	4.35162
H-353-7	4.12852	4.67864	0.47046	4.39498	4.31506	4.42893	4.40631
H-353-8	4.09111	4.87155	0.54290	4.46431	4.53607	4.40301	4.48774

Table 8.10: Major, a, and minor, b, axis beam widths, $\varepsilon = \sqrt{1-(b/a)^2}$ and FWHM, given by elliptical Gaussian fitting to the modelled beam in the LOS coordinate system measured FWHM from (1) Mars_DatV32, and (2) Mars_DataV41.

Appendix A

Mathematical reference

There follow some basic definitions and statements from the elementary theory of functional analysis, concerning Hilbert spaces in particular, that are used in the text. The necessary background can be found in any introduction to functional analysis or Hilbert spaces; for example the following references, with the most elementary given first: [62], [74], [29].

A.1 Hilbert spaces and definitions of mathematical terms

A.1.1. A linear space (vector space) is said to be *complete* if every Cauchy sequence has a limit *in that space.*

A.1.2. A mapping, S, between linear spaces is said to be *a contraction* if $\|S\| \leq 1$; that is $\|S\boldsymbol{x}\| \leq \|S\|\|\boldsymbol{x}\|$ for all $\boldsymbol{x}$ in the domain of S. It is a strict contraction if $\|S\| < 1$.

A.1.3. A sequence $\{x_n\}_{n\in\mathbb{N}}$ in a linear space is said to be *Cauchy* if, given $\varepsilon > 0$, $\exists\, K \in \mathbb{N}$ such that $n, m > K \implies |x_n - x_m| < \varepsilon$.

A.1.4. An *inner product* on a complex linear space V is a bilinear mapping $\langle \cdot \,|\, \cdot \rangle$:

$V \times V \to \mathbb{C}$ such that, for all x, y, $z \in V$ and $\lambda \in \mathbb{C}$

$$\begin{aligned} (i) &\quad \langle x|y\rangle = \overline{\langle y|x\rangle}, \\ (ii) &\quad \langle \lambda x|y\rangle = \lambda \langle x|y\rangle , \\ (iii) &\quad \langle x+y|z\rangle = \langle x|z\rangle + \langle y|z\rangle , \\ (iv) &\quad \langle x|x\rangle \geq 0 \text{ with equality iff } x=0. \end{aligned}$$

That is, the mapping is a complex valued, conjugate symmetric, non-degenerate bilinear pairing.

A.1.5. Given an inner product space $(V, \langle \cdot \,|\, \cdot \rangle)$, V is (an orthogonal) *direct sum* of inner product spaces $\mathcal{E}$ an $\mathcal{M}$ if $\mathcal{E} \cap \mathcal{M} = \{0\}$, every $\boldsymbol{v} \in V$ can be written uniquely as $\boldsymbol{v} = \boldsymbol{e} + \boldsymbol{m}$ for $\boldsymbol{e} \in \mathcal{E}$ and $\boldsymbol{m} \in \mathcal{M}$ with $\langle \boldsymbol{e} \,|\, \boldsymbol{m} \rangle = 0$. The direct sum is written $V = \mathcal{E} \oplus \mathcal{M}$.

A.1.6. A *pre-Hilbert space* or *inner product space* is pair $\big(V, \langle \cdot \,|\, \cdot \rangle\big)$ where V is a linear space and $\langle \cdot \,|\, \cdot \rangle$ is an inner product.

A.1.7. A *Hilbert space* is a complete normed linear space in which the norm is defined be an inner product. Equivalently, the space has a complete metric determined by the inner product.

Every pre-Hilbert space has a completion in which the points of the completed space are equivalence classes of Cauchy sequences, two sequences being equivalent if they have the same limit. Strictly speaking the direct sums of Hilbert spaces that are referred to in this thesis are pre-Hilbert, not Hilbert. That distinction will be glossed over throughout the thesis; firstly because the completions will always exist, and secondly because the main interest is in the finite dimensional models of the spaces rather than the spaces themselves, and the finite models are just complex vector spaces which are necessarily complete.

If a Hilbert space has a denumerable, orthonormal basis, then that basis will be complete. It is a theorem of Hilbert space theory that, given any denumerably infinite orthonormal sequence $\{\boldsymbol{e}_n\}_{n \in \mathbb{N}}$ in a Hilbert space $\mathcal{H}$, the following three statements are equivalent: (i) $\{\boldsymbol{e}_n\}_{n \in \mathbb{N}}$ is complete, (ii) $\mathcal{H}$ is the complex linear span of $\{\boldsymbol{e}_n\}_{n \in \mathbb{N}}$, and (iii) the squared norm $\| \boldsymbol{x} \|^2$ is $\sum_{n \in \mathbb{N}} | \langle \boldsymbol{x} \,|\, \boldsymbol{e}_n \rangle |^2$ for all $\boldsymbol{x} \in \mathcal{H}$.

A.2 Operator, matrix and vector norms

The vector spaces that are considered for the numerical models are all finite dimensional over either the real or the complex fields. In this appendix the field will be denoted $\mathbb{F}$ and a vector in the n-dimensional vector space $\mathbb{F}^n$ will be denoted by $\mathbf{x}$ while matrices will be denote by A.

The vector norms used in the document are:

$$\text{The } L^1 \text{ norm} \qquad \| \mathbf{x} \|_1 = |x_i| + |x_2| + \ldots + |x_n| \tag{A.2.1}$$

$$\text{The } L^2 \text{ norm} \qquad \| \mathbf{x} \|_2 = \left(|x_i|^2, |x_2|^2, \ldots, |x_n|^2\right)^{1/2} \tag{A.2.2}$$

$$\text{The } L^\infty \text{ norm} \qquad \| \mathbf{x} \|_\infty = \max\{|x_i|, |x_2|, \ldots, |x_n|\} \tag{A.2.3}$$

A matrix norm, $\| \cdot \|_m$, is said to be subordinate to a vector norm, $\| \cdot \|_v$, if

$$\| A \|_m = \max_{\mathbf{x} \neq 0} \frac{\| A\mathbf{x} \|_v}{\| \mathbf{x} \|_v} = \max_{\|\mathbf{x}\|_v = 1} \| A\mathbf{x} \|_v \, . \tag{A.2.4}$$

The matrix norms used are the Frobenius, uniform and maximum norms:

$$\|A\|_F = \left(\sum_i \sum_j |a_{ij}|^2 \right)^{\frac{1}{2}}, \tag{A.2.5}$$

$$\|A\|_\infty = \max_i \sum_j |a_{ij}|, \tag{A.2.6}$$

$$\|A\|_{\max} = \max_i \max_j \{|a_{ij}|\}. \tag{A.2.7}$$

The Frobenius norm (also called the Euclidean norm) is the matrix equivalence of the L^2 operator norm:

$$\|g\|_2 = \langle g|g \rangle = \left(\int |g|^2 \, \mathrm{d}x \right)^{\frac{1}{2}},$$

where the integral is taken over the entire domain of g.

A.3 Green's identities

If V is a volume with boundary $S = \partial V$, then Green's first identity for two scalar functions, u and v, defined and differentiable over the closure of V is

$$\int_V \left(\nabla v \cdot \nabla u + u \nabla^2 v\right) \, \mathrm{d}V = \int_S u \nabla v \cdot \, \mathrm{d}\boldsymbol{S}. \tag{A.3.1}$$

Green's second identity takes the form

$$\int_V \left(u \nabla^2 v - v \nabla^2 u\right) \, \mathrm{d}V = \int_S (u \nabla v - v \nabla u) \cdot \, \mathrm{d}\boldsymbol{S}. \tag{A.3.2}$$

A.4 Fourier spectra

The basis functions for the Fourier expansion in the circular aperture are the discrete set indexed by $(n, m) \in \mathbb{Z}_{\geq 0} \times \mathbb{N}$

$$\Psi_{nm} = \frac{1}{\sqrt{N_{nm}}} J_n(k_{nm} r) \Phi_n(\varphi),$$

where the N_{nm} are the normalisation factors appropriate to the field type and the Φ_n are sines or cosines fitting n times onto the unit circle. In free space, retaining the polar coordinates, the spectrum becomes

$$\tilde{\Psi}_{nk} = \sqrt{k} J_n(kr) \Phi_n(\varphi),$$

for a continuous parameter k. Thus, calculating the Fourier expansion of the waveguide modes for a waveguide of radius a, and thus the far field pattern of the idealised telescope, amounts to computing the Fourier coefficients

$$P_{nmk} = \sqrt{\frac{k}{N_{nm}}} \int_0^a J_n(kr) J_n(k_{nm} r) r \, \mathrm{d}r$$

the set of which, for each mode Φ_{nm}, is a mapping $(0, \infty) \to \mathbb{C}$. In practise only a finite set of waveguide modes is used, and a discrete subset of the Fourier coefficients for each would be calculated. The integgral has the analytic solution (4.2.2) on page 84, the first term in the square brackets vanishing for the magnetic field, and the second term for the electric field. To get a reasonable idea of the main beam k need only extend out to an angle given by the aperture radius times the plate scale.

An alternative would be write the modes over the aperture in Cartesian coordinate and to use the Hermite function expansion of the modes over the aperture. These functions (see [69], [22]) are the eigenfunctions of the 1D Fourier transform and form a basis for the space of Schwartz functions. A function, f, is Schwartz, denoted $f \in \mathcal{S}(\mathbb{R})$ if, for all integers $k, l \geq 0$,

$$\sup_{x\in\mathbb{R}} |x|^k |f^{(l)}(x)| < \infty.$$

Since these are related to Gaussian beam modes which are useful for the preliminary analysis of optical systems, these two subjects will be discussed in the following three sections.

A.5 Bolometer theory

To discuss what is observable in a measurement made with a horn of the type used in Planck it is necessary to outline the function of the detectors; bolometers in this case.

Radiation incident upon the bolometer is absorbed by a mass of heat capacity $C(T)$, where T is the absolute temperature. The absorber is thermally linked to a heat sink at temperature T_s through a thermal conductor with conductance function $G(T)$. The absorber in the Planck bolometers is a metallised mesh. Absorption of radiation increases the temperature of the mesh, the change being measured as as a change in the resistance, $R(T)$, of a semi-conducting neutron transmutation doped germanium thermometer biased by constant current I_b.

Denote the electrical bias power by $P_b = R(T)I_b^2$ and the load resistance by R_l, chosen so that $R_l \gg R(T)$. Denote the ratio of the infinitesimal change in resistance to infinitesimal change in temperature, normalised by the resistance, by

$$\alpha = \frac{1}{R(T)} \frac{\mathrm{d}R}{\mathrm{d}T}.$$

Define the *voltage responsivity* function to be the ratio of a small voltage change to a small input power from a signal with angular frequency ω; the voltage responsivity is given by

$$S(T) = \frac{\alpha\sqrt{P_b(T)R(T)}}{G(T) + i\omega C(T) - \alpha P_b(T)\left[1 - 2R(T)/(R(T) - R_l)\right]} \qquad \text{in V/W.}$$

The response time function of the bolometer is given by

$$\tau(T) = \frac{C(T)}{G(T) - \alpha P_b(T)\left[1 - 2R(T)/(R(T) - R_l)\right]}.$$

Define the *noise equivalent power* (NEP) of the bolometer to be the optical signal power equal to the noise in a 1 Hz amplifier bandwidth at the output. Then, multiplying the product of the optical efficiency, ε, and the total noise by the voltage responsivity, gives

$$\begin{aligned}\varepsilon^2\mathrm{NEP}^2 = \gamma_1 4kT^2G(T) + \gamma_2\frac{4kTR(T)}{S^2} \\ + \frac{e_n^2 + \gamma_3 R^2(T)i_n^2}{S^2} + \mathrm{NEP}^2_{phot} + \frac{v^2_{excess}}{S^2}.\end{aligned}$$

The expression and definition of the parameters γ_j are from [41], to which reference should be made for details: $\gamma_1 4kT^2G(T)$ represents thermal fluctuation noise from phonon exchange between the absorber and the heat sink; $\gamma_2 4kTR(T)/S^2$ is the Johnson noise of the thermometer referred to the bolometer input by dividing by S^2; the term $\left(e_n^2 + \gamma_3 R^2(T)i_n^2\right)/S^2$ is the amplifier noise referred to the bolometer input, in units of V^2/Hz, e_n and i_n being voltage and current noises; the term NEP^2_{phot} is the incident power photon noise in W^2/Hz, while the final term represents all excess noise in the bolometer and the readout, in V^2/Hz, divided by the voltage responsivity. The design and calibration of the Planck bolometers is summarised in [31].

In the Planck multi-mode horns the absorbing mesh is a silicon nitride spider web coated with gold on a bonding layer of titanium, the disc of the web having a diameter of 2λ at the band centre and the equivalent sheet impedance is designed for optimum absorption when it is placed in a cavity with a 1/4-wave back-short, at mid band. The web-like structure reduces the probability of cosmic ray interaction and, in addition, renders the structure inefficient as an absorber of high frequency photons, thus helping to filter out radiation that is out of the band. The absorption efficiency will be frequency dependent.

A.6 The paraxial wave equation

Assume that a function $V \in C^2(\mathbb{R}^4; \mathbb{C})$ describes a strictly monochromatic scalar wave of the form $V(x, y, z, t) = F(x, y, z) \exp(-j\omega t)$ so that the function V satisfies the wave equation. Then the spacial component, F, satisfies the Helmholtz equation

$$(\nabla^2 + k^2)F = 0 \tag{A.6.1}$$

where $k = \omega/c$, $F \in C^2(\mathbb{R}^3; \mathbb{C})$ and, if V is to represent any physical field, $F \in L^2(\mathbb{R}^3)$ so that it has finite total power. The paraxial wave equation [27] [26] [49],

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} - 2jk\frac{\partial u}{\partial z} = 0 \tag{A.6.2}$$

is derived from the Helmholtz equation, so the association with a scalar electric or other field is strictly irrelevant; it is enough to assume that we have an L^2 function that satisfies the Helmholtz equation. Nevertheless, since the Helmholtz equation describes a process of propagation, the function will be referred to as a beam or a field. It may be supposed to describe, for example, a scalar field, a component of a vector field or of a correlation tensor.

A.7 Modes of the paraxial wave equation

There are various sets of solutions to the paraxial wave equation, one of which is the family of Gauss-Hermite functions. The idea is the following: the beam is propagating in the z-direction, so we attempt to describe the field as a one-parameter family of functions in $L^2(\mathbb{R}^2)$, parametrised by z. Since $L^2(\mathbb{R}^2) \cong L^2(\mathbb{R}) \otimes L^2(\mathbb{R})$ consider the paraxial wave equation in the plane; then one solution is the one-parameter family of functions (see [60][26])

$$\begin{aligned} F_n(x, z) = &\left[\frac{\sqrt{2}}{\sqrt{\pi} w(z) 2^n n!}\right]^{1/2} H_n\left(\frac{x\sqrt{2}}{w(z)}\right) \\ &\exp\left[-\frac{x^2}{w(z)^2} - jk\left(z + \frac{x^2}{2R(z)} - (\frac{2n+1}{2k})\,\boldsymbol{\varphi}(z)\right)\right] \end{aligned} \tag{A.7.1}$$

where w_0 is the beam waist radius at $z = 0$; $R(z) = z[1 + (\pi w_0^2/\lambda z)^2]$ is the beam radius of curvature, $w(z) = w_0\sqrt{1 + (\lambda z/\pi w_0^2)^2}$ is the beam waist radius, and $\boldsymbol{\varphi}(z) = \arctan(\lambda z/\pi w_0^2)$ is the phase shift.

This is the so called 'normal set'. The set $\{F_{nm}(x, y, z) = F_n(x, z)F_m(y, z) : n, m \in \mathbb{N} \cup \{0\}\}$ spans the space of solutions to the paraxial wave equation. The higher the order of the mode function, the more rapidly the beam described by the mode diverges. Consequently the high order modes are only valid approximations to a beam close to the z axis. This equation can be written in terms of the eigen-functions of the Fourier transform as follows: set

$$h_n(x) = \frac{(-1)^n}{n!}\exp(\pi x^2)D^n \exp(-2\pi x^2). \tag{A.7.2}$$

The h_n are eigen-functions of the Fourier transform operator and the set of functions $\{e_n : n \geq 0\}$, where

$$e_n(x) = \left[\frac{\sqrt{2}\cdot n!}{(4\pi)^n}\right]^{1/2} h_n(x) \quad \text{for all} \quad n \in \mathbb{Z}_{\geq 0} \tag{A.7.3}$$

forms an orthonormal basis for the space of Schwartz functions, $\mathcal{S}$, with the $L^2(\mathbb{R})$ inner product, (see section A.8 on page 235 and references [64][22]). Write

$$e_n(x; z) = \left[\frac{(2\pi)^n}{w(z)\sqrt{\pi}}\right]^{\frac{1}{2}} e_n\left(\frac{x}{w(z)\sqrt{\pi}}\right) \tag{A.7.4}$$

In terms of this basis, (A.7.1) becomes

$$F_n(x, z) = e_n(x; z)\exp\left[-jk\psi(x; z)\right]\exp[-jkC_n(z)] \tag{A.7.5}$$

where

$$C_n(z) = z - (n + \frac{1}{2})\frac{\boldsymbol{\varphi}(z)}{k} \tag{A.7.6}$$

and

$$k \cdot \psi(x; z) = \frac{\lambda z}{w_0^2}\left(\frac{x}{w\sqrt{\pi}}\right)^2 = k \cdot \frac{x^2}{2R(z)}. \tag{A.7.7}$$

The equations (A.7.1) and (A.7.5) have the form $F_n(x, z) = e_n(x, z)\exp(-jk\Psi_n(x, z))$, where the amplitude function, $e_n(x, z)$, and the phase function, $\Psi_n(x, z) = \psi(x; z) + C_n(z)$, are real valued. Since the waist function, w, is an even function of z, e_n is even in z. Also, $\boldsymbol{\varphi}(z)$ is odd, so the phase terms, ψ and C_n, are odd in z. Furthermore, ψ is an even function of x and e_n is an even function of x if n is even and an odd function if n is odd. Therefore, $\Psi_n(x, z)$ is an odd function of z and an even function of x, whilst $e_n(x; z)$ is even or odd according to whether n is even or odd.

For fixed z and all n it is clear from (A.7.5) that $F \in L^2(\mathbb{R})$, but less obvious that $F_n(x; z) \in \mathcal{S}$; that it is a Schwartz function is proven in section A.8. It is clear from (A.7.4) and the definition of $k \cdot \psi$ in (A.7.7) that F_n does not change form as it propagates, only the scale and phase changes.

At any point in the plane at which a Gauss-Hermite mode describes the paraxial propagation of the beam, the direction of propagation of the mode is given by the gradient of the phase function.

A.8 Schwartz functions and Gauss-Hermite modes

The space of Schwartz functions on $\mathbb{R}$, denoted by $\mathcal{S} = \mathcal{S}(\mathbb{R})$ is the set of all indefinitely differentiable functions on the real line with all derivatives rapidly decreasing. That is a function $f \in \mathcal{S}(\mathbb{R})$ if, for all integers $k, l \geq 0$,

$$\sup_{x \in \mathbb{R}} |x|^k |f^{(l)}(x)| < \infty$$

For all $a \in \mathbb{R}_{>0}$, the weighted Gaussian $\exp(-ax^2)$ is a Schwartz function. It follows, by the the closure of $\mathcal{S}$ under pointwise multiplication of functions, that the functions

$$e_n(x) = \left[\frac{\sqrt{2}n!}{(4\pi)^n}\right]^{1/2} h_n(x) \quad \text{for all} \quad n \in \mathbb{Z}_{\geq 0}$$

and therefore also the functions $e(x; z)$ defined in (A.7.4), are Schwartz functions for all fixed $z \in \mathbb{R}$. The functions $h_n(x)$ are eigen-functions of the Hermite operator $Hf(x) = -D^2 f + 4\pi^2 x^2 f$ with eigenvalue $4\pi(n + 1/2)$ (see [22]). Since

$$4\pi(n + \frac{1}{2})\langle h_n, h_m\rangle = \langle Hh_n, h_m\rangle = \langle h_n, Hh_m\rangle = 4\pi(m + \frac{1}{2})\langle h_n, h_m\rangle$$

if and only if $n = m$, it follows that h_n is orthogonal to h_m for all $n, m \geq 0$. Therefore the set $\{e_n : n \geq 0\}$ forms an orthonormal set in $\mathcal{S}$.

The functions given in equation (A.7.5), for all $n \geq 0$ and any fixed $z \in \mathbb{R}$ are Schwartz functions. To see this proceed as follows: Holding z constant, the term $\exp[jkC_n(z)]$ can be ignored as it is simply a constant of unit modulus. Furthermore, setting $\xi = x/w\sqrt{\pi}$, differentiation with respect to ξ yields the following two equations:

$$\begin{aligned} D_\xi^k e_n(x;z) &= \frac{1}{(w\sqrt{\pi})^k} D_x^k e_n(x;z), \\ D_\xi^l \exp(-jk\psi(\xi)) &= \frac{1}{(w\sqrt{\pi})^l} D_x^k \exp(-jk\psi(\xi)) \end{aligned}$$

Thus, differentiating with respect to ξ or x changes only the scaling, not the result we wish to prove. For all $k, l, n \geq 0$, and all $x \in \mathbb{R}$, Leibniz's rule gives

$$\begin{aligned} |x|^k |D^l F_n(x)| &= |x|^k \left| \sum_{p=0}^{l} \binom{l}{p} D^{l-p} e_n(x) D^p e^{-jk\psi} \right| \\ &= |x|^k \left| \sum_{p=0}^{l} \binom{l}{p} D^{l-p} e_n(x) Q_p(x) e^{-jk\psi} \right| \end{aligned}$$

where $Q_p(x)$ is a polynomial in x of order p with complex coefficients, the details of which do not concern us. Write $Q_p(x) = q_0 + q_1 x + q_2 x^2 + \cdots + q_p x^p$ and $q = \max\{|q_0|, |q_1|, \ldots, |q_p|\}$, then

$$\begin{aligned} |x|^k |D^l F_n(x)| &< l! \cdot q \cdot |x|^k \sum_{p=0}^{l} (1 + |x| + |x|^2 + \cdots + |x|^p) \left| D^{l-p} e_n(x) \right| \\ &= l! \cdot q \cdot \sum_{p=0}^{l} (|x|^k + |x|^{1+k} + \cdots + |x|^{p+k}) \left| D^{l-p} e_n(x) \right| \\ &< l! \cdot q \cdot (p+1) \cdot \sum_{p=0}^{l} \max\{|x|^k, |x|^{p+k}\} \left| D^{l-p} e_n(x) \right| < \infty \end{aligned}$$

where we have used $|\exp(-jk\psi)| = 1$ and the fact that $e_n \in \mathcal{S}$. More generally, $\mathcal{S}$ is an ideal in the ring $C^\infty(\mathbb{R})$.

For all $n, m \geq 0$, F_n is orthonormal to F_m. This is almost immediate from the orthonormality of e_n and e_m:

$$\langle F_n, F_m \rangle = \int_{\mathbb{R}} F_n(x;z) \bar{F}_m(x;z) \, \mathrm{d}x$$

$$
\begin{aligned}
&= \int_{\mathbb{R}} e_n(x;z)\exp\left[-jk\psi(x;z)\right]\exp[-jkC_n(z)]e_m(x;z)\exp\left[jk\psi(x;z)\right] \\
&\qquad\qquad \exp[jkC_m(z)]\,\mathrm{d}x \\
&= \exp[(m-n)\frac{\boldsymbol{\varphi}(z)}{k}]\cdot\int_{\mathbb{R}} e_n(x;z)e_m(x;z)\,\mathrm{d}x \\
&= \exp[(m-n)\frac{\boldsymbol{\varphi}(z)}{k}]\left[\frac{(2\pi)^{n+m}}{w^4\pi^2}\right]^{\frac{1}{2}}\cdot\int_{\mathbb{R}} e_n(\xi)e_m(\xi)\,\mathrm{d}\xi \\
&= \exp[(m-n)\frac{\boldsymbol{\varphi}(z)}{k}]\left[\frac{(2\pi)^{n+m}}{w^4\pi^2}\right]^{\frac{1}{2}}\cdot\delta^n_m \\
&= \begin{cases} 0 \quad : \quad n \neq m \\ 1 \quad : \quad n = m \end{cases}
\end{aligned}
$$

Starting from the definition of h_{n+1} given in (A.7.2), write

$$
D^{n+1}\exp(-2\pi x^2) = D^n\{-4\pi x\exp(-2\pi x^2)\}
$$

and expand the right hand side as a polynomial using the Leibniz rule. After rearranging, the resulting expression for h_{n+1} is the recurrence relation $h_{n+1} = 4\pi\{xh_n - h_{n-1}\}/(n+1)$. Then

$$
\begin{aligned}
F_{n+1}(x;z) &= \left[\frac{(2\pi)^{n+1}}{w\sqrt{\pi}}\right]^{\frac{1}{2}}\cdot\left[\frac{\sqrt{2}\cdot(n+1)!}{(4\pi)^{n+1}}\right]^{\frac{1}{2}} h_{n+1}\left(\frac{x}{w\sqrt{\pi}}\right) \\
&\qquad \times\exp(-jk\psi(x;z))\exp(jkC_{n+1}(z)) \\
&= \frac{4\pi}{n+1}\Bigg\{\frac{x}{w\sqrt{\pi}}\left[\frac{(2\pi)^n}{w\sqrt{\pi}}\right]^{\frac{1}{2}}\left[\frac{\sqrt{2}\cdot n!}{(4\pi)^n}\right]^{\frac{1}{2}} h_n\left(\frac{x}{w\sqrt{\pi}}\right)\left[\frac{(n+1)}{2}\right]^{\frac{1}{2}} \\
&\qquad - \left[\frac{(2\pi)^{n-1}}{w\sqrt{\pi}}\right]^{\frac{1}{2}}\left[\frac{\sqrt{2}\cdot(n-1)!}{(4\pi)^{n-1}}\right]^{\frac{1}{2}} h_{n-1}\left(\frac{x}{w\sqrt{\pi}}\right)\left[\frac{(n+1)\cdot n}{8}\right]^{\frac{1}{2}}\Bigg\} \\
&\qquad \times\exp(-jk\psi(x;z))\exp(jkC_{n+1}(z)) \\
&= \frac{4\pi}{n+1}\Bigg\{\frac{x}{w\sqrt{\pi}}\left[\frac{(n+1)}{2}\right]^{\frac{1}{2}}\exp(-jk\psi(x;z))\exp(jk[C_n - C_n + C_{n+1}]) \\
&\qquad \times e_n(x;z) - \left[\frac{(n+1)\cdot n}{8}\right]^{\frac{1}{2}} e_{n-1}(x;z)\exp(-jk\psi(x;z)) \\
&\qquad \times\exp(jk[C_{n-1} - C_{n-1} + C_{n+1}])\Bigg\} \\
&= \frac{4\pi}{\sqrt{2(n+1)}}\left\{\frac{x}{w\sqrt{\pi}}F_n(x;z) - \exp(-j\,\boldsymbol{\varphi}(z))\frac{\sqrt{n}}{2}F_{n-1}(x;z)\right\}
\end{aligned}
$$

$$\times \exp(-2jk\,\boldsymbol{\varphi}(z))$$

having used $C_{n+1}(z) - C_m(z) = (m - n - 1)\,\boldsymbol{\varphi}(z)/k$ from the definition in equation (A.7.6).

Thus it is seen that the Gauss-Hermite modes satisfy the recurrence relation

$$\begin{aligned} F_{n+1}(x;z) = &\frac{4\pi}{\sqrt{2(n+1)}}\left\{\frac{x}{w\sqrt{\pi}}F_n(x;z) - \exp(-j\,\boldsymbol{\varphi}(z))\frac{\sqrt{n}}{2}F_{n-1}(x;z)\right\} \\ &\times \exp(-2jk\,\boldsymbol{\varphi}(z)) \end{aligned} \tag{A.8.1}$$

that can be used in numerical routines for their calculation because the recurrence is stable.

Bibliography

[1] M. Abramowitz and I. A. Stegun, editors. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover, 9th printing edition, 1972.

[2] P. A. Ade, R. J. Wylde, and J. Zhang. Ultra-Gaussian Horns for ClOVER – a B-Mode CMB Experiment. In *20th International Symposium on Space Terahertz Technology, Charlottesville*, April 2009.

[3] H. Akima. A method of bivariate interpolation and smooth surface fitting for values given at irregularly distributed points. *ACM Trans. on Math. Software*, 4(2), June 1978.

[4] H. Akima. On estimating partial derivatives for bivariate interpolation of scattered data. *Rocky Mountain J. of Math.*, 14(1), 1984.

[5] J. Allison and F.A. Benson. Surface roughness and attenuation of precision drawn, chemically polished, electropolished, electroplated and electroformed waveguides. *Proc. IEEE (London)*, 102:251 – 259, 1955.

[6] G.B. Arfken and H.J. Weber. *Mathematical Methods for Physicists.* Harcourt Academic Press, 5 edition, 1996.

[7] S. R. Bell. *The Cauch Transform, Potential Theory, and Conformal Mapping.* Studies in advanced mathematics. CRC Press, 1992.

[8] G. Birkhoff and G-C. Rota. *Ordinary differential equations.* John Wiley and Sons, 4 edition, 1989.

[9] P.T. Boggs, H.R. Byrd, and R.B. Schnabel. A Stable and Efficient Algorithm for Nonlinear Orthogonal Distance Regression. *SIAM J. Sci. and Stat. Comput.*, 8(6):1052 – 1078, 1987.

[10] P.T. Boggs and J.E. Rogers. Orthogonal Distance Regression. *Contemporary Mathematics*, 112:183 – 194, 1990.

[11] M. Born and E. Wolfe. *Principles of Optics.* Cambridge University Press, sixth edition, 1997.

[12] Long Chen, Pengtao Sun, and Jinchao Xu. Optimal anisotropic meshes for minimizing interpolation errors in L^p-norm. *Mathematics of Computation*, 76(257):179–204, 2007.

[13] P. J. B. Clarricoats and A. D. Olver. *Corrugated horns for microwave antennas.* Number 18 in IEE Electromagnetic Waves Series. Peter Peregrinus Ltd., 1984.

[14] R. Colgan. *Electromagnetic and Quasi-optical Modelling of Horn Antennas for Far-IR Space Applications.* PhD thesis, National University of Ireland, Maynooth, 2001.

[15] R. E. Collin. *Field Theory of Guided Waves.* IEEE Press, 2 edition, 1991.

[16] R. J. Daddato. Optical Metrology of the Planck Reflectors. Technical report, ESTEC TEC-MMO Planck Optics Workshop, December 2006.

[17] M. de Berg, M. van Kreveld, M. Overmars, and O. Schartzcopf. *Computational Geometry.* Springer, 2 edition, 2000.

[18] L. Diaz and T. Milligan. *Antenna engineering using physical optics.* Artech House, 1996.

[19] J. J. Dongarra *et al. LINPACK users' guide.* SIAM, 1993.

[20] D. Dubruel. HERSCHEL/PLANCK Inputs for the RFFM flight prediction numerical model. Industrial contract report (confidential), Version 2,, ThalesAlenia Space, February 2009.

[21] C. Dumesnil. HFI ALIGNMENT ICD TN-PH293-700917-IAS. Instrument interface control document, Version 2, IAS, Paris, April 2007.

[22] H. Dym and H. P. McKean. *Fourier Series and Integrals*, volume 14 of *Probability and Mathematical Statistics.* Accademic Press, 1972.

[23] E. G. Effros and Z-J. Ruan. *Operator Spaces.* Number 23 in London Mathematical Society Monographs. Oxford University Press, 2005.

[24] R. Fletcher. *Practical methods of optimisation.* John Wiley & Sons, 2 edition, 2000.

[25] E. Gleeson. *Single and Multi-moded Corrugated Horn Design for Cosmic Microwave Background Experiments.* PhD thesis, National University of Ireland, Maynooth, 2004.

[26] P. F. Goldsmith. *Quasioptical Systems.* IEEE Press, 1998.

[27] G. Goubau and F. R. Schwering. On the guided propagation of electromagnetic wave beams. *IRE Trans. on Antennas and Propagation*, pages 248–256, May 1961.

[28] BLAST Technical Group. *Basic Linear Algebra Subprograms Technical (BLAST) Forum Standard.* University of Tennessee, Knoxville, Tenessee, August 2001.

[29] E. Hewitt and K. Stromberg. *Real and Abstract Analysis.* Number 25 in Graduate Texts in Mathematics. Springer-Verlag, 1965.

[30] F. S. Hill. *Computer Graphics Using Open GL.* Prentice Hall, 2 edition, 2001.

[31] W. A. Holmes, J. J. Bock, B. P. Crill, T. C. Koch, W. C. Jones, A. E. Lange, and C. G. Paine. Initial test results on bolometers for the planck high frequency instrument. *Applied Optics*, 47(32):5996 – 6008, November 2008.

[32] E. Isaacson and H. B. Keller. *Analysis of Numerical Methods.* John Wiley, 1966.

[33] J. D. Jackson. *Classical Electrodynamics.* John Wiley & Sons, Inc., third edition, 1999.

[34] A. Jeffrey, editor. *Table of Integrals, Series and Products.* Academic Press, 5 edition, 1994.

[35] B. Joe. Shape measures for quadrilaterals, pyramids, wedges, and hexahedra. Technical report, Zhou Computing Services Inc., 3 2008.

[36] L. W. Johnson and R. D. Riess. *Numerical Analysis.* Addison-Wesley, 2 edition, 1982.

[37] D. E. Knuth. *The art of computer programming.* Addison-Wesley, 1997.

[38] L. Lewin. *Advanced theory of waveguides.* Wireless Engineer, 1951.

[39] B. Maffei *et al.* Planck-HFI optical design and pre-flight performances. In *EuCAP 2010 - The 4th European Conference on Antennas and Propagation*, 2010.

[40] B. Maffei *et al.* Planck pre-launch status: HFI beam expectations from the optical optimisation of the focal plane. *Astronomy and Astrophysics*, 520(1), 2010.

[41] J. C. Mather. Bolometer noise: non-equilibrium theory. *Applied Optics*, 21:1125 – 1129, 1982.

[42] K. A. Milton and J. Schwinger. *Electromagnetic Radiation: Variational Methods, Waveguides and Accelerators.* Particle Acceleration and Detection. Springer, 2006.

[43] A. Murphy. Cylindrical waveguide transitions. Development of the equations for mode scattering amplitudes in cylidrical waveguides., 1990.

[44] A. Murphy, T. Peacocke, B. Maffei, I. McAuley, F. Noviello, V. Yurchenko, P.R. Ade, G. Savini, J.-M. Lamarre, J. Brossard, R. Colgan, E. Gleeson, E. Lange, Y. Longval, G. Pisano, J.-L. Puget, I. Ristorcelli, R. Sudiwala, and J. Wylde.

Multi-mode horn design and beam characteristics for the planck satellite. *Journal of Instrumentation*, 2010.

[45] J. A. Murphy, R. Colgan, C. O'Sullivan, B. Maffei, and P. Ade. Radiation patterns of multi-moded corrugated horns for far-ir space applications. *Infrared Physics and Technology*, 41:515 – 528, 2001.

[46] J. A. Murphy *et al.* Millimeter-wave profiled corrugated horns for the QUaD Cosmic Background Polarization Experiment. *Int. J. of Infrared and Millimeter Waves*, 26(4), 2005.

[47] A. D. Olver, P. J. B. Clarricoats, A. A. Kishk, and L. Shafai. *Michrowave Horns and Feeds*. Number 39 in IEEE Elecromagnetic Waves Series. IEEE Press, 1994.

[48] P. T. Peacocke and A. Koch. Pipeline processing for stochastic studies of telescope performance effected by thermo-elastic deformation. In *Photonics West*, volume 6893, pages 68930E–1 – 10. SPIE, 2008.

[49] T. Peacocke. Optical Analysis of the Heterodyne Instrument for the Far Infra-Red: Methods and Results. Master's thesis, Experimental Physics, National University of Ireland Maynooth, 2007.

[50] T. Peacocke. Planck reflectors at operating temperature: Quilted surface form error maps for GRASP models. Industrial contract report, ESTEC TEC-MMO and Experimental Physics, National University of Ireland Maynooth, November 2008.

[51] T. Peacocke. Planck SRFM & PRFM surface fitting results. Industrial contract report, ESTEC TEC-MMO and Experimental Physics, National University of Ireland Maynooth, March 2008.

[52] T. Peacocke. Planck SRFM preliminary surface fitting results. Industrial contract report, ESTEC TEC-MMO and Experimental Physics, National University of Ireland Maynooth, March 2008.

[53] T. Peacocke. Reverse Engineering the Planck Telescope: Phase 4. Industrial contract report, ESTEC TEC-MMO and Experimental Physics, National University of Ireland Maynooth, October 2010.

[54] T. Peacocke and A. Murphy. Self-adapting quadrature method for efficient physical optics propagation of complex field distributions through electrically large reflecting telescopes. In *31st ESA Anetnna Workshop*, volume 2, pages 815 – 818. ESA, ESA Publications Division, May 2009.

[55] K. Pontoppidan, editor. *GRASP9 Reference Manual (9.6).* TICRA Engineering Consultants, Copenhagen, Denmark, 2009.

[56] K. Pontoppidan, editor. *GRASP9 Technical Description.* TICRA Engineering Consultants, Copenhagen, Denmark, 2009.

[57] W. H. Press, S. A. Teukolesky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipies in Fortran 77*, volume 1. Cambridge University Press, 2 edition, 2001.

[58] F. Riesz and B. Szökefalvi-Nagy. *Functional Analysis.* Frederick Ungar Publishing Co., 1955.

[59] H. Samelson. *Notes on Lie algebras.* Universitext. Springer-Verlag, 2 edition, 1990.

[60] A. E. Seigman. *Lasers.* University Science Books, 1986.

[61] J-P Serre. *Lie Algebras and Lie Groups.* Lecture notes in mathematics. Springer-Verlag, 2 edition, 1992.

[62] G. F. Simmons. *Introduction to topology and modern analysis.* McGraw-Hill, 1963.

[63] R. Sippel and S. Lucarelli. Plank Reflectors Mechanical Analysis report. Technical Report Issue 2,, Astrium, June 2002.

[64] E. M. Stein and R. Shakarchi. *Fourier Analysis.* Princeton Lectures in Analysis 1. Princeton University Press, 2003.

[65] J. Tauber *et al.* Planck pre-launch status: The optical system. *Astronomy and Astrophysics*, 520(1), 2010.

[66] Hershel Planck Team. Planck FM telescope assembly ICD. Interface control document, ThalesAlenia Space, April 2008.

[67] Plank Reflector Team. Plank Reflectors ICD for the Secondary Reflector FM. Industry interface control document Issue 4,, EADS Astrium, June 2006.

[68] Plank Reflector Team. Primary Reflector Flight Model. Industry interface control document Issue 4,, EADS Astrium, June 2006.

[69] N. Wiener. *The Fourier integral and certain of its applications.* Dover, 1958.

[70] S. Withington, M. Hobson, and E.S. Campbell. Modal foundations of close-packed optical arrays with particular application to infrared and millimeter-wave astronomy. *J. Applied Physics*, 96:1794 – 1802, 2005.

[71] S. Withington and G. Yassin. Dyadic analysis of partially coherent submillimeter-wave antenna systems. *IEEE Transactions on Antenna and Propagation*, 49(8), August 2001.

[72] S. Withington and G. Yassin. Power coupled between partially coherent vector fields in different states of coherence. *J. Opt. Soc. Am. A*, 18(12), December 2001.

[73] S. Withington and G. Yassin. Analyzing the power coupled between partially coherent waveguide fields in different states of coherence. *Optical Soc. of America*, 19(7):1376 – 1382, July 2002.

[74] N. Young. *An Introduction to Hilbert Space.* Cambridge University Press, 1988.