UNIVERSITY OF LONDON THESIS

Degree PhD Year 2007 Name of Author ERIK PAUL JOHAN DE WITTE

Design and Development of Spherical Array Antennas

Erik De Witte

A thesis submitted for the degree of
Doctor of Philosophy
of the
University of London.

Department of Electronic & Electrical Engineering
University College London

January, 2007

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

Array antennas have been developed that can achieve electronic beamsteering over wide angles. With circular array antennas, a radiation pattern can be steered over 360 azimuthal degrees without pattern deterioration. Some applications may require an omnidirectional beamsteering capability in azimuth as well as in elevation, with antenna characteristics that are independent from the direction of the beam. To achieve this without mechanical actuators requires an antenna that is homeomorphic to the sphere. Spherical array antennas are the 3-dimensional equivalent of circular array antennas and offer a full omnidirectional beamsteering coverage.

Using standard Fourier techniques, the analysis and synthesis of a circular array can be simplified. The technique is known as phase mode theory and allows a number of signal processing techniques such as electronic direction finding to be demonstrated. The symmetry found in spherical array antennas can equally be exploited using spherical harmonics as a basis for Fourier analysis. Spherical array antennas have not enjoyed the same attention as circular arrays, nor have they been the subject, with the exception of SONAR arrays, of a profound study that is equivalent to phase mode theory. This thesis aims to put spherical arrays on the same foot as circular arrays, by using spherical modes to build the theoretical framework for the analysis, synthesis and design of spherical array antennas.

Using a technique for conformal array antennas, an holistic model for the spherical array antenna is derived to analyse the radiation from complex spherical arrays, based on analytically obtained expressions. Mutual coupling effects are quantified and included in the calculations. Also some typical antenna characteristics and design considerations are reviewed in the light of modal theory, in order to obtain design formula's for the spherical array antenna.

Apart from the electromagnetic approach, attention is given to the processing of the signals. Modal theory lends itself for the development of smart antennas, and the foundations to do that with spherical arrays are laid down. A particular problem arises when one wants to distribute the array elements evenly over the spherical surface. This topic is also given an in-depth treatment, for distributions of single elements as well as for more advanced subarraying techniques.

To back up the findings from theory and simulation, as well as to demonstrate the unique capabilities of the spherical array antenna, a prototype array has been designed and partially constructed. The design and development of this antenna is covered to the point where a partially completed spherical array could be measured. Results of these measurements are encouraging and sufficient to show the potential of spherical array antennas. Finally, suggestions are given for future experiments and how to further advance spherical array technology.

Acknowledgements

This PhD programme was funded by the Department of Electronic & Electrical Engineering of UCL, the EPSRC and the Leslie H. Paddle Scholarship from the IEE (now called the IET). It would have been impossible to write this thesis without the much appreciated help and support from a lot of people. I want to acknowledge at least some of them:

I am much obliged to Dr. Paul Brennan, my first supervisor, for his expert advice, his guidance and support, and the confidence he showed towards me. Working with him was not just a fruitful but also an enjoyable experience.

I would also like to thank Prof. Hugh Griffiths, my second supervisor. He supported the idea behind this project since day one, and suggested me to develop it further in the context of a PhD degree. I thank him for keeping to show interest in this work and reading this thesis.

Also Prof. Chris Baker was involved early on. I am grateful for the interest he showed towards this project, but also the general support he has given in his capacity of PhD tutor and head of the research group. I also thank him for the valuable feedback on my thesis.

Prof. Kin-Fai (Kenneth) Tong gave a significant boost to this project with his expert advice on the design and development of patch-type antennas. I thank him for providing his J-shaped antenna as a starting point for the design of the array elements.

Many useful suggestions and valuable remarks were raised by Prof. Ralph Benjamin. His regular visits as a visiting Professor to our Department are much appreciated.

Thanks to Dr. Paul Radmore for his help with some of the maths. Also thanks to Sir D.E.N. Davies, who invented phase mode theory, for the interesting discussion on that matter.

A friendly work environment goes a long way. Many thanks to my dear colleagues and friends from the Microwave, Radar and Optics research group, for the good times we enjoyed in and around (especially around) the Department. I specifically want to mention Richard Bullock for his know-how and hands-on experience in electronics design.

Special thanks go out to Leonidas Marantis. We have worked closely together on the development of LISA and I got a whole lot further thanks to him being there to help. It was a real pleasure to work together. Leonidas will continue working on this project and I wish him all the best!

The world would stop turning without those who provide hardware support. Many thanks to Jim Vansickle and Trevor Hamer for the excellent pieces of work.

Thanks to Mark Eidem and Dr. Carlo Rizzo for the valuable advice regarding the anechoic chamber.

I reserve a warm thank you for my family, especially my parents. Their continued involvement has been a source of both stimulation and joy. Their visits to London were a welcome diversion from all the hard work.

Loving thanks to (Dr.) Katrien Jacobs, my soon-to-be wife. She has been my strength and inspiration and I am utterly grateful for her understanding and support.

Contents

List of Figures

List of Tables

List of Acronyms

ADC	**Analog to Digital Convertor**
AEP	**Array Element Pattern** The radiation pattern obtained when one element of an array antenna is excited and all other elements are not directly excited
ARP	**Annular Ring Patch**
AUT	**Antenna Under Test**
COTS	**Commercial Of The Shelf**
CP	**Circular Polarisation**
CPST	**Circularly Polarised Spherical Transform** The CPST is a modal expansion developed in this thesis to describe the tangential field of a spherical circularly polarised wave.
CPW	**Co-Planar Waveguide**
CST	CST STUDIO SUITE 2006™ version 2006.0.0, a product from CST GmbH
DAC	**Digital to Analog Convertor**
EM	**ElectroMagnetic**
FDTD	**Finite Differential Time Domain**
FEM	**Finite Element Method**
FFT	**Fast Fourier Transform**
FPGA	**Field Programmable Gate Array**
LEO	**Low Earth Orbit(-ing)**
LHCP	**Left Hand Circular Polarisation**
LISA	**LISA Is a Spherical Array** Name of the prototype antenna, the development of which is described in this thesis
LSBR	**Least Squares Boundary Residual**
MATLAB	MATLAB® version 6.5 release 13, a product from The Mathworks,Inc.
MEMS	**MicroElectroMechanical System**
MoM	**Method of Moments** A numerical method to solve an integral equation typically encountered when solving electromagnetic problems. It is explained in more detail in section 3.3.

NOSCIPRA	**Circularly Polarised, Non-Sequentially Coupled, Single-Feed, Annular-Ring Microstrip Antenna** A new type of antenna presented in chapter 3 of this thesis to serve as array element
NP	**North Pole - The point on the unit sphere where** $z = 1$
NRP	**Numerical Refinement Procedure** A specific procedure introduced in chapter 3 to numerically refine the solution for the SWE of a single radiating element on a conducting sphere.
PCB	**Printed Circuit Board**
PEC	**Perfectly Electric Conductor**
RADAR	**RAdio Detection And Ranging**
RCS	**Radar Cross Section**
RF	**Radio Frequency**
RGP	**Realised Gain Pattern**
RHCP	**Right Hand Circular Polarisation**
Rx	**Receive**
SAR	**Shorted Annular Ring**
SFT	**Spherical Fourier Transform** The SFT is a scalar expansion of functions of spherical coordinates into a basis of spherical harmonics.
SMT	**Spherical Modal Theory** Theory to solve problems with the aid of scalar or vectorial spherical modes, as explained in chapter 2 of this thesis
SONAR	**SOund Navigation And Ranging**
SP	**South Pole - The point on the unit sphere where** $z = -1$
STL	**Spherical Transmission Line**
SWE	**Spherical Wave Expansion** A modal expansion of spherical waves, for which the modes obey the Maxwell equations
SWG	**Spherical WaveGuide**
Tx	**Transmit**

Definition of Terms

archimedian solid Semi-regular convex polyhedron, with two or more types of congruent regular convex faces. See also appendix A. on page: 18, 19, 215, 216, 218, 219, 221–223

array RGP The array RGP is the superposition of all element RGPs according to the free excitation. on page: 128, 131, 262, 263, 273

available spatial bandwidth The highest degree of spherical modes that can be used to synthesise a radiation pattern while avoiding aliasing effects. Notation: B_s on page: 23, 162, 177, 197

average available dimension The average available element dimensions is an approximate measure for the maximum diameter of a number of circular elements on the sphere. Notation: Aed on page: 23, 205

average element density Number of radiating elements on the sphere divided by its surface area in square wavelengths. Notation: d on page: 23, 163

axial ratio A measure for polarisation purity, formally defined in equation (2.57) on page: 18, 58, 137, 249, 252, 258

backprojection Using data of a spherical wavefront to calculate the wavefront at a smaller radius on page: 168, 169, 247

boresight The polar direction from the antenna viewpoint, typically the normal direction to a planar antenna on page: 82, 258

boresight on-axis AR The axial ratio of an antenna, typically shown as a function of frequency, obtained from measurement data that are gathered while the AUT revolves around its boresight axis. The axial ratio is calculated from the parameters of a sinusoidal function that is fitted to the gathered data. on page: 258

broadside The equatorial direction from the antenna viewpoint, typically a direction in the plane of a planar antenna on page: 82

catalan solid A dual solid of an archimedian solid. See also appendix A. on page: 216, 221–223

circumradius The radius of the circumscribed sphere on page: 18, 221

CP bandwidth Operational bandwidth for which the CP is acceptable. An often used criterion is an upper bound of 3 dB for the axial ratio on page: 100, 102, 105, 108, 233, 252

dihedral angle The angle between two planes, as specified by the angle between their normals. on page: 225, 226

directivity The directivity $D(\theta, \phi)$ of an antenna is a function describing the spatial distribution of power radiated by a lossless antenna. It is defined in section 2.4.2. See also gain. on page: 18, 56, 234, 256, 262

distribution bandwidth A function of spatial bandwidth B can be reconstructed exactly from the samples of a sampling distribution of bandwidth B. Defined on page 174. on page: 18, 19, 177, 195, 197, 199, 201

distribution efficiency A measure for how many spherical modes can be exactly reconstructed from and relative to a number of samples. Notation: η_d on page: 23, 163, 164, 177, 178, 200

dual solid The solid obtained by replacing the vertices of a solid with faces and vice versa. More precisely defined on page 221. on page: 18, 215, 216, 219, 221–223

element RGP The radiation pattern obtained when one element of an array antenna is excited and all other elements are disconnected from their feedlines and match terminated on page: 18, 128, 131, 262, 263

equiangular distribution A distribution of points on the sphere on a (θ, ϕ)-grid with uniform angles on page: 195, 201, 202, 212, 302

evanescent An evanescent wave decays rapidly (exponentially) and does not carry any real power in a lossless medium (unless power is extracted from it). It exists typically to satisfy boundary conditions around a discontinuity in an electromagnetic environment. on page: 51, 52, 54, 61, 73, 78, 86, 127, 171

excitation A function or vector that describes the phases and amplitudes of the feeding signals to the elements of an array antenna. More specific definitions are excitation vector and free excitation. on page: 19, 262, 271, 275, 277, 297

excitation vector The set of user-generated incident signals to the element terminals of an array antenna, without taking into account mutual coupling effects. Also called steering vector. on page: 18, 128, 131, 252

extremal system A fundamental system that has been optimised so that detG is maximal. on page: 204

face-uniform Property of a polyhedron with all its faces congruent on page: 216, 219

free excitation The set of incident signals that carry real power to the elements of an array antenna as a result of the excitation vector and both external and internal mutual coupling on page: 18, 122, 128, 131, 252

fundamental system A sampling distribution with an equal number of elements as a minimal design and a regular Gram matrix G. on page: 18, 202, 204

gain The gain $G(\theta, \phi)$ of an antenna is related to the directivity, but also takes losses into account. on page: 18, 26, 38, 205, 230, 233, 234

geodesic A geodesic is the shortest path between two points over the surface of a sphere. A geodesic construction is a family of approximately spherical constructions popularised by Buckminster Fuller, that is based on consecutive subdivision and radial projection of triangles. on page: 39, 121, 122, 214, 221, 223

hypersampling Sampling on a sampling distribution that can mathematically be proven to allow Nyquist sampling. All known hypersampling distributions are also oversampling distributions. on page: 18, 201, 202

impedance bandwidth Operational bandwidth for which the return loss is acceptable on page: 101, 104, 252

inradius The radius of the inscribed sphere on page: 18, 221, 223

insertion loss Ratio of voltage amplitudes of the transmitted and incident waves from one port to another of a microwave network. Usually given in dB and denoted by S_{ij}, where i and j are the port numbers. See also return loss on page: 19, 105

integration strength A sampling distribution has an integration strength t if it also a t-design. on page: 219, 223

microwave network A black-box description of a microwave component such as an antenna, filter or amplifier, that is fully determined by a scattering matrix on page: 18, 19, 125

midradius The geometric mean of the inradius and the circumradius. See also equation (8.17). on page: 222

minimal design A sampling distribution with an unattainable minimum number of sampling points as defined in equation (8.7), that allows an exact quadrature rule up to a degree of twice the distribution bandwidth. See also t-design. on page: 18, 19, 199, 204, 234

multipole expansion An expansion that is suitable to describe vector fields, as opposed to the scalar or monopole expansion on page: 41, 61, 200

Nyquist sampling Sampling a function of spatial bandwidth B on a sampling distribution with sampling degree 2B. See also hypersampling and oversampling. on page: 18, 201

oversampling Sampling on a denser sampling distribution than is required for Nyquist sampling. See also hypersampling. on page: 18, 201

phase mode A term of the Fourier series, often used in the context of circular array antennas to eg. synthesise an excitation on page: 26, 28, 33, 37, 42

platonic solid Regular convex polyhedron, with congruent regular convex faces. See also appendix A. on page: 200, 215, 216, 221–223, 230, 276

polarisation A single component of a vector field, according to a suitably chosen vectorial basis on page: 18, 23, 54, 56

quadrature rule A weighted summation of a discrete number of sample points that serves as a numerical approximation to an integral. See also t-design. on page: 18, 19, 199

regular vs singular A matrix $\mathbf{A}$ is regular iff $\det \mathbf{A} \neq 0$ and singular iff $\det \mathbf{A} = 0$ on page: 18, 204

return loss Ratio of voltage amplitudes of the reflected and incident waves from and to a port of a microwave network. Usually given in dB and denoted by S_{ii}, where i is the port number. See also insertion loss on page: 18, 249, 252

sampling degree Twice the distribution bandwidth. Notation: L_s on page: 18, 23, 177, 195, 197, 201, 276

sampling distribution A mathematical tool to describe the action of sampling. Defined in section 7.2. on page: 18, 19, 177, 276

sampling efficiency A measure for how directive the radiation pattern of an array of discrete antennas can be, relative to a continuous aperture of the same size. Notation: η_s on page: 23, 162, 201

scattering matrix A matrix with all scattering coefficients between the ports of a microwave network on page: 18, 121, 125, 252

semi-regular Synonymous to archimedian on page: 219, 221, 223, 228

sidelobes Peaks in the radiation pattern of an antenna, that are separated from the main beam by one or more crests. on page: 149, 153, 166, 262, 277

spatial bandwidth The highest degree of spherical mode in the SFT or SWE of a function. on page: 18, 19, 172, 177, 276

spherical harmonics Doubly parameterised scalar functions of the two angular spherical coordinates that form an orthogonal and complete basis for all complex functions on the sphere. They are formally defined in equation (2.12). on page: 16, 23, 27, 41–43, 45, 112, 147, 172, 179, 200, 203, 235, 298

spinning SFT An alternative for the SFT, defined in equation (2.60) on page: 59, 61

STL expansion A modal expansion based on an STL interpretation of free space on page: 52, 61, 111, 112, 170, 184

SWG expansion A modal expansion based on an SWG interpretation of free space on page: 23, 61, 170, 171

t-design A sampling distribution that facilitates an equal-weight and exact quadrature rule for functions on the sphere with a spatial bandwidth up to degree t. See also minimal design. on page: 18, 19, 199, 219, 230, 276

TE and TM modes Transverse electric modes (or magnetic modes) have no electric field component in the direction of propagation. Transverse magnetic modes (or electric modes) have no magnetic field component in the direction of propagation. See also TE_{lm}. on page: 50, 51, 64, 65, 76, 105, 136, 233

uniform resolution If a function can be perfectly reconstructed from its Fourier expansion coefficients in one direction, than the same is true in any other direction. Therefore, if an array antenna has a sufficiently large number of elements, a radiation pattern can just as accurately be synthesised in any direction. on page: 198, 200, 214, 263

Notations and Conventions

$\mathbf{A}$	Notation for matrices
$\vec{A}$	Notation for vector fields
$\hat{e}_x$	Notation for a field of unity vectors, that is part of a basis for polarisation
$\overline{\mathbf{A}}$	Notation for the complex conjugate of $\mathbf{A}$
$\mathbf{A}^T$	Notation for the transpose of $\mathbf{A}$
$\mathbf{A}^*$	$\overline{\mathbf{A}}^T$
Aed	See average available dimension.
B_s	Notation for available spatial bandwidth. See also L_s and L_m.
d	See average element density.
η	Electromagnetic impedance
η_d	See distribution efficiency.
η_s	See sampling efficiency.
(ξ, θ, ϕ)	The z-y-z convention is used for the **Euler angles** throughout this text, with Euler angles ξ, θ and ϕ illustrated on page 113.
$e^{-j\omega t}$	The time convention in this thesis is such that at any observation point, the phase evolution is described by the factor $e^{-j\omega t}$.
$\hat{f}$	The Spherical Fourier Transform (SFT) of f.
$\{\vec{F}\}_\mathbf{T}$	The tangential part of field $\vec{F}$.
$\hat{f}$	The SFT of f
G	The aliasing function, defined on page 204.
$\mathbf{G}$	Gram matrix $\mathbf{G} = (\mathbf{A}\mathbf{A}^*)$
$h_l^{(i)}(r)$	Spherical Hankel function of the first ($i = 1$) or second ($i = 2$) kind. Without a subscript, the first kind is assumed: $h_l = h_l^{(1)}$
$\mathbf{I_n}$	The identity matrix of size n, defined as the n-by-n square matrix with ones on the main diagonal and zeros elsewhere.
$L^2(S)$	The Hilbert space of square integrable functions on the domain S
L_m	The highest degree of spherical modes that can propagate to the far-field, from a spherical surface of a given size. See also B_s and L_s.
l, m	l and m are used to denote respectively the degree and the order of the spherical harmonics.
L_s	Notation for sampling degree. See also B_s and L_m.
N	Number of radiating elements in an array antenna

N_e	See N
N_l^m	Normalisation for the spherical harmonics, defined in equation (2.13)
N_{lm}	Normalisation for the SWG expansion, defined in equation (2.35)
N_s	Number of subarrays in an array antenna
ω	A point (θ, ϕ) on the sphere
p, n	p and n are used to denote respectively the radial and the azimuthal cycles in a discrete set of circular modes.
R	Unless defined otherwise, R is the radius of a sphere.
$\Re(z)$	Real part of complex number z.
S^2	The 2-sphere is a two-dimensional surface in a three-dimensional euclidean space, for which all points are equidistant to the origin. Often just called "the sphere".
$SO(2)$	The special orthogonal group of rotations in two dimensions
$SO(3)$	The special orthogonal group of rotations in three dimensions
$\mathbf{TE}_{lm}$	For spherical modes TE_{lm}, the azimuthal parameter m comes second in the notation, whereas for circular modes TE_{np} the azimuthal parameter n comes first. See also l, m and p, n.
$z_l(r)$	The general solution for the radial variation of a spherical wave, as defined in equation (2.19).

Chapter 1

Introduction

1.1 Spherical Array Antennas

What are spherical array antennas and why do we need them? We will answer these questions first, to provide a backdrop against which to place the research reported in this thesis. We will therefore start with a formal definition of spherical array antennas, and describe their context in terms of other efforts that have been made to develop spherical array antennas and applications where they can, have or will be used.

1.1.1 Definition

Spherical array antennas are array antennas that are conformal to a spherical surface. In the IEEE Standard Definitions of Terms for Antennas (IEEE Std. 145-1993) they are defined as follows:

> **2.375 spherical array.** A two-dimensional array of elements whose corresponding points lie on a spherical surface.

To understand this definition, it is necessary to understand what is meant by an array of elements. The IEEE Standard Definitions of Terms for Antennas defines an array antenna as:

> **2.27 array antenna.** An antenna comprised of a number of identical radiating elements in a regular arrangement and excited to obtain a prescribed radiation pattern.
> - The regular arrangements possible include ones in which the elements can be made congruent by simple translation or rotation.
> - This term is sometimes applied to cases where the elements are not identical or arranged in a regular fashion. For those cases qualifiers shall be added to distinguish from the usage implied in this definition. For example, if the elements are randomly located, one may use the term random array antenna.

Spherical arrays are often classified as conformal arrays. The IEEE definition for conformal arrays is:

> **2.74 conformal array.** An array that conforms to a surface whose shape is determined by considerations other than electromagnetic; for example, aerodynamic or hydrodynamic.

According to this definition, spherical arrays can be conformal arrays or not, depending on the reason they were chosen to be used in a particular application. In fact, in this thesis it will be shown that there are good electromagnetic reasons to use spherical array antennas, even in the absence of any other requirements of conformity. However, the general consensus seems to be that spherical arrays do classify as conformal array antennas[1].

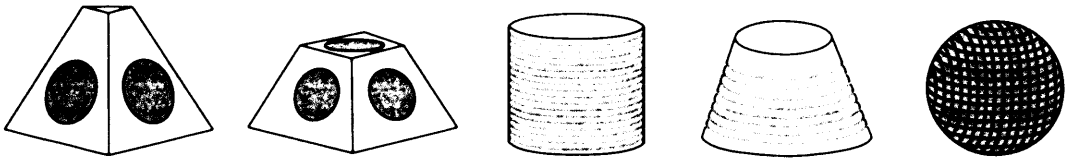
1.1.2 Context

Array antennas in general have been used widely for applications where any or several of the following requirements are to be met:

- Fast electronic beamsteering
- Easy reconfiguration of the antenna
- Adaptive signal processing for radiation pattern synthesis
- Direction finding
- Robustness and protection against environmental influences
- Conformal design
- Multiple independently controlled beams

In particular, array antennas of a simple geometry, such as the planar array antenna, have received enormous attention over a 20 year span from the 50s through to the 70s. As technology advanced, and with the advent of integrated circuit (IC) and digital technology, electronically steered array antennas have made their way to a well matured technology that can be found in applications ranging from high-tech naval combat radars to mass products such as mobile phone base stations.

Linear or planar array antennas do however suffer from a key limitation: as their radiation pattern is steered away from the boresight direction, the radiation pattern deteriorates and the antenna gain is reduced. The angular coverage of electronic beam steering with linear or planar arrays can therefore not be achieved for more than 60° off the boresight direction[1]. An array that does not suffer from this limitation is the circular array antenna. The circular array antenna can steer its beam through 360° with little change in beamwidth or sidelobe levels[2]. Although circular antennas were presented in 1936[3], they did receive far less attention then their linear or planar counterparts. Significant developments on circular array antennas were made in the 60's, when the concept of phase modes was introduced[4][2]. The concept of phase modes is a central starting point for this work. A phase mode is a term of the Fourier series expansion of the excitation of a circular array, and the far-field radiation pattern corresponding to each phase mode has the same characteristic variation of phase with direction, maintained over potentially very broad instantaneous bandwidth[4][2]. Since the advent of phase mode theory, circular array antennas have enjoyed a steady amount of attention over the past several decades, and



3-face planar array (a) 5-face planar array (b) cylindrical array (c) conical array (d) spherical array (e)

Figure 1.1: Arrays with hemispherical coverage of a steerable beam need either a combination of multiple planar arrays as shown in (a) and (b), a singly curved surface as in (c) and (d), or a doubly curved surface as in (e). Configurations (a), (b), (c) and (d) suffer from pattern deterioration as the beam is scanned away from one of the boresight directions.

Table 1.1: Comparison of hemispherical array antennas with multiple faces, based on the calculations in [10]. The performance characteristics become better as the number of faces is increased. The limiting case is a spherical array antenna.

Number of Faces	3	4	5	6
Max. Scan Angle	63°	55°	47°	41°
Element Spacing (avoiding grating lobes)	0.628λ	0.691λ	0.679λ	0.7λ
Max. Scan Aperture Reflection	14%	7%	4%	2%
Max. Scan Beamwidth Broadening	222%	175%	147%	132%
Max. Scan Gain loss	4.1 dB	2.8 dB	1.9 dB	1.3 dB

their properties have allowed a number of applications to be demonstrated, including broadband pattern synthesis, null steering, direction finding and superresolution[5][6][7][8][9].

When applications do not only require beam steering capability over the azimuthal angle but also in elevation, for instance to cover the full hemisphere, the form of the antenna will need to extend in the third dimension, as illustrated in figure 1.1. One way of doing this is by combining a number of planar array antennas on the faces of a 3-dimensional body. This approach is discussed in [10]. In the referenced paper, a comparison is made between array antennas with different numbers of faces. The antennas are designed to achieve hemispherical coverage, and their performance characteristics are estimated based on a few simplifying assumptions and design formula's. The minimum number of planar faces to achieve hemispherical coverage is three. In that case, each planar array needs to scan its beam to a maximum of 63° degrees off the boresight direction, which causes a broadening of the beamwidth with a resulting loss of gain. In addition, increased mutual coupling effects between the array elements at large scan angles will cause a further gain reduction. If more planar faces are used, these effects become less severe. Table 1.1 resumes the calculations in [10]. It can be seen that as more faces are used, less scanning is required from the planar arrays which results in less scanning loss. It is also noticed that on average, there is a slight reduction in the number of elements that is required to avoid grating lobes. As the spherical array antenna is the limiting case for an increasing number of array faces, it has arguably the optimal shape to cover the hemisphere or even the entire spherical surrounding of the antenna with an electronically steerable beam.

The ability of the spherical array antenna to keep its radiation pattern invariant with respect to scan angle led to several studies of spherical array antennas near the end of the

60's[11][12][13][14][15]. For systems requiring wide-angle coverage with uniform beams, it was found that spherical arrays have distinct advantages over their cylindrical or multi-planar counterparts. Other unique properties of the spherical array were soon discovered. The relation between the element distribution and grating lobes was not entirely understood yet, but in [12] a quasi-uniform distribution was proposed to place 162 elements on the sphere to suppress grating lobes. It was also found that the spherical array can maintain circular polarisation in all directions. Large numbers of elements can be placed on the sphere with nearly uniform spacings. These spacings can approach a wavelength before grating lobes become objectionable. Sidelobe levels are a little higher - levels of -22 dB are found to be attainable[11]. The analytical tools used to study the performance and operation of spherical arrays remained fairly simple. The far-field of the array was numerically calculated from a far-field approximation of a summation of scalar contributions from sources on a spherical surface, as presented in [16]. In [17] an expansion of the far-field was calculated for the simplified specific case of dipole rings on the sphere, making use of an interesting recursive calculation technique. Simplified element patterns were sometimes taken into account[12][14], but it remained difficult to accurately take polarisation into account. Simple beam cophasal excitations (definition can be found in [9]) were used to obtain maximal directivity patterns. In [18], there was an early mentioning of spherical harmonics for the calculation of the element density of spherical arrays.

As had been the case for circular array antennas, there was only a marginal interest in building practical applications with spherical array antennas, mainly due to the complexity of a potential implementation as well as the perceived mathematical difficulty to analyse them thoroughly. One extensive design study called ESSA (Electronically Switched Spherical Array) was carried out for NASA by the Ball Brothers Research Corporation[19]. A simplified approach was taken by electronic switching of subarrays of high-gain patches on the sphere. The array was built and radiation patterns were obtained empirically. A spherical array of monopoles has been investigated in [20]. Interestingly, the polarisation issue was avoided by using radial monopoles.

In parallel, the electromagnetics community has been extending its use of modal theory, mainly to analyse the radiation from cylindrical array antennas. A systematic adoption of a modal methodology in all aspects of the analysis, synthesis, design and development of spherical array antennas has however not yet been taken, despite the fact that phase mode theory has previously revolutionised the applicability of circular array antennas. Taking this holistic modal approach towards spherical array antennas is the central idea of this thesis. Section 1.2 will expand further on this matter and will briefly review the methods for analysis of antenna arrays.

During this PhD programme it emerged that there are several ongoing research efforts related to spherical array antennas. While working in parallel, it seems that each of them is concentrating on a different aspect of spherical array antennas:

- A group around Dr. Zvonimir Sipus from the University of Zagreb, Croatia have developed code using Galerkin's Method of Moments (MoM) with the spherical modes as basis functions. They have used this code to analyse the radiation and input impedance of curved patch type antennas above a spherical groundplane[21][22][23]. They reported

Figure 1.2: A small spherical array demonstrator, as presented in [24]. It is interesting to note that without mutual awareness, their design approach for the subarrays has converged to a solution that is very similar to the design of our prototype spherical array. Illustration used with kind permission from Dr. Peter Knott from the Research Institute for High Frequency Physics and Radar (FGAN-FHR) in Germany.

a large sensitivity of the resonant frequency on the patch radius of curvature, which is a difficult parameter to control.

- Dr. Peter Knott from the Research Institute for High Frequency Physics and Radar (FGAN-FHR) in Germany has developed a very elegant spherical array antenna that was presented at EUCAP 2006[24]. Interestingly, their design efforts seem to have converged independently towards a very similar subarraying solution as the one we use for our prototype spherical array antenna. So far, the focus of their work has been on building a conformal array demonstrator, without developing analysis tools or beamforming capabilities. The antenna is shown in figure 1.2.

- Mingda Huang and Soon Yim Tan from the School of Electrical and Electronic Engineering at Nanyang Technological University in Singapore have also been working on modal signal processing with spherical array antennas[25][26]. They have also extended the concept to spheroidal phase modes[27]. Their work has built further on the spherical harmonic phase modes presented in [28]. These scalar modes cannot be used for accurate analysis or synthesis of electromagnetic fields, but are sufficient to develop signal processing concepts.

1.1.3 Applications of Today and Tomorrow

Spherical arrays are most commonly found in SONAR applications (see figure 1.3). Spherical modal theory is being introduced in the signal processing of SONAR arrays, which is focused on ranging and direction finding[29]. Perhaps one of the reasons why spherical arrays have found more applications in sonar than in radar for instance, is that the sound waves in sonar applications have a scalar nature, which simplifies the problem at a fundamental level compared to electromagnetic antenna applications.

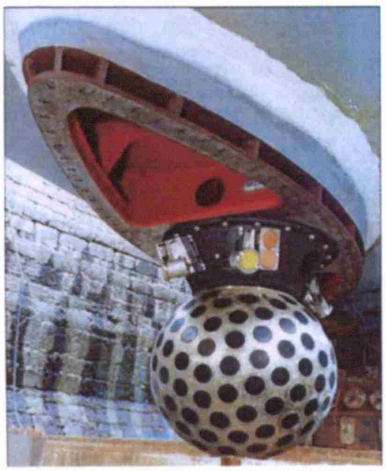


Figure 1.3: A spherical array for a SONAR application. SONAR works with acoustic waves, so that the pattern from a SONAR array can be described with scalar functions. Although the velocity of sound is much slower than the propagation velocity of an electromagnetic wave, the used wavelengths are often comparable. The size of the array of pressure sensors shown above is therefore representative for what could be expected to be the size of a spherical antenna array for RADAR applications. Courtesy of Thales

For electromagnetic antenna applications, very few spherical arrays ever made it past the preliminary design stage. The increased complexity of the spherical array, in part due to polarisation requirements, may explain in part their unpopularity. The past few years however, have seen a renewed interest in spherical array antennas and their applications, as will be illustrated in the following paragraphs.

Perhaps the most typical applications to benefit from the omnidirectional beamsteering capabilities of spherical array antennas are to be found in satellite communications. Communication with LEO satellites requires a rapid beam-steering capability, but the mechanical steering of a sometimes large reflector antenna may be undesirable. Some of the problems associated with mechanical steering are the vibrations from the mechanical subsystem, the outspoken profile of the reflector dishes and most importantly, its slow speed[30]. This project originally grew out of an investigation to replace a mechanically steered reflector antenna from the ground segment of the PROBA satellite with an electronically steered array antenna.

The Air Force Research Laboratory (AFRL) of the United States is currently developing a large spherical array antenna for the next generation of satellite ground stations of the Air Force Satellite Control Network (AFSCN)[31][32][33]. It is foreseen that large spherical array antennas will take over from the mechanically steered reflector antennas to improve operating efficiency and mission capabilities. Reflector antennas have a relatively low initial cost, but their operation and maintenance cost is high. They are susceptible to single point of failure, have a low antenna utilisation and are inflexible towards new mission and operational requirements[33]. Phased array antennas offer a solution to these problems, but are traditionally a lot more expensive. However, the rapid expansion of the wireless communications industry has rendered the cost of components such as low noise amplifiers, phase shifters, connectors, etc. less expensive. Together with the better operational efficiency, it is thought that phased array antennas

are emerging as a commercially viable alternative for mechanically steered reflector antennas in satellite operations applications[33]. The spherical array antenna under development at the AFRL will combine 675 flat panel facets into a geodesic construction, which further reduces the overall cost. Each panel integrates 78 broadband radiating elements and supports two independent Tx beams and two independent Rx beams. The radiating elements are of the double-stacked patch type. The antenna will have a $10m$ diameter. An active sector of 120° degrees will be commuted on the sphere. From the same authors some design considerations for spherical array antennas were presented in [34]. In the paper it is confirmed that spherical array antennas are optimal if scan angles of more than 60° are required and that spherical arrays can use up to 50% fewer elements than a multi-planar array with the same capabilities. The derivations in the paper could have benefitted from a modal approach. The simple proposed beam steering technique of simple on-off switching of array sectors may offer a low-cost solution but seems a little outdated in the light of smarter modal processing techniques.

In [35] the military need is reviewed for a technology that provides small tactical platforms with a greatly enhanced Beyond Line Of Sight (BLOS) communications capability. Currently most UK military BLOS communications to such platforms is provided by HF radio. UHF satcom capabilities are limited by capacity and solutions in the SHF X-band need to be sought. The data rates require a high-gain antenna that is capable of keeping its beam pointed towards the satellite. A small spherical array antenna has been developed at Roke Manor Research to achieve this[36]. The antenna system is shown in figure 1.4. Direction finding is achieved with an open-loop system using inertial navigation sensors so that the antenna can compensate for the movement of the platform. Steering is based on a complicated scheme of phase compensation, as a function of element position and orientation in the array antenna, a strategy that could be greatly simplified with the aid of modal processing. One cause of error in the system is attributable to the use of a magnetometer in armoured vehicles. A solution to this problem could be offered by a spherical array that is capable of real-time electronic direction finding at signal processing level.

Not only the ground platform from the satellite communication link may benefit from spherical array antennas. Another potentially interesting application comes from smaller LEO micro-satellites. One of the tasks of the attitude control subsystem of a satellite is to make sure that its antenna will always point towards the ground station. A small satellite could have a spherical array antenna on its exterior, that can at all times electronically steer its beam towards the ground station. Small satellites with a low power consumption may not have requirements on the attitude control from the power subsystem (to keep a solar array facing the sun) and could be implemented without any attitude control subsystem, resulting in a saving on the mass budget.

Sometimes applications emerge where one least expects them. One such example is the hemi-spherical wideband antenna array for breast cancer imaging[37]. A conformal array antenna is being developed for the detection of breast cancer with non-ionising microwaves, much in the same way as ground penetrating radars are being used for landmine detection. The dielectric properties of tumour tissue are sufficiently different from healthy tissue, for a tumour to be detected in the reflected wave. The concave hemispherical array, developed at the University of

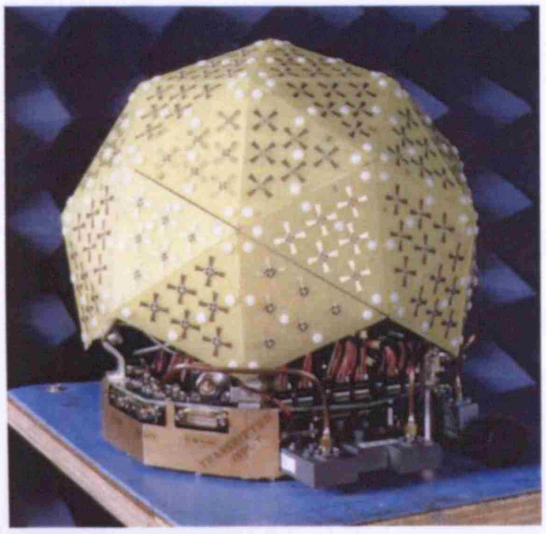


Figure 1.4: A small spherical array for mobile satellite communications. Illustration used with kind permission from Bryan Rickett from Roke Manor Research Ltd.

Bristol in the UK is shown in figure 1.5.

In the future, spherical phased array antennas may find their way into RADAR applications. The SAMPSON multi-function radar (MFR) developed by AMS is a state-of-the-art naval radar system[38]. It employs two enormous planar array faces that are mounted on a mechanically rotating stand inside a spherical housing. It is a good example of a system where spherical array antennas can be deployed to get rid of the mechanical actuators. For this kind of applications, spherical arrays may not offer an economically viable alternative today, but they might well do in the future.

In [39] an assessment is published from the Jet Propulsion Laboratory (JPL) on the long-term capability of the antennas for the Deep Space Network (DSN). Various alternatives for the present 70-meter antennas are considered in the report. The report concludes that in the next 10 years, phased array antennas will become an economically attractive option. Phased array antennas would offer a high beam agility, multi-beam multi-target applications, high reliability and easy maintainability. Disadvantages include a much higher cost at present. The array is expected to be implemented as an array of smaller flat panel modules. Regarding the candidature of the spherical array to govern the overall geometry, the report states that these are in an exploratory stage and that further research is necessary.

1.2 Methodology

The research undertaken in this thesis is as much about the methodology itself as it is about its subject of study. The methodology presented and used in this thesis to investigate spherical array antennas takes a central place.

Figure 1.5: A concave spherical array prototype for breast cancer detection. Illustration used with kind permission from Dr. Ian Craddock from the Dept. Electrical & Electronic Engineering of the University of Bristol.

1.2.1 Modal Theory as the Central Driver

In section 1.3.1 the hypotheses are stated that lay at the start of this project. From the formulations of the hypotheses it can be expected that the research work reported in this thesis did not only target an improved understanding of spherical array antennas, it aimed more specifically at an improved understanding of spherical array antennas by developing a methodology based on spherical modal theory.

Chapter 2 of this thesis contains an in-depth treatment of spherical modal theory and develops the theoretical framework for the other chapters of this thesis. In short, modal theory is applied here to simplify 3-dimensional radiation problems with a periodic boundary condition. The periodic boundary condition is generated by the presence of a symmetrical body such as a sphere, cylinder or cone and it allows a Fourier analysis of the field that exists in the solution domain. Thereby the original problem gets rewritten from a continuous 3-dimensional formulation into a discrete spectrum of (vectorial) basisfunctions. The method thus starts with writing the unknown field that is the solution to the problem at hand as a linear combination of spectral modes, called the spherical wave expansion. The problem is solved when the coefficients of the expansion are known. Sometimes it is possible to obtain analytical expressions for the coefficients, so that numerical methods to find the solution are avoided.

The rationale for taking this approach can be summarised as follows:

- The periodicity of a the sphere must somehow be exploited to simplify whatever problem we want to solve about spherical array antennas. This is the concept of Fourier analysis and is the role of modal theory.

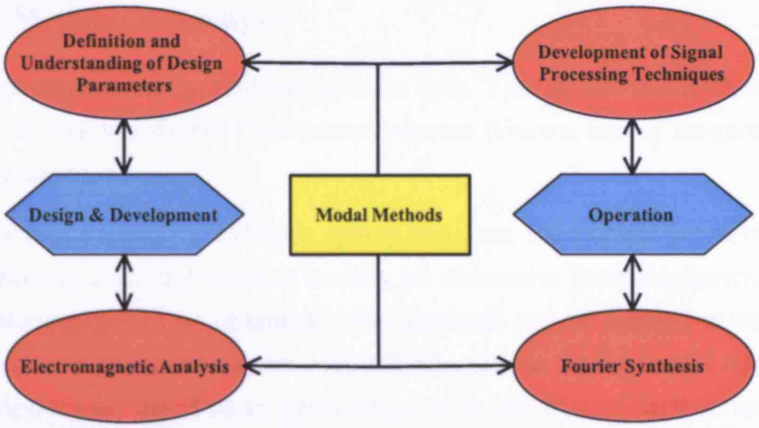


Figure 1.6: Between the first concept or idea of a spherical array antenna and an operational system lie a vast amount of tasks and studies that need to be addressed. We have structured these tasks in four categories, shown as red ovals in the diagram above. Firstly, we can relate the categories vertically, two by two. The two categories on the left contribute mostly to the design and development of the hardware and RF subsystems, by predicting the performance of a design even before it is being build. Tasks in the categories on the right aim to give answers to the question on how to excite the array so that it radiates according to some desired pattern. This knowledge is required to operate the antenna. Secondly, we consider the horizontal relationships. The two groups below are fundamental and recurring problems in the field of electromagnetics. Analysis of an antenna solves its response in the far-field to a known excitation. Synthesis of an antenna solves the inverse problem. The groups on top have to do with the creation of a deeper understanding of the spherical array characteristics, as opposed to merely interpreting the results from computer simulations that were built with the techniques in the groups below. In this thesis, modal theory will be be applied to simplify all of these tasks. Part of the scope of this project (see section 1.3.2) is to build a prototype antenna to demonstrate some capabilities of the spherical array and to develop signal processing techniques for it. Therefore, more attention has been given to tasks that fit in the left categories. Nevertheless, the foundations have also been laid for a number of tasks that fit in the categories on the right.

- Phase mode theory has been applied with great success in circular array antennas. Long after the method was first introduced, new applications and derivations were developed based on this theory.

- Taking an analytical approach to the analysis of spherical array antennas deepens our understanding of the electromagnetic interactions that govern their radiation properties.

- As reported in section 1.1, spherical antenna arrays have been studied with other methods in the past. For some time, the sounds of research on spherical array antennas have quietened and not many applications have ever seen the light. Recently there seems to be a renowned interest in applications of spherical array antennas. It is hoped that by placing the study of spherical array antennas in a new context, an impetus will be given to the development of new techniques and methods for spherical array applications.

As is generally the case with conformal array antennas, the study of spherical array antennas can be broken down into several smaller areas of study, as illustrated in figure 1.6. Spherical modal theory takes a central place in this work because it will be applied to solve problems belonging to each of these smaller subdomains.

1.2.2 Other Methods of Analysis

This section concentrates on the analysis problem only. That is, find the electromagnetic field, (extending all the way to infinity) from known sources (electric and/or magnetic currents) on the surface of an antenna.

For conformal arrays such as the cylinder, sphere and cone, the exterior problem can be solved with modal methods. Although there is usually no alternative for the spherical modal theory if analytical solutions are to be obtained, other methods can be applied to find numerically obtained solutions. In this section some common alternatives for the modal method are listed. The brief descriptions are based on an overview given in [1]. More details on and references to the described methods can be found in the same reference.

A whole family of methods can be distinguished by their ability to cope with electrically large and complex structures. They are the so-called asymptotic methods, because they use a high-frequency approximation to find a solution. Perhaps the best well known asymptotic method is the uniform theory of diffraction (UTD). UTD combines concepts such as phase, polarisation and diffraction in an otherwise classical optical ray approach. The field is assumed to propagate along rays that are either straight lines in free space or geodesic curves along metal surfaces.

If the antenna is not too large and too complex, the high-frequency approximation is not necessary to deal with the structure. Any of the following methods can be used:

- Method of Moments (MoM). This is a frequency domain solver for the integral equation (IE). The IE typically expresses an equality between a tangential incident field and a tangential scattered field on a conducting boundary (the incident field is caused by known sources). The scattered field originates from unknown surface currents on that boundary and can be written as a linear operator on those currents. The operator is an integration over the product of the source terms with a Green's function that incorporates geometrical information of the problem at hand. The MoM turns the IE into a matrix equation so that the unknowns can be found from a matrix inversion. More details can be found on page 75 of this thesis.

- Finite Differential Time Domain (FDTD). The FDTD discretises the domain with a fine hexahedral mesh. It then applies the differential formulation of Maxwell's equations on the edges of the formed cuboids. A broadband time-domain excitation signal is propagated through the structure and the time-domain response is recorded. An inverse Fourier transform will then give a potentially broadband frequency response. This method has gained popularity over the past few years because it scales better with mesh size compared to some other solvers that use discretisation and because it is efficient for broadband analysis. It is very memory hungry and the meshing requirements make it less suitable to analyse large conformal array antennas.

- Finite Element Method (FEM). The FEM also finds the solution to the differential equation (DE) that describes wave propagation. The solution is determined by this equation,

together with the boundary conditions. The FEM discretises the domain in smaller subregions (elements) on which the solution can locally be approximated with a simple function such as a first-order polynomial. It forms a system of equations by imposing the boundary conditions on all external and internal boundaries. That system is then solved. A limitation is that radiation from the external boundary cannot be calculated. The calculation domain must therefore extend far from the structure or be fitted with absorbing boundaries. Conformal array antennas are often too large and complex for this method. It is however, a good candidate to combine with other methods to form a hybrid method. An idea on this, specific for spherical arrays, is given in chapter 10.2 on future work.

- Unit Cell Approach. This method has often been used to analyse cylindrical arrays. It relies on a periodicity of the array antenna elements. For spherical arrays it would therefore impose requirements on the placement of the elements that may not be acceptable. More information on the distribution of radiating elements on the sphere is given in chapter 8.

The antenna under investigation in this thesis is a large spherical array antenna. The ratio of its largest dimension to the dimension of its smallest detail is big, which excludes the EM solvers that require discretisation of the solution domain. The MoM, especially Galerkin's formulation together with spherical modal basisfunctions (see page 75), seems to offer a good numerical approach. Also the UTD could be used, but is less attractive in the contextual approach taken in this thesis, as explained in section 1.2.1.

1.3 Essence of this Thesis

1.3.1 Hypothesis

The hypotheses that lay at the start of this work are listed hereafter. This thesis reports the steps that have been taken towards an acceptance or rejection of the hypotheses below and on the outcome of the research that was necessary to investigate these hypotheses. Some parts have remained standing as an hypothesis when this project drew to a closure and will be picked up in chapter 10.2 on future work. Where appropriate, the answers that have led to an acceptance or rejection of one or more of the following conjectures will be formulated in the closing sections of each chapter.

- By using, and where necessary extending, Spherical Modal Theory (SMT), the electromagnetic analysis of spherical antennas can be simplified. The simpler description can be used in a computer model to simulate the radiation properties of a large spherical array antenna. Because the model is based on analytically obtained formula's, the simulations are faster than can be achieved with numerical EM solvers.
- Spherical Modal Theory (SMT) allows an accurate integrated approach towards the analysis and synthesis of radiation patterns from a spherical array antenna, with inclusion of mutual coupling effects. With SMT the mutual coupling effects can be taken into account at once, in both the analysis and

synthesis of a radiation pattern, and therefore disappear as unwanted, unknown, and compromising electromagnetic effects.

- Using SMT, the design and performance anticipation of spherical array antennas can be reduced to a few simple formula's, as a quicker and easier alternative for detailed analysis.
- Based on SMT, a framework can be constructed in analogy with phase mode theory for circular array antennas, that facilitates the development of advanced signal processing techniques for spherical array antennas.
- Spherical array antennas offer an economical way of achieving true omnidirectional beamsteering. With methods based on SMT the costs of spherical array antennas can be kept low by concentrating information of all array elements in a fewer number of channels. With the intelligence of the antenna residing in the signal processing stage, a cost reduction and reliability gain is achieved over switching or commutating solutions with complex hardware and/or expensive MicroElectroMechanical System (MEMS) technology.

1.3.2 Scope

The work undertaken in this PhD programme is the first cornerstone of a continuing research project on spherical array antennas. Ultimately, it is the intention to design and build one or more fully functional spherical array antenna demonstrators to improve our understanding of spherical array antenna capabilities, modelling and operation and back these up with a thorough testing programme on the performance of real-world spherical array antenna systems, including the signal-processing back-end.

The scope of this PhD programme is insufficient to fully develop the concepts highlighted in section 1.2 in all areas defined in figure 1.6 and build demonstrators for the technology. With reference to the same figure, most of the work is concentrated on the left side of the diagram, as it makes chronological sense to understand a technology before making it. Nevertheless, the work started originally with the investigation of the extension of phase mode theory to spherical geometries, which belongs to the right hand side of the diagram. In addition, a significant amount of work has been added to the areas on the right.

More in particular, the scope of this programme can be outlined as follows:

- Development of electromagnetic analysis tools. The tools can be implemented as academical research code with the use of high-level scripting languages such as MATLAB. All the core components for the analysis of spherical array antennas must be implemented. Perhaps with slight adaptations to different geometries, future users should be able to use the code, without getting to grips with the details such as the implementation of modal expansion rotations.
- The analysis models must be accurate and comprehensive. Mutual coupling effects should be taken into account.

- Outline the methodology for spherical array signal processing that is equivalent to phase mode theory for circular arrays. Previously performed work with scalar spherical harmonics is inadequate and the methodology must use vectorial modes to incorporate the more complicated polarisation issues.

- Increase the understanding of design parameters for spherical arrays by defining them in the modal context.

- Perform an in-depth study on the various possibilities for spherical element distributions, as there does not seem to be an agreement on what is the best distribution, if any exists. There does not even seem to be an established criterion in use to measure the quality of a distribution. Clear design goals need to be defined and solutions found accordingly.

- Start the development of a hardware demonstrator. This antenna must stand out as a major technological achievement and a proof-of-concept prototype. The completion of a complete advanced antenna system, with inclusion of advanced digital processing stages does not fit in the scope of this programme. However, the design can be worked out to the details, and the RF front stage can be build.

The following further restrictions are imposed on the nature of the spherical arrays that will be investigated:

- Spherical arrays with a metal spherical groundplane are considered, rather than electromagnetically transparent array antennas. In [30] we already considered transparent spherical array antennas. The difference in the theoretical handling is minor and requires only slightly different radial functions to be used, as will be explained on page 47. However, transparent spherical array antennas have bandwidth issues, as explained in appendix C. Despite of that, non-transparent spherical arrays are sometimes analysed with the wrong kind of radial functions. Therefore we will stick to one kind of radial functions in this thesis - those necessary for modelling spherical arrays with a metal groundplane.

- Only circular polarisation is considered. There is a good reason to use circular polarisation with spherical array antennas, as will be explained in section 6.1. In case that linear polarisation is required in conjunction with spherical array antennas, the array elements will need to be dual polarised.

- The focus lies on the development of large high-gain spherical array antennas. This follows from the fact that the applications that can benefit from spherical array technology typically require high-gain antennas.

1.3.3 Major Contributions

In the author's opinion, the most significant contributions of this work to the field of research can be summarised as follows:

- **Extension of phase mode processing for spherical array antennas.** Since we firstly proposed to use spherical harmonics as the equivalent of phase modes in IEE Electronics Letters[28], the idea has been picked up by other researchers and the work has been referenced in at least the following papers: [27][25][26][29][40][41].

- **Introduction of a new spherical wave expansion.** Using the scalar spherical harmonics to expand an electric field is incorrect, because a multi-pole expansion is needed for this purpose. The classical multi-pole spherical wave expansion is cumbersome to use because of its vectorial nature. In section 2.4 of this thesis two sets of scalar expansions are presented that can each represent one of two polarisations: one for the left hand and one for the right hand circular polarisation. Using this expansions to describe the far-field of an antenna is both uncomplicated and exact, and allows the generalisation of signal processing techniques such as spherical convolution for vector fields (presented in section 7.4.4).

- **Further advance the state of the art on electromagnetic analysis.** Theoretical work is carried out to enrich the theoretical toolset for the analysis of spherical array antennas. Complex models are built that use the modal method to characterise a variety of complex spherical array antennas, including mutual coupling effects.

- **Determination of how to distribute array elements on a sphere.** The distribution of the elements on the sphere is thoroughly investigated. Optimal strategies for different array-sizes are proposed, with and without subarraying. The question on how to sample on the sphere is of broader interest for spherical signal processing, e.g. for near-field antenna measurement systems.

- **New developments in antenna design.** Several design solutions are offered that can be applied in the general context of conformal array antennas. Examples of these improvements are the proposed novel element distribution for spherical arrays, but also the design of the radiating elements themselves. A very low-cost tile with three circularly polarised elements has been designed. This tile can be used in any geodesic construction.

1.4 Publications

The plan is to publish the work reported in this thesis in the form of one or two journal papers upon completion of this thesis. Parts of the work have already been published (or been accepted for publication) in the following papers:

- Phase mode processing for spherical antenna arrays[28]
- Mutual Coupling Analysis of a Spherical Array Antenna[42]
- Spherical Wave Expansion for a Circular Aperture in a Large Conducting Sphere[43]
- Sampling and Interpolation on the Sphere[44]
- A Spherical Near-Field System for Testing a Spherical Array of Patch Antennas[45]

- Design and Development of a Spherical Array Antenna[46]
- Circularly Polarized Array of Ring Antennas using Novel Proximity-Coupled Feeds[47]
- Spherical Near-Field Measurements of a Spherical Array Antenna[48]

1.5 Thesis Outline

In chapter 2 the theoretical foundations are laid that are used throughout the other chapters. This chapter can be digested easily by anyone who is familiar with spherical harmonics and Fourier analysis on the sphere, but is still recommended reading as a number of conventions and definitions adhered to throughout this text are outlined in this chapter. For those who are not familiar with spherical wave expansions for instance, this chapter is intended to offer an easily accessible theoretical introduction.

In chapter 3 the antenna pattern of a single radiating element placed on a large metal sphere is analytically obtained as a linear combination of spherical functions. This forms the basis for a simulation model that is developed in subsequent chapters.

In chapter 4 several radiating elements are placed on the sphere. The array pattern is not simply a superposition of single element patterns due to the interactions between the elements. In this chapter these interactions are investigated and expressions are derived to quantify these electromagnetic interactions in terms of mutual admittances.

The mutual admittances obtained in chapter 4 are used in chapter 5 to develop the scattering matrix model for the spherical array antenna. This model makes it possible to accurately simulate the behaviour of the spherical array, with inclusion of the mutual coupling effects between its elements.

As with any antenna, the spherical array needs to be designed so that it meets certain performance parameters. In chapter 6 it is investigated how typical performance parameters for antennas such as gain and sidelobe levels can be met with the spherical array antenna.

The modal approach that is taken in the analysis of the spherical array can also be applied to process the signals from the array elements. In chapter 7 signal processing techniques suitable to deal with spherical problems are studied. The chapter provides a framework for the development of more advanced signal processing techniques, by reviewing and extending the concept of convolution on the sphere.

It will be shown in chapter 8 that it is not a straightforward problem to choose a suitable array element distribution over the sphere. The details of this problem are explained and used to choose the best array element distribution.

A prototype spherical array antenna is being developed in hardware. Chapter 9 explains the development of this device, as well as the measurements that have been and will be performed on the prototype antenna. Finally, the results from the first set of measurements are evaluated.

Chapter 2

Spherical Wave Theory

This chapter will present the theoretical framework on which the following chapters are built. Most of this chapter is distilled from existing literature and some readers may be familiar with the material, while some of the theory has been developed specifically for this thesis. Spherical Modal Theory (SMT) theory was already applied to scattering problems involving spheres by Gustav Mie and Paul Debye in the beginning of the 20^{th} century[49][50]. Since then, several authors have used spherical harmonics and their multipole expansion extensions in radiation problems[51][52][53]. More recently, the efficient computation of the scalar Spherical Fourier Transform (SFT) received a lot of attention by mathematicians and physicists[54][55][56][57]. Application to EM problems has been well covered in the following standard reference books: [58][59][60][61]. The material in this chapter is compiled mainly from those books and from the more applied textbooks on spherical scattering[62] and spherical antenna measurements[63]. The latter are especially useful for the many formula's in their appendices.

In section 2.1 we will briefly highlight the concept of phase modes as used to analyse and synthesise the radiation pattern of circular arrays, and mathematically address their nature. This section serves as a starting point for a generalisation towards the modal approach we will take to address spherical array antennas.

Section 2.2 intends to be a crash course for the reader who is not familiar with spherical harmonics and the SFT or wants to refresh on the subject. The SFT is often used to solve scalar problems with spherical symmetry. As this theory is central to the rest of this thesis it is developed here from scratch and should be self-contained for anyone with a general engineering background.

In section 2.3, the Spherical Wave Expansion (SWE) is constructed on top of the SFT, to accommodate the vectorial wave solutions one encounters when addressing electromagnetic problems. Several formulations for the SWE exist in the literature. In this thesis, different formulations of the SWE will be used, depending on the problem at hand. Also in this section, the relations between the different expansions are explained and some definitions and conventions specific to this text are introduced. It is therefore recommended for all readers to have at least skimmed through this section before reading subsequent chapters of this thesis.

In the far-field of an antenna, certain simplifications can be made so that the radiated field can be sufficiently described with scalar functions. In section 2.4 a new kind of basis functions are

presented that will be used throughout this text to expand the far-field of circularly polarised antennas. As the basisfunctions are scalar but different from the commonly used spherical harmonics, this section will also relate the proposed far-field expansion to the SFT.

2.1 The Modal Approach

The concept of phase modes was introduced to analyse circular array antennas[2][64]. Circular phase modes are functions of constant amplitude, defined over the unit circle, with their phase cycling m times through 2π radians per period, where m is the order of the phase mode. These functions exploit the symmetry of the circle, or else, the periodicity of all functions defined over it. We use these functions to analyse recurring phenomena within periodical signals, and indeed, they form the basis for the circular Fourier transform every engineer is familiar with. The projection of an excitation $F(\phi)$ function over the circle into this basis offers a number of advantages to process the signals of the array[2][64]. The Fourier transform of F is given as:

$$F(\phi) = \sum_{m=-N}^{N} C_m \exp(jm\phi) \tag{2.1}$$

Essentially the advantages of phase mode processing are based on two invariance properties:

- Invariance under radiation: each phase mode present in the array excitation gives rise to a corresponding mode in the far-field pattern.
- Rotational invariance: each phase mode can be rotated on the circle by a simple multiplication with a phase factor.

In the next section we will write these invariance properties in the form of two eigen-equations. We will identify an operator with both the actions of rotation and radiation. These operators are such that their eigenfunctions are further subject to two other important properties:

- Orthogonality
- Completeness

This means that we can use these functions to serve as a basis in which any periodical function (a function defined over a circle) can be represented without loss of information. This is the cornerstone of the Fourier transform in equation (2.1).

2.2 The Spherical Fourier Transform

The Spherical Fourier Transform (SFT) of a function f on the sphere S^2 expands that function in a linear combination of the so-called spherical harmonics Y_l^m[62]:

$$f(\theta,\phi) = \sum_{l\in\mathbf{N}} \sum_{|m|\leqslant l} \hat{f}(l,m) Y_l^m(\theta,\phi) \tag{2.2a}$$

$$\hat{f}(l,m) = \int_{S^2} f \cdot \overline{Y_l^m} d\omega \tag{2.2b}$$

where the Fourier coefficients $\hat{f}(l,m)$ are doubly parameterised in the degree l and order m and the spherical harmonics Y_l^m constitute a basis for Fourier decomposition. This section explains this transform in more detail and shows how it comes about as a logical extension of the more familiar Fourier transform of periodic functions.

2.2.1 Understanding Spherical Harmonics

Spherical harmonics are best understood in the context of Fourier transforms. Fourier transforms are used to exploit the symmetry of a domain over which functions are defined. Periodic signals for instance, can be regarded as functions defined over the circle. The spectral components in their Fourier spectrum posses symmetry when plotted on the circle. A physical process is mathematically represented by an operator that acts on the function. Many physical processes can be modelled as linear and coordinate-invariant operators. On the circle or the sphere, rotationally invariant operators commute with rotations. Hence, the rotational symmetry of the circle and the sphere can be used to simplify the action of these linear rotationally invariant operators.

Let S denote the domain on which our functions of interest are defined, e.g.. the circle S^1 or the sphere S^2. We construct a Hilbert space $L^2(S)$ on S with the square-integrable functions f for which $\|f\| < \infty$. Here the L^2-norm is used with the L^2 inner product:

$$\langle f, h \rangle = \int_S f\overline{h}\, d\omega \tag{2.3a}$$

$$\|f\| = \sqrt{\langle f, f \rangle} \tag{2.3b}$$

The symmetry we seek to exploit is formally described by the symmetry group of S. For the circle that is the special orthogonal group in two dimensions $SO(2)$ whereas for the sphere we have $SO(3)$. The main complication when moving from the circular to the spherical Fourier transform lies in the fact that $SO(2)$ is an abelian group, whereas $SO(3)$ is not. Now let ω be a point on the sphere and g a rotation of the rotation group. The left regular representation of the rotation group in $L^2(S)$ maps each rotation g to an operator $\Lambda(g)$ that rotates the graph of a function f defined over S as follows[54]:

$$\Lambda(g)f(\omega) = f\left(g^{-1}\omega\right) \tag{2.4}$$

We can divide the space $L^2(S)$ into orthogonal rotationally invariant subrepresentations Ψ_i so that:

$$f \in \Psi_i \Rightarrow \forall g : \Lambda(g)f \in \Psi_i \tag{2.5}$$

Fourier analysis boils down to the orthogonal decomposition of the left regular representation into irreducible subrepresentations. These subspaces are spanned by the Fourier basis functions. In case of the circular Fourier transform, the subspaces are one dimensional and each has a basisfunction of the form $\exp(jm\phi)$, where m is called the order of the subspace. In the case of the spherical Fourier transform, the spaces have dimension $2l+1$, where l is the degree of the subspace. The beauty of the Fourier transform is seen in the action of linear rotationally invariant

operators on a Fourier expanded function as they act on each subrepresentation separately. The basisfunctions for the subrepresentations can be found as eigenfunctions of hermitian operators. This ensures that the basisfunctions are orthogonal and complete.

The basisfunctions that span the subrepresentations can be obtained as the eigenfunctions of the Laplace-Beltrami operator, or the Laplace operator restricted to S. They are also homogeneous polynomials that are a solution to the Laplace equation with the unrestricted Cartesian Laplace operator, and therefore harmonic functions. Before we obtain the spherical harmonic functions Y_l^m in the next section, we express the rotational invariance of the eigenspaces they form, which is the central idea in Fourier analysis:

$$\Lambda(g)Y_l^m(\omega) = \sum_{|k|\leqslant l} Y_l^k(\omega)D_{k,m}^{(l)}(g) \tag{2.6}$$

This formula states that any rotation will take a spherical harmonic of degree l back into the space of spherical harmonics of degree l. The rotation matrices $D_{k,m}^{(l)}(g)$ between the spherical harmonics are known as the Wigner's D-matrices for SO(3). Several algorithms have been published to compute these matrices efficiently for applications in chemical physics[65][66][67][68]. We will revisit the computation of these matrices in section 4.2 of chapter 4.

2.2.2 Spherical Harmonics as Eigenfunctions

In the previous section it was stated that the basis functions for the Fourier transform are obtained as the eigenfunctions from linear hermitian operators. We can use the operators to impose physical boundary conditions of periodicity on the wave functions by writing down eigenequations and hence derive which solutions are physically possible on the domain on which they are defined.

We define the operator $\boldsymbol{L}$ as:

$$\boldsymbol{L} = -j(\vec{r}\times\nabla) = j\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\phi}\hat{e}_\theta - \frac{\partial}{\partial\theta}\hat{e}_\phi\right) \tag{2.7}$$

Up to a constant, this operator is identical to the angular momentum operator, well known in quantum mechanics[69]. The rectangular components of $\boldsymbol{L}$ in Cartesian and spherical coordinates are:

$$\begin{aligned}
\boldsymbol{L_x} &= -j\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) = j\left(\sin\phi\frac{\partial}{\partial\theta} + \frac{\cos\phi}{\tan\theta}\frac{\partial}{\partial\phi}\right)\\
\boldsymbol{L_y} &= -j\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) = j\left(-\cos\phi\frac{\partial}{\partial\theta} + \frac{\sin\phi}{\tan\theta}\frac{\partial}{\partial\phi}\right)\\
\boldsymbol{L_z} &= -j\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) = -j\frac{\partial}{\partial\phi}
\end{aligned} \tag{2.8}$$

The components of $\boldsymbol{L}$ cannot simultaneously be observed and they have no simultaneous eigenfunctions. It is easy to show they do not commute. However, we can find simultaneous eigenfunctions for $\boldsymbol{L}^2 = \boldsymbol{L_x}^2 + \boldsymbol{L_y}^2 + \boldsymbol{L_z}^2$ and one component of $\boldsymbol{L}$, typically $\boldsymbol{L_z}$[69].

In the case of circular phase modes, there is only one component $\boldsymbol{L_z}$. The eigenfunctions of

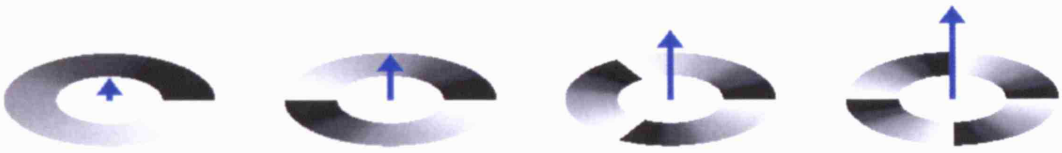


Figure 2.1: The figure shows the familiar basis functions for the Fourier transformation, with their associated momentum vectors for m ranging from 1 (left) to 4 (right). The periodical functions are plotted over the circle with constant amplitude. Their varying phase is coded in shades of grey, with black and white representing the extremities of the interval 0 to 2π.

this operator are given by:

$$\Phi_m(\phi) \;=\; \frac{1}{\sqrt{2\pi}} \exp(jm\phi) \tag{2.9}$$

and the eigen-equation with eigenvalues are:

$$\boldsymbol{L_z}\Phi_m \;=\; m\Phi_m \quad \text{with} \quad m = 0, \pm 1, \pm 2, \ldots \tag{2.10}$$

The eigenfunctions are indeed the basis functions for the circular Fourier transform. An intuitive semi-classical interpretation of angular momentum is obtained by visualising the operator as a vector in the z-direction, whose admitted lengths are the eigenvalues of (2.10). This representation and the corresponding phase modes are depicted in figure 2.1.

The generalisation to spherical geometry can now be seen most clearly: one just has to drop the assumption of having only one component $\boldsymbol{L_z}$. Hence, eigenfunctions have to be sought that obey simultaneously the eigen-equations of both operators $\boldsymbol{L_z}$ and $\boldsymbol{L}^2$. In spherical coordinates, $\boldsymbol{L}^2$ is given by:

$$\boldsymbol{L}^2 \;=\; -\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right) \tag{2.11}$$

Once the orientation of the z-axis is chosen, and thus the North Pole - The point on the unit sphere where $z = 1$ (NP) and South Pole - The point on the unit sphere where $z = -1$ (SP) are known, the spherical harmonics can be constructed by exploiting spherical symmetry. Singularities at the poles should thereby be avoided. The eigenfunctions are then found to be:

$$\Phi_l^m(\theta,\phi) \;=\; Y_l^m = N_l^m \cdot P_l^m(\cos\theta) \cdot \exp(jm\phi) \tag{2.12}$$

where N_l^m is a normalisation factor:

$$N_l^m \;=\; \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}(-1)^m \tag{2.13}$$

and where P_l^m is an associated Legendre function[62] of degree l and order m. The functions Y_l^m are called spherical harmonics[62]. The spherical harmonic functions combine the azimuthal phase variation of the circular modes with an associated Legendre function that is a function of elevation only.

Figure 2.2: The figure shows the angular momentum vectors associated with spherical harmonics. Due to the uncertainty principle, each vector is specified by its length and its z-component only, which is why the vectors are represented by conical surfaces. The figure shows all 5 vectors for which $l = 2$ and all 9 vectors for which $l = 4$.

The eigen-equation with eigenvalues are:

$$\boldsymbol{L}^2 Y_l^m \;=\; l(l+1) Y_l^m \quad \text{with} \quad -l \leqslant m \leqslant l \tag{2.14}$$

Also here the results can be interpreted from a semi-classical point of view. The possible values for the length of the momentum vector are given by $|\boldsymbol{L}| = \sqrt{l(l+1)}$, and their z-component can take on all integer values for which $-l \leqslant m \leqslant l$. The momentum vectors are shown in figure 2.2. It follows that for $m \neq 0$, $\boldsymbol{L_z}$ is always smaller than the length of the momentum vector, except for $l = 0$ and $m = 0$, the only harmonic with uniform amplitude.

The spherical harmonic functions can be applied to find solutions for the scalar wave equation[59], described in a spherical coordinate system. Each frequency component in the solution obeys the Helmholtz wave equation:

$$\nabla^2 \Psi + k^2 \Psi \;=\; 0 \tag{2.15}$$

The first step is to separate the angular variables from the radial variable. The details of this standard procedure can be found in [60]. The ∇ operator can be rewritten by recognising the operator $\boldsymbol{L}^2$ as the angular part of the Laplacian in spherical coordinates:

$$\nabla^2 \;=\; \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{\boldsymbol{L}^2}{r^2} \tag{2.16}$$

Subsequently the solution is expanded in the spherical harmonics (2.12):

$$\Psi \;=\; \sum_{l,m} C_l^m \, R_l^m(r) \, Y_l^m(\theta, \phi) \tag{2.17}$$

By substitution of (2.17) and (2.16) into (2.15) and using (2.14), the radial function is obtained

as a solution to the radial equation:

$$\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l(l+1)}{r^2} + k^2\right] z_l(kr) = 0 \tag{2.18}$$

An important observation is that the order m does not appear in this equation. The independence of the radial function from the order m allows description of the propagating characteristics of the spherical modes based on their degree l alone. The general solution for the radial function is[60]:

$$z_l(kr) = C_l^{(1)} h_l^{(1)}(kr) + C_l^{(2)} h_l^{(2)}(kr) \tag{2.19}$$

The radial function is a linear combination of the spherical Hankel functions h_l of the first and second kind. The spherical Hankel functions of the first kind represent outbound waves, originating in the centre of the sphere and travelling towards infinite radius. The spherical Hankel functions of the second kind travel inbound, originating at infinite radius and travelling towards the centre of the sphere. The linear combination of the two can represent any superimposed standing wave as well. The sum of a Hankel function of the first and the second kind results in a spherical Bessel function. These are typically used when the origin is part of the solution space. Indeed, any spherical mode that travels inward towards the centre, would pass through the centre and reappear as an outward travelling mode. The inward and outward travelling modes combine into a standing wave that is described by the spherical Bessel function. Spherical Bessel function have no singularities, whereas spherical Hankel functions of the first kind have a singularity at the origin (the origin of the spherical wave).

Finally, the complete expansion of the the solution to (2.15) can be written down:

$$\Psi = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left(C_{lm}^{(1)} h_l^{(1)}(kr) + C_{lm}^{(2)} h_l^{(2)}(kr) \right) Y_l^m \tag{2.20}$$

Typically, the coefficients in this expansion are obtained from a boundary condition over two concentric spherical surfaces (with one possibly of infinite radius), and the solution is known everywhere within the source-free space.

2.3 Vectorial Spherical Harmonics

2.3.1 Problem Formulation

We want to construct vector functions $\vec{F}$ that:

- Obey the vector wave equation:

$$\begin{aligned} &\nabla^2 \vec{F} + k^2 \vec{F} = 0 \\ \text{with:}\quad &\nabla^2 \vec{F} = \nabla(\nabla \cdot \vec{F}) - \nabla \times (\nabla \times \vec{F}) \end{aligned} \tag{2.21}$$

- Show the same invariance properties as the scalar spherical harmonics (as well as being orthogonal and complete).

- Can represent electromagnetic fields by obeying Maxwell's equations.

The first two requirements are achieved by constructing a set of independent vector functions through the action of specific differential operators on the scalar spherical harmonics. We will see that by taking suitable combinations of these independent vector functions, solutions can be found for Maxwell's equations.

2.3.2 Spherical Wave Expansions for Electromagnetism

A Separation of radial and tangential components

Solutions to the vector wave equation in Cartesian coordinates can easily be separated and solved with three independent scalar equations. This is not true in spherical systems. However, there are several ways to split a vector field defined in a spherical system into three fields that can be treated separately.

A first method, partially described in [70], relates closely to the the approach taken previously for the scalar solution. The radial component of the field is separated and expanded like the scalar solution. The tangential vector field that is left, is projected in a basis of vector spherical harmonics. As for the scalar case, these are obtained as eigenfunctions from two operators: the angular part of the Laplacian (or the squared momentum operator) and the z-directed component of the momentum operator. Only, these operators are generalised so that they can act on tangential vector fields:

$$\tilde{L}^2\vec{v} = \left[-\left(\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} - \frac{1}{\sin^2\theta}\right)v_\theta + \left(\frac{2\cot\theta}{\sin\theta}\frac{\partial}{\partial\phi}\right)v_\phi\right]\hat{e}_\theta + \left[-\left(\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} - \frac{1}{\sin^2\theta}\right)v_\phi - \left(\frac{2\cot\theta}{\sin\theta}\frac{\partial}{\partial\phi}\right)v_\theta\right]\hat{e}_\phi \quad (2.22)$$

and

$$\tilde{L}_z\vec{v} = \frac{1}{j}\frac{\partial v_\theta}{\partial\phi}\hat{e}_\theta + \frac{1}{j}\frac{\partial v_\phi}{\partial\theta}\hat{e}_\phi \quad (2.23)$$

The same eigen-equations can now be used to find the vector spherical harmonics:

$$\tilde{L}^2\vec{A}_l^m = l(l+1)\vec{A}_l^m \quad (2.24)$$

$$\tilde{L}_z\vec{A}_l^m = m\vec{A}_l^m \quad (2.25)$$

Unlike the scalar case, there are now two independent vector harmonics associated with each (l,m) pair.

$$\vec{B}_l^m = \hat{e}_r \times C_l^m \quad (2.26a)$$

$$\vec{C}_l^m = \frac{1}{j\sqrt{l(l+1)}}LY_l^m \quad (2.26b)$$

The vector spherical harmonics can now be applied in the expansion of a solution to the vector

wave equation (2.21). Taking account of the radial factors, the tangential field solution becomes:

$$\vec{\Psi}^m_{l,\tan} = C_{blm} R_{blm}(r) \vec{B}^m_l(\theta,\phi) + C_{clm} R_{clm}(r) \vec{C}^m_l(\theta,\phi) \tag{2.27}$$

where the C_{xlm} are constants and the R_{xlm} are functions of the radius. The radial factors take the same form as in equation (2.19). The full solution, including the radial field component, becomes:

$$\begin{aligned} \vec{\Psi}^m_l = &\left(C^{(1)}_{ylm} h^{(1)}_l(kr) + C^{(2)}_{ylm} h^{(2)}_l(kr)\right) Y^m_l(\theta,\phi) \cdot \hat{e}_r + \\ &\left(C^{(1)}_{blm} h^{(1)}_l(kr) + C^{(2)}_{blm} h^{(2)}_l(kr)\right) \vec{B}^m_l(\theta,\phi) + \\ &\left(C^{(1)}_{clm} h^{(1)}_l(kr) + C^{(2)}_{clm} h^{(2)}_l(kr)\right) \vec{C}^m_l(\theta,\phi) + \end{aligned} \tag{2.28}$$

By postulating the Maxwell's equations, extra relations between the constants can be found, but the formulation above is impractical to find these. In the next section the radial term will be absorbed in the two other terms of the solution, so that the obtained solution obeys the Maxwell equations. Expansions that obey the Maxwell equations are suited to describe waves with spherical symmetry, and are called Spherical Wave Expansion (SWE).

B Expansion for the spherical waveguide

In a source-free region of empty space, Maxwell's equations are:

$$\nabla \times \vec{E} = jk\eta \vec{H} \tag{2.29a}$$

$$\nabla \cdot \vec{E} = 0 \tag{2.29b}$$

$$\nabla \times \vec{H} = \frac{-jk}{\eta} \vec{E} \tag{2.29c}$$

$$\nabla \cdot \vec{H} = 0 \tag{2.29d}$$

This system of equations can be transformed into either of the following systems of three equations:

$$\begin{aligned} \left(\nabla^2 + k^2\right)\vec{H} &= 0 \\ \nabla \cdot \vec{H} &= 0 \\ \vec{E} &= \frac{j\eta}{k} \nabla \times \vec{H} \end{aligned} \quad \Leftrightarrow \quad \begin{aligned} \left(\nabla^2 + k^2\right)\vec{E} &= 0 \\ \nabla \cdot \vec{E} &= 0 \\ \vec{H} &= \frac{-j}{\eta k} \nabla \times \vec{E} \end{aligned} \tag{2.30}$$

The set of fields presented in the previous section would obey the wave equation. However, it is difficult to govern the coefficients so that the divergence remains zero and not violate Maxwell's equations. A better suited set of solutions was proposed by Hansen[51][52][59]. He proposed a system of three vector fields that can be solved independently in any coordinate system. The fields can be constructed from the scalar solution Ψ to the Helmholtz wave equation (2.15). For

spherical coordinate systems, the three vector fields are constructed as follows:

$$\vec{L} = \nabla\Psi \tag{2.31a}$$

$$\vec{M}_l^m = \vec{L} \times \vec{r} \tag{2.31b}$$

$$\vec{N}_l^m = \frac{1}{k}\nabla \times \vec{M}_l^m \tag{2.31c}$$

The definitions in (2.31) obey equation (2.21), subject to (2.15). The field $\vec{L}$ is a divergent field, whereas $\vec{M}_l^m$ and $\vec{N}_l^m$ are purely solenoidal fields. It is easy to show that

$$\vec{M}_l^m = \frac{1}{k}\nabla \times \vec{N}_l^m \tag{2.32}$$

which makes $\vec{M}_l^m$ and $\vec{N}_l^m$ particularly well suited to represent electromagnetic fields in a source-free homogeneous medium.

With a suitable combination of the vector fields $\vec{M}_l^m$ and $\vec{N}_l^m$, an arbitrary vector field can be represented that satisfies (2.15) and has zero divergence. The constructed field will therefore be a solution of Maxwell's equations. This can be performed for either $\vec{E}$ or $\vec{H}$, the corresponding $\vec{H}$ or $\vec{E}$ field is than found by applying the curl operator in equation (2.30), which, by virtue of (2.31) and (2.32) results in a similar combination of $\vec{M}_l^m$ and $\vec{N}_l^m$.

The expansion used in [63] is most practical, as it further power normalises the coefficients in the expansion so that they have the dimension of $\sqrt{Watt}$:

$$\vec{E} = k\sqrt{\eta}\sum_{l,m}\left(C_{lm}^{\mathrm{TE}}\vec{M}_l^m + C_{lm}^{\mathrm{TM}}\vec{N}_l^m\right) \tag{2.33a}$$

$$\vec{H} = \frac{-jk}{\sqrt{\eta}}\sum_{l,m}\left(C_{lm}^{\mathrm{TE}}\vec{N}_l^m + C_{lm}^{\mathrm{TM}}\vec{M}_l^m\right) \tag{2.33b}$$

Written out in full, the expressions for $\vec{M}_l^m$ and $\vec{N}_l^m$ are:

$$\vec{M}_l^m = N_{lm}\left(z_l(kr)\frac{jmP_l^m(\cos\theta)}{\sin\theta}e^{jm\phi}\,\hat{e}_\theta - z_l(kr)\frac{dP_l^m(\cos\theta)}{d\theta}e^{jm\phi}\,\hat{e}_\phi\right) \tag{2.34a}$$

$$\vec{N}_l^m = N_{lm}\left(\frac{l(l+1)}{kr}z_l(kr)P_l^m(\cos\theta)e^{jm\phi}\,\hat{e}_r + \frac{1}{kr}\frac{d}{d(kr)}\left[krz_l(kr)\right]\frac{dP_l^m(\cos\theta)}{d\theta}e^{jm\phi}\,\hat{e}_\theta + \frac{1}{kr}\frac{d}{d(kr)}\left[krz_l(kr)\right]\frac{jmP_l^m(\cos\theta)}{\sin\theta}e^{jm\phi}\,\hat{e}_\phi\right) \tag{2.34b}$$

with the normalisation factor:

$$N_{lm} = \frac{1}{\sqrt{l(l+1)}}\,N_l^m = \frac{1}{\sqrt{l(l+1)}}(-1)^m\sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}} \tag{2.35}$$

The field $\vec{M}_l^m$ is the tangential field component corresponding to (2.26b). The $\vec{N}_l^m$ field has both tangential and radial components. By splitting the electromagnetic field in transverse electric (TE) and a transverse magnetic (TM) fields, the electric and respectively the magnetic

field contains only the $\vec{M}_l^m$-field, whereas the corresponding magnetic and respectively electric field contains only the $\vec{N}_l^m$-field. This simplifies the expressions for the outbound TE field as follows:

$$\begin{aligned}
\vec{E}_l^m &= k\sqrt{\eta}C_{lm}^{\mathrm{TE}}N_{lm}\Bigg(\left(h_l(kr)\frac{jm}{\sin\theta}P_l^m(\cos\theta)e^{jm\phi}\right)\hat{e}_\theta - \\
&\qquad \left(h_l(kr)\frac{dP_l^m(\cos\theta)}{d\theta}e^{jm\phi}\right)\hat{e}_\phi\Bigg) \\
\vec{H}_l^m &= \frac{-jk}{\sqrt{\eta}}C_{lm}^{\mathrm{TE}}N_{lm}\Bigg(\left(h_l(kr)\frac{l(l+1)}{kr}P_l^m(\cos\theta)e^{jm\phi}\right)\hat{e}_r + \\
&\qquad \left(\frac{1}{kr}\frac{d}{d(kr)}\Big[krh_l(kr)\Big]\frac{dP_l^m(\cos\theta)}{d\theta}e^{jm\phi}\right)\hat{e}_\theta + \\
&\qquad \left(\frac{1}{kr}\frac{d}{d(kr)}\Big[krh_l(kr)\Big]\frac{jm}{\sin\theta}P_l^m(\cos\theta)e^{jm\phi}\right)\hat{e}_\phi\Bigg)
\end{aligned} \tag{2.36a}$$

and the TM field:

$$\begin{aligned}
\vec{E}_l^m &= k\sqrt{\eta}C_{lm}^{\mathrm{TM}}N_{lm}\Bigg(\left(h_l(kr)\frac{l(l+1)}{kr}P_l^m(\cos\theta)e^{jm\phi}\right)\hat{e}_r + \\
&\qquad \left(\frac{1}{kr}\frac{d}{d(kr)}\Big[krh_l(kr)\Big]\frac{dP_l^m(\cos\theta)}{d\theta}e^{jm\phi}\right)\hat{e}_\theta + \\
&\qquad \left(\frac{1}{kr}\frac{d}{d(kr)}\Big[krh_l(kr)\Big]\frac{jm}{\sin\theta}P_l^m(\cos\theta)e^{jm\phi}\right)\hat{e}_\phi\Bigg) \\
\vec{H}_l^m &= \frac{-jk}{\sqrt{\eta}}C_{lm}^{\mathrm{TM}}N_{lm}\Bigg(\left(h_l(kr)\frac{jm}{\sin\theta}P_l^m(\cos\theta)e^{jm\phi}\right)\hat{e}_\theta - \\
&\qquad \left(h_l(kr)\frac{dP_l^m(\cos\theta)}{d\theta}e^{jm\phi}\right)\hat{e}_\phi\Bigg)
\end{aligned} \tag{2.36b}$$

Note that we have substituted the general solution z_l for the radial functions with Hankel functions of the first kind h_l to describe outbound waves.

C Applicability of spherical waveguide theory

The spherical wave expansion above is most useful for solving problems associated with spherical waveguides. Spherical waveguides should be interpreted here in a broad sense and the theory should be applied whenever it is practical to write the boundary conditions in a spherical coordinate system. For instance, a spherical conducting body in free space can be regarded as a spherical waveguide. The waveguide is then bounded by the surface of the spherical body and an imaginary sphere with infinite radius. Because of the orthogonality of the spherical modes, the full solution to the problem is known when the expansion coefficients can be derived from the boundary conditions. Although spherical waveguides can take other forms, the spherical waveguide referred to in this text will refer to the outer space of a conducting sphere of radius R.

It is generally true that waveguides with a finite cross-section support only a limited number of modes that can propagate through them. Although spherical waveguides do not have a constant cross-section, the same rule applies. The smallest cross-section of the waveguide, the sphere at R, sets a limit on the number of modes that can propagate through the waveguide towards the far-field. The transition from propagating to evanescent modes occurs gradually around the degree $l = kR$. The higher the degree of a mode exceeds this level, the more that mode will be attenuated. In the formulation for the spherical wave expansion, the cut-off characteristics are contained in the definition of the radial function (2.19).

D Expansion for the spherical transmission line

Sometimes, the cut-off characteristics that are inherently embedded in the spherical wave expansion (2.33) are undesirable. That is the case if one is interested in the near-field characteristics around a spherical surface, including the evanescent modes. Although it is true that evanescent modes are not capable of transferring power to the far-field, they are capable of doing that within the near-field. More in particular, when studying mutual coupling effects between elements of a spherical array, the evanescent modes have to be taken into account. Therefore, a slightly different formulation of the spherical wave expansion is more appropriate, in which the radial variation is absorbed into the modal coefficients, apart from the "normal" r^{-1} variation. The spherical system is now viewed as a non-uniform transmission line, and the coefficients for the electric field are called modal voltages, and those of the magnetic field modal currents. They are obtained from the transverse field components only, and the radial field components are derived from them. In this formulation, called STL expansion, the radial variation of the modes is obtained by solving a transmission line equation separately, and cut-off characteristics are contained in the impedance properties of the modes. Also, any superposition of inbound and outbound waves can now be expressed as usual with the voltage standing wave ratio or the local reflection coefficient.

The spherical waveguide formulation in (2.33) is transformed into a transmission line formulation by separating the radial factor in the transverse field components:

$$\vec{M}_l^m = z_l(kr) \underbrace{\left(N_{lm} \frac{jmP_l^m(\cos\theta)}{\sin\theta} e^{jm\phi}\,\hat{e}_\theta - N_{lm} \frac{dP_l^m(\cos\theta)}{d\theta} e^{jm\phi}\,\hat{e}_\phi \right)}_{-\vec{e}_{lm}^{\,\mathrm{TE}} = \vec{h}_{lm}^{\,\mathrm{TM}}} \tag{2.37a}$$

$$\vec{N}_l^m = \vec{N}_r + \{\vec{N}_l^m\}_{\mathrm{T}} \tag{2.37b}$$

$$\vec{N}_r = \frac{\sqrt{l(l+1)}}{kr} z_l(kr) Y_l^m \hat{e}_r \tag{2.37c}$$

$$\{\vec{N}_l^m\}_{\mathrm{T}} = \frac{1}{kr}\frac{d}{d(kr)}[krz_l(kr)] \underbrace{\left(N_{lm} \frac{dP_l^m(\cos\theta)}{d\theta} e^{jm\phi}\,\hat{e}_\theta + N_{lm} \frac{jmP_l^m(\cos\theta)}{\sin\theta} e^{jm\phi}\,\hat{e}_\phi \right)}_{-\vec{e}_{lm}^{\,\mathrm{TM}} = -\vec{h}_{lm}^{\,\mathrm{TE}}} \tag{2.37d}$$

The tangential part of the field can then be expanded as:

$$\vec{E}_t = \sum_{lm} V_{lm}^{\mathrm{TE}}(r)\frac{\vec{e_{lm}^{\mathrm{TE}}}}{r} + \sum_{lm} V_{lm}^{\mathrm{TM}}(r)\frac{\vec{e_{lm}^{\mathrm{TM}}}}{r} \tag{2.38a}$$

$$\vec{H}_t = \sum_{lm} I_{lm}^{\mathrm{TE}}(r)\frac{\vec{h_{lm}^{\mathrm{TE}}}}{r} + \sum_{lm} I_{lm}^{\mathrm{TM}}(r)\frac{\vec{h_{lm}^{\mathrm{TM}}}}{r} \tag{2.38b}$$

where the radial factors are absorbed by the coefficients of the expansion, that have become radial functions:

$$V_{lm}^{\mathrm{TE}}(r=R) = -\sqrt{\eta}\, krz_l(kr)\, C_{lm}^{\mathrm{TE}} \tag{2.39a}$$

$$V_{lm}^{\mathrm{TM}}(r=R) = -\sqrt{\eta}\,\frac{d}{d(kr)}\left[krz_l(kr)\right] C_{lm}^{\mathrm{TM}} \tag{2.39b}$$

$$I_{lm}^{\mathrm{TE}}(r=R) = \frac{j}{\sqrt{\eta}}\,\frac{d}{d(kr)}\left[krz_l(kr)\right] C_{lm}^{\mathrm{TE}} \tag{2.39c}$$

$$I_{lm}^{\mathrm{TM}}(r=R) = \frac{-j}{\sqrt{\eta}}\, krz_l(kr)\, C_{lm}^{\mathrm{TM}} \tag{2.39d}$$

These formula's can be used to convert a solution for the spherical waveguide expansion (2.33), found over a sphere of radius R, to a solution for the spherical transmission line expansion (2.38). The modal fields in this expansion are still orthogonal over the sphere. The modal voltages and currents can equally be obtained by projecting a tangential field onto them:

$$\begin{aligned} V_{lm} &= \iint r\vec{E}_t \cdot \vec{e}_{lm}\, dS \\ I_{lm} &= \iint r\vec{H}_t \cdot \vec{h}_{lm}\, dS \end{aligned} \tag{2.40}$$

The knowledge of $\vec{E}_t$ suffices as the modal currents can be obtained from the modal voltages if the input impedances are known.

Once launched into the spherical transmission line, the modal fields propagate according to the system of equations:

$$\begin{aligned} \frac{dV_{lm}}{dr} &= j\kappa Z_{lm}^c I_{lm} \\ \frac{dI_{lm}}{dr} &= j\kappa Y_{lm}^c V_{lm} \end{aligned} \tag{2.41}$$

in which Z_{lm}^c is the characteristic impedance of the spherical mode, Y_{lm}^c the characteristic admittance and κ the propagation number:

$$\begin{aligned} \kappa &= \sqrt{k^2 - k_c^2} \qquad \text{where} \quad k_c^2 = \frac{l(l+1)}{r^2} \\ Z_{lm}^c &= \eta\frac{k}{\kappa} \qquad \text{for TE-modes} \\ Z_{lm}^c &= \eta\frac{\kappa}{k} \qquad \text{for TM-modes} \end{aligned} \tag{2.42}$$

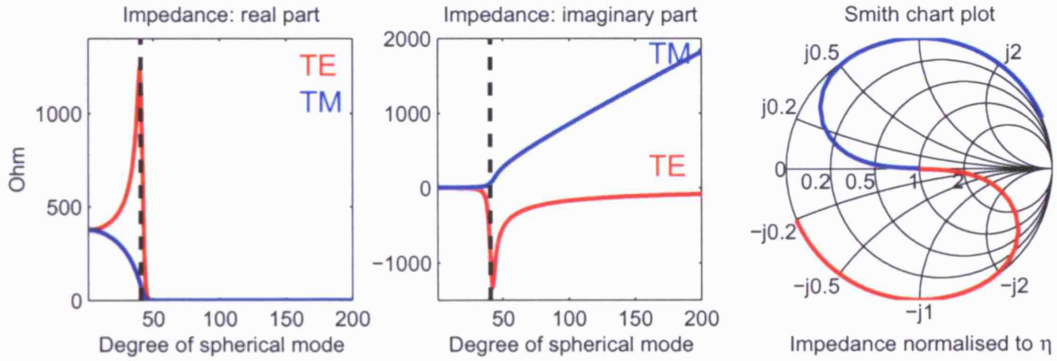


Figure 2.3: The input impedances of spherical modes quickly become highly reactive or inductive for degrees $l > kr$ (indicated by the dashed line). The corresponding evanescent field modes do not propagate to the far-field. However, these modes exist in the immediate vicinity of the spherical array and can contribute to mutual coupling effects.

As the cross-section of the spherical waveguide increases with the radius, there is no sharp cut-off wavelength above which no propagation is possible. However, for wavelengths that are longer than

$$\lambda_c = \frac{2\pi r}{\sqrt{l(l+1)}} \tag{2.43}$$

the fields decay more rapidly than $\frac{1}{r}$. For a fixed wavelength, the modes with a degree $L \gg kr$ decay rapidly and do not contribute to the far-field. Hence these modes are called evanescent modes.

The radial functions in (2.39) are only one solution set to the governing equations in (2.41). Generally there is also a reflected wave, for which $z_l(kr) = h_l^{(2)}(kr)$. We will assume the spherical transmission line to be matched so that no reflected waves have to be taken into account in the calculation of the input impedance.

Because a spherical transmission line is non-uniform, the input impedance for a matched spherical transmission line is not the same as its characteristic impedance. These impedances are given in equation (2.44) and plotted in figure 2.3.

$$\begin{aligned} Z^i_{lm\mathrm{TE}} &= j\eta \left(\frac{h_l(kr)}{h_l'(kr) + \frac{h_l(kr)}{kr}} \right) && \text{for TE-modes.} \\ Z^i_{lm\mathrm{TM}} &= -j\eta \left(\frac{h_l'(kr) + \frac{h_l(kr)}{kr}}{h_l(kr)} \right) && \text{for TM-modes.} \end{aligned} \tag{2.44}$$

Finally, the radial (longitudinal) field components are easily obtained from:

$$\frac{jk}{\eta} E_r = \sum_{lm} I^{\mathrm{TM}}_{lm}(r) \, \frac{l(l+1)}{r^2} k_c^2 \, Y_l^m(\theta, \phi) \tag{2.45a}$$

$$jk\eta H_r = \sum_{lm} V^{\mathrm{TE}}_{lm}(r) \, \frac{l(l+1)}{r^2} k_c^2 \, Y_l^m(\theta, \phi) \tag{2.45b}$$

2.4 Far-Field Expansions

The two vectorial spherical wave expansions from section 2.3 given in (2.33) and (2.38) are useful in the near and reactive near fields respectively. In the far-field however, the fields lose their radial component and become tangential to the wavefront. Also, the relation between the $\vec{E}$ and $\vec{H}$ fields simplifies to one being perpendicular to the other and scaled by a real impedance. Both fields can be described with the same vector function. Apart from a trivial phase variation with distance, a factor that is usually left out from the equations, this function is two-dimensional and independent from the radius. The two vectorial components can be separated in two scalar functions, the co- and cross-polar components. The way these polarisations are defined, depends on the vectorial orientation of the fields one is interested in, and on the coordinate system in which they are defined.

2.4.1 An Expression for the Far-Field

In this section we will take the spherical modes of the previous section to the far-field, and see what they become. We will establish a suitable definition for the polarisations and use them to obtain two scalar expansions for the far-field. Hereto the radial factors in the modal fields should be replaced by their large-argument approximations:

$$h_l^{(1)} \quad \longrightarrow \quad (-j)^{l+1}\,\frac{e^{(jkr)}}{kr} \tag{2.46a}$$

$$\frac{1}{kr}\frac{d}{d(kr)}\left[krh_l^{(1)}\right] \quad \longrightarrow \quad (-j)^{l}\,\frac{e^{(jkr)}}{kr} \tag{2.46b}$$

The factor $\frac{e^{(jkr)}}{kr}$ is commonly left out of the far-field definition, as it represents normal decay and phase development with distance. Following [63], the far-field definition for the spherical modes is multiplied by an extra factor $\sqrt{4\pi}$. The building blocks $\vec{M}_l^m$ and $\vec{N}_l^m$ from (2.34) are redefined for the far-field expansion as:

$$\begin{aligned}\vec{\mathcal{M}}_l^m(\theta,\phi) &= \sqrt{4\pi}\frac{kr}{e^{jkr}}\lim_{kr\to\infty}\left(\vec{M}_l^m(r,\theta,\phi)\right)\\ \vec{\mathcal{N}}_l^m(\theta,\phi) &= \sqrt{4\pi}\frac{kr}{e^{jkr}}\lim_{kr\to\infty}\left(\vec{N}_l^m(r,\theta,\phi)\right)\end{aligned} \tag{2.47}$$

The radial component of $\vec{\mathcal{N}}_l^m$ falls of as $\frac{1}{r^2}$, so that the expressions for $\vec{\mathcal{M}}_l^m$ and $\vec{\mathcal{N}}_l^m$ become:

$$\begin{aligned}\vec{\mathcal{M}}_l^m(\theta,\phi) &= \sqrt{4\pi}(-j)^{l+1}N_{lm}\left(\frac{jmP_l^m(\cos\theta)}{\sin\theta}e^{jm\phi}\,\hat{e}_\theta - \frac{dP_l^m(\cos\theta)}{d\theta}e^{jm\phi}\,\hat{e}_\phi\right)\\ \vec{\mathcal{N}}_l^m(\theta,\phi) &= \sqrt{4\pi}(-j)^{l}\;N_{lm}\left(\frac{dP_l^m(\cos\theta)}{d\theta}e^{jm\phi}\,\hat{e}_\theta + \frac{jmP_l^m(\cos\theta)}{\sin\theta}e^{jm\phi}\,\hat{e}_\phi\right)\end{aligned} \tag{2.48}$$

The far-field is now defined relative to the driving voltage V as:

$$\vec{\mathcal{F}}(\theta,\phi) = \frac{1}{V}\sum_{lm}\left(C_{lm}^{\mathrm{TE}}\vec{\mathcal{M}}_l^m + C_{lm}^{\mathrm{TM}}\vec{\mathcal{N}}_l^m\right) \tag{2.49}$$

and it follows that in the far-field, the electric and magnetic fields are given by:

$$\vec{E}(\theta,\phi) = V\frac{k\sqrt{\eta}}{\sqrt{4\pi}}\frac{e^{jkr}}{kr}\vec{\mathcal{F}}(\theta,\phi) \tag{2.50a}$$

$$\vec{H}(\theta,\phi) = V\frac{k}{\sqrt{\eta}\sqrt{4\pi}}\frac{e^{jkr}}{kr}\hat{e}_r \times \vec{\mathcal{F}}(\theta,\phi) \tag{2.50b}$$

2.4.2 Directivity

The directivity $D(\theta,\phi)$ is defined as the ratio between the power radiated per unit solid angle in the direction (θ,ϕ) and the power radiated per unit solid angle if the antenna radiated isotropically.

With the power density given by the time-average Poynting vector $\vec{S}_{\text{av}} = \frac{1}{2}\Re\{\vec{E}\times\vec{H}^*\}$, the power radiated per unit solid angle becomes:

$$r^2\,\vec{S}_{\text{av}}\cdot\hat{e}_r = r^2\frac{1}{2\eta}|\vec{E}|^2 \tag{2.51}$$

$$= \frac{1}{8\pi}V^2|\vec{\mathcal{F}}(\theta,\phi)|^2 \tag{2.52}$$

$$= \frac{1}{4\pi}V_{rms}^2|\vec{\mathcal{F}}(\theta,\phi)|^2 \tag{2.53}$$

For an isotropic radiator, the power density is $\frac{P}{4\pi}$, where $P = V_{rms}^2$ is the total radiated power by the antenna. The directivity is then obtained as:

$$D(\theta,\phi) = \frac{V_{rms}^2|\vec{\mathcal{F}}(\theta,\phi)|^2}{P} = |\vec{\mathcal{F}}(\theta,\phi)|^2 \tag{2.54}$$

2.4.3 Polarisation

Using equation (2.54) to obtain the directivity of an antenna will result in a loss of the vectorial information contained in the far-field. However, the directivity can equally be defined for two orthogonal polarisations in the far-field, so that the vectorial information of the field is preserved to match, for example, the polarisation of a second antenna. The far-field pattern is then usually presented by intersecting plots of its co-polar and cross-polar components. Projecting the field onto a co-polar and cross-polar component is a straightforward procedure when dealing with planar phase fronts, both for linear and circular polarisation. With spherical phase fronts, the definition of these entities deserves a little attention.

For linear polarisation, we will use "definition 3" by Ludwig[71]. It is defined in equation (2.55) and illustrated in figure 2.4. This system of unit vectors has the advantage of being continuous at boresight, and it has only one discontinuity at the other pole. The usual system of spherical unit vectors (θ,ϕ) is discontinuous at both poles. Also the system can readily be adapted to elliptical polarisation, by adding the unit vectors for linear polarisation out of phase. The axes must be oriented so that this system of unit vectors is optimally placed to capture the co-polar and cross-polar components. Hereto the z-axis should be at boresight and the x-axis parallel to a linearly polarised source. The latter requirement can be adjusted by setting the angle ϕ_0 in

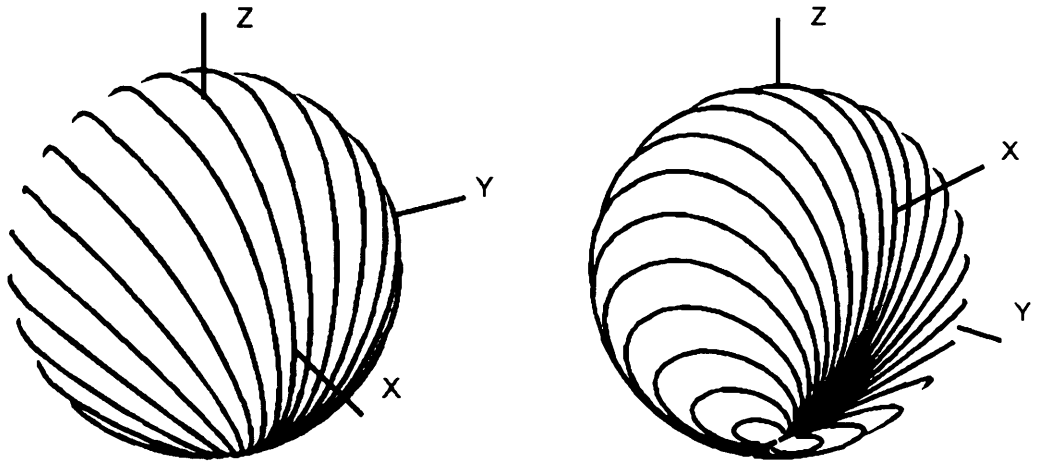


Figure 2.4: The convention for co-polar polarisation in a spherical wavefront is visualised by field lines for the unit vectors. The orientation is such that the visible lines run downwards. The cross-polar component is obtained by a 90° turn, according to the right-hand rule in the direction of the z-axis.

the definition to another reference angle than zero. The co-polar and cross-polar unit vectors for circular polarisation are also defined in (2.55).

for linear polarisation:

$$\hat{e}_{co} = \cos(\phi - \phi_0)\hat{e}_\theta - \sin(\phi - \phi_0)\hat{e}_\phi \tag{2.55a}$$

$$\hat{e}_{cross} = \sin(\phi - \phi_0)\hat{e}_\theta + \cos(\phi - \phi_0)\hat{e}_\phi \tag{2.55b}$$

for circular polarisation:

$$\hat{e}_{rhcp} = \frac{1}{\sqrt{2}} e^{j\phi}(\hat{e}_\theta + j\hat{e}_\phi) \tag{2.55c}$$

$$\hat{e}_{lhcp} = \frac{1}{\sqrt{2}} e^{-j\phi}(\hat{e}_\theta - j\hat{e}_\phi) \tag{2.55d}$$

The exponential factors in the definition of the circular polarised unit vectors cause the singularity of the spherical coordinate system at the north pole to vanish, so that only one singularity remains at the south pole. This can be seen in figure 2.4. The far-field $\vec{\mathcal{F}}(\theta, \phi)$ can now be projected onto the two orthogonal components. For spherical modes this results in the simple expressions:

$$\vec{\mathcal{F}}(\theta, \phi) = \mathcal{F}_{rhcp}(\theta, \phi)\, \hat{e}_{rhcp} + \mathcal{F}_{lhcp}(\theta, \phi)\, \hat{e}_{lhcp} \tag{2.56a}$$

with:

$$\begin{aligned}
\mathcal{F}_{rhcp}(\theta,\phi) &= \vec{\mathcal{F}}(\theta,\phi)\cdot\hat{e}^{*}_{rhcp}(\theta,\phi) \\
&= \frac{1}{V}\sum_{lm}\sqrt{2\pi}\,(C^{\mathrm{TE}}_{lm}+C^{\mathrm{TM}}_{lm})\,R^{m}_{l}(\theta,\phi) && \text{(2.56b)}\\
&= \frac{1}{V}\sum_{lm} r_{lm}R^{m}_{l}(\theta,\phi) && \text{(2.56c)}\\
\mathcal{F}_{lhcp}(\theta,\phi) &= \vec{\mathcal{F}}(\theta,\phi)\cdot\hat{e}^{*}_{lhcp}(\theta,\phi) && \text{(2.56d)}\\
&= \frac{1}{V}\sum_{lm}\sqrt{2\pi}\,(C^{\mathrm{TE}}_{lm}-C^{\mathrm{TM}}_{lm})\,L^{m}_{l}(\theta,\phi) && \text{(2.56e)}\\
&= \frac{1}{V}\sum_{lm} l_{lm}L^{m}_{l}(\theta,\phi) && \text{(2.56f)}
\end{aligned}$$

where the scalar functions are:

$$\begin{aligned}
R^{m}_{l} &= (-j)^{l}N_{lm}e^{j(m-1)\phi}\left(\frac{mP^{m}_{l}(\cos\theta)}{\sin\theta}+\frac{dP^{m}_{l}(\cos\theta)}{d\theta}\right)\\
L^{m}_{l} &= (-j)^{l}N_{lm}e^{j(m+1)\phi}\left(\frac{mP^{m}_{l}(\cos\theta)}{\sin\theta}-\frac{dP^{m}_{l}(\cos\theta)}{d\theta}\right) && \text{(2.56g)}
\end{aligned}$$

and

$$\begin{aligned}
r_{lm} &= \sqrt{2\pi}(C^{\mathrm{TE}}_{lm}+C^{\mathrm{TM}}_{lm})\\
l_{lm} &= \sqrt{2\pi}(C^{\mathrm{TE}}_{lm}-C^{\mathrm{TM}}_{lm}) && \text{(2.56h)}
\end{aligned}$$

The functions R^{m}_{l} and L^{m}_{l} form an orthonormal basis. We will refer to the projection of a function onto the R^{m}_{l} or L^{m}_{l} basis functions as the Circularly Polarised Spherical Transform (CPST). The quality of the polarisation is determined from the ratio between the co-polar and cross-polar components, which is quantified by the axial ratio. The axial ratio is defined as:

$$\tan\alpha = \frac{|Q|-1}{|Q|+1} \qquad (2.57)$$

with the circular polarisation ratio Q:

$$Q = \frac{\mathcal{F}_{rhcp}}{\mathcal{F}_{lhcp}} \qquad (2.58)$$

For a purely right hand circularly polarised wavefront, the axial ratio is 1. It is 0 for linear polarisation and -1 for left hand circular polarisation.

To calculate the Right Hand Circular Polarisation (RHCP) and Left Hand Circular Polarisation (LHCP) components from the θ and ϕ field components, for instance when measured data is

available on a (θ, ϕ)-grid, than the following formula's can be used:

$$\mathcal{F}_{rhcp}(\theta,\phi) = \vec{\mathcal{F}}(\theta,\phi)\cdot\hat{e}^*_{rhcp}(\theta,\phi) \tag{2.59a}$$

$$= \frac{e^{-j\phi}}{\sqrt{2}}(\mathcal{F}_\theta + j\mathcal{F}_\phi) \tag{2.59b}$$

$$\mathcal{F}_{lhcp}(\theta,\phi) = \vec{\mathcal{F}}(\theta,\phi)\cdot\hat{e}^*_{lhcp}(\theta,\phi) \tag{2.59c}$$

$$= \frac{e^{j\phi}}{\sqrt{2}}(\mathcal{F}_\theta - j\mathcal{F}_\phi) \tag{2.59d}$$

2.4.4 Spherical Fourier Transform of Scalar Far-Field Functions

By taking the limit of the vectorial spherical wave expansion to the far-field and consecutively projecting the tangential field on two orthogonal polarisations, two scalar expansions were obtained that completely describe the far-field.

The orthonormal scalar basis functions R_l^m and L_l^m that were used in the far-field expansions (2.56) are not generally used in scalar spherical problems. The Fourier basis of choice for scalar problems with spherical symmetry contains the spherical harmonics that were given in (2.12) and the spherical harmonic expansion was given in (2.20). Several algorithms have been presented in the literature to efficiently compute the SFT (e.g.. [72][73],[57]). It is therefore an interesting exercise to relate the newly presented transform in (2.56) to the well known SFT.

In this section we will introduce yet another set of basis functions, which we will call spinning SFT, to provide the missing link between the scalar spherical harmonic transform as we know it and the circularly polarised harmonic transform that was introduced in section 2.4.3. The objective here is to take the left or right hand circularly polarised component of a desired vectorfield, say $\mathcal{F}_{rhcp}$, and obtain the circularly polarised spherical transform $r_{lm} = \langle \mathcal{F}_{rhcp}, R_l^m \rangle$ by calculating the SFT $\langle \mathcal{F}^+, Y_l^m \rangle$ of some function $\mathcal{F}^+$.

We define the spinning SFT basis functions as follows:

$$S_l^{m^+} = N_l^m \left(\sin\theta \frac{dP_l^m(\cos\theta)}{d\theta} + mP_l^m(\cos\theta) \right) \exp(jm\phi) \tag{2.60a}$$

$$S_l^{m^-} = N_l^m \left(\sin\theta \frac{dP_l^m(\cos\theta)}{d\theta} - mP_l^m(\cos\theta) \right) \exp(jm\phi) \tag{2.60b}$$

From the definitions of R_l^m and L_l^m in (2.56) and from (2.60), we have:

$$\begin{aligned} R_l^m &= \frac{(-j)^l}{\sqrt{l(l+1)}} \frac{e^{-j\phi}}{\sin\theta} S_l^{m^+} \\ L_l^m &= \frac{(-j)^l}{\sqrt{l(l+1)}} \frac{e^{j\phi}}{\sin\theta} S_l^{m^-} \end{aligned} \tag{2.61}$$

so that:

$$\begin{aligned} r_{lm} &= \langle \mathcal{F}_{rhcp}, R_l^m \rangle \\ &= \frac{(-j)^l}{\sqrt{l(l+1)}} \langle \mathcal{F}_{rhcp}, \frac{e^{-j\phi}}{\sin\theta} S_l^{m^+} \rangle \\ &= \frac{(-j)^l}{\sqrt{l(l+1)}} \langle \underbrace{\mathcal{F}_{rhcp} \frac{e^{j\phi}}{\sin\theta}}, S_l^{m^+} \rangle \\ &= \frac{(-j)^l}{\sqrt{l(l+1)}} \langle \quad \mathcal{F}^+ \quad , S_l^{m^+} \rangle \end{aligned} \tag{2.62a}$$

$$\begin{aligned} l_{lm} &= \langle \mathcal{F}_{lhcp}, L_l^m \rangle \\ &= \frac{(-j)^l}{\sqrt{l(l+1)}} \langle \mathcal{F}_{lhcp}, \frac{e^{j\phi}}{\sin\theta} S_l^{m^-} \rangle \\ &= \frac{(-j)^l}{\sqrt{l(l+1)}} \langle \underbrace{\mathcal{F}_{lhcp} \frac{e^{-j\phi}}{\sin\theta}}, S_l^{m^-} \rangle \\ &= \frac{(-j)^l}{\sqrt{l(l+1)}} \langle \quad \mathcal{F}^- \quad , S_l^{m^-} \rangle \end{aligned} \tag{2.62b}$$

The next step is to write the projection of $\mathcal{F}^{\pm}$ on the spinning harmonics as a spectral recombination of its spherical harmonic transform. We start by using the recurrence properties of the associated Legendre functions (with $z = \cos\theta$):

$$\begin{aligned} \frac{dP_l^m(\cos\theta)}{d\theta} &= -\sin\theta P_l^{m\,\prime}(z) \\ &= \frac{1}{\sin\theta}(z^2-1) P_l^{m\,\prime}(z) \\ &= \frac{1}{\sin\theta}\left[-(l+1) z P_l^m(z) + (l-m+1) P_{l+1}^m(z)\right] \\ &= \frac{1}{\sin\theta}\left[\frac{-(l+1)(l+m)}{2l+1} P_{l-1}^m(z) + \frac{l(l-m+1)}{2l+1} P_{l+1}^m(z)\right] \end{aligned} \tag{2.63}$$

where we also used

$$z P_l^m(z) = \frac{l+m}{2l+1} P_{l-1}^m(z) + \frac{l-m+1}{2l+1} P_{l+1}^m(z)$$

We apply the obtained result for the derivative of the associated Legendre function to rewrite the expressions for the spinning spherical harmonics:

$$\begin{aligned} S_l^{m^+} &= N_l^m \left(\sin\theta \frac{dP_l^m(\cos\theta)}{d\theta} + m P_l^m(\cos\theta)\right) \exp(jm\phi) \\ &= N_l^m \left(\frac{-(l+1)(l+m)}{2l+1} P_{l-1}^m(\cos\theta) + m P_l^m(\cos\theta) + \frac{l(l-m+1)}{2l+1} P_{l+1}^m(\cos\theta)\right) \exp(jm\phi \\ &= -(l+1)\sqrt{\frac{(l+m)(l-m)}{(2l+1)(2l-1)}} Y_{l-1}^m + m Y_l^m + l\sqrt{\frac{(l-m+1)(l+m+1)}{(2l+1)(2l+3)}} Y_{l+1}^m \end{aligned} \tag{2.64a}$$

where we have used (follows directly from the definitions in (2.13) and (2.35)):

$$N_l^m = N_{l-1}^m \sqrt{\frac{(2l+1)}{(2l-1)}\frac{(l-m)}{(l+m)}}$$

$$N_l^m = N_{l+1}^m \sqrt{\frac{(2l+1)}{(2l+3)}\frac{(l+m+1)}{(l-m+1)}}$$

Likewise, we obtain for $S_l^{m^-}$:

$$S_l^{m^-} = -(l+1)\sqrt{\frac{(l+m)(l-m)}{(2l+1)(2l-1)}}Y_{l-1}^m + mY_l^m + l\sqrt{\frac{(l-m+1)(l+m+1)}{(2l+1)(2l+3)}}Y_{l+1}^m \quad (2.64b)$$

which completes the task of writing the spinning spherical harmonics as a superposition of spherical harmonics. This recombination can subsequently be used in (2.62), so that the spectral components r_{lm} and l_{lm} are obtained as a recombination of the spherical harmonic transform of $\mathcal{F}^+$ and $\mathcal{F}^-$ respectively.

The nature of the functions $\mathcal{F}^{\pm}$ may need some clarification. The received polarisation is adjusted in phase by a factor $e^{\mp j\phi}$ and amplitude tapered with a factor $\frac{1}{\sin\theta}$. The phase factor removes a component we have introduced in the definition of the RHCP and LHCP unit vector distributions of (2.55). That factor was introduced to smooth out the otherwise inelegant discontinuity at the north pole of the spherical coordinate system. With this definition of circular polarisation, the vectorial basis functions were found to be the circular polarised spherical harmonics R_l^m and L_l^m. The spinning spherical harmonics were defined without such a "smoothing factor" and would correspond with the more conventional definition of circular polarisation $\hat{e}_{rhcp} = \hat{e}_\theta + j\hat{e}_\phi$, which we will refer to as a "spinning polarisation" in order to differentiate from the circular polarisation as we defined it. Equations (2.62) imply that we are neutralising the smoothing factor and hence reverting to spinning polarisation. The reader should not be misled by this interpretation. The projections of the circularly polarised components as we have defined them on the circularly polarised spherical harmonics are the same as the projection of the spinning $\mathcal{F}^{\pm}$ on the spinning spherical harmonics. The results we get for the r_{lm} however still correspond with the circularly polarised harmonic components.

In a hardware implementation, the importance of the observations made above amounts to the correct setting of the phase reference for the different elements that sample the field over a spherical surface. We will come back to this in section 7.2.

2.5 Summary and Conclusions

In this chapter Spherical Modal Theory (SMT) was reviewed. Several modal expansions were reviewed that can be found in the literature, and the relations between them have been shown. The different formulations for the multipole expansion are equivalent but depending on the problem, one is more practical to use than the other. Throughout this thesis the most practical formulation has been used, depending on the problem at hand, and this chapter provides the theory that bonds everything together. Table 2.1 gives an overview of the different expansions

that were presented.

Table 2.1: Overview of Spherical Expansions

expansion	**use it when...**	**page**
SFT	...the expansion is of a scalar field.	42
vectorial expansion	...the expansion is of a vectorfield of which the radial and tangential components can be dealt with separately. There is no physical relationship imposed between these components, such as that of the wave equation.	49
SWG expansion	...studying the radiated field as a whole. The evanescence of higher degree modes aids to limit the computing requirements to calculate this expansion. The field obeys the Maxwell equations and the expansion is an SWE.	51
STL expansion	...the evanescent near-field or the propagation characteristics of the modes itself are the subject of study. The modes are defined tangentially to a spherical surface. The field obeys the Maxwell equations and the expansion is an SWE.	53
far-field expansion	...the expansion is for the tangential far-field of an SWE.	55
CPST	...the expansion is for a circularly polarised tangential far-field of an SWE.	57
spinning SFT	...the expansion is for a differently defined circular polarised tangential far-field of an SWE or as an auxiliary theoretical tool.	59

An entirely new expansion, the CPST, was developed that will mostly be used in this thesis. The newly introduced set of multipole basis functions are automatically associated with either the co- or the crosspolar field for a particular choice of polarisation that was defined in section 2.4.3. This simplifies the use of SWEs for circularly polarised spherical array antennas. In section 2.4.4 yet another expansion was introduced that solely plays an auxiliary role: that of relating the calculation of a CPST to the calculation of an SFT.

Chapter 3

Array Elements and their Radiation Pattern

The spherical Array Antenna consists of a number of radiating elements, placed on a conducting sphere. Before studying how all these elements work together in an array configuration, the radiation patterns of individual elements are investigated. This proceeds in two steps:

- Obtain modal expansion for the radiated field, using the spherical wave expansion defined in chapter 2.
- Reconstruct the field from its expansion coefficients.

For the first step a mathematical model for the radiating element is required. As will be shown in section 3.1, two types of element can be distinguished, each slightly different in the way they are represented mathematically. The first type of element is dealt with in greater detail in section 3.2. The second type of element is covered in section 3.3.

The second step, the reconstruction of the field, is shown in section 3.4. We will see that the far-field can be accurately described by the analytically obtained modal expansions. For the reactive near-field however, it is necessary to refine the expansion numerically. In chapter 4 the mutual coupling effects between radiating elements is studied, for which an accurate description of the reactive near-field is necessary. Numerical refinement of the expansions is performed in section 3.5. In the same section the results obtained in previous sections will be backed up with numerical verification.

Finally, in section 3.6 attention is given to the selection of a practical element to use on the spherical array antenna. This section is of general interest for conformal antennas that employ circularly polarised elements.

3.1 Type 1 and Type 2 Elements

We will follow the orthogonality relation method presented in [74] to solve the so-called exterior electromagnetic boundary value problem. An alternative method, the recurrence relation method, is presented in [75]. A solution is sought for the radiated field from an element on the sphere by considering the volume that exists between the sphere and a sphere of infinite radius. The fields inside that volume are completely specified by the sources in the volume and their boundary conditions. This will allow us to derive formula's for the spherical wave expansion

coefficients. A first and easy boundary condition is obtained by forcing the fields to vanish at infinity. The other boundary condition comes from the conducting sphere on which the radiating element is mounted. The boundary conditions are enforced with the Lorentz reciprocity theorem[74] on the volume of space that is bounded by the conducting sphere and the sphere at infinity:

$$\int_S \left(\vec{E}_1 \times \vec{H}_2 - \vec{E}_2 \times \vec{H}_1\right) = \int_V \left(\vec{E}_2\vec{J}_1 - \vec{E}_1\vec{J}_2 - \vec{H}_2\vec{M}_1 + \vec{H}_1\vec{M}_2\right) \tag{3.1}$$

This theorem is valid for any two fields $(\vec{E}_1, \vec{H}_1)$ and $(\vec{E}_2, \vec{H}_2)$ with their electric and magnetic sources $(\vec{J},\vec{M})$, present in the volume V bounded by S. Let $(\vec{E}_1, \vec{H}_1)$ be the source field outside the sphere for which the spherical wave expansion will be sought. On the sphere, the tangential component of $\vec{E}_1$ is $\vec{E}_t$. Let $(\vec{E}_2, \vec{H}_2)$ be a secondary field of a spherical harmonic wave, parameterised by:

- Its degree (l)
- Its order (m)
- TM or TE wave
- Outbound or inbound wave

As we consider the antenna to be placed in a source-free medium, there will only be outbound waves. The last parameter is dropped, and in the rest of this chapter the radial function is replaced by a Hankel function of the first kind. By sweeping the parameters over all possible values, the source field is effectively projected into the spherical harmonic base, and the expansion coefficients are found.

With reference to figure 3.1, equation (3.1) can be simplified in one of the following ways:

- A type 1 radiator:
  - Vanishing source terms
  - Known tangential electric field
- A type 2 radiator:
  - Known source terms
  - Vanishing tangential electric field

The surface integral may be performed over the conducting sphere only, because the fields have to vanish at infinity.

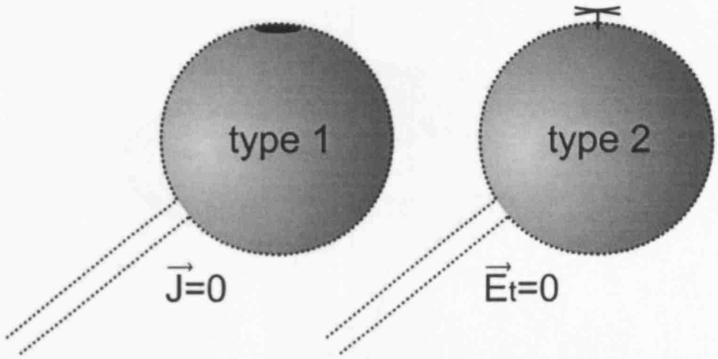


Figure 3.1: Lorentz reciprocity theorem is applied to the exterior space of the sphere, over the surface indicated by the dotted line. Depending on the sort of radiator on the sphere, an aperture or a surmounted element, the theorem simplifies in two different ways.

3.2 Modelling Type 1 Elements

Type 1 elements consist of a small aperture in a large conducting sphere. The aperture is fed from behind with a waveguide or cavity. It suffices to write down the field in the aperture to work out the radiation outside the spherical surface of the antenna. We will consider circular apertures, fed by circular waveguides. In section 3.2.1, the field in the circular aperture is obtained as a function of the modes of propagation that can exist in the circular waveguide. It will be assumed that the apertures are small relative to the sphere, so that their curvature can be neglected. This is a reasonable assumption since we are focussing on large high-gain spherical array antennas. In section 3.2.2 the radiated field is obtained from the field in the aperture. This is done by obtaining the expansion coefficients of the spherical wave expansion. In section 3.2.3 the expansion is rewritten in a form that better describes the reactive near-field.

3.2.1 Source Field in the Aperture

In this section the source field that excites the aperture will be developed and expressed in the spherical coordinate system. This will be done for a circular aperture, formed by a circular waveguide that protrudes a large spherical conducting shell.

Also in the circular waveguide there are two sets of solutions: one for the TE modes and one for the TM modes. The subscript n indicates the number of full-wave patterns around the circumference of the waveguide, whereas p indicates the number of half-wave patterns across the diameter. The fundamental mode is the TE_{11}-mode. For a TE_{np}-mode, the mode is defined

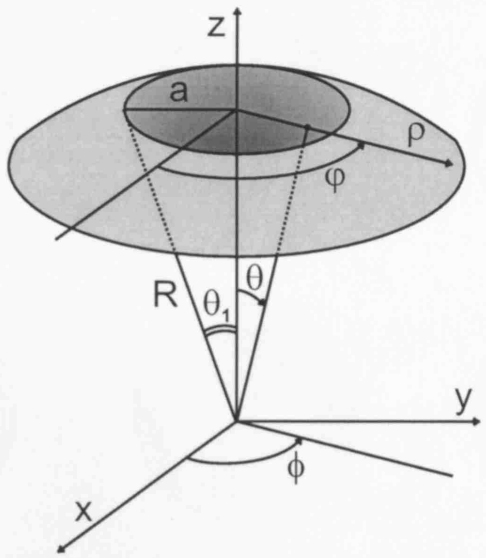


Figure 3.2: The coordinates in the circular aperture are transformed to the spherical coordinate system so that the circular modes can be written in spherical coordinates.

as:

$$
\begin{aligned}
E_\rho &= \mp V_{np}^{\mathrm{TE}} \sqrt{\frac{\epsilon_n}{\pi}} \frac{n}{\sqrt{s_{np}^{\mathrm{TE}\,2} - n^2}} \frac{1}{J_n(s_{np}^{\mathrm{TE}})} \frac{1}{\rho} J_n\left(\frac{s_{np}^{\mathrm{TE}}\rho}{a}\right) \begin{matrix}\sin\\ \cos\end{matrix}(n\varphi) \\
E_\varphi &= V_{np}^{\mathrm{TE}} \sqrt{\frac{\epsilon_n}{\pi}} \frac{s_{np}^{\mathrm{TE}}}{\sqrt{s_{np}^{\mathrm{TE}\,2} - n^2}} \frac{1}{J_n(s_{np}^{\mathrm{TE}})} \frac{1}{a} J_n'\left(\frac{s_{np}^{\mathrm{TE}}\rho}{a}\right) \begin{matrix}\cos\\ \sin\end{matrix}(n\varphi) \\
E_z &= 0 \\
H_\rho &= -I_{np}^{\mathrm{TE}} \sqrt{\frac{\epsilon_n}{\pi}} \frac{s_{np}^{\mathrm{TE}}}{\sqrt{s_{np}^{\mathrm{TE}\,2} - n^2}} \frac{1}{J_n(s_{np}^{\mathrm{TE}})} \frac{1}{a} J_n'\left(\frac{s_{np}^{\mathrm{TE}}\rho}{a}\right) \begin{matrix}\cos\\ \sin\end{matrix}(n\varphi) \\
H_\varphi &= \mp\ I_{np}^{\mathrm{TE}} \sqrt{\frac{\epsilon_n}{\pi}} \frac{n}{\sqrt{s_{np}^{\mathrm{TE}\,2} - n^2}} \frac{1}{J_n(s_{np}^{\mathrm{TE}})} \frac{1}{\rho} J_n\left(\frac{s_{np}^{\mathrm{TE}}\rho}{a}\right) \begin{matrix}\sin\\ \cos\end{matrix}(n\varphi) \\
H_z &= -jV_{np}^{\mathrm{TE}} \frac{\lambda s_{np}^{\mathrm{TE}}}{2\pi a \eta} \sqrt{\frac{\epsilon_n}{\pi}} \frac{s_{np}^{\mathrm{TE}}}{\sqrt{s_{np}^{\mathrm{TE}\,2} - n^2}} \frac{1}{J_n(s_{np}^{\mathrm{TE}})} \frac{1}{a} J_n\left(\frac{s_{np}^{\mathrm{TE}}\rho}{a}\right) \begin{matrix}\cos\\ \sin\end{matrix}(n\varphi)
\end{aligned}
\tag{3.2}
$$

This definition is valid for two orthogonal linear polarisations, depending on whether the top or bottom trigonometric functions are used. The definition is identical to the one used in [58], if one takes into account that a different time convention is used. The modal fields are dimensionless. The coordinates ρ and φ are shown in figure 3.2; s_{np}^{TE} is the p^{th} root of the derivative of the Bessel function J_n'; ϵ_n is 1 for $n = 0$ and 2 otherwise. Some modes are illustrated in figure 3.3.

Using figure 3.2, the variables in (3.2) will be converted to those of the spherical coordinate system. Here an approximation is made that is valid as long as the radius a of the aperture is

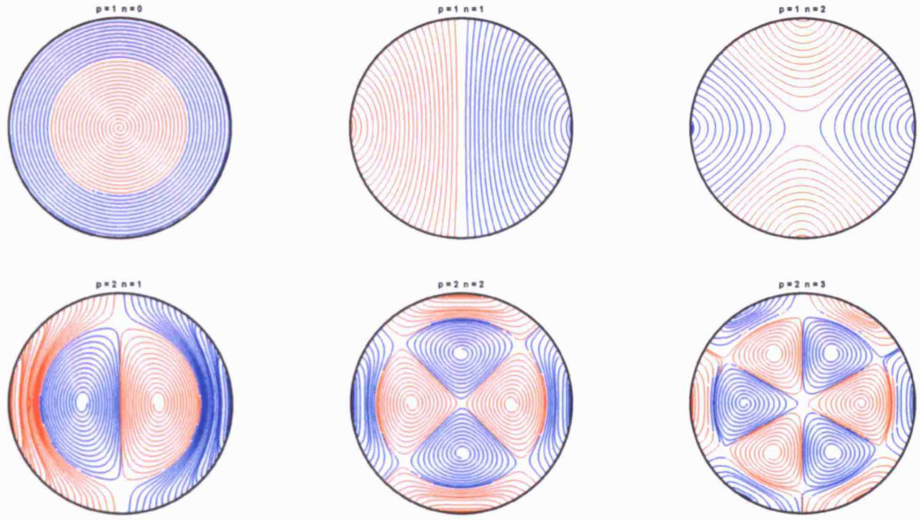


Figure 3.3: Electric field lines for a few TE_{np} circular modes. The colour gives a sense of rotation (curl) of the field lines. E.g.. for the TE_{11}-mode the field lines are all oriented upwards.

small compared to the radius R of the sphere:

$$\rho \simeq R\theta \tag{3.3a}$$

$$\theta_1 = \arcsin\left(\frac{a}{R}\right) \simeq \frac{a}{R} \tag{3.3b}$$

Hence, the following substitutions can be made:

$$\begin{cases} \phi \rightarrow \varphi \\ \dfrac{\theta}{\theta_1} \rightarrow \dfrac{\rho}{a} \end{cases} \tag{3.4}$$

Also the co-polar and crosspolar components are combined (90° out of phase) into RHCP and LHCP components. For a TE mode, the tangential electric field becomes:

$$E_\theta = \mp V_{np}^{\mathrm{TE}} N_{apn}^{\mathrm{TE}}\, jn \left(\frac{\theta_1}{\theta s_{np}^{\mathrm{TE}}}\right) J_n\left(\frac{s_{np}^{\mathrm{TE}}\theta}{\theta_1}\right) \exp(\pm jn\phi) \tag{3.5a}$$

$$E_\phi = V_{np}^{\mathrm{TE}} N_{apn}^{\mathrm{TE}} J_n'\left(\frac{s_{np}^{\mathrm{TE}}\theta}{\theta_1}\right) \exp(\pm jn\phi) \tag{3.5b}$$

with the aperture constant:

$$N_{apn}^{\mathrm{TE}} = \sqrt{\frac{1}{\pi}} \frac{s_{np}^{\mathrm{TE}}}{a\sqrt{{s_{np}^{\mathrm{TE}}}^2 - n^2} J_n(s_{np}^{\mathrm{TE}})} \tag{3.5c}$$

and the top signs for RHCP and the bottom signs for LHCP. For $n < 0$ we will adopt the conventions that $J_{-n} = J_n$ and that negative values for n correspond to LHCP, so that the definition for RHCP becomes valid for both polarisations. The tangential electric field for a

TM-mode is given by:

$$E_\theta = -V_{np}^{\mathsf{TM}} N_{apn}^{\mathsf{TM}} J_n'\left(\frac{s_{np}^{\mathsf{TM}}\theta}{\theta_1}\right)\exp(jn\phi) \tag{3.6a}$$

$$E_\phi = -V_{np}^{\mathsf{TM}} N_{apn}^{\mathsf{TM}} jn\left(\frac{\theta_1}{\theta s_{np}^{\mathsf{TM}}}\right) J_n\left(\frac{s_{np}^{\mathsf{TM}}\theta}{\theta_1}\right)\exp(jn\phi) \tag{3.6b}$$

with the aperture constant:

$$N_{apn}^{\mathsf{TM}} = \sqrt{\frac{1}{\pi}}\frac{1}{aJ_{n+1}(s_{np}^{\mathsf{TM}})} \tag{3.6c}$$

and where s_{np}^{TM} is the p^{th} root of J_n.

From here on, the modal voltage V_{np} is assumed to be normalised so that the power carried by the mode in a perfectly matched waveguide would be $P_{np} = \frac{1}{2}V_{np}I_{np}^* = 1$ Watt:

$$\begin{aligned}
V_{np}^{\mathsf{TE}} &= \sqrt{\frac{2\eta}{\sqrt{1-\left(\frac{f_c}{f}\right)^2}}} & V_{np}^{\mathsf{TM}} &= \sqrt{2\eta\sqrt{1-\left(\frac{f_c}{f}\right)^2}} \\
I_{np}^{\mathsf{TE}} &= \frac{V_{np}^{\mathsf{TE}}}{Z_{np}^{\mathsf{TE}}} & \text{with} \quad Z_{np}^{\mathsf{TE}} &= \frac{j\eta}{\sqrt{\left(\frac{f_c^{\mathsf{TE}}}{f}\right)^2-1}} \\
I_{np}^{\mathsf{TM}} &= \frac{V_{np}^{\mathsf{TM}}}{Z_{np}^{\mathsf{TM}}} & \text{with} \quad Z_{np}^{\mathsf{TM}} &= -j\eta\sqrt{\left(\frac{f_c^{\mathsf{TM}}}{f}\right)^2-1}
\end{aligned} \tag{3.7}$$

and with

$$f_c^{\mathsf{TE}} = \frac{s_{np}^{\mathsf{TE}}}{2\pi a\sqrt{\mu\varepsilon}} \qquad f_c^{\mathsf{TM}} = \frac{s_{np}^{\mathsf{TM}}}{2\pi a\sqrt{\mu\varepsilon}}$$

The characteristic impedances are formulated to be valid both above and below cut-off.

3.2.2 Solving the Reciprocity Theorem

As the right hand side of (3.1) is zero for a type 1 radiator, the solution is obtained by solving the following integral equation:

$$\int_S \left(\vec{E}_1 \times \vec{H}_2\right)\cdot \hat{e}_r dS = \int_S \left(\vec{E}_2 \times \vec{H}_1\right)\cdot \hat{e}_r dS \tag{3.8}$$

A Solving the left integral

The left integral takes the form:

$$\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\theta_1} \left(E_{1_\theta}H_{2_\phi} - E_{1_\phi}H_{2_\theta}\right)\sin\theta d\theta d\phi \tag{3.9}$$

By taking one spherical mode at a time for the secondary field, an expression will be arrived for all coefficients in the expansion. four cases can be distinguished:

- TE-mode in the aperture, secondary field is a spherical TM-mode.
- TM-mode in the aperture, secondary field is a spherical TE-mode.
- TE-mode in the aperture, secondary field is a spherical TE-mode.
- TM-mode in the aperture, secondary field is a spherical TM-mode.

In any of the cases above, the outer integral can be removed by the orthogonality of the azimuthal phase functions:

$$\int_{\phi=0}^{2\pi} e^{jn\phi} e^{jm\phi} d\phi = \begin{cases} 2\pi & \text{for } m = -n \\ 0 & \text{for } m \neq -n \end{cases}$$

Hence, the left integral vanishes but for the spherical modes that have an order $m = -n$. For these modes, the factors $N_l^{-n} P_l^{-n}$ are replaced by $(-1)^n N_{ln} P_l^n$.

TE to TM By substituting the tangential fields (3.5) for $\vec{E}_1$ and a spherical $(l, -n)$-mode from (2.36b) for $\vec{H}_2$ into (3.9), the integral becomes:

$$\kappa_1 \int_{\theta=0}^{\theta_1} \frac{\theta_1}{s_{np}^{\mathrm{TE}}\,\theta} J_n\left(\frac{s_{np}^{\mathrm{TE}}\theta}{\theta_1}\right) \frac{dP_l^n(\cos\theta)}{d\theta} \sin\theta d\theta + \kappa_1 \int_{\theta=0}^{\theta_1} J_n'\left(\frac{s_{np}^{\mathrm{TE}}\theta}{\theta_1}\right) P_l^n(\cos\theta) d\theta$$

with

$$\kappa_1 = \frac{-2\pi kn}{\sqrt{\eta}} (-1)^n V_{np}^{\mathrm{TE}} N_{apn}^{\mathrm{TE}} N_{ln} h_l(kr)$$

For small angles we have $\sin\theta \approx \theta$. Integration by parts then reveals a closed expression for the integral:

$$\begin{aligned} &\kappa_1 \frac{\theta_1}{s_{np}^{\mathrm{TE}}} \int_{\theta=0}^{\theta_1} \left(J_n\left(\frac{s_{np}^{\mathrm{TE}}\theta}{\theta_1}\right) \frac{dP_l^n(\cos\theta)}{d\theta} + \frac{dJ_n\left(\frac{s_{np}^{\mathrm{TE}}\theta}{\theta_1}\right)}{d\theta} P_l^n(\cos\theta) \right) d\theta \\ =&\kappa_1 \frac{\theta_1}{s_{np}^{\mathrm{TE}}} J_n\left(s_{np}^{\mathrm{TE}}\right) P_l^n(\cos\theta_1) \end{aligned} \tag{3.10}$$

TM to TE By substituting the tangential fields (3.6) for $\vec{E}_1$ and a spherical $(l, -n)$-mode from (2.36a) for $\vec{H}_2$ into (3.9), the integral takes the same form as for TE to TM coupling.

The solution becomes:

$$\kappa_2 \frac{\theta_1}{s_{np}^{\text{TM}}} J_n\left(s_{np}^{\text{TM}}\right) P_l^n(\cos\theta_1) = 0 \tag{3.11}$$

$$\left(\text{with} \quad \kappa_2 = \frac{-2\pi kn}{\sqrt{\eta}}(-1)^n V_{np}^{\text{TM}} N_{apn}^{\text{TM}} N_{ln} \frac{1}{kr} \frac{d}{d(kr)} \left[krh_l(kr)\right]\right)$$

TE to TE By substituting the tangential fields (3.5) for $\vec{E}_1$ and a spherical $(l,-n)$-mode from (2.36a) for $\vec{H}_2$ into (3.9), the integral becomes:

$$\kappa_3 \int_{\theta=0}^{\theta_1} n^2 \frac{\theta_1}{s_{np}^{\text{TE}}\theta} J_n\left(\frac{s_{np}^{\text{TE}}\theta}{\theta_1}\right) P_l^n(\cos\theta)d\theta + \kappa_3 \int_{\theta=0}^{\theta_1} J_n'\left(\frac{s_{np}^{\text{TE}}\theta}{\theta_1}\right) \frac{dP_l^n(\cos\theta)}{d\theta} \sin\theta d\theta$$

with

$$\kappa_3 = \frac{2\pi jk}{\sqrt{\eta}}(-1)^n V_{np}^{\text{TE}} N_{apn}^{\text{TE}} N_{ln} \frac{1}{kr} \frac{d}{d(kr)} \left[krh_l(kr)\right]$$

With the notations $x = \frac{\theta s_{np}^{\text{TE}}}{\theta_1}$ and $f' = \frac{df}{dx}$, and leaving out the subscripts and κ_3, the integral simplifies to:

$$\int_{\theta=0}^{\theta_1} \frac{n^2}{x} J(x)P(\cos\theta)d\theta + \int_{\theta=0}^{\theta_1} J'(x)\frac{s}{\theta_1}P'(\cos\theta)\sin\theta d\theta$$

By adding and subtracting $xJ(x)P(\cos\theta)$, this becomes:

$$\int_{\theta=0}^{\theta_1} \left(\frac{n^2}{x} - x\right) J(x)P(\cos\theta)d\theta + \int_{\theta=0}^{\theta_1} J'(x)\frac{s}{\theta_1}P'(\cos\theta)\sin\theta d\theta + \int_{\theta=0}^{\theta_1} xJ(x)P(\cos\theta)d\theta$$

Using the Bessel equation $xJ''(x) + J'(x) + \left(x - \frac{n^2}{x}\right) J(x) = 0$ transforms it into:

$$\int_{\theta=0}^{\theta_1} xJ''P + \int_{\theta=0}^{\theta_1} J'P + \int_{\theta=0}^{\theta_1} J'\frac{s}{\theta_1}P'\sin\theta + \int_{\theta=0}^{\theta_1} xJP$$

Integration by parts on the first term then gives:

$$\cancel{[J'xP]_0^{\theta_1}} - \cancel{\int_{\theta=0}^{\theta_1} J'P} - \int_{\theta=0}^{\theta_1} J'xP' + \cancel{\int_{\theta=0}^{\theta_1} J'P} + \int_{\theta=0}^{\theta_1} J'\frac{s}{\theta_1}P'\sin\theta + \int_{\theta=0}^{\theta_1} xJP$$

Now the small aperture approximation $\sin\theta \approx \theta$ is made, so that the third and fifth terms cancel out and there is only one term left. Written out in full, the integral is simplified to:

$$\kappa_3 \int_{\theta=0}^{\theta_1} \frac{s_{np}^{\text{TE}}\theta}{\theta_1} J_n\left(\frac{s_{np}^{\text{TE}}\theta}{\theta_1}\right) P_l^n(\cos\theta)d\theta \tag{3.12}$$

TM to TM By substituting the tangential fields (3.6) for $\vec{E}_1$ and a spherical $(l,-n)$-mode from (2.36b) for $\vec{H}_2$ into (3.9), the integral takes the same form as for TE to TE coupling:

$$\kappa_4 \int_{\theta=0}^{\theta_1} \frac{s_{np}^{\mathrm{TM}}\theta}{\theta_1} J_n\left(\frac{s_{np}^{\mathrm{TM}}\theta}{\theta_1}\right) P_l^n(\cos\theta)d\theta \tag{3.13}$$

with

$$\kappa_4 = \frac{-2\pi jk}{\sqrt{\eta}}(-1)^n V_{np}^{\mathrm{TM}} N_{apn}^{\mathrm{TM}} N_{ln} h_l(kr)$$

B Solving the right integral

In the right hand side of (3.8), the source field $\vec{H}_1$ is expanded according to (2.33). The purpose is to write down an expression for the coefficients of this expansion. For $\vec{E}_2$ the same spherical mode is used as the one that was used for $\vec{H}_2$ in the left integral. The integral on the right side simplifies by the orthogonality property of the spherical vector harmonics[63].

The orthogonality used in the right integral could have been applied at the beginning, together with the reciprocity theorem. The circular modal fields are then projected onto the spherical modal fields by their internal product. This alternative approach to derive the coefficients results in the same integrals to be solved.

Coupling into spherical TE-modes In two of the four cases of the previous section, the secondary field was taken to be a transverse electric spherical mode of order $m = -n$. For these cases the transverse electric expansion coefficients for the primary field are found:

$$\begin{aligned}
\vec{E}_2 &= k\sqrt{\eta}\sum_{l,m}\vec{M}_l^m \\
\int_S \left(\vec{E}_2 \times \vec{H}_1\right)\cdot\hat{e}_r dS &= -jk^2\int_S\left(\vec{M}_l^m \times \sum_{l'm'}\left(C_{l'm'}^{\mathrm{TE}}\vec{N}_{l'}^{m'} + C_{l'm'}^{\mathrm{TM}}\vec{M}_{l'}^{m'}\right)\right)\hat{e}_r dS \\
&= -jk^2\sum_{l'm'} C_{l'm'}^{\mathrm{TE}}(-1)^{m'+2}h_l(kr)\frac{1}{kr}\frac{d}{d(kr)}\left[krh_{l'}(kr)\right]\delta_{m',-m}\delta_{l',l} \\
&= -C_{ln}^{\mathrm{TE}}\, jk^2\,(-1)^n\, h_l(kr)\,\frac{1}{kr}\frac{d}{d(kr)}\left[krh_l(kr)\right]
\end{aligned} \tag{3.14}$$

Coupling into spherical TM-modes In the two other cases, the secondary field was taken to be a transverse magnetic spherical mode of order $m = -n$:

$$\begin{aligned}
\vec{E}_2 &= k\sqrt{\eta}\sum_{l,m}\vec{N}_l^m \\
\int_S \left(\vec{E}_2 \times \vec{H}_1\right)\cdot\hat{e}_r dS &= -jk^2\int_S\left(\vec{N}_l^m \times \sum_{l'm'}\left(C_{l'm'}^{\mathrm{TE}}\vec{N}_{l'}^{m'} + C_{l'm'}^{\mathrm{TM}}\vec{M}_{l'}^{m'}\right)\right)\hat{e}_r dS \\
&= -jk^2\sum_{l'm'} C_{l'm'}^{\mathrm{TM}}(-1)^{m'+1}\frac{1}{kr}\frac{d}{d(kr)}\left[krh_l(kr)\right]h_{l'}^1(kr)\delta_{m',-m}\delta_{l',l} \\
&= C_{ln}^{\mathrm{TM}}\, jk^2\,(-1)^n\, h_l(kr)\,\frac{1}{kr}\frac{d}{d(kr)}\left[krh_l(kr)\right] \qquad (3.15)
\end{aligned}$$

C Solutions for the SWE coefficients

Finally, by equating the appropriate expressions for the left and the right integral, a closed expression is obtained for all coefficients in the spherical wave expansion of the field outside of the sphere, caused by an arbitrary mode in the circular aperture. An arbitrary excitation in the aperture can be represented by superposition of circular modes. Let $R_l^\alpha(r)$ and $R_l^\beta(r)$ be defined as:

$$R_l^\alpha(r) = h_l(kr) \qquad (3.16a)$$

$$R_l^\beta(r) = \frac{1}{kr}\frac{d}{d(kr)}\left[krh_l(kr)\right] = \frac{h_l(kr)}{kr} + h_l'(kr) \qquad (3.16b)$$

For a **circular transverse electric mode in the aperture**, the coefficients for the SWE are obtained from (3.10) and (3.15):

$$C_{lm}^{\mathrm{TM}} = \begin{cases} \dfrac{-2\pi jn\theta_1 V_{np}^{\mathrm{TE}} N_{apn}^{\mathrm{TE}} N_{ln}}{s_{np}^{\mathrm{TE}}\sqrt{\eta}k\, R_l^\beta(R)} J_n\left(s_{np}^{\mathrm{TE}}\right) P_l^n(\cos\theta_1) & \text{for } m = n \\ 0 & \text{for } m \neq n \end{cases} \qquad (3.17)$$

The nominator is zero for $n = 0$. For TE_{0x} circular modes, the electric field in the aperture has only a ϕ-component. The electric field of spherical TM modes of zero order do not have a ϕ-component. Hence the fact that circular TE_{0x}-modes do not couple into spherical TM-modes. From (3.12) and (3.14):

$$C_{lm}^{\mathrm{TE}} = \begin{cases} \dfrac{-2\pi V_{np}^{\mathrm{TE}} N_{apn}^{\mathrm{TE}} N_{ln}}{\sqrt{\eta}k\, R_l^\alpha(r)} \displaystyle\int_{\theta=0}^{\theta_1} \frac{s_{np}^{\mathrm{TE}}\theta}{\theta_1} J_n\left(\frac{s_{np}^{\mathrm{TE}}\theta}{\theta_1}\right) P_l^n(\cos\theta)d\theta & \text{for } m = n \\ 0 & \text{for } m \neq n \end{cases} \qquad (3.18)$$

A circular transverse magnetic mode in the aperture cannot support spherical transverse electric modes. As the magnetic field has no radial component over the conducting sphere, nor in the aperture (because of the transverse magnetic mode), and the radial component of the electric field is also zero over a complete sphere (as would be required for a transverse electric spherical mode), only a TEM mode could exist. Spherical waveguides do not support divergenceless TEM modes. That is why expression (3.11) equated to zero. Consequently:

$$C_{lm}^{\mathrm{TE}} = 0 \tag{3.19}$$

The coefficients for the transverse magnetic modes are obtained from (3.13) and (3.15):

$$C_{lm}^{\mathrm{TM}} = \begin{cases} \dfrac{2\pi V_{np}^{\mathrm{TM}} N_{apn}^{\mathrm{TM}} N_{ln}}{\sqrt{\eta} k \, R_l^{\beta}(r)} \displaystyle\int_{\theta=0}^{\theta_1} \frac{s_{np}^{\mathrm{TM}}\theta}{\theta_1} J_n\left(\frac{s_{np}^{\mathrm{TM}}\theta}{\theta_1}\right) P_l^n(\cos\theta) d\theta & \text{for } m = n \\ 0 & \text{for } m \neq n \end{cases} \tag{3.20}$$

For a right hand circularly polarised aperture, the circular modes of order n only couple into spherical modes of the same order. If LHCP circular modes are considered, only circular modes of order $-n$ are excited. A linear polarised field in the aperture couples equally into both spherical modes of order $m = n$ and $m = -n$.

3.2.3 Transmission Line Description

The expansion that has been used in section 3.2.2 to describe the near- and far-field from a single element is not suitable for calculating the reactive near field. The used expansion, which we termed "spherical waveguide expansion" in chapter 2 takes the cut-off characteristics in the spherical waveguide into account when the modal coefficients are calculated. This offers the advantage that the expansion into modal coefficients can be truncated at a reasonably low degree. The other expansion, termed "spherical transmission line expansion" in chapter 2, expands the excitation field into all modes, including the evanescent modes and incorporates the propagation behaviour into the modal impedances. It is necessary to include a lot more modes to reconstruct the true field on the sphere. If the field that exists on the sphere could be represented with modes that can propagate to the far-field, a pencil beam could be formed with the same angle as the aperture, when viewed from the centre of the sphere. That is obviously impossible, and an infinite number of modes is needed to reconstruct the reactive near-field exactly.

Theoretically the resultant fields calculated with each of the aforementioned expansions are equivalent, provided they are truncated at the same degree. Any of the two expansions can be converted in the other one. Nevertheless, when one is interested in the near-field, it does not make any sense to introduce numerical difficulties by consecutively dividing and multiplying by a very large number for the higher degrees. The "spherical transmission line expansion" avoids these numerical difficulties. In chapter 4 we are interested in the reactive near-field to work out the mutual coupling effects between different array elements. We will therefore also develop the transmission line expansion solution for the field of a single element.

The results for the spherical waveguide coefficients in section 3.2.2 can be converted to the modal voltages and currents of (2.38) by using (2.39). For a **circular transverse electric mode in the aperture**:

$$V_{lm}^{\mathrm{TE}}(r=R) = 2\pi R V_{np}^{\mathrm{TE}} N_{apn}^{\mathrm{TE}} N_{ln} \int_{\theta=0}^{\theta_1} \frac{s_{np}^{\mathrm{TE}}\theta}{\theta_1} J_n\left(\frac{s_{np}^{\mathrm{TE}}\theta}{\theta_1}\right) P_l^n(\cos\theta) d\theta \tag{3.21a}$$

$$V_{lm}^{\mathrm{TM}}(r=R) = 2\pi R j n \frac{\theta_1}{s_{np}^{\mathrm{TE}}} V_{np}^{\mathrm{TE}} N_{apn}^{\mathrm{TE}} N_{ln} J_n\left(s_{np}^{\mathrm{TE}}\right) P_l^n(\cos\theta_1) \tag{3.21b}$$

$$I_{lm}^{\mathrm{TE}}(r=R) = Y_{lm\mathrm{TE}}^{i}(R) V_{lm}^{\mathrm{TE}}(r=R) \tag{3.21c}$$

$$I_{lm}^{\mathrm{TM}}(r=R) = Y_{lm\mathrm{TM}}^{i}(R) V_{lm}^{\mathrm{TM}}(r=R) \tag{3.21d}$$

where the Y_{lm}^{i} are the input admittances that can be obtained from equation (2.44).

Whereas for a **circular transverse magnetic mode in the aperture**, the modal voltages are:

$$V_{lm}^{\mathrm{TE}}(r=R) = 0 \tag{3.22a}$$

$$V_{lm}^{\mathrm{TM}}(r=R) = -2\pi R V_{np}^{\mathrm{TM}} N_{apn}^{\mathrm{TM}} N_{ln} \int_{\theta=0}^{\theta_1} \frac{s_{np}^{\mathrm{TM}}\theta}{\theta_1} J_n\left(\frac{s_{np}^{\mathrm{TM}}\theta}{\theta_1}\right) P_l^n(\cos\theta) d\theta \tag{3.22b}$$

Reconstruction of the reactive near-field with the expansion above is illustrated further on in section 3.4.3.

3.3 Modelling Type 2 Elements

Type 2 elements consist of a radiating element and a feed that extends outside the conducting sphere. A typical example is the patch antenna, which is often preferred over aperture radiators because the latter need to be fed by bulky waveguides. Patch type antennas have the practical advantage that they can be implemented as a printed circuit, which greatly improves manufacturing accuracy and repeatability.

Provided that the element can be sufficiently well represented by an equivalent magnetic or electric source term, the reciprocity theorem can be solved along the same lines as was done for the type 1 elements. The difficulty here is to obtain the equivalent source term. The exact solution for this problem, called the full-wave solution[76], can be obtained by solving for the surface currents on the patch by using the boundary condition that the total tangential electric field on the patch must be zero (the method as summarised here is described in detail in [77] and [76]). The total tangential field has two terms: the scattered electric field radiated by the currents in the patch and the incident field from the feedline or probe. It is assumed that the latter is known or can be obtained. Taking the spherical geometry into account, the scattered field can be written as an integral over the product of the unknown currents with appropriate Green functions, or else as a linear operator L that acts on the patch currents. The result of the operator acting on the patch currents needs to cancel out the incident electric field:

$$L(\vec{J}) = \vec{E}^{i} \tag{3.23}$$

Because the operator cannot easily be inverted, the Method of Moments (MoM) can be used to extract the unknown patch currents. In the MoM, the patch currents are expanded in basis functions and the expansion coefficients I_n become the unknowns:

$$\vec{J} = \sum_n I_n \vec{J}_n \tag{3.24}$$

The equation to solve becomes:

$$\sum_n I_n \boldsymbol{L}(\vec{J}_n) = \vec{E}^i \tag{3.25}$$

The problem of finding the I_n is turned into a linear system of equations by projecting both sides of the equation on a set of weighting functions. In Galerkin's method, the basis currents J_n are used as weighting functions, and the linear projections of the electric field on the basis currents are reactions:

$$\sum_n I_n \langle \vec{J}_m, \boldsymbol{L}(\vec{J}_n) \rangle = \langle \vec{J}_m, \vec{E}^i \rangle \tag{3.26}$$

The quantities in this matrix equation can be interpreted as generalised impedances, currents and voltages, so that the equation can be written simply as:

$$\mathbf{ZI} = \mathbf{V} \tag{3.27}$$

This method is either very involved or, if the integrations are done numerically, computationally expensive. Furthermore, solving for the full-wave solution is only desirable for small spheres, where the curvature is not negligible. In such cases, the patches would probably be curved as well. Curved patch antennas over curved groundplanes attracted a lot of attention in the 90's[78][79][80][81] and the interested reader will find a thorough overview in [76]. As was already mentioned in the introduction to this thesis, the applications we have in mind for spherical array antennas do require a reasonable amount of antenna gain, which in turn requires the spherical array antenna to be large. Therefore it was deemed out of scope to consider curved patches, just as large apertures were excluded in section 3.2.

The question is then how to obtain a sufficiently accurate equivalent source term for the type 2 elements, without having to calculate the full-wave solution. Two simplified models can be used: the cavity model and the transmission line model[76]. We will use the cavity model to represent the patch antennas above the sphere. As will be discussed in section 3.3.1, the cavity model is based on some assumptions that are more accurate for some elements than for others. The Shorted Annular Ring (SAR) antenna can well be represented by the cavity model and we will use this antenna to illustrate how to find a spherical wave expansion for the field radiated by a type 2 element. The actual choice of radiating elements for practical applications is left to section 3.6.

Once a suitable equivalent source term is obtained, finding the rest of the solution proceeds more or less in the same way as for the type 1 elements. For the SAR element this is done in section 3.3.2.

3.3.1 The Source Current of Shorted Annular Ring Antennas

In this section we use the cavity model to derive a simple equivalent source term for the SAR element. For a more in-depth discussion on the design and performance of the SAR element, refer to section 3.6.2.

The SAR element belongs to the family of circular patch antennas. These antennas have a cylindrical volume closed by a conducting sheet of metal at the top and the groundplane at the bottom. The sidewall of the cylinder is a radiating aperture. A circular patch antenna has one open sidewall, whereas an annular ring antenna has an inner and an outer sidewall. The shorted annular ring also has an open outer sidewall, whereas its inner sidewall is shorted to the groundplane. In the cavity model, the open sidewalls are replaced with sheets of magnetic conductors, so that a cavity is formed by the electric and magnetic walls. This assumption is valid for thin substrates, so that the following observations can be made:

- The field between the patch and the groundplane does not vary with height.
- The electric field has only a pure outward component, orthogonal to the groundplane and the patch, whereas the magnetic field is purely parallel to the groundplane.
- At the edges, there is no current normal to the edge, so that the magnetic field vanishes at the edge. Therefore, the open edge can be replaced with a perfect magnetic conductor.
- Fringing fields are negligible.

In the closed cavity a finite number of modes can exist. These modes have well known fields in the cavity. To work out how the modal fields in the cavity radiate, the equivalence principle is used to replace the magnetic conductor with a sheet of magnetic currents[82]. The currents are readily derived from the modes that can exist in the cavity. Once the currents are known, the radiated field can be obtained as in section 3.3.2.

By the assumptions taken on the boundary conditions, the modes that can exist in the cavity are TM_{np} modes. It would be inefficient to make the patch large enough to support modes for which $p > 1$. Generally, the TM_{11} mode is chosen for boresight radiation. The TM_{01} mode is sometimes used for broadside radiation, as this mode excites surface waves and lateral waves along the groundplane interface. Usually this is undesirable, but generally it is not possible to avoid excitation of the TM_{01} mode with patch-type antennas such as the circular disk or annular ring. The SAR antenna however, can be designed so that it does not excite the TM_{01} mode, more on this will be covered in section 3.6.2. Theoretically the gain of the antenna can be improved further by using higher order modes, with $n > 1$ and making the radius of the ring larger. Some attempts in this direction were made, but combining a clear mode separation with a well matched feed proved to be difficult.

As explained in [83], the sheet of magnetic current can be folded up by superposition into one single loop of magnetic current. The equivalent magnetic current source for a SAR patch

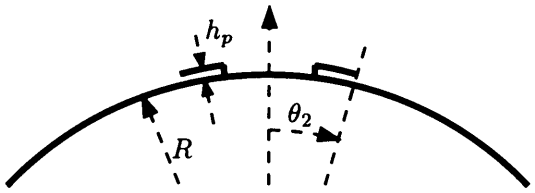


Figure 3.4: Geometry of the Shorted Annular Ring (SAR) element on the sphere.

antenna excited with a TM_{np} mode is given by [80][84]:

$$\vec{M}(\phi) = \frac{M_0}{R}\delta(\theta-\theta_2)\delta(r-r_p)e^{jn\phi} \tag{3.28a}$$

$$M_0 = E_0 h_p \left(J_n(k_{pn}R\theta_2)Y_n'(k_{pn}R\theta_1) - J_n'(k_{pn}R\theta_1)Y_n(k_{pn}R\theta_2)\right) \tag{3.28b}$$

where $r_p = R + h_p$, h_p the height of the patch above the sphere and R the radius of the sphere. J_n and Y_n are the Bessel functions of the first and second kind respectively and k_{pn} are the roots of the following characteristic equation:

$$J_n'(k_{pn}R\theta_2)Y_n'(k_{pn}R\theta_1) - J_n'(k_{pn}R\theta_1)Y_n'(k_{pn}R\theta_2) = 0 \tag{3.29}$$

The cavity model used to derive equation (3.28) is only valid if h_p is small compared to the wavelength, so it can be safely assumed that $r_p \approx R$. In figure 3.4 the geometry of the SAR element on the sphere is illustrated.

3.3.2 Solving the Reciprocity Theorem

We will now solve the reciprocity integral in equation (3.1) for a type 2 radiator such as the patch antenna. The secondary field remains the same as before, one (l,m)-spherical mode for which we seek an expression for its coefficient. The primary field is again the source field. The source field is no longer determined by its tangential electric field on the surface of the sphere (the tangential field on the surface of the sphere is zero everywhere) but by a source term that exists in the integration volume, which is the space extending from the surface of the sphere to infinity. For the SAR antenna, we can use equation (3.28) for the source term. The reciprocity theorem becomes:

$$\int_S \vec{E}_2 \times \vec{H}_1 = \int_V \vec{H}_2 \cdot \vec{M}_1 \tag{3.30}$$

The left side of this equation has been solved before. The solutions for spherical TE and TM modes were given in equations (3.14) and (3.15) respectively and repeated here:

$$\int_S \left(\vec{E}_2 \times \vec{H}_1\right) \cdot \hat{e}_r dS = \mp C_{ln}^{\substack{\text{TE}\\\text{TM}}} jk^2 (-1)^n h_l(kr) \frac{1}{kr}\frac{d}{d(kr)}[krh_l(kr)] \tag{3.31}$$

Remains the right hand integral to be solved:

$$I = \int_V \vec{H}_2 \cdot \vec{M}_1 = \int_V H_{2_\phi} \cdot M_{1_\phi} \tag{3.32}$$

We get expressions for the $\vec{H}_2$ spherical mode from equations (2.36a) and (2.36b), depending on whether $\vec{H}_2$ is a TE or TM mode respectively:

$$H_{2_\phi} = \frac{-jk}{\sqrt{\eta}} N_{lm} \frac{1}{kr} \frac{d}{d(kr)} \left[kr h_l(kr) \right] \frac{jm}{\sin\theta} P_l^m(\cos\theta) e^{jm\phi} \quad \text{for TE mode} \tag{3.33}$$

$$H_{2_\phi} = \frac{jk}{\sqrt{\eta}} N_{lm} h_l(kr) \frac{dP_l^m(\cos\theta)}{d\theta} e^{jm\phi} \quad \text{for TM mode} \tag{3.34}$$

As was the case in the derivations for circular apertures, only when $m = -n$, the integral will be different from zero. By substituting (3.28) for $\vec{M}_1$ and an $(l, -n)$-mode for H_2 in (3.30), we get:

$$I = \frac{M_0}{a} \frac{(-1)^n 2\pi k}{\sqrt{\eta}} N_{ln} \frac{1}{kr_p} \frac{d}{d(kr_p)} [kr_p h_l(kr_p)] \frac{n}{\sin\theta_2} P_l^n(\cos\theta_2) \quad \text{for TE mode} \tag{3.35}$$

$$= \frac{M_0}{a} \frac{(-1)^n 2\pi jk}{\sqrt{\eta}} N_{ln} h_l(kr_p) \frac{dP_l^n(\cos\theta_2)}{d\theta} \quad \text{for TM mode} \tag{3.36}$$

By equating (3.35) and (3.36) with (3.31) we get the solutions for the expansion coefficients. The expression was simplified with the assumption $r_p \approx R$ and with equations (3.16):

$$\begin{aligned} C_{ln}^{\mathrm{TE}} &= \frac{M_0}{a\sqrt{\eta}} \frac{2\pi j}{k} N_{ln} \frac{1}{R_l^\alpha(R)} \frac{nP_l^n(\cos\theta_2)}{\sin\theta_2} \\ C_{ln}^{\mathrm{TM}} &= \frac{M_0}{a\sqrt{\eta}} \frac{2\pi}{k} N_{ln} \frac{1}{R_l^\beta(R)} \frac{dP_l^n(\cos\theta_2)}{d\theta_2} \end{aligned} \tag{3.37}$$

3.3.3 Transmission Line Description

As in section 3.2.3 for type 1 elements, the modal voltages for the transmission line expansion are obtained directly from those for the waveguide expansion in (3.37) by using (2.39):

$$\begin{aligned} V_{ln}^{\mathrm{TE}}(r = R) &= \frac{-jM_0}{a} 2\pi R N_{ln} \frac{nP_l^n(\cos\theta_2)}{\sin\theta_2} \\ V_{ln}^{\mathrm{TM}}(r = R) &= \frac{-M_0}{a} 2\pi R N_{ln} \frac{dP_l^n(\cos\theta_2)}{d\theta_2} \end{aligned} \tag{3.38}$$

3.4 Field Reconstruction

With the spherical wave expansions obtained in the previous sections for both type 1 and type 2 radiators, it is now possible to reconstruct the field at any radius outside the sphere. Both the near-field and the far-field can be visualised.

3.4.1 Near Field Reconstruction

The solutions for the expansion coefficients are plotted in figure 3.5. It was explained in chapter 2 that these expansion coefficients incorporate the evanescent behaviour of higher order modes. Around the degree $l_c = kR$ and above, the modal waves become evanescent, which is shown by the dark tinted areas in the plots. The modes up to $l_c + 10$ contain almost all the radiated power.

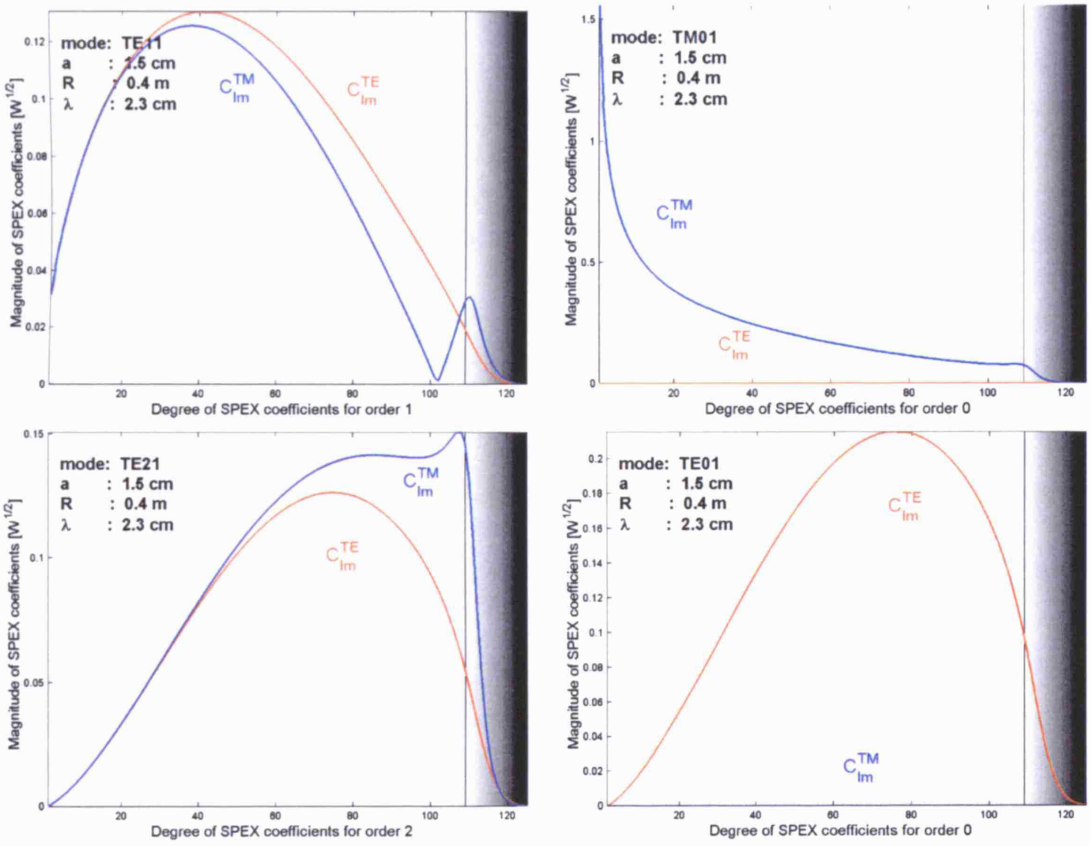


Figure 3.5: The coefficients of the spherical wave expansion are shown for the first 4 circular modes. The frequency was chosen to be 13 GHz, so that it exceeds the cut-off frequency in the circular waveguide for all these modes. The dark shaded region indicates cut-off characteristics of the spherical waveguide.

Also clear from the figures is that higher order circular modes in the aperture are less able to couple into the spherical modes. TM circular modes do not couple into TE spherical modes, whereas TE circular modes couple more strongly into TM spherical modes than they do in TE spherical modes.

By filling in the obtained coefficients into the linear combination in (2.33), the total field can be reconstructed at any distance from the sphere, including the far-field. However, as the obtained solutions for the modal fields exclude the non-propagating modes, an incomplete picture is obtained in the vicinity of the sphere. The reconstruction of the field on the surface of the sphere is the far-field information, back-projected to the source.

Field reconstruction through spherical modes is illustrated in figure 3.6. It is shown that, as with any Fourier transform, when more modes are used, the excitation function can be represented with higher resolution. Small apertures in a sphere cause sharp discontinuities in the electric field, and a large amount of spherical modes is needed to build up such an excitation. As can be seen in the plots of figure 3.5, the spatial bandwidth of one aperture in the domain of spherical modes is higher than the cut-off region for the corresponding modes. Therefore, even when all modes up to the cut-off region are used to reconstruct the field, the sharp transition in electric field strength at the boundary of the aperture can not exactly be reconstructed by these propagating modes. This effect shows in the bottom plots of figure 3.6 by a light ripple in the tangential electric field strength around the aperture, where in reality it should be zero at that radius. To obtain a more accurate near-field reconstruction, the STL expansion from section 3.2.3 should be used, as illustrated in section 3.4.3.

3.4.2 Far-Field Reconstruction

In section 2.4, we obtained the far-field of an electromagnetic wave that is expanded in spherical modes. The far-field was formulated so that, when the field is generated by an antenna, its magnitude corresponds with the directivity function of the antenna (see equation (2.54)). As we obtained the expansion coefficients of an excited aperture in a conducting sphere, the corresponding far-field directly follows from (2.49). In figure 3.7 the directivity is plotted for two different circular modes in the aperture. The directivity was normalised relative to its maximum. The pattern is shown as the magnitude of its real part, so that a snapshot in time is taken of the far-field. This is necessary because we are dealing with circularly polarised modes in the aperture and the directivity is plotted regardless of any polarisation definition. The plots therefore are not constant in time, but will oscillate between the plotted curves. A better way to deal with polarisation was developed in section 2.4.3 and is used in figure 3.8, where the directivity plots of figure 3.7 for the TE_{11} and TE_{21} modes are repeated with the directivity separated into the co-polar (RHCP) and cross-polar (LHCP) components.

The obtained co-polar and cross-polar patterns agree with what could be expected from circularly polarised fields in the aperture. Indeed, assuming a right hand circular polarisation in the aperture, we have $n > 0$ for all circular modes, and hence also the spherical modes are right

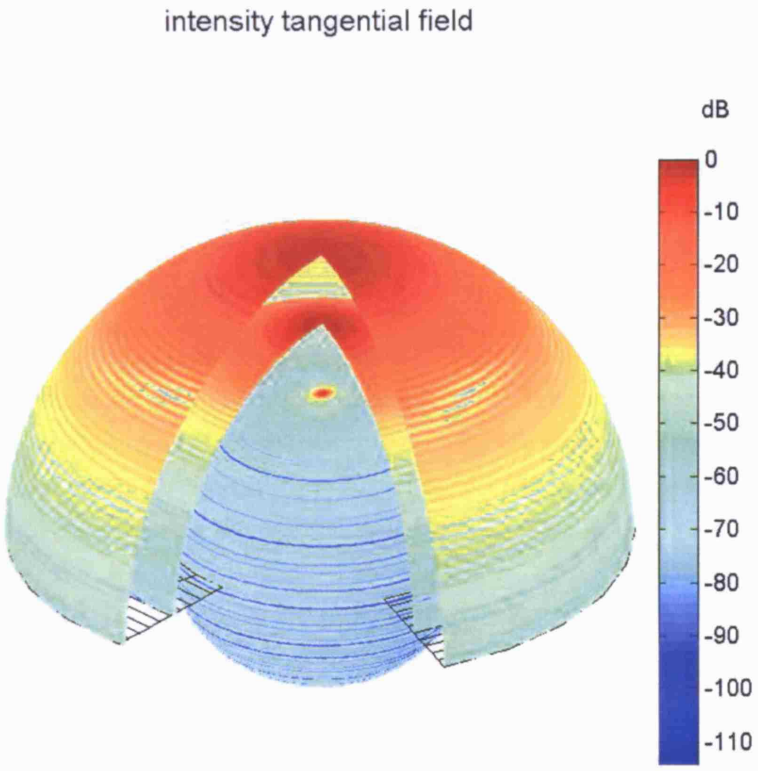


Figure 3.6: The figure shows a reconstruction of the field of an aperture of 3cm in a conducting sphere of 80cm diameter, excited by a TE_{11} circular mode. 125 spherical modes were needed to represent the field accurately.

hand circularly polarised: $m > 0$. For $m > 0$, the following recurrence relations are valid:

$$\begin{aligned}\frac{dP_l^m(\cos\theta)}{d\theta} &= \frac{1}{2}\left(P_l^{m+1}(\cos\theta) - (l-m+1)(l+m)P_l^{m-1}(\cos\theta)\right)\\ \frac{mP_l^m(\cos\theta)}{\sin\theta} &= \frac{1}{2}\cos\theta\left((l-m+1)(l+m)P_l^{m-1}(\cos\theta) + P_l^{m+1}(\cos\theta)\right) + m\sin\theta P_l^m(\cos\theta)\end{aligned}$$

We see that for small angles θ we have:

$$\frac{dP_l^m(\cos\theta)}{d\theta} \approx \frac{mP_l^m(\cos\theta)}{\sin\theta} \qquad \text{for} \quad \theta \to 0$$

so that R_l^m reaches a maximum and L_l^m tends to zero at boresight of the right hand circularly polarised aperture. The opposite is true at the south pole, as the two terms above take opposite signs, so that L_l^m reaches a maximum and R_l^m tends to zero.

In equation (2.57) the axial ratio was defined as a measure of how well the field can be represented with one polarisation. In figure 3.9, the axial ratio is plotted as a function of polar angle θ for the TE_{11} mode with the same parameters as before.

3.4.3 Reactive Near Field Reconstruction

For the reactive near-field, a separate expansion was developed in section 3.2.3. The expansion coefficients there obtained are illustrated in figure 3.10. The reconstructed fields for the waveguide expansion and the transmission line expansion are illustrated in figure 3.11. The results shown are for an aperture in a conducting sphere. For a type 2 radiator the same reconstructions can be obtained by simply using the expansion coefficients obtained in equation (3.38).

3.5 Numerical Validation and Refinement

3.5.1 Refinement of the Expansions

The analytic procedures followed in the previous sections of this chapter were in fact first order approximations to the real world scenario. This was clear for the type 2 radiators, but also in the derivations for the type 1 radiators inexact assumptions were taken. In particular, it was assumed that the field in the aperture was determined by the modes that propagate in the feeding waveguide. Subsequently, this field was linearly projected onto the spherical modes that can exist in the spherical waveguide formed by the space outside the spherical body. In reality however, a more complicated, partially reactive field will exist at the boundary, with circular modes that are beyond cut-off in the circular waveguide and therefore only present near the boundary but not deeper in the feeding waveguide. This unknown reactive component of the aperture field won't affect the far-field of a single element, as it can not contribute any energy towards it. However, it may excite reactive spherical modes that exist in the reactive near field surrounding the sphere. This becomes important in an array configuration where multiple elements can couple into the same reactive field. In order to use the obtained expansions for the analysis of mutual coupling effects in chapter 4, the expansions need to be refined to accurately represent the reactive near-field. This can be achieved with the Least Squares Boundary Residual

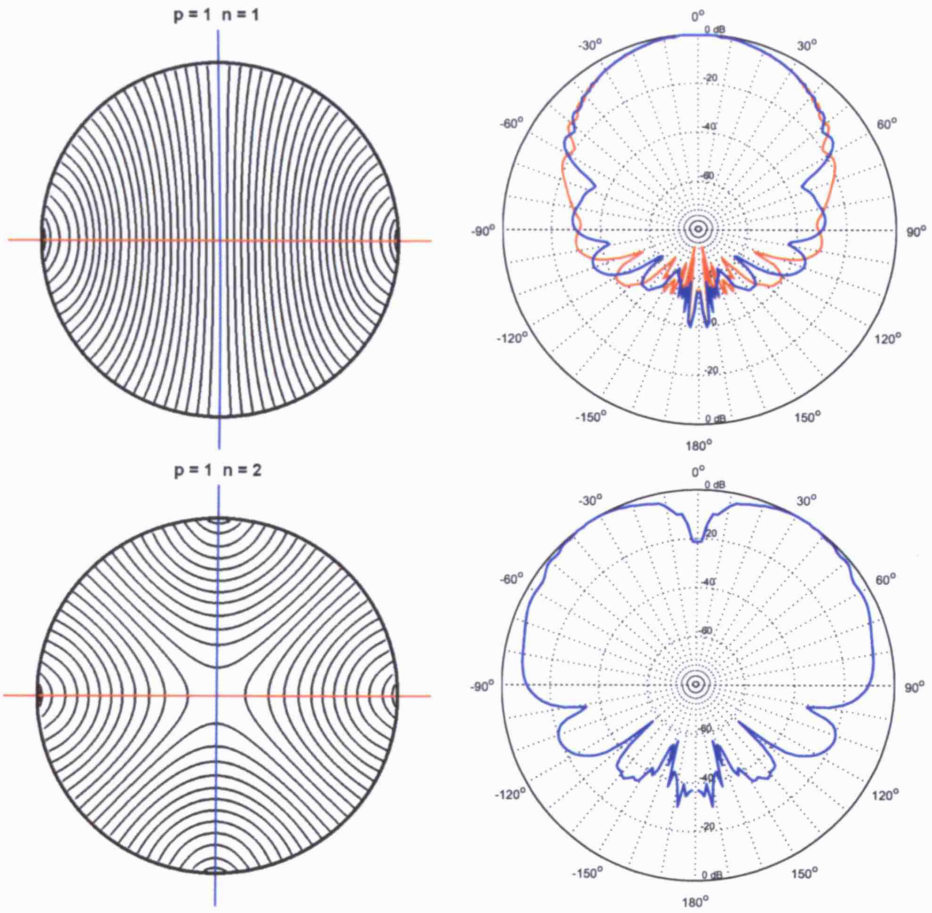


Figure 3.7: The plots on the right show the normalised directivity of two different circular modes in an aperture of 3cm diameter in a sphere of 80cm diameter for a wavelength of 2.3cm. The far-field is intersected in two perpendicular planes as shown on the left. The symmetry of the TE_{21}-mode makes the two intersections overlap. In the directivity plots, 0° is at boresight of the apertures, or the direction of the z-axis.

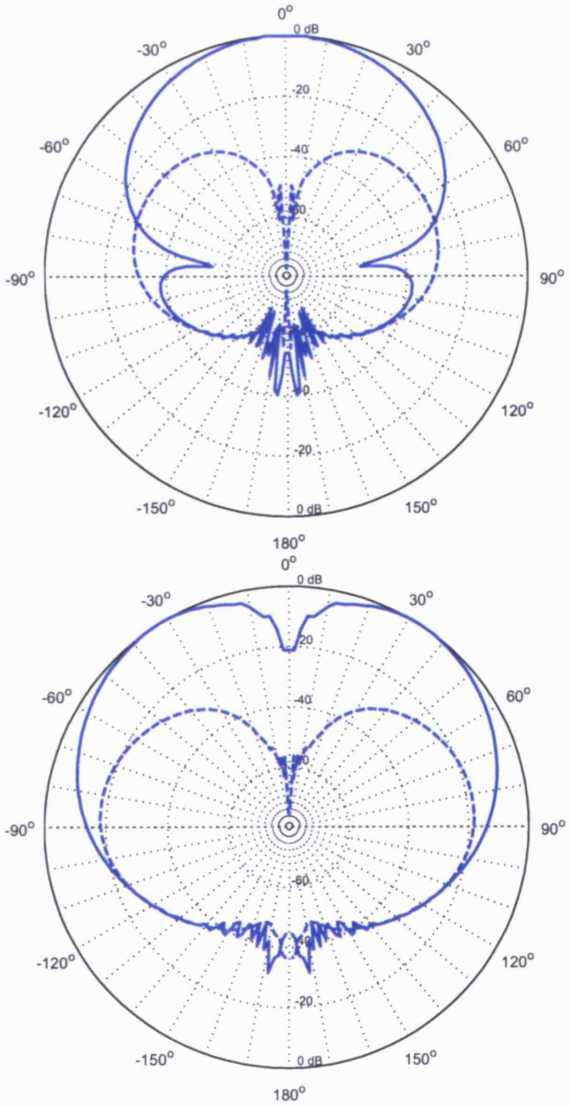


Figure 3.8: The same far-fields are illustrated as in figure 3.7, but now the co-polar and cross-polar components of the far-fields are separated. The solid lines show the RHCP component, whereas the dashed lines represent the LHCP component. The directivity of the co-polar and cross-polar components is now independent of azimuthal angle ϕ and the directivity plots in any other plane rotated around the z-axis coincide with the ones shown. The periodical variation of the field that exists in the aperture is incorporated in the phase of the far-field, but by definition the directivity only shows the magnitude of the far-field.

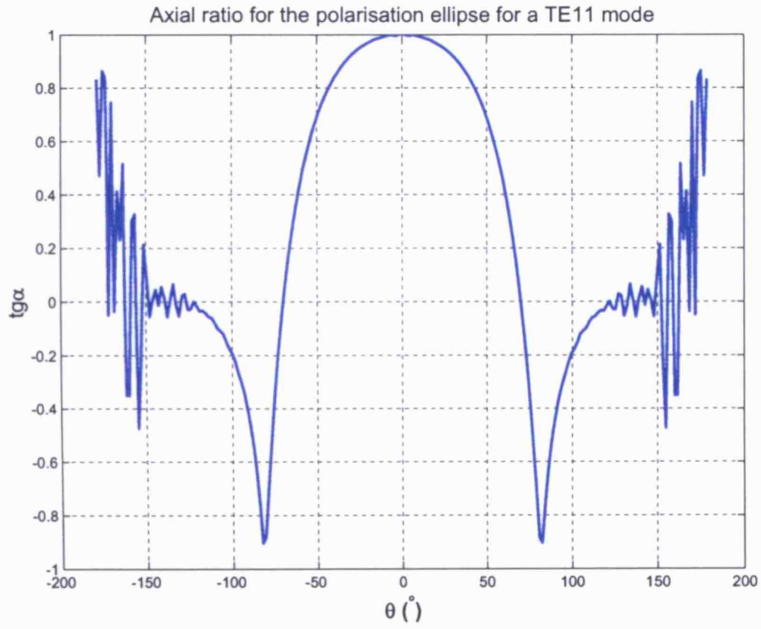


Figure 3.9: The axial ratio of the polarisation ellipse is plotted in the vertical plane for the same TE_{11} mode as before. From the figure it can be seen that right hand polarisation is well maintained up to $40°$ off boresight, but falls rapidly to left hand polarisation for wider angles. At broadside, the polarisation is not linear, but left hand. There is linear polarisation towards the south pole.

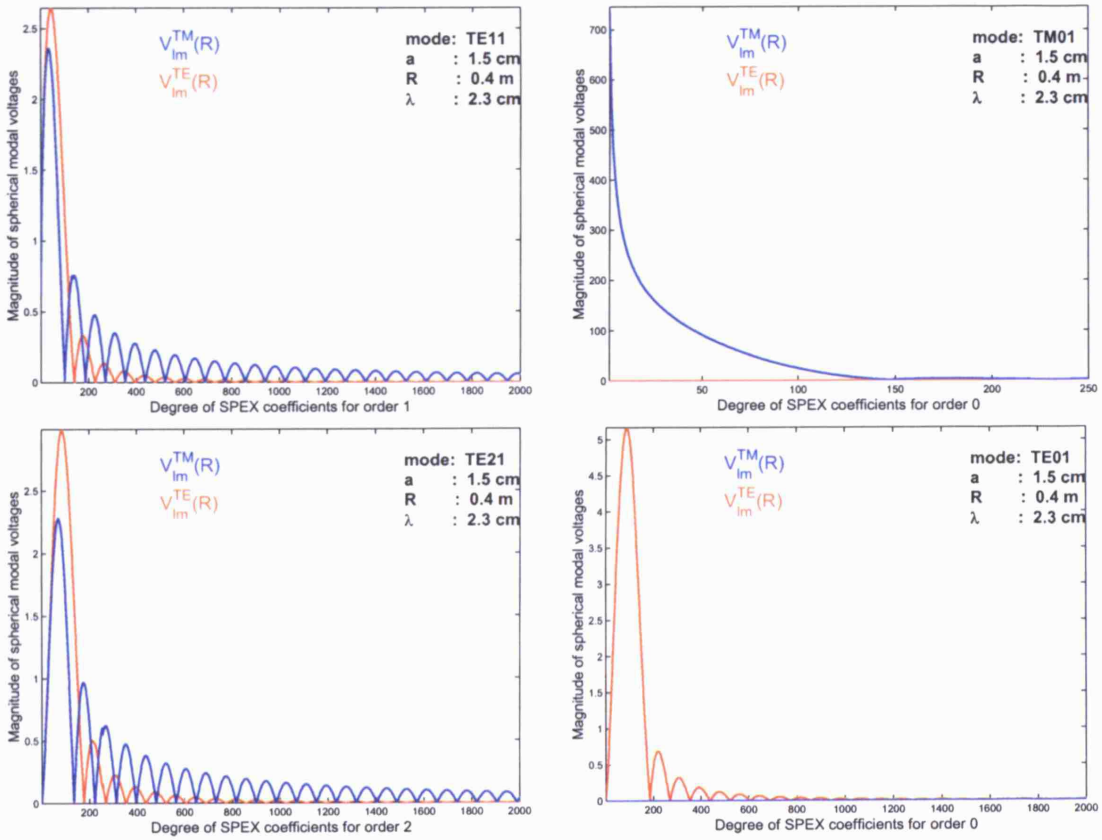


Figure 3.10: Expansion coefficients of a circular aperture for the Transmission Line Expansion.

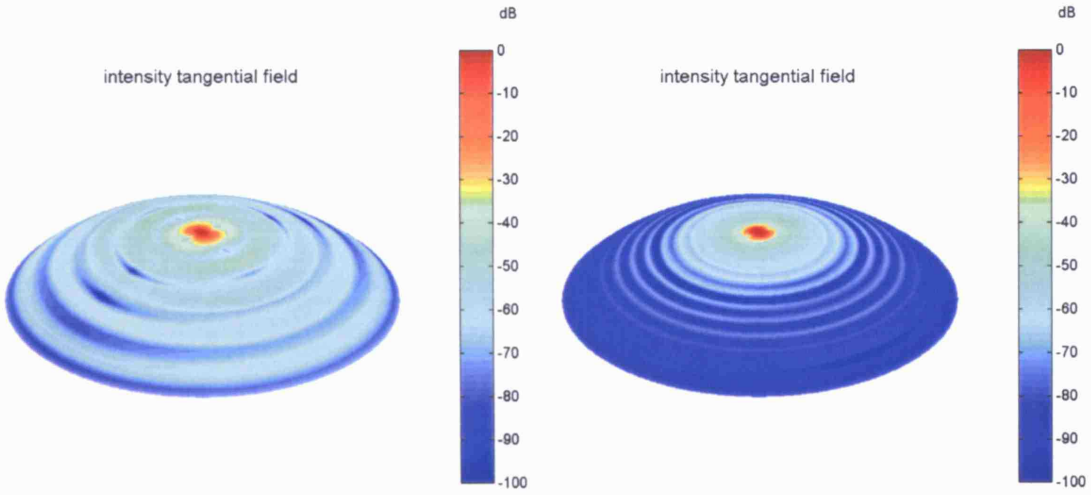


Figure 3.11: The figures show the reconstructed near-field on a spherical cap around an aperture, excited by a TE_{11} mode. On the left the "spherical waveguide expansion" was used, truncated above cut-off at degree 125. For the reconstruction on the right, the "spherical transmission line expansion" was used and modes up to degree 5000 were included.

(LSBR) procedure[85][86][87]. We will establish the Numerical Refinement Procedure (NRP) as a mixture of the LSBR with an initialisation procedure.

For type 2 radiators, such as the SAR that was discussed in section 3.3, the same method could be applied by choosing the aperture to be the vertical wall that surrounds the cavity formed by the patch and the sphere. This is a true physical aperture and forms the boundary between the spherical waveguide and a cylindrical cavity. It is uncertain whether this could improve the accuracy of the obtained expansion coefficients, as the cavity model is an approximation in itself. It is expected that the method would converge to the full-wave solution of the patch above the sphere. So far, refinement of the modal expansion of a patch type element has not been investigated any further.

A Initialisation procedure

The LSBR procedure boils down to a minimisation procedure on the difference between the aperture field from the point of view of the waveguide at one side of the aperture to the same aperture field when constructed in the other waveguide. Assuming nothing is known about the stability and condition of this minimisation procedure, it makes sense to start searching for the optimum from a set of variables that is close to the solution. In this section the numerical problem is defined in terms of variables and a convergence criterion. An initialisation procedure is then started, that should get the variables near the solution, but without any guarantees on convergence. The next section will pick up the values for the variables that were found in the initialisation procedure and perform a full-blown LSBR procedure on them. We have found this combination of procedures yields a rapid solution, backed-up by the thoroughness of the LSBR procedure.

In figure 3.12 the problem to be solved is shown schematically. The unknown is the true field in the aperture, including any evanescent waves that may extend on both sides of the aperture.

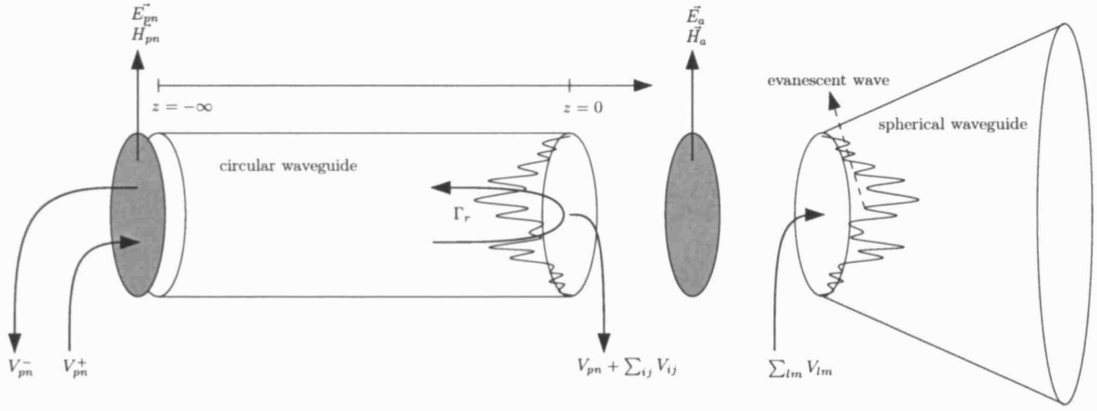


Figure 3.12: The numerical refinement procedure strives to solve the problem that is depicted by this model. The field in the aperture is partially unknown as there are unknown evanescent waves superimposed on the known propagating modes. The conical waveguide shown on the right schematically represents the spherical waveguide, consisting of the free space outside the conducting sphere.

In the aperture however, there must be a continuity of the field, so that the same aperture field can be written as a linear combination of circular modal voltages $V_{ij}(z)$ as well as a linear combination of spherical modal voltages $V_{lm}(r)$. Without lack of generalisation, we assume one propagating mode in the circular waveguide, with modal voltage $V_{pn}(z)$. The unknown electric and magnetic fields in the aperture are written as $\vec{E}_a$ and $\vec{H}_a$ respectively. Note that outside the aperture, on the rest of the surface of the sphere, the tangential electric field is zero, but not the tangential magnetic field. With $\vec{E}_a$ and $\vec{H}_a$ we will refer to those parts of the electric and magnetic fields that exist on a spherical surface of radius R and that is truncated to the solid angle of the aperture. The reflection coefficient in the aperture is Γ_r, so that in the circular waveguide the modal voltage is:

$$V_{pn}(z) = V_{pn}^+ e^{-j\beta z} + V_{pn}^- e^{j\beta z} \tag{3.39}$$

$$= V_{pn}^+ (e^{-j\beta z} + \Gamma_r e^{j\beta z}) \tag{3.40}$$

In the aperture, this voltage becomes:

$$V_{pn} = V_{pn}^+ + V_{pn}^- = V_{pn}^+(1 + \Gamma_r) \tag{3.41}$$

With the characteristic impedance of the mode equal to Y_{pn}^c, the current associated with this mode becomes:

$$I_{pn} = I_{pn}^+ + I_{pn}^- \tag{3.42}$$

$$= Y_{pn}^c(V_{pn}^+ - V_{pn}^-) \tag{3.43}$$

$$= Y_{pn}^c \left(\frac{1 - \Gamma_r}{1 + \Gamma_r}\right) V_{pn} \tag{3.44}$$

The power carried down the pipe is:

$$P_+ = V_{pn}^+ (I_{pn}^+)^* = |V_{pn}^+|^2 Y_{pn}^c = |V_{pn}|^2 Y_{pn}^c \left| \frac{1}{1+\Gamma_r} \right|^2 \tag{3.45}$$

Where we have taken into account that Y_{pn}^c is real for the propagating pn-mode. The reflected power is:

$$P_- = V_{pn}^- (I_{pn}^-)^* = -|V_{pn}^-|^2 Y_{pn}^c = -|V_{pn}|^2 Y_{pn}^c \left| \frac{\Gamma_r}{1+\Gamma_r} \right|^2 \tag{3.46}$$

The net instantaneous power flow through the aperture is:

$$P_{pn} = V_{pn} I_{pn}^* = |V_{pn}|^2 Y_{pn}^c \left(\frac{1-\Gamma_r}{1+\Gamma_r} \right)^* \tag{3.47}$$

The time-average power flow through the aperture is given by

$$P_{av} = \Re(P_{pn}) = P_+ + P_- \tag{3.48}$$

The evanescent waves in the aperture exist as non-propagating -z-directed modes and there is no reflection to be taken into account for these modes:

$$V_{ij}^+ = 0 \tag{3.49}$$

$$I_{ij}^+ = 0 \tag{3.50}$$

$$I_{ij} = -Y_{ij}^c V_{ij} \tag{3.51}$$

$$P_{ij} = V_{ij} I_{ij}^* = -|V_{ij}|^2 \left(Y_{ij}^c\right)^* \tag{3.52}$$

Figure 3.13 illustrates the refinement procedure to obtain the unknown reactive voltages V_{ij}, and with them a more accurate set of modal voltages V_{lm} at the input of the spherical waveguide. Starting by setting $V_{ij} = 0$, the V_{lm} were obtained in equations (3.21) and (3.22). Using these modal voltages at the input of the spherical waveguide, the input admittances (as obtained from the input impedances given in equations (2.44)) are calculated and used to obtain the input modal currents. These in turn specify the magnetic field over the surface of the sphere, including the aperture, via the STL expansion formula's in equation (2.38). This yields a numerical evaluation for $\vec{H}_a$ in a large number of sample points of the aperture.

Going round in the other direction (anti - clockwise) in figure 3.13, the magnetic field can be evaluated from the perspective of the circular waveguide. Using equations (3.44) and (3.51):

$$\vec{H}_a = Y_{pn}^c \frac{1-\Gamma_r}{1+\Gamma_r} V_{pn} \vec{h_{pn}} - \sum_{ij} Y_{ij}^c V_{ij} \vec{h_{ij}} \tag{3.53}$$

Provided that we know Γ_r, the only unknowns in equation (3.53) are V_{ij}, the modal voltages of the evanescent waves in the circular waveguide. Defining the difference field as:

$$\vec{D} = Y_{pn}^c \frac{1-\Gamma_r}{1+\Gamma_r} V_{pn} \vec{h_{pn}} - \vec{H}_a \tag{3.54}$$

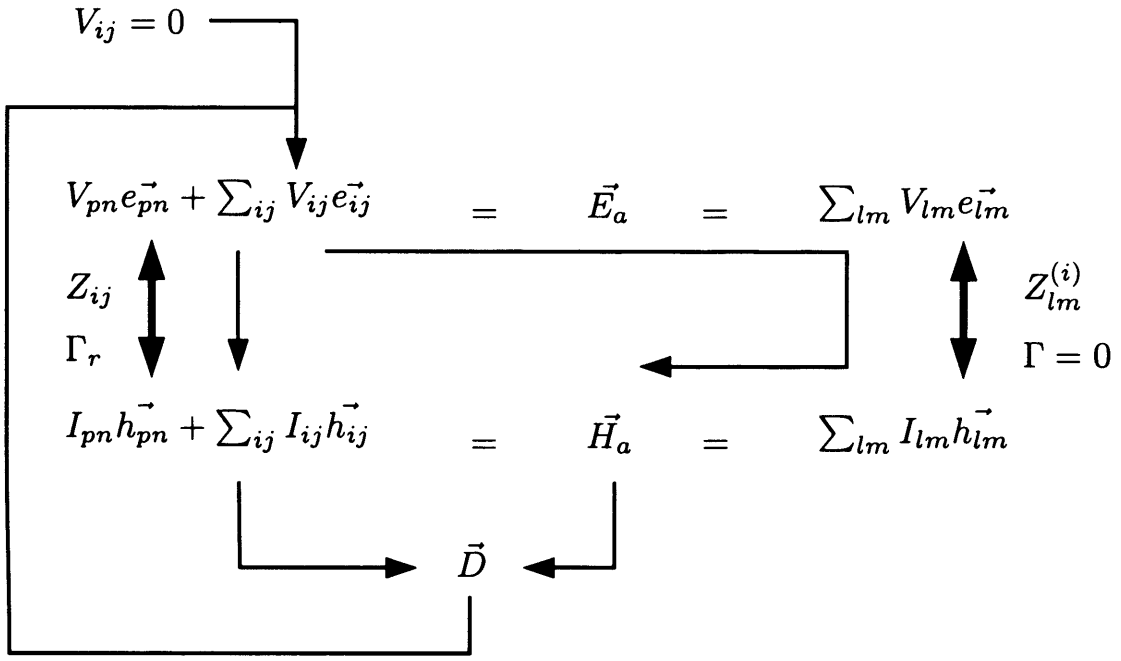


Figure 3.13: Schematic of the initialisation procedure that performs a circular refinement on the field in the aperture. The thin arrows indicate the sequence of calculations in each iteration.

we can project the difference field on each ij-mode with:

$$V_{ij} = \frac{1}{Y_{ij}^c}\left(\vec{D} \cdot \vec{h_{ij}}^*\right) \tag{3.55}$$

Subsequently, we can update $\vec{E_a}$ with:

$$\vec{E_a} = V_{pn}\vec{e_{pn}} + \sum_{ij} V_{ij}\vec{e_{ij}} \tag{3.56}$$

which can then be numerically projected onto the spherical modes to obtain the V_{lm} for another iteration.

The one missing link to complete the previous calculations is the reflection coefficient Γ_r. This can be obtained by requiring that at each iteration, no power can leak away in the aperture. The reactive field that exists in the aperture gives it a certain admittance, and the mismatch with the modal admittance of the propagating mode causes reflection back in the feeding waveguide. This power is not taken on by the spherical modes that exist outside the sphere. The conservation of power is formally described by the equation:

$$|V_{pn}|^2 Ypn^c \left(\frac{1-\Gamma_r}{1+\Gamma_r}\right)^* - \sum_{i,j\neq p,n} |V_{ij}|^2 \left(Y_{ij}^c\right)^* = \sum_{lm} |V_{lm}|^2 \left(Y_{lm}^i\right)^* \tag{3.57}$$

Where the Y_{lm}^i are the input admittances of the spherical modes. The first term of the left side of the equation must have a positive real part for any power to be radiated. The second term is entirely imaginary, and as the V_{ij} get updated, only the reactive part of the right side of the

equation changes. This won't affect the radiated far-field, but it will affect the reactive near-field. The relative aperture admittance, as seen from within the feeding waveguide becomes:

$$y = \frac{1-\Gamma_r}{1+\Gamma_r} = \frac{\sum_{lm}|V_{lm}|^2 Y^i_{lm} + \sum_{i,j\neq p,n}|V_{ij}|^2 Y^c_{ij}}{|V_{pn}|^2 Ypn^c} \tag{3.58}$$

The denominator usually equals 1, as V_{pn} is normalised in the set-up phase to deliver $1W$ of power in the waveguide.

The results of three iterations of the described refinement procedure are shown in figure 3.14. The method does not seem to converge to a solution, and the reactive components keep growing from one iteration to the next. The reactive field becomes stronger in each step, and while the delivered power is kept constant by the algorithm, the power that is launched in the waveguide, P_+ and the reflection from the aperture keep increasing as well. Possible reasons for this non-convergence may be:

- The spherical harmonic expansion may be cut off at a too low degree, as suggested by figure 3.15. This injects a high spatial frequency error in each step of the iteration. The injected errors may cause aliasing in the lower spatial resolutions that are resolved by the circular modes in the aperture. As these errors cannot be corrected for in subsequent runs of the iteration, they could grow without bound and cause the algorithm to diverge from a solution.

- The numerical projection in the aperture is evaluated on a sampling grid. Because of the high density of the sampling grid, and the relative small size of the aperture compared to the whole sphere, an equiangular sampling grid was adopted. In chapter 8 better ways of interpolation are suggested, and the chosen numerical projection method may introduce extra errors.

- For numerical reasons, a small region around the pole is left out from the calculations. This may also adversely affect the results.

Efforts to stabilise the algorithm by minimising the aforementioned errors have not been successful. A closer examination of the numerical stability of the whole method seems appropriate. However, rather than investigating the lack of convergence of the proposed algorithm, it was decided to bolt on a second stage of refinement, using the LSBR method that is generally adopted for this kind of problem. A single iteration of the procedure described above is used as an initialisation procedure for the LSBR.

B Mode matching with LSBR

The Least Squares Boundary Residual is a rigorous and widely used method to solve waveguide discontinuities[85][86][87]. The idea behind the method is fairly straightforward. The implementation however, may require the implementation of a powerful search algorithm. The initialisation procedure that was described in the previous section allows us to implement the LSBR method with a basic numerical minimisation algorithm described in this section. The

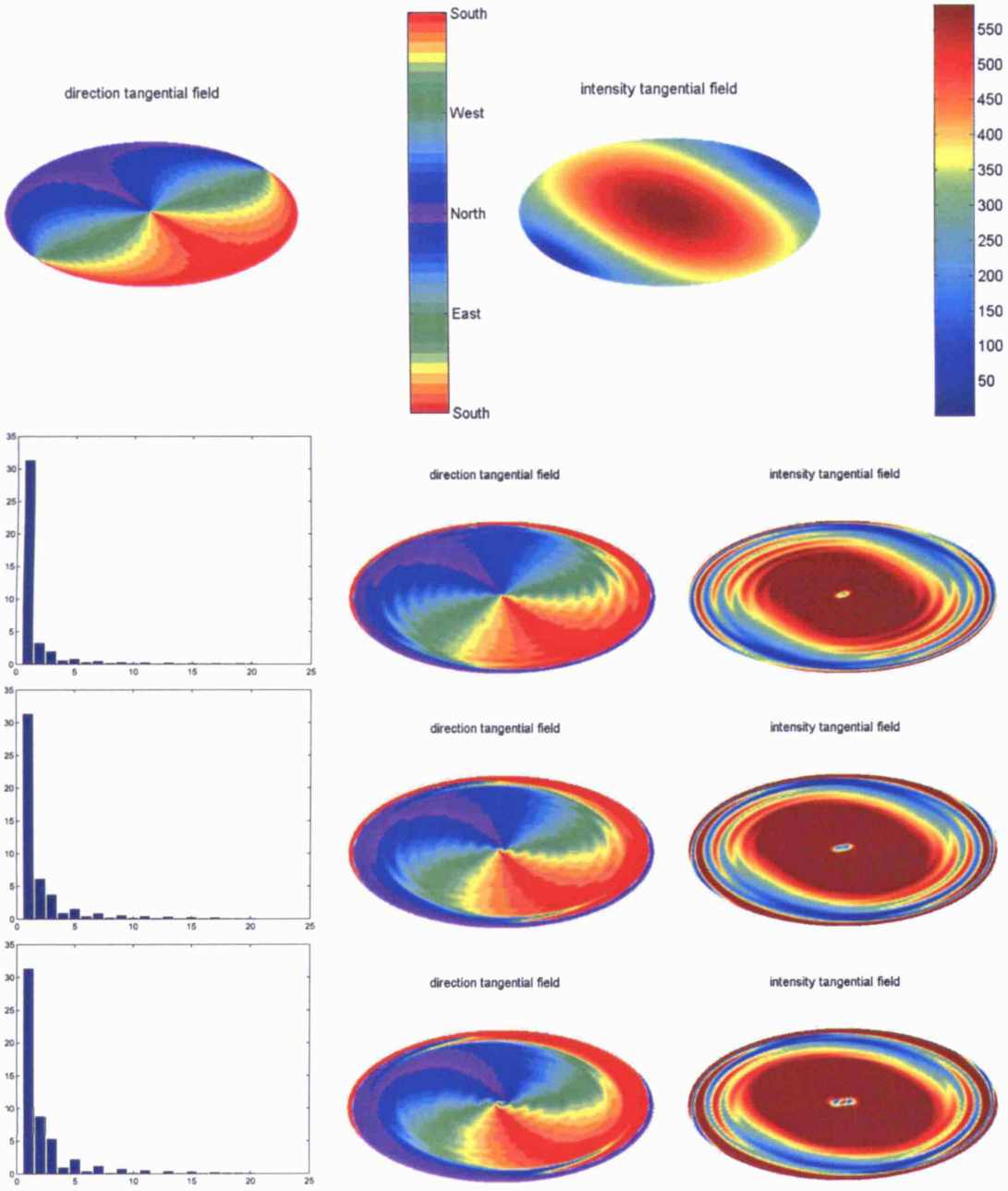


Figure 3.14: The tangential electric field in the aperture is shown after each of three iterations of the initialisation procedure. The modal coefficients keep growing with each refinement, and the algorithm is stopped without convergence. The process is started with the field of the fundamental mode only, which is illustrated in the top row of plots. The plot that show the direction of the tangential field have to be interpreted with the colour codes for wind directions on the globe. Note that the aperture is a spherical cap at North Pole position, rather than a flat circular disk.

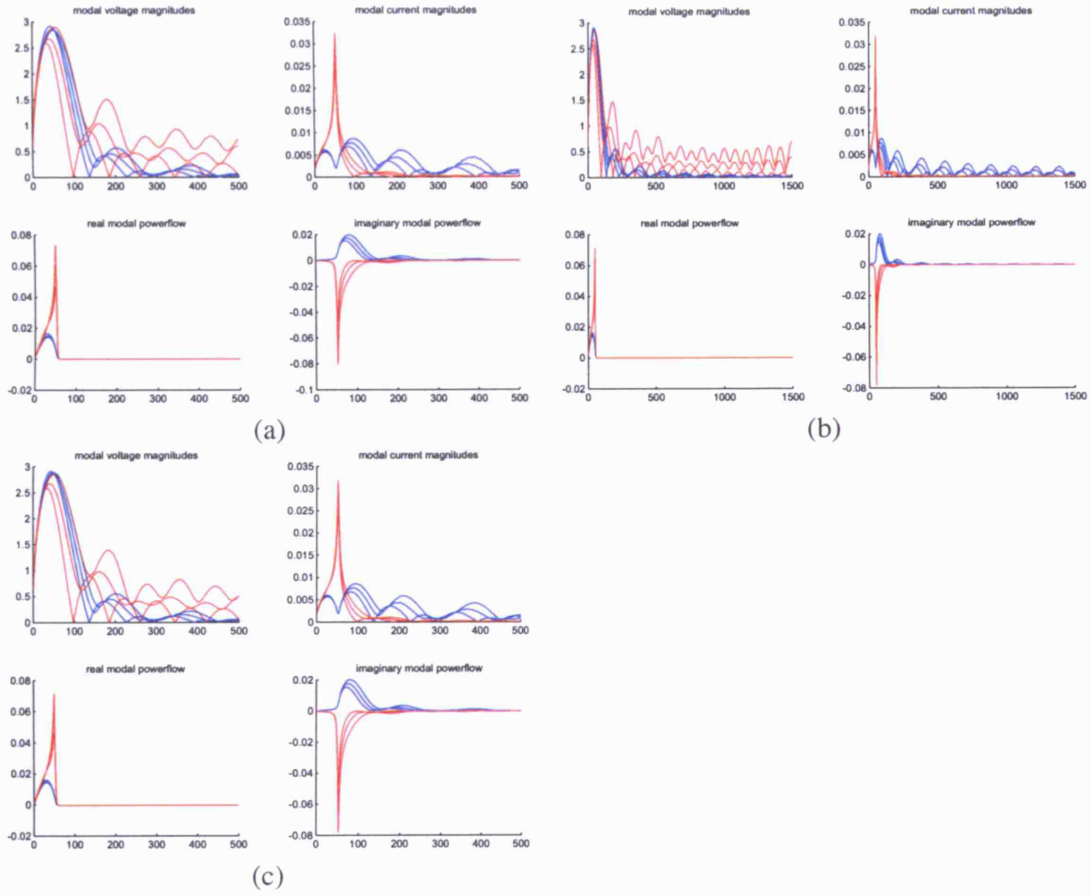


Figure 3.15: The obtained spherical modal voltages after each of 3 iterations of the initialisation procedure. The reactive components have increased substantially but the propagating modes remain roughly unaltered as it should be. The four plots in (a) show that the spherical modes have been cut off at a too low degree, which injects an error in each step of the iteration. The four plots in (b) show that increasing the maximum degree of the spherical harmonic expansion does not remove the errors. The four plots in (c) were obtained with a denser sampling grid in the aperture. Neither of the two changes seem to improve the convergence of the algorithm.

advantage of this approach is that a commonly accepted numerical procedure is used, and no further investigation is required into the stability of the initialisation procedure as described in section A. A drawback of using this method is the computation time required by the optimisation process.

In essence, the LSBR uses the same approach as we used in the initialisation procedure, and two estimates for the aperture field are calculated based on the modal voltages of the circular and the spherical waveguide respectively. Rather than circularly updating the modal voltages as before, a target value is calculated based on the distance between the two estimates. Now all modal voltages, in both waveguides are regarded as variables in each iteration, and the target is minimised by a search algorithm. Let $\vec{E}_a^{\mathrm{L}}$ be the left estimate of the electric field in the aperture, $\vec{E}_a^{\mathrm{R}}$ the right estimate, $\vec{H}_a^{\mathrm{L}}$ the left estimate for the magnetic field, and $\vec{H}_a^{\mathrm{R}}$ the right estimate. The target distance to be minimised is then defined as:

$$M = \int\int_{ap} \left(|\vec{E}_a^{\mathrm{R}} - \vec{E}_a^{\mathrm{L}}|^2 + \eta_e |\vec{H}_a^{\mathrm{R}} - \vec{H}_a^{\mathrm{L}}|^2 \right) ds \tag{3.59}$$

In figure 3.16 the LSBR procedure is outlined. Although in itself the method is rather simple, it proved to be a challenge to implement a speedy algorithm for large variable spaces. This was achieved by combining a simple multivariate steepest descent search algorithm with a one-dimensional line search. In each iteration, after calculation of the local gradient, a line search is started to find the optimum step size for each variable. This line search was implemented with a Golden Section technique. The Golden Section technique is an interval-reducing optimisation method for non-differentiable functions that guarantees to reach a preset accuracy, while the number of function evaluations in each step of the iteration is just one. With the complex computations at hand, the latter offers a speed advantage over other interval-reducing methods. The bulk of the computation time of the LSBR procedure goes to the calculation of the gradients. All modal fields are stored in memory. The LSBR procedure may require several days of computing time and all available memory on today's typical desktop PCs. The memory requirement to store all the modal fields (the alternative would be to re-compute them in each iteration, which would result in a prohibitively high computing time) is given by:

$$O = Nt \times Np \times (Lc + Ls) \times 128 \quad \text{bytes} \tag{3.60}$$

Where $Nt \times Np$ is the number of sample points in the aperture and Lc and Ls are the lengths of the circular and spherical modal expansions. For $Nt = 75$, $Np = 50$, $Lc = 20$ and $Ls = 1000$, the memory load is roughly 500 MB. A workaround for this high load would be to use a sparser grid for the lower degree modes and a high density grid for the high degree modes, which could further enhance the performance of the algorithm. Figure 3.16 illustrates the LSBR refinement procedure, and figures 3.17 to 3.20 illustrate the results obtained with it.

3.5.2 Validation of the Expansions

In order to validate the previously obtained expansion we used a commercial electromagnetic solver to obtain the far-field from a circular aperture in a conducting sphere and compared it

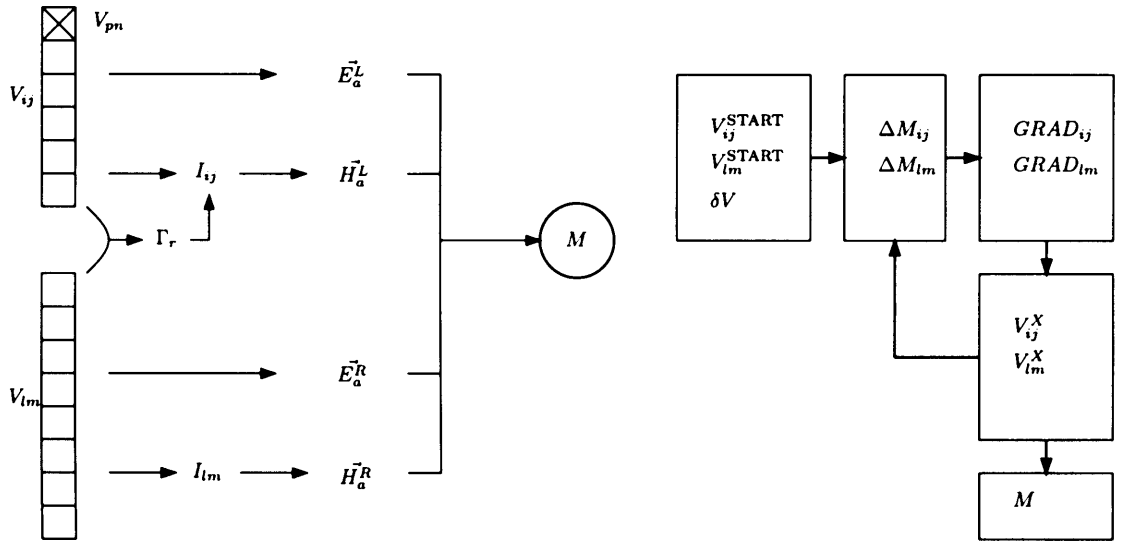


Figure 3.16: The Least Squares Boundary Residual assigns a distance M between the left and right estimates of an aperture field. The variables are the modal voltages from both sides of the aperture, except the voltages of the propagating modes in the feeding waveguide, which are kept fixed. The search algorithm starts by applying small perturbations to the voltages and numerically calculating the gradient of the target distance in the space of voltages. A steepest descent search algorithm searches for the minimum distance M. To speed up the algorithm, we added a 1-dimensional line search to each step according to the Golden Section Technique.

with the analytically obtained results obtained above. The software used is CST, one of the well established producers of EM simulation software. The model for the simulation is shown in figure 3.21. A solid perfectly electrically conducting sphere was modelled, with a circular waveguide cut out from it. The waveguide port is then excited with the fundamental mode and the far-field is obtained automatically. For this large structure (the model was chosen to reflect real world dimensions, with a sphere of 80 cm diameter and excited with a frequency of 3 GHz) it pays to simulate with a Finite Differential Time Domain (FDTD) solver, because with this kind of solvers the simulation time increases more slowly with increasing mesh size than any other EM solver.

Figure 3.22 shows an impressive agreement between the results obtained with the analytical methods that were described in this chapter and the pure numerical simulation described above. From this agreement it is safe to assume that the derived formula's for the expansion coefficients are indeed correct. The numerical refinement procedure was not used here, as it is irrelevant for far-field reconstruction.

3.6 Practical Element Selection

The previous sections of this chapter were devoted to the modelling of a radiating element on a conducting sphere. Most attention was given to the modelling of so-called type 1 elements that are essentially waveguide-fed apertures through the spherical shell. It is very likely that the radiation pattern of these elements have been modelled with great accuracy, as they were in perfect agreement with simulations from an EM solver. The models were extended to cover

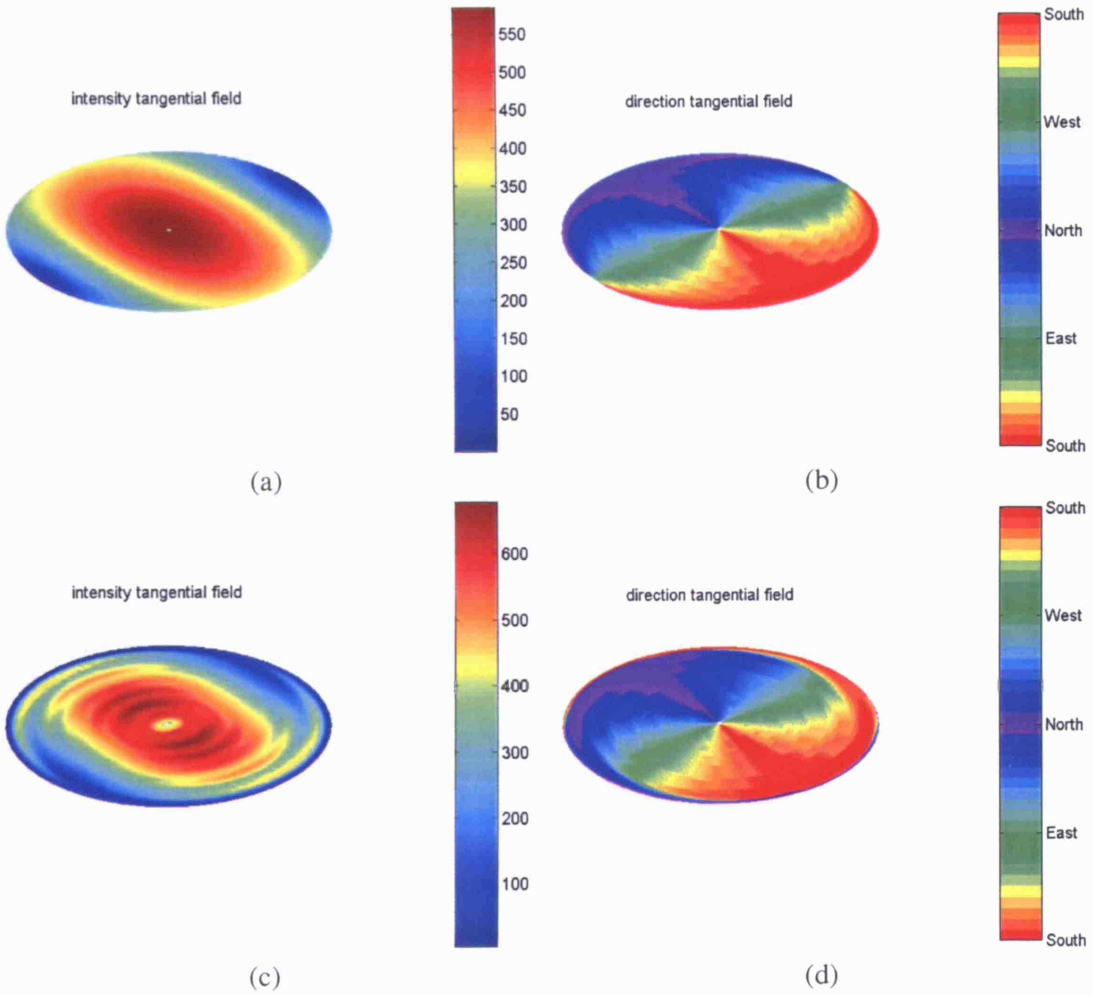


Figure 3.17: The electric field in the aperture is shown before the numerical LSBR refinement (a) and (b) and after completion of the two refinement stages (c) and (d).

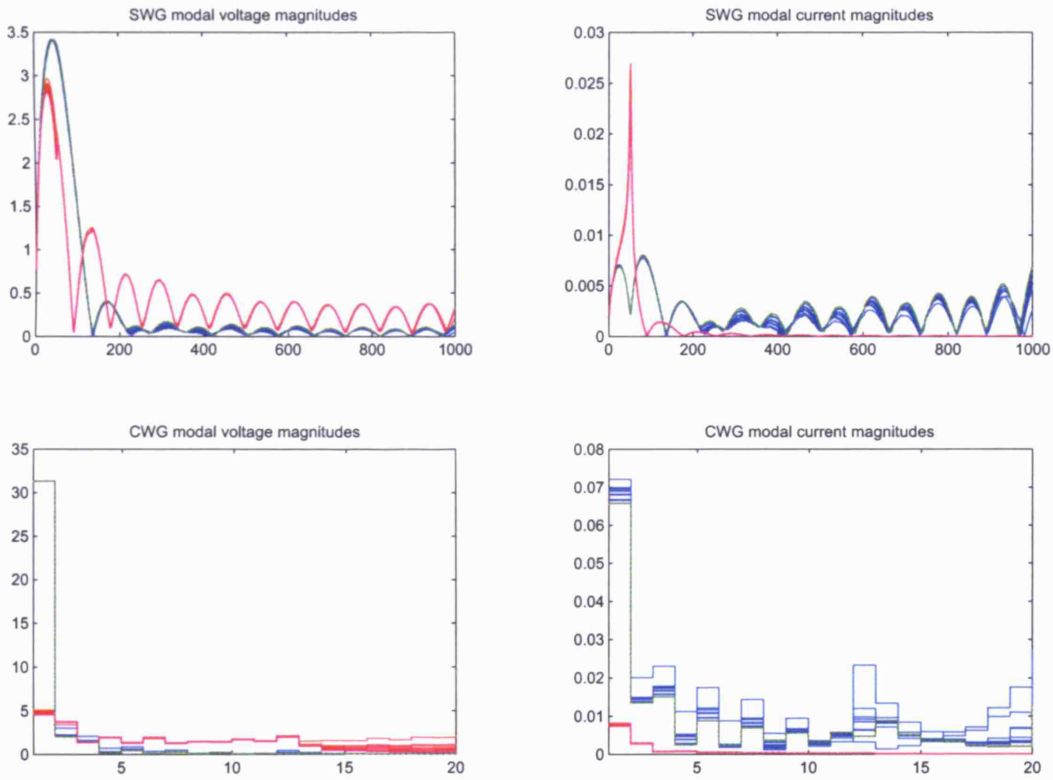


Figure 3.18: The plots show the coefficients for both the spherical (top row) and the circular (bottom row) expansions of the aperture field. The coefficients are plotted for each iteration of the LSBR procedure. They are shown in solid red and blue for the TM and TE modes respectively, and in dashed magenta and green for the final step of the iteration. For this particular refinement procedure, ideally more modes should have been taken into account. Memory restrictions however limit the number of modes that can be included when running the calculations on a regular desktop machine. It can be seen that the LSBR method does converge to a solution.

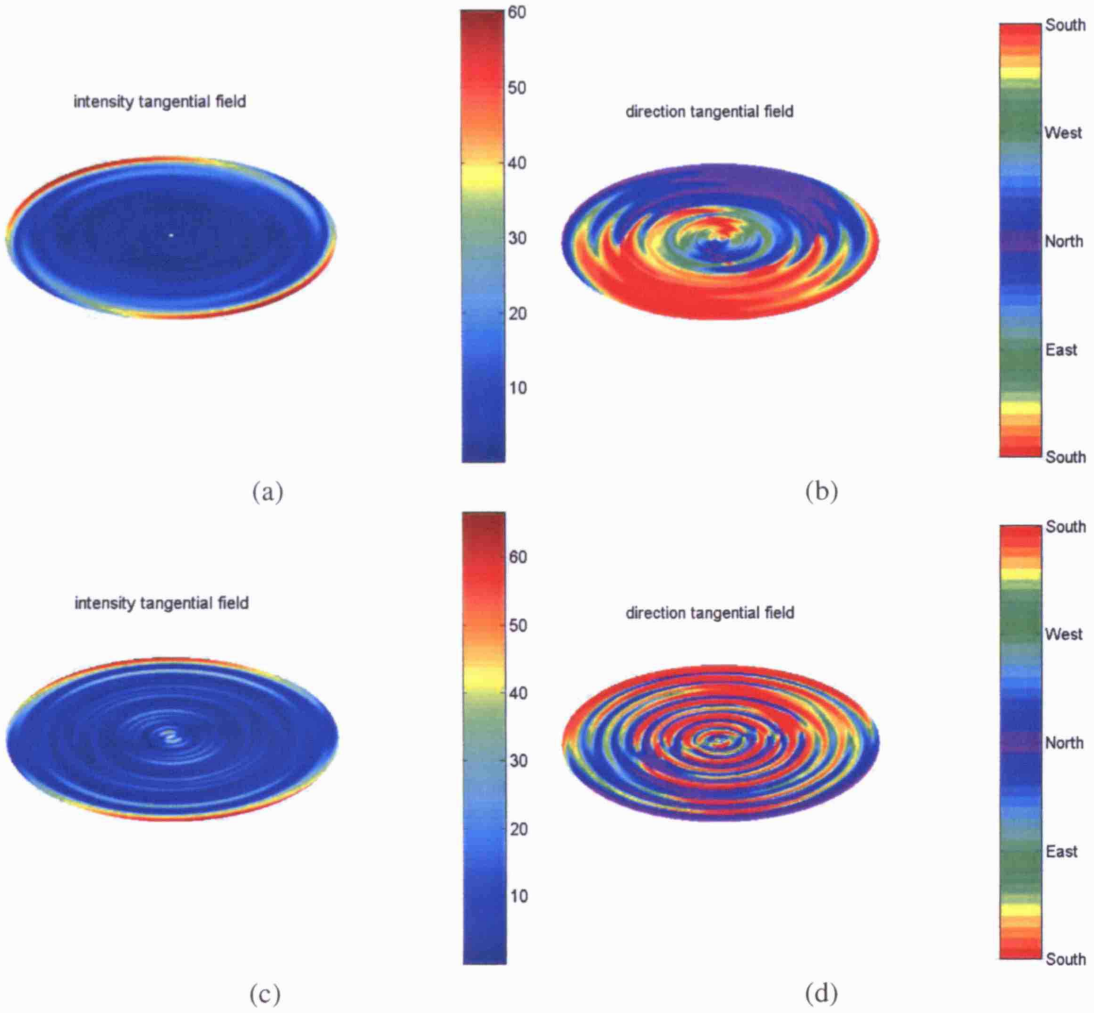


Figure 3.19: The differences between the two aperture field expansions are shown after the two stages of numerical refinement. Figures (a) and (b) show the difference for the E-field and figures (c) and (d) show the difference for the H-field. The difference in intensity between the field according to the refined circular mode expansion (the "left" field) and the field according to the refined spherical mode expansion (the "right" field) are shown relative to the maximum field intensity of the "left" field. The differences are small over most of the aperture. At the edge of the aperture, the fields are not in correspondence, which is the most significant contributor to the remaining error shown in figure 3.20. A possible explanation for the poor match at the edges is that the Spherical equivalent of the Gibb's phenomenon causes a ripple around the discontinuity at the edge of the aperture. It is also likely that higher degree modes need to be included to obtain a better match, as figure 3.18 suggests that the expansions did not include modes of a sufficiently high degree.

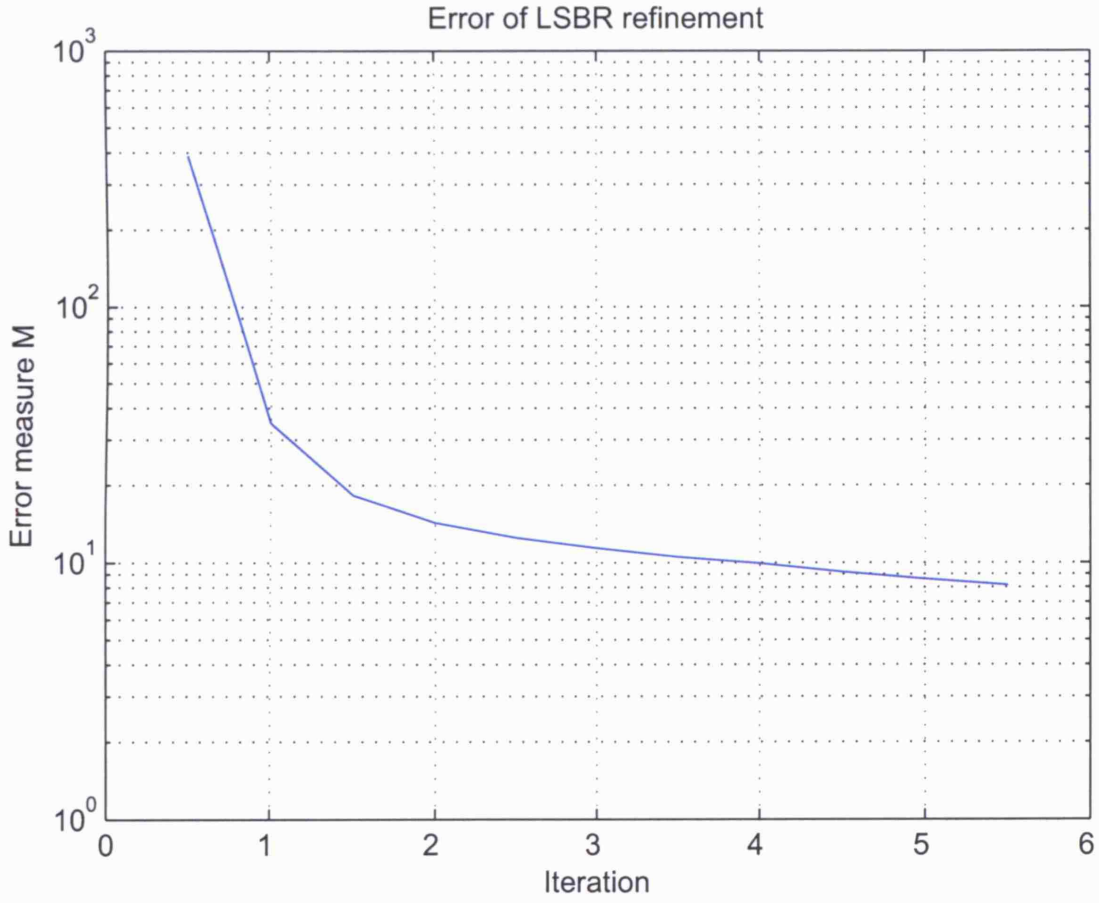


Figure 3.20: The target function used in the LSBR refinement is plotted as a function of the number of iterations. This function is a measure for the error between the left and right expanded aperture fields, as defined in equation (3.59). The plot is on a logarithmic scale. It is clear that the LSBR procedure was able to significantly reduce the mismatch between the two expansions for the aperture field. Over 100 hours of computing time were used in the optimisation procedure, the bulk of which was used to calculate the gradients in the LSBR refinement step.

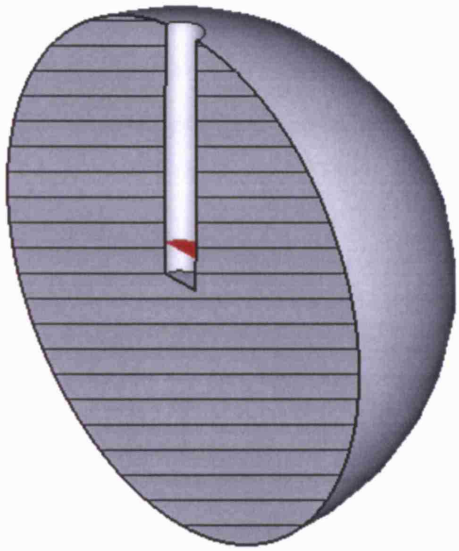


Figure 3.21: The model that was built and modelled with the CST simulation software is shown. A solid metal sphere was modelled to speed up computations. From this solid sphere, a circular waveguide was cut out, that ends on the surface of the sphere. The waveguide is fed by a waveguide port, so that the resonant modes can be precalculated and excited as desired.

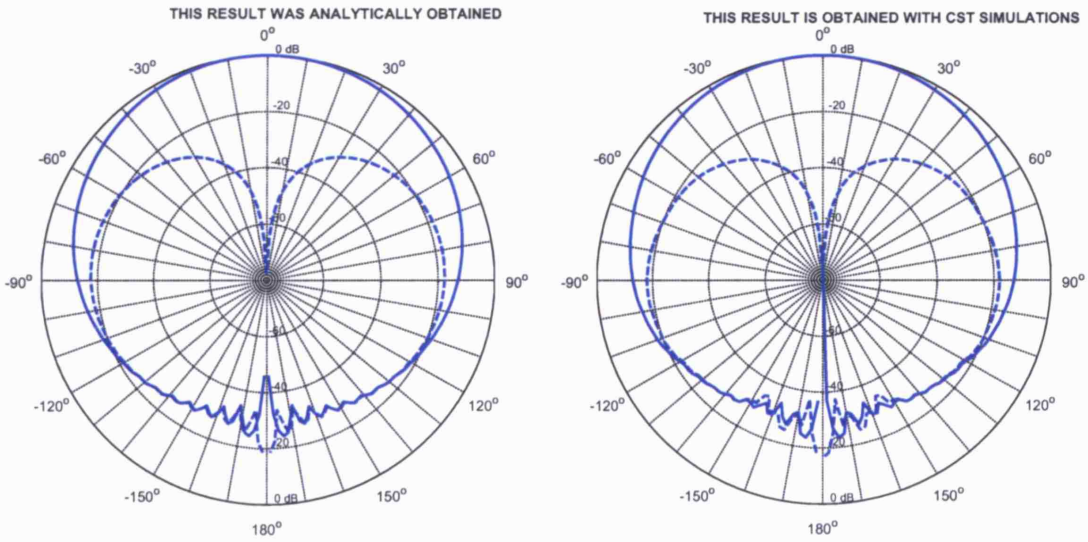


Figure 3.22: The analytically obtained spherical wave expansion is used to calculate the far-field of the fundamental mode in a circular aperture in a metal sphere. The result, for the two polarisations, is shown in the left plot. The right plot shows the simulated far-field of the solid model shown in figure 3.21. The far-field data from the simulation software was converted to the same polarisation definition using MATLAB. The numerical simulations took several hours to complete, whereas evaluation of the theoretical expressions only takes minutes.

so-called type 2 elements, the category that contains patch-type elements. The extension was made with the adoption of some simplifying assumptions, that seemed appropriate for SAR elements. A comparison between simulated results and the theoretical model of various patch-type elements is left for future work, as described in chapter 10.

In this section, we turn our attention to the design of the radiating elements of LISA, a prototype spherical array antenna. The design of the radiating elements was ultimately governed as much by practical and financial considerations as theoretical aspirations. This has led the design of the elements to evolve from aperture radiators over SAR elements to a range of printed antennas until it settled on the design described in section 3.6.3. This evolution is indeed reflected somewhat in the efforts that were put into mathematical modelling - with older ideas having received a lot of theoretical attention and the latest design a lot of more pragmatic simulations.

In our search for the most suitable array element, the greatest challenge was to combine a simple and low-cost design with excellent circular polarisation and a symmetrical radiation pattern. A simple and affordable design is bound to be printed on a Printed Circuit Board (PCB). Printed antennas are cost-effective, easy to implement, and for a large number of elements they benefit from economies of scale.

If allowed by the application, circular polarisation is the most obvious choice for conformal array antennas, as explained in section 6.1. If not, elements with dual polarisation control are required, but this has not been considered further in the scope of this project.

A variety of circularly polarised printed antennas that were available in the literature have been modelled and simulated using CST. It is generally very difficult to achieve good polarisation purity with patch-type antennas. The simplest and oldest designs of circularly polarised patch-type antennas consist of a probe-fed patch, a rectangle or a circular disk for instance, with some irregularity on the edge, so that two orthogonal modes are excited in quadrature. These antennas typically have a very narrow CP bandwidth, in the order of 1% or less. Although our prototype array is designed to work at a spot-frequency of 3 GHz, such a narrow-bandwidth is unacceptable because manufacturing tolerances and handling of the components during assembly could change the frequency of optimal Circular Polarisation (CP) by more than this bandwidth. An array with a large number of elements would therefore quickly become subject to unreliable performance.

The search for the ideal array element was further complicated by the fact that the performance claimed in a great abundance of papers is hard to reproduce, especially if no hardware design has been developed and measured. Often it suffices to include a feeding circuit to a proposed antenna design to make the circular polarisation disappear in some directions or to vanish alltogether. Unsurprisingly, designs where the feed connects from below, through the groundplane, suffer less from the effect of the feed, which results in a more robust circular polarisation.

For a while small printed loop antennas were considered as candidates for the array elements. They result in seemingly simple low-cost single-layer designs. By loading a ring with a reactance, a travelling wave can be generated in the loop, so that circular polarisation is achieved with a single feed[88]. The reactance can be implemented simply as a gap in the loop. It was

found that loop antennas are in reality more complex than they seem at first sight. One major difficulty is the interference between the radiated fields and the feed line. The feed line needs a balun which takes up considerable space. Even with a Co-Planar Waveguide (CPW) implementation the effect of the feed line could not be undone. For the latter a compact double-Y balun design was used[89]. An interesting alternative are loops that are also fed with a probe through the groundplane, as presented in [90].

Eventually, the investigation of existing designs and adaptations thereof, led to the design of a new radiating element. The somewhat unusual but elegant design was called Circularly Polarised, Non-Sequentially Coupled, Single-Feed, Annular-Ring Microstrip Antenna (NOSCIPRA). It is described in detail in section 3.6.3. The element combines a proximity feed with an invisible feedline and further improves the circular polarisation bandwidth for patch type radiators, while achieving a symmetrical radiation pattern and good directivity.

To place the new NOSCIPRA antenna into context and to highlight its innovation relative to existing designs, an attempt was made to classify some commonly known circularly polarised patch-type antennas. The classification, together with some advantages and disadvantages, is given in section 3.6.1.

3.6.1 A Survey of Circularly Polarised Patch Antennas

Since the invention of microstrip technology, the production of accurately printed antennas has become simple and inexpensive. This, and their inherent low profile, have made this type of antenna a popular choice to use as radiating elements in array configurations. The most common type of printed antenna is the patch-type. It has a radiating metallic sheet, cut or edged into some shape, on one side of a dielectric substrate and a ground plane on the other. One of the disadvantages of patch-type antennas is that their radiated fields are more complex to model analytically. Nowadays, the enormous advances in simulation software for electromagnetic problems has made the performance prediction of such antennas a lot easier.

A Circularly Polarised Patch Antennas: Feed Systems

An important part of the patch-type antenna design concerns the feeding method. Over the years a plethora of feeding methods has been presented. The most conventional feeding methods however, belong to one of the following three categories[76][9]:

- **probe feed**: The patch is fed by a vertical probe through the groundplane. Generally the designer first finds a point on the patch where the impedance is 50Ω so that the feed is matched. This type of feed results in a narrow-band antenna. One way of improving the impedance bandwidth is to increase the thickness of the substrate. However, the increased inductance of the vertical probe then makes it difficult to match the feed to the antenna.

- **slot-coupled feed**: The antenna consists of two substrates, separated by a ground plane in the middle. The feeding of the patch on the upper substrate happens by proximity coupling through a slot in the groundplane. The feedline that excites the slot in the groundplane and the feeding circuit can be implemented on the lower substrate, entirely

shielded from the radiating structure by the groundplane. There are a lot more degrees of freedom in this kind of design, which makes the design more complicated, but also allows better matching and broadband operation than is the case with the probe-fed patch.

- **edge feed**: A patch can be fed directly by connecting a microstrip line to the edge. This results in a high input-impedance that can be reduced with an inset feed that connects to the bottom of a gap in the patch. An impedance transformer can also be used to match the line to the antenna.

More recently, the **proximity-coupled** feeding structure enjoyed some attention in the light of circularly polarised patch antennas.

Achieving good CP with microstrip antennas has always been, and still is, a technological challenge. There are two ways of achieving circular polarisation with patch antennas:

- **dual-feed**: Two feeding points each excite a linearly polarised mode. The two modes exist along orthogonal lines, for instance between the two facing pairs of edges of a square patch. By exciting the two orthogonal modes in phase quadrature, circular polarisation is obtained. This method is often avoided because it is complex, costly and results in larger antennas as compared to the single-feed method. However, usually an antenna with dual-feed can maintain its CP over a much wider bandwidth than a single-feed antenna can.

- **single-feed**: As stated in the previous point, the drawback of single-feed patch antennas is that they typically have an extremely narrow CP bandwidth. Most proposed designs for single-feed CP patch antennas achieve around 1% CP bandwidth.

We will further concentrate on single-feed designs, for their simplicity makes them most attractive in an array configuration with a large number of elements. We have simulated a relatively large number of proposed single-feed designs from the literature, as well as variants on them. A major difficulty to achieve CP is formed by the feedline itself. With today's advanced simulation software, we found that numerous papers report results on CP bandwidth from simulations only. Often these simulations are reproducible only to the extent of the feeding point and the performance gets destroyed when a more elaborate feeding circuit is included in the simulation. In reality of course, a proper feeding circuit, or at least a connector with feeding line, cannot be avoided. It seems obvious that a feeding circuit that is less visible to the radiation environment will have less effect on the achieved radiation pattern and its polarisation. Visibility is here used in an electromagnetic context, with a line being visible from the radiation environment if it is radiating into that environment. Our experience seems to suggest however, that any non-compatible feeding structure that is not covered by the metal sheet of the groundplane usually has a large effect on the CP. With non-compatible we mean that the feeding structure is by itself not compatible with CP. Non-compatible structures include all straight microstrip lines. An example of a compatible structure could be an L-shaped microstrip line, fed at a proper location. We will show a design based on this principle later in this section.

Based on the feeding method and the shape of the patch antenna, we can further distinguish between the following circularly polarised patch antennas:

- **parallel mode excitation**: The geometry of feed and patch has been designed so that two orthogonal modes, resonant on the patch shape, are excited with 90° phase difference. Hereto a patch design that would otherwise be linearly polarised is adapted with some perturbed segments. Examples are square patches with two diagonal corners cut away or a circular disk with two diametrically positioned squares chipped away from the edge. The feedpoint to excite CP then needs to be carefully established. The axial ratio that can be achieved in this way is usually poor, especially off boresight.

- **serial mode excitation**: A recent technique that can be applied with both proximity-coupled[91] and slot-coupled[92] feeds. A single bent feedline couples into two orthogonal modes of the patch, while no disturbance of the original patch shape is necessary.

Recently, a new design was presented in [93], which we shall conveniently call the J-feed design. It is a generalisation of the serial mode excitation and consists of a single J-shaped feedline that induces a rotating TM_{11} modal current in the ring above it. The design no longer focuses on the excitation of two orthogonal TM_{11} modes in phase-quadrature, although the excited field can obviously still be decomposed in these orthogonal modes. A major achievement of this design is that it can be fed directly with a 50Ω line, and does not need an impedance transformer. A drawback is that the feedline has an impact on the radiated field and deteriorates the CP performance.

B Circularly Polarised Patch Antennas: Patch Shapes

Another, obviously important, characteristic of patch antennas is their geometry. For circular polarisation the existence of two orthogonal modes on the surface of the patch is required. Therefore geometries that can support two orthogonal modes seem to be optimal. The support for orthogonal modes suggests patch shapes of square or circular symmetry. Triangular patch designs with CP have been reported, with as main advantage their compactness, but we have not investigated these further.

In an array environment, it is also desirable to have array elements with axially symmetric radiation patterns. We therefore intuitively prefer patch shapes that reflect this axial symmetry. Furthermore, it is helpful for theoretical modelling of the antenna element to have angular symmetry so that the structure can easily be described in spherical coordinates. A ring structure is therefore favourable over a crossed dipole for instance, and the selection of the right patch shape narrows down to circular disks and annular rings or combinations thereof.

Another requirement for array elements is that they are as compact as possible, for a given antenna gain. It is easily seen that if the centre part of a circular disk is taken away to form an annular ring, the paths of the currents associated with the fundamental TM_{11}-mode (that extend from one pole to a diametrical pole) will be longer on the average than those on a circular disk. This explains intuitively the observation that indeed an Annular Ring Patch (ARP) antenna can be made smaller than its circular disk counterpart[94][95][96]. Other interesting characteristics of the ARP antenna have attracted much attention. It has been shown that their bandwidth can be significantly increased by operating in higher order modes[97]. Also, they can be made to

operate broadband or multi-band by loading them with concentric rings[98]. A disadvantage though, is that the ARP generally has no 50Ω point and can therefore not be fed with a probe feed. Proposed feeding methods range from an edge-feed with impedance transformer[94] to the attractive J-feed design mentioned above[93].

A common disadvantage to most patch-type antennas is the excitation of surface waves. In array antenna applications it is important to avoid surface wave excitation as much as possible. Surface waves will not contribute to the radiated field, but instead contribute strongly to the mutual coupling effects. An improvement to the ARP antenna is the SAR antenna, which is able not to excite the mode that causes surface waves. This will be discussed further in section 3.6.2.

3.6.2 The Shorted Annular Ring Antenna (SAR)

A common problem with microstrip antennas is that they generally excite surface waves over the groundplane, in addition to the intended space wave. With spherical array antennas, the energy that is put into these surface waves is trapped around the spherical groundplane and will significantly contribute to mutual coupling effects because the surface waves decay slower with distance than space waves do. In this section we will turn our attention to a particular kind of antenna, the SAR that can avoid surface wave excitation and minimise lateral space wave excitation. The antenna is therefore very interesting to use as radiating element on the spherical array.

SAR antennas have a relatively high gain compared to other microstrip designs. On a spherical array it is admissible to use high-gain array elements. For a planar or linear array, this would result in severe pattern deterioration when the beam is scanned off boresight. For the spherical array this is not an issue. The effect of the element pattern on the overall array performance will be revisited in chapter 7.

Moreover, the SAR elements can be well represented by the cavity model, which makes it easy to obtain the field radiated from such an element on a conducting sphere. This was demonstrated already in section 3.3.

Before continuing, it should be mentioned that the idea of using the SAR antenna as array elements on LISA, our prototype spherical array antenna, was ultimately abandoned. The SAR antenna is fairly complex to manufacture and it is larger in size than the ARP antenna. The technical difficulties to overcome with SAR elements arise mainly from the required shorting of the top layer of the circuit that contains the ring to the groundplane. The top layer on which the annular ring is printed is usually a thicker and soft substrate with low dielectric constant. This is necessary to improve respectively the impedance bandwidth and the radiation efficiency. A good choice for instance is foam-based substrate (see section 3.6.3), but the shorting to the ground of the inner edge of the ring poses technical difficulties with this kind of substrate. Also, the feed design needs special attention, as the SAR requires a dual quadrature feed to achieve CP. For a smaller number of array elements, and without subarraying, each array element has its own demodulator. Therefore a design was made that integrated the SAR antenna, its feed, and the demodulator all in one circuit. The demodulator was placed in the centre, for which

the standard feeding method was adapted to a microstrip line that fed the SAR from the centre, through the shorting pins. The design is shown in figure 3.23. The SAR design was dropped when the decision was taken to increase the number of radiating elements on LISA to 720, and cluster them in subarrays. The antenna that made it as the array element to use on LISA is discussed in section 3.6.3. The subarraying technique gets attention in chapter 8.

In section 3.3.1 it was explained how the SAR element and other circular patch type elements can be modelled as a cavity in which TM modes can be excited. For the circular disk and the annular ring, the fundamental mode is the TM_{11} mode. Apart from smaller energy leakages into higher order modes, energy is also inadvertently put in the TM_{01} mode, because this mode has a zero cut-off frequency. It is the TM_{01} mode that excites surface waves over the groundplane[83]. An expression for this surface wave is derived in [83]. In the same paper it is shown that if the ring of equivalent magnetic current for a circular patch antenna, as we used in section 3.3.1, has a particular critical radius, then no TM_{01} mode is excited and hence there is no surface wave radiation (apart from small contributions from the feed and higher order modes). The problem is that the radius is too large to make the patch antenna resonant at its fundamental mode TM_{11}. The solution is to make the antenna appear electrically smaller, while keeping the same outer radius. This is done by using the shorted annular ring structure. Easy design formula's for the SAR antenna can be found in [99]. For the SAR antenna, the TM_{01} mode becomes better tunable, with a resonant frequency below that of the TM_{11}. The latter is thus no longer the fundamental mode. We can therefore clearly distinguish how the different modes radiate by simulating the structure at their respective resonant frequencies. This is shown in figures 3.24 and 3.25. The design shown in figure 3.23 is the result of further developments, where the space inside the SAR antenna was maximised to fit the feeding circuit. This was achieved by choosing a higher value out of the set of discrete radii that suppresses the surface waves. The width of the ring has to remain approximately constant (approximately because of the fringing fields) to keep it resonant. Further simulations on this design have shown that it is possible to reduce the surface waves to a great extent while achieving a pure circular polarisation over wide angles. The latest version of the design, shown in figure 3.23, still has a very narrow CP bandwidth and there is some remaining surface wave excitation. It is harder to remove the surface wave excitation completely because for the larger antenna the resonances of the modes are closer together (in the frequency domain) so that the TM_{01} mode separates less clearly from the adjacent TM_{11} mode. However, the design was never finished and it would be interesting to continue this work, perhaps in the context of a smaller array antenna.

3.6.3 The Non-Serially Proximity Coupled Circularly Polarised Ring Antenna (NOSCIPRA)

The SAR design was abandoned at the same time that the decision was taken to increase the number of elements on the array surface to 720 and cluster them together in subarrays. At that point the SAR design with its shorting pins was found to be too complicated. Also, the feeding point was no longer going to be in the centre, but rather there would be three feeding lines going from the centre of a flat tile to three radiating elements. Some very simple and promising designs were based on small loop antennas. It proved to be very difficult however to

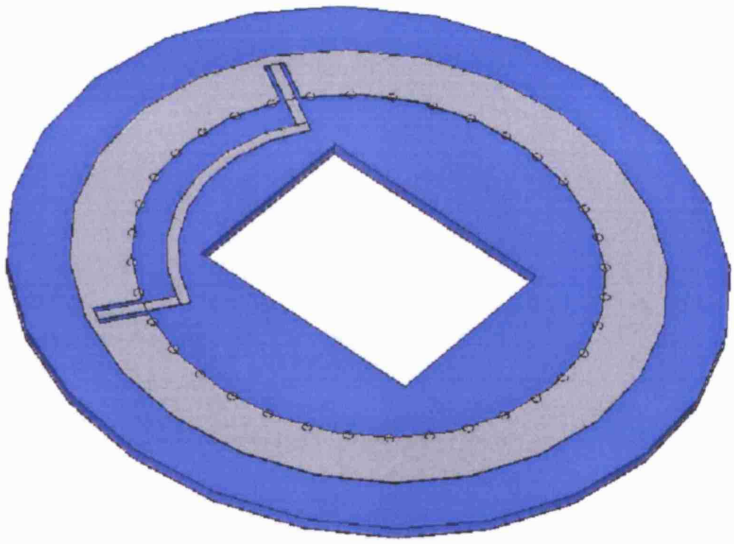


Figure 3.23: A design for an array element with the Shorted Annular Ring (SAR) antenna. The gap in the centre would hold a modulator/demodulator circuit. Note the shorting pins at the inner radius. The feeding point on the bended microstrip line is to be replaced with a microstrip T-junction, which is then fed from the circuit in the middle.

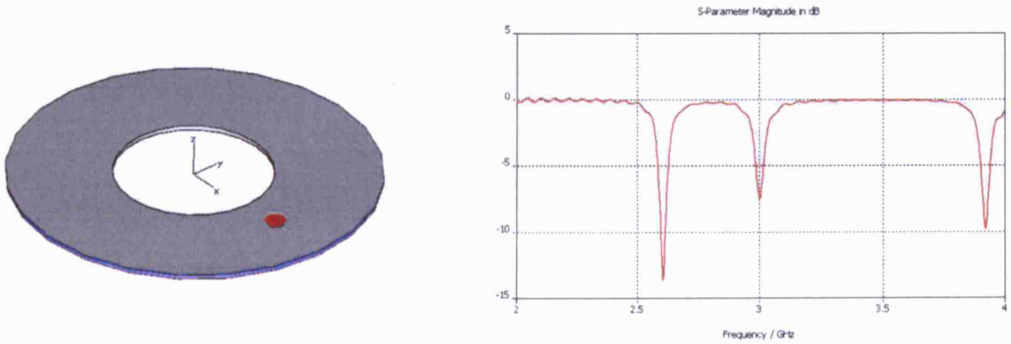


Figure 3.24: The figure on the left shows a basic design of a SAR antenna. The red dot shows the feedpoint between the annular ring and the groundplane. On the right the insertion loss S_{11} is shown. The resonant modes are clearly distinguishable. The far-fields at these resonant frequencies are shown in figure 3.25.

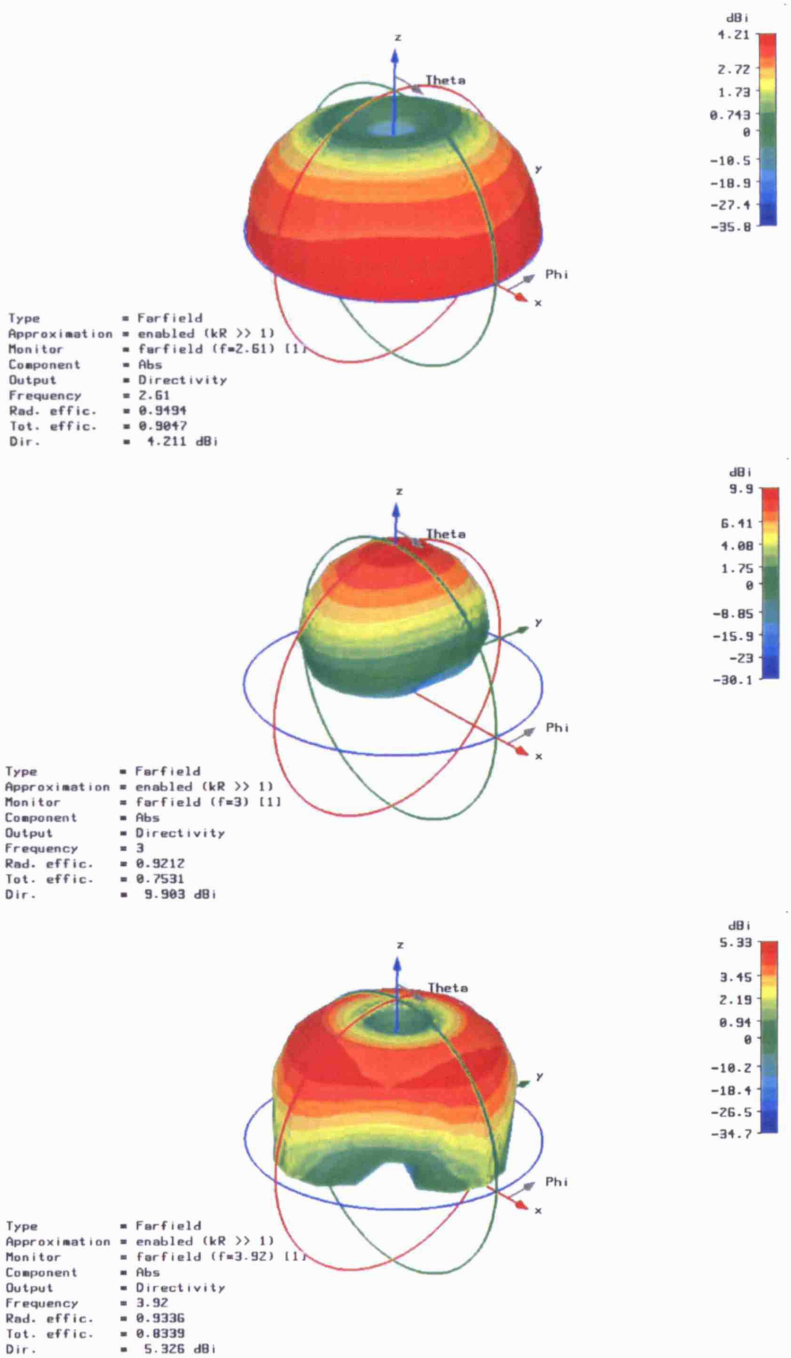


Figure 3.25: Top: far-field of the design in figure 3.24 at the resonant frequency of the TM_{01} mode. Middle: The same for the TM_{11} mode. Bottom: TM_{21} mode. At 3 GHz (the design frequency), there are virtually no surface or lateral waves, due to the good separation with the resonance of the TM_{01} mode. The lateral waves excited by the latter can be seen in the top plot.

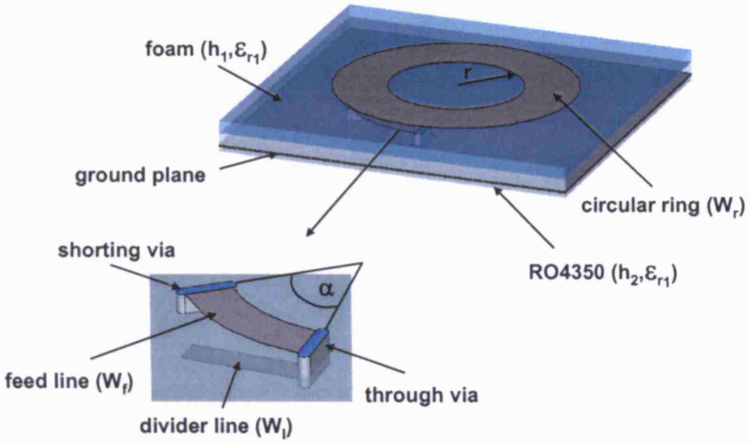
integrate the loop antennas, the matching circuits and baluns on the same PCB without letting the feedlines destroy the circular polarisation.

Instead a new antenna was designed, that can be seen as the successor to the J-feed design described in section 3.6.1. In fact, a prototype of the J-feed antenna was build and measured first. The results were slightly disappointing, in part because the manufacturing was done rather crudely. Obtaining a good CP appeared to be difficult with a real world antenna, which is why the design was further developed to achieve a more robust CP.

The geometrical details of the NOSCIPRA elements are illustrated in figure 3.26. The antenna also induces the current in the annular ring that is associated with a rotating TM_{11} mode. But rather than having the feedline crossing the ring twice and hence performing serial proximity coupling, we realised that this was not a necessary requirement to launch a circular wave into the ring (hence the name). In fact the feed and the ring above act like coupled transmission lines, with the shorted feedline coupling all its energy into the annular ring above. We consider the J-shape of the feed to be merely a leftover from previous designs that focused on a serial excitation of orthogonal modes. This improvement is not just of academic interest. It can actually enhance the performance of the antenna because the feedline is now entirely covered by the annular ring. This must arguably be more efficient, but above all, by exciting the curved arc-feedline from below through a via, the radiation environment is not disturbed by the feedline at all. Moreover, the shape of the feedline is more compatible with the circularly polarised field, which is intuitively better. Hence we were able to achieve a remarkably good CP, unseen with this type of antennas. The CP is better at boresight, but also remains good over a wider range of angles off boresight and also the CP bandwidth has improved to exceed 1%. Because of the symmetry in the visible structure, the radiation pattern is also more symmetrical. As stated before, further improvement of the CP bandwidth is hard to achieve with patch type elements. However, when several elements in an array environment work together in an appropriate way, further improvement of the CP bandwidth is possible. This discussion will be continued in chapter 9.

Apart from the simple ring current, no attempt has been made yet to model the NOSCIPRA antenna analytically with inclusion of the feedline. Although this would no doubt be an interesting exercise that would improve our understanding of how this antenna works, the approach taken in the design process of this antenna has been more pragmatic and was based on a large number of computer simulations. In the structure of this thesis, a further discussion of the design considerations, development issues and obtained results fits better in chapter 9 that covers the development of our prototype array antenna.

It remains true that the antenna excites the TM_{11} mode, as can be clearly seen from figure 3.27. The excitation of surface waves is surprisingly low with this antenna, as can be seen from the purity of the excited mode, and from the radiation pattern depicted in figure 3.28. This suggests that also the NOSCIPRA element can be modelled analytically with the formula's in section 3.3. The gain is somewhat lower than that of the SAR element, but after some design iterations it is possible to get almost no radiation in the lateral direction. An analytical study of why this is true has not been carried out yet.



parameter	value
W_r	7.65 mm
W_f	4 mm
W_l	2.187 mm
h_1	1.8 mm
ϵ_{r_1}	1.2 mm
h_2	1.524 mm
ϵ_{r_1}	3.48 mm
r	9.43 mm
α	50°

Figure 3.26: Illustration of the NOSCIPRA elements. The elements are implemented as a 4-layer PCB. The top layer, edged from a foam substrate, contains the ring. The second layer contains the feedline, the third layer contains the groundplane and the fourth layer contains the feeding circuit (later a power divider will be implemented on the fourth layer). There is a plated through-hole via from one end of the feedline all the way to the fourth layer on the other side of the board and a shorting via to the groundplane. For more details on the substrates etc. please refer to section 9.4.7.

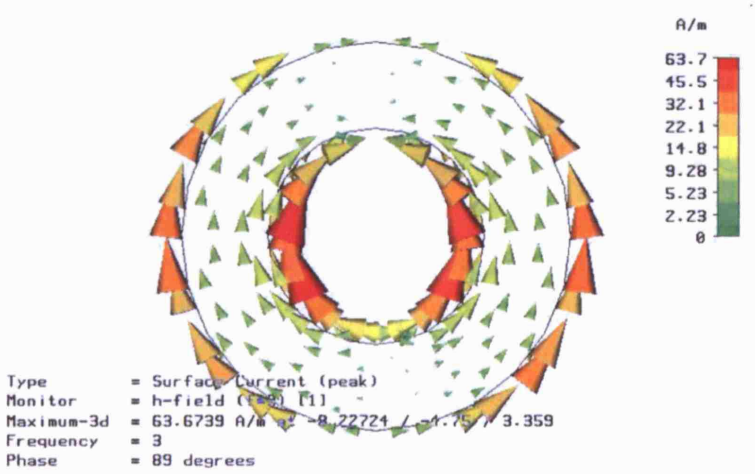


Figure 3.27: From the surface current it can be seen that the TM_{11} mode is excited in the ring. Animation of this plot shows the current distribution undergo a nice and smooth rotation as a function of input signal phase. The purity of the TM_{11} current distribution explains why there is so little lateral radiation in the far-field plot of figure 3.28.


```
Type           = Farfield
Approximation  = enabled (kR >> 1)
Monitor        = farfield (f=3) [1]
Component      = Abs
Output         = Directivity
Frequency      = 3
Rad. effic.    =  1.000
Tot. effic.    = 0.9799
Dir.           =  8.192 dBi
Phase center   = (0.742939, 0, 52.4805) Sigma 13.9752
```

Figure 3.28: The simulated far-field (using CST Microwave Studio) from the NOSCIPRA element. Element gain is slightly less than for the SAR element, but probably more important is the absence of lateral radiation. The importance of that was explained in section 3.6.2.

3.7 Summary and Conclusions

This chapter focused on the electromagnetic properties of one single antenna element of the array antenna. The relation with the general context of this thesis is that the antenna element is placed on or in the proximity of a large conducting sphere. In sections 3.1 to 3.5 various antenna elements were studied theoretically. In section 3.6 an overview of practical radiation elements was given and the design of a new kind of antenna was presented that was developed in the context of this thesis to serve as radiating element of LISA, a prototype spherical array antenna.

Of particular importance are the analytical expressions for the SWE coefficients of a single antenna on a conducting sphere. These results will be used in subsequent chapters in the more complex spherical array environment. An impressive agreement between the theoretical expressions and numerical simulations was illustrated in figure 3.22 on page 99.

The fast and accurate analysis of a single radiating element on a conducting sphere is a first example where SMT can be successfully applied to simplify and rapidly solve an electromagnetic problem regarding a spherical antenna. The first point of the hypothesis in section 1.3.1, that stated that the electromagnetic analysis of spherical antennas can be simplified by using SMT, can thus already be accepted. The point will be made clearer though, as more complex structures will be analysed in the next two chapters.

Chapter 4

Mutual Coupling

4.1 Introduction

Mutual coupling between the radiating elements in an array antenna is generally regarded as an undesirable effect whereby one element couples some energy into its neighbouring elements, who in turn can couple energy back and so forth. The array antenna becomes a complicated coupled structure that is hard to predict and whose radiating characteristics deviate from their design values. Mutual coupling effects can cause losses because some of the energy is coupled back into the feeding network and dissipated there, rather than being radiated.

When mutual coupling is appropriately dealt with, it is possible to take these effects into account in the excitation of the array, and thereby removing the problems associated with them. In chapter 5 this technique will be developed for the spherical array antenna. But first it is necessary to zoom in on the electromagnetic interaction between the radiating elements on the sphere and derive the formula's to be used later in chapter 5. We will concentrate in this chapter on the mutual coupling between circular apertures in a conducting spherical shell (type 1 radiators), although the formulation is just as valid for type 2 radiators provided they can be modelled with an equivalent source above a conducting sphere. As was explained in section 3.3, the patch type element should therefore be designed to minimise surface wave excitation.

The approach taken in this chapter will be to take the spherical expansion of the near-field of a first aperture at the north pole and use the rotational invariance of such an expansion to calculate the fields at the arbitrary position of a second aperture. From the knowledge of the magnetic field in the second aperture, due to an excitation in the first aperture, the mutual admittance can be obtained. The followed method is similar to that presented in [100] and [101] albeit with different basisfunctions and for circular rather than linearly polarised apertures. Also, the generalisation to other modes than the fundamental is made. It will be shown in section 4.4 that it is not sufficient to consider the fundamental mode alone.

4.2 Field Rotation

In section 3.2.3 we have obtained the STL expansion for the reactive near-field from one radiating element on a conducting sphere. It is necessary to use this expansion because it gives a full representation of the field, including the evanescent modes that only exist in the immediate

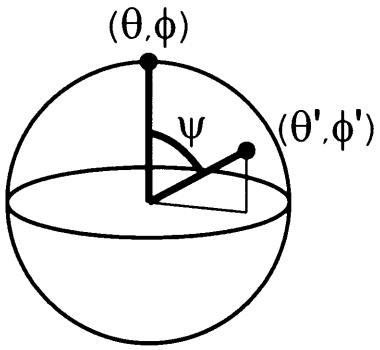


Figure 4.1: Angles used in the addition theorem (4.1)

vicinity of the sphere. Also via these modes, elements can couple reactively into one another and perturb the overall performance of the array antenna. The expansions in equations (3.21) and (3.22) were obtained with the element placed at the north pole. Our goal is to calculate how this near-field couples into another element that is placed on the sphere. In order to do that, we place the new element on the sphere and rotate the sphere so that it is positioned at the north pole. The action of rotating the sphere will be performed by an Euler transformation on the axes. We also want to take polarisation into account. Therefore we will also rotate the axes so that the polarisation of the second element matches that of the first element when it was positioned at the north pole.

With the second element at the north pole, it is receiving energy from the same spherical modes as the first element was coupling its energy into while in the same position. Hence we can work out the coupling from the first element to the second. However, the axes have undergone a rotation, and the expansion of the first element must be known in the system of rotated axes to be useful. That is where the rotational invariance of the spherical modes comes into play. Every spherical mode of degree l will transform into a linear combination of spherical modes of the same degree under rotation of the axes. There are $2l + 1$ spherical modes of degree l, so that each eigenspace of $2l + 1$ modes is transformed by a rotation matrix with dimensions $(2l + 1) \times (2l + 1)$. By comparison, for circular modes the eigenspaces were one-dimensional, and indeed each term in the Fourier-series can be rotated on the circle by multiplication with a 1×1 phase factor.

4.2.1 Rotation of the Expansion

The invariance of the eigenspace of spherical harmonics of degree l is enforced by the Legendre addition theorem:

$$P_l(\cos\psi) = \frac{4\pi}{2l+1} \sum_{m=-l}^{l} Y_l^{m^*}(\theta,\phi) Y_l^m(\theta',\phi') \tag{4.1}$$

with:

$$\cos\psi = \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\phi - \phi')$$

This theorem can be used to reference the Legendre polynomial to another z-axis. It can be reformulated to cover the associated Legendre functions and its derivatives[102]. Important here

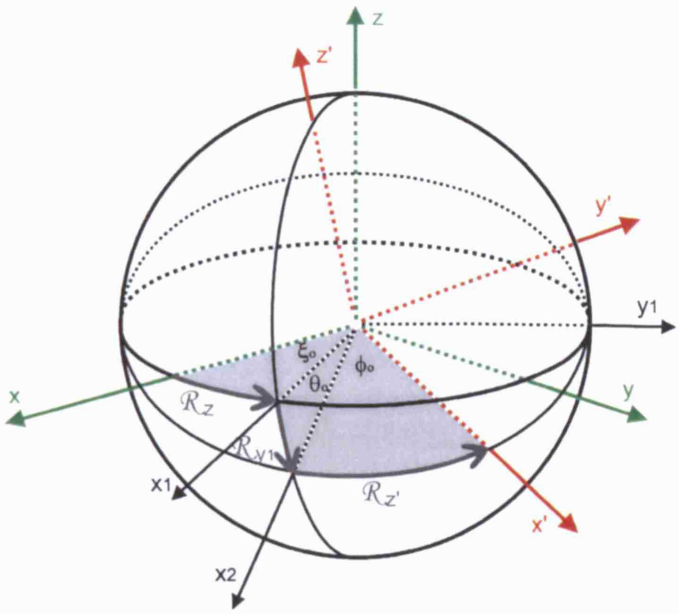


Figure 4.2: An arbitrary rotation of the axes xyz to the axes x'y'z' is described by three consecutive Euler rotations $R_z(\xi_0)$, $R_{y1}(\theta_0)$ and $R_{z'}(\phi_0)$.

is that only Legendre functions of the same degree occur in the linear combination when the z-axis is tilted. As with circular modes, rotations around the z-axis can still be performed by multiplication with a phase-term. Euler's theorem dictates that any arbitrary rotation can be described by three independent parameters, for which usually three angles of rotation, called the Euler angles, are used. Several conventions exist for the Euler angles, the convention adopted in this text is often used and depicted in figure 4.2. We identify a triplet of angles $(\xi_0, \theta_0, \phi_0)$, where θ_0 and ϕ_0 describe a rotation in the spherical coordinate system, and ξ_0 a rotation around the z-axis to align the element's polarisation with an arbitrary direction. We will describe functions that are known in one coordinate system in a second coordinate system, and therefore transform the coordinate system. The first and the third Euler rotations are rotations around the old and the new z-axis, and are achieved by multiplication with a phase-term. The second Euler rotation is a rotation of the z-axis itself, for which a suitable version of the addition theorem is needed. All together, an arbitrary rotation of the axes is performed by a linear recombination within each eigenspace of spherical harmonics of degree l. With the vector spherical wave functions defined through the action of linear operators on the spherical harmonics, this invariance still holds and each spherical wave function of the STL expansion in (2.38) can be written as a linear combination of $2l + 1$ spherical wave functions of degree l, that are defined in a rotated system of axes. Let (θ', ϕ') be the spherical coordinates in a rotated system of axes x'y'z'. The wave function $\vec{e}_{lm}(\theta, \phi)$ known in the system xyz is then known in the rotated system as:

$$\vec{e}_{lm}(\theta, \phi) = \sum_{m'=-l}^{l} D^l_{m'm}(\xi_0, \theta_0, \phi_0)\vec{e}_{lm'}(\theta', \phi') \tag{4.2}$$

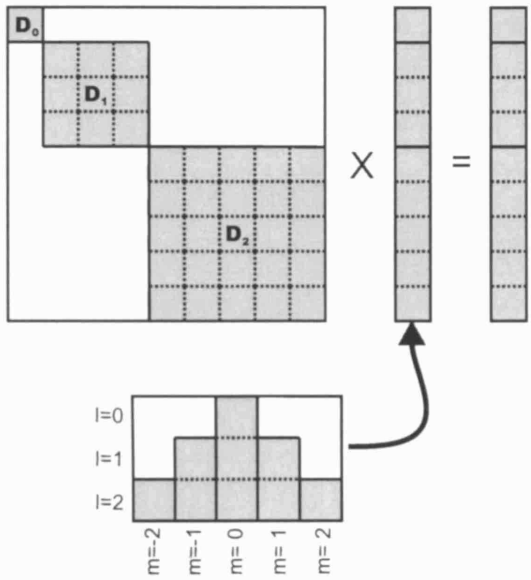


Figure 4.3: With the coefficients of the spherical wave expansion organised in a vector as shown, a rotation of the axes appears as a block diagonal rotation matrix. The spherical wave functions of the same degree recombine, but do not mix up with those of another degree. The rotation blocks D_l are calculated as a function of three Euler angles.

with the coefficients $D^l_{m'm}$ yet to be determined. With this knowledge, a spherical wave expansion that was determined in system xyz can be transferred to system x'y'z':

$$\vec{E}_t(r,\theta,\phi) = \sum_{lm} V^{\mathrm{TE}}_{lm}(r)\frac{\vec{e}^{\,\mathrm{TE}}_{lm}}{r} + \sum_{lm} V^{\mathrm{TM}}_{lm}(r)\frac{\vec{e}^{\,\mathrm{TM}}_{lm}}{r} \tag{4.3a}$$

$$= \sum_{l=1}^{\infty}\sum_{m=-l}^{l}\sum_{m'=-l}^{l} D^l_{m'm}(\xi_0,\theta_0,\phi_0)\left(V^{\mathrm{TE}}_{lm}(r)\frac{\vec{e}^{\,\mathrm{TE}}_{lm'}(\theta',\phi')}{r} + V^{\mathrm{TM}}_{lm}(r)\frac{\vec{e}^{\,\mathrm{TM}}_{lm'}(\theta',\phi')}{r}\right) \tag{4.3b}$$

$$= \sum_{lm'} V'^{\mathrm{TE}}_{lm'}(r)\frac{\vec{e}^{\,\mathrm{TE}}_{lm'}(\theta',\phi')}{r} + \sum_{lm} V'^{\mathrm{TM}}_{lm'}(r)\frac{\vec{e}^{\,\mathrm{TM}}_{lm'}(\theta',\phi')}{r} \tag{4.3c}$$

with:

$$V'_{lm'} = \sum_{m=-l}^{l} D^l_{m'm}(\xi_0,\theta_0,\phi_0)V_{lm} \tag{4.4}$$

So the expansion coefficients in the rotated system of axes are obtained from the original expansion coefficients by left multiplication with a block diagonal matrix as illustrated in figure 4.3. Obtaining the blocks D_l of the rotation matrix is the subject of next section.

4.2.2 Calculating the Rotation Matrix

The rotation matrix D is a block matrix with rotation matrices D_l on its diagonal and all other elements zero. The calculation of the elements $D^l_{m'm}$ of the rotation matrices D_l is the subject of this section. We know that by decomposing the rotation in three consecutive Euler rotations,

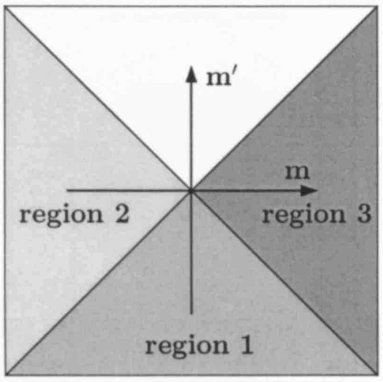


Figure 4.4: The Jacobi polynomial is undefined in the shaded regions, but symmetry relations exist that relate the rotation coefficients in these regions to those in the unshaded region.

the elements of the rotation matrix can be written as[56]:

$$D^l_{m'm}(\xi_0, \theta_0, \phi_0) \;=\; e^{-jm\xi_0} d^l_{m'm}(\theta_0) e^{-jm'\phi_0} \tag{4.5}$$

where the yet unknown elements $d^l_{m'm}(\theta_0)$ have to account for the rotation of the z-axis. Determination of these elements involves generalising the addition theorem to the associated Legendre functions and can be found in the literature[103][102][104][63]. An expression for the rotation coefficients $d^l_{m'm}$ is:

$$d^l_{m'm} \;=\; \sqrt{\frac{(l+m')!(l-m')!}{(l+m)!(l-m)!}} \left(\cos\frac{\theta}{2}\right)^{m'+m} \left(\sin\frac{\theta}{2}\right)^{m'-m} P^{m'-m,m'+m}_{l-m'}(\cos\theta)$$

$$\text{for} \begin{cases} -m \leqslant m' \leqslant m \\ 0 \leqslant m' \end{cases} \tag{4.6}$$

$$\begin{aligned} d^l_{m'm} &= (-1)^{(m'+m)} d^l_{-m'-m} && \text{for region 1} \\ d^l_{m'm} &= d^l_{-m-m'} && \text{for region 2} \\ d^l_{m'm} &= (-1)^{(m'+m)} d^l_{mm'} && \text{for region 3} \end{aligned}$$

where $P^{\alpha,\beta}_l(x)$ is the Jacobi polynomial, defined for $\alpha \geqslant 0$ and $\beta \geqslant 0$, and the regions for the symmetry relations are indicated in figure 4.4.

The rotation coefficients $d^l_{m'm}$ can be obtained from a left and a right recursion relation:

left recursion start:

$$d^l_{lm}(\theta_0) \;=\; \sqrt{\frac{(2l)!}{(l+m)!(l-m)!}} \left(\cos\frac{\theta_0}{2}\right)^{l+m} \left(\sin\frac{\theta_0}{2}\right)^{l-m} \tag{4.7a}$$

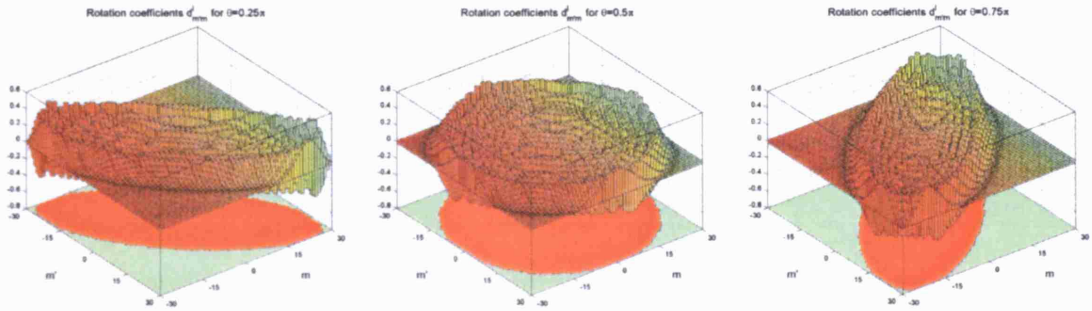


Figure 4.5: The magnitude of the coefficients in the rotation matrix $d(\theta)$ are illustrated for 3 angles: $\theta = \frac{\pi}{4}$, $\theta = \frac{\pi}{2}$ and $\theta = \frac{3\pi}{4}$. For stability, the recurrences have to start at the edge and run to the middle of the oval domain. The stability domain calculated by the algorithm is shown as a red shade below the bar-plots.

left recursion:

$$d^l_{(m'-1)m}(\theta_0) = -\frac{(2m - 2m'\cos\theta_0)}{\sqrt{(l+m')(l-m'+1)}\sin\theta_0}\, d^l_{m'm}(\theta_0) - \sqrt{\frac{(l-m')(l+m'+1)}{(l+m')(l-m'+1)}}\, d^l_{(m'+1)m}(\theta_0) \quad (4.7b)$$

right recursion start:

$$d^l_{(-l)m}(\theta_0) = \sqrt{\frac{(2l)!}{(l+m)!(l-m)!}}(-1)^{l+m}\left(\cos\frac{\theta_0}{2}\right)^{l-m}\left(\sin\frac{\theta_0}{2}\right)^{l+m} \quad (4.8a)$$

right recursion:

$$d^l_{(m'+1)m}(\theta_0) = -\frac{(2m - 2m'\cos\theta_0)}{\sqrt{(l-m')(l+m'+1)}\sin\theta_0}\, d^l_{m'm}(\theta_0) - \sqrt{\frac{(l+m')(l-m'+1)}{(l-m')(l+m'+1)}}\, d^l_{(m'-1)m}(\theta_0) \quad (4.8b)$$

We will now seek a function $d_{m'}(\theta)$ for each line of m' in the two-dimensional grid of (m', m)-points. We have the choice of starting the recurrence on the left with $m' = -l$ and move to the right with (4.8) or we can start on the right and use (4.7) to move to the left. As we will see, we will need to start on either side and let the left and right recurrence meet in the middle for stability reasons.

When dealing with linear three-term recurrences, one must bear in mind that there are two independent solutions, one of which is to be found by setting the starting values. However, one of the solutions may be dominant over the other, so that the inevitable slightest inaccuracy will cause the recurrence to move towards the dominant solution, away from the minimal solution. We will investigate the stability of the recurrence with an heuristic approach as in [105] and [63]. The three-term recurrence has the form $y_{m'+1} + b_{m'}y_{m'} + c_{m'}y_{m'-1} = 0$. To investigate the

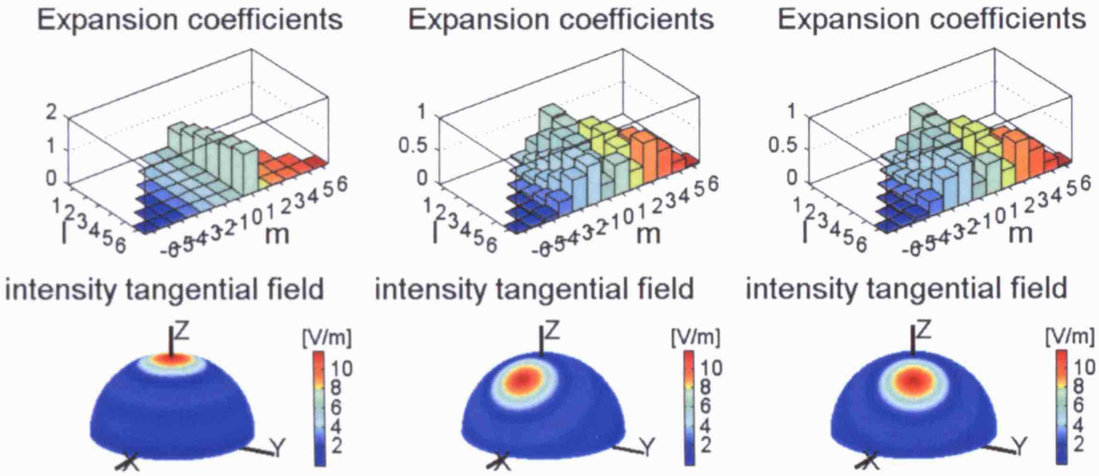


Figure 4.6: The rotation of a spherical expansion is illustrated. In the left column, the original expansion is shown. There are 12 modes of order 1 in the field shown below, but only the coefficients for the TE-modes are shown above. The plots in the second column show the expansion in a set of axes rotated over an angle $\theta_0 = \frac{\pi}{6}$, obtained with (4.4) and the corresponding field below. The third column adds another rotation over $\phi_0 = \frac{\pi}{6}$. Only the magnitude of the expansion coefficients is plotted, which explains why no difference is observed between the coefficients in column two and three. Note that the fields illustrated in the second row were obtained with a true superposition of the expansions illustrated in the first row.

stability of the recurrence, we replace it at all points m' with one that is locally similar, but with constant coefficients b and c. Stability is a property of the recurrence alone, not of the starting values. Hence we can use a suitable function and see how the two solutions behave. We try the recurrence with a solution of the form $y_{m'} = (k)^{m'}$, so that we get $k^{m'-1}(k^2 + bk + c) = 0$, which indeed has two solutions $y_{m'} = (k_1)^{m'}$ and $y_{m'} = (k_2)^{m'}$, where k_1 and k_2 are the roots of the quadratic equation. If there are two real roots for the mantissa, one of smaller magnitude than the other, the one with the larger mantissa is dominant over the other. Imaginary roots cause an oscillatory behaviour. The determinant of the quadratic equation is:

$$D = \frac{(2m - 2m'\cos\theta)^2}{(l + m' + 1)(l - m')\sin^2\theta} - 4\sqrt{\frac{(l + m')(l - m' + 1)}{(l - m')(l + m' + 1)}}$$

The region where $D < 0$ is indicated in figure 4.5 by the red surface below the plots of the rotation coefficients $d^l_{m'm}$. When a recurrence is started on one side it starts in a yellow region, and the solution that ramps up is the one to be found. The recurrence then enters a region where it becomes oscillatory. If the recurrence were to be continued beyond the second boundary, it would ramp up again towards infinity, rather then decay back to zero. Therefore we stop the recurrence in the middle of the red region, and move in from the other side with the other recurrence relation. Where the two recurrences meet, the same coefficient is obtained within machine precision. The obtained results for the rotation coefficients are shown in figure 4.5.

The obtained rotation coefficients can now be filled in in the expression for the rotation matrix in (4.5), so that the expansion coefficients can be expressed in the rotated system of axes as in (4.4). This is illustrated in figure 4.6.

4.3 Calculating the Coupling Coefficients

In this section we will calculate the mutual admittance between two circular apertures in the conducting sphere. We place an active primary element at NP, excited with a voltage V_{np}^{TX} and a secondary short-circuited element at some other position on the sphere. We refer to this system of axes as the unprimed system. Let $\vec{H}_1^{\mathrm{TX}_{np}}$ be the magnetic field over the secondary aperture, caused by the circular mode TX_{np} in primary aperture. We now rotate the sphere so that the secondary aperture is in NP position and call the obtained system of axes the primed system. Let $\vec{E}_2^{\mathrm{TX}_{qw}}$ be the electric field of a circular mode in the secondary aperture, excited with a voltage V_{qw}^{TX}. The mutual admittance is then given by[1]:

$$Y_{\mathrm{TX}_{np}\mathrm{TX}_{qw}} = \frac{1}{V_{np}^{\mathrm{TX}} V_{qw}^{\mathrm{TX}}} \int_{\text{aperture 2}} \left(\vec{E}_2^{\mathrm{TX}_{qw}} \times \vec{H}_1^{\mathrm{TX}_{np}}\right) \cdot \hat{e}_r \, \mathrm{dS'} \tag{4.9}$$

Where the integration is over the secondary aperture in the primed system of axes. The field $\vec{E}_2^{\mathrm{TX}_{qw}}(\theta', \phi')$ can be expanded as in (2.38) and the modal voltages are obtained from (3.21) and (3.22) upon substitution of n with q and p with w:

$$\vec{E}_2^{\mathrm{TX}_{qw}}(\theta', \phi') = \sum_{l,m} V_{2,\mathrm{TX}_{qw}}^{'\,\mathrm{TE}_{lm}}(r) \frac{\vec{e}_{lm}^{\mathrm{TE}}}{r} + \sum_{l} V_{2,\mathrm{TX}_{qw}}^{'\,\mathrm{TM}_{lm}}(r) \frac{\vec{e}_{lm}^{\mathrm{TM}}}{r} \tag{4.10}$$

The summation over m hereby disappears, as we consider one circular mode TX_{qw} and the spherical modes with $m \neq q$ vanish:

$$\begin{aligned}
\vec{E}_2^{\mathrm{TX}_{qw}}(\theta', \phi') &= \sum_{l} V_{2,\mathrm{TX}_{qw}}^{'\,\mathrm{TE}_{lq}}(r) \frac{\vec{e}_{lm}^{\mathrm{TE}}}{r} + \sum_{l} V_{2,\mathrm{TX}_{qw}}^{'\,\mathrm{TM}_{lq}}(r) \frac{\vec{e}_{lm}^{\mathrm{TM}}}{r} \qquad (4.11)\\
V_{2,\mathrm{TE}_{qw}}^{'\,\mathrm{TE}_{lq}}(r) &= 2\pi r V_{qw}^{\mathrm{TE}} N_{awq}^{\mathrm{TE}} N_{lq} \int_{\theta'=0}^{\theta_2} \frac{s_{qw}^{\mathrm{TE}}\theta'}{\theta_2} J_q\left(\frac{s_{qw}^{\mathrm{TE}}\theta'}{\theta_2}\right) P_l^q(\cos\theta') d\theta'\\
V_{2,\mathrm{TE}_{qw}}^{'\,\mathrm{TM}_{lq}}(r) &= 2\pi r j q \frac{\theta_2}{s_{qw}^{\mathrm{TE}}} V_{qw}^{\mathrm{TE}} N_{awq}^{\mathrm{TE}} N_{ln} J_q\left(s_{np}^{\mathrm{TE}}\right) P_l^n(\cos\theta_2)\\
V_{2,\mathrm{TM}_{qw}}^{'\,\mathrm{TE}_{lq}}(r) &= 0\\
V_{2,\mathrm{TM}_{qw}}^{'\,\mathrm{TM}_{lq}}(r) &= -2\pi r V_{qw}^{\mathrm{TM}} N_{awq}^{\mathrm{TM}} N_{lq} \int_{\theta'=0}^{\theta_2} \frac{s_{qw}^{\mathrm{TM}}\theta'}{\theta_2} J_q\left(\frac{s_{qw}^{\mathrm{TM}}\theta'}{\theta_2}\right) P_l^q(\cos\theta') d\theta'
\end{aligned}$$

The field $\vec{H}_1^{\mathrm{TX}_{np}}$ is obtained by using the same expansion, which was valid prior to rotation of the axes, and rotating the modal currents to the new system of axes as we have done in (4.4). Also here the summation over m disappears, as we consider only a circular mode TX_{np} and the spherical modes with $m \neq n$ vanish:

$$\vec{H}_1^{\mathrm{TX}_{np}}(\theta', \phi') = \sum_{lm'} I_{1,\mathrm{TX}_{np}}^{'\,\mathrm{TE}_{lm'}}(r) \frac{\vec{h}_{lm'}^{\mathrm{TE}}}{r} + \sum_{lm'} I_{1,\mathrm{TX}_{np}}^{'\,\mathrm{TM}_{lm'}}(r) \frac{\vec{h}_{lm'}^{\mathrm{TM}}}{r} \tag{4.12}$$

with:

$$\begin{aligned} I'^{\mathrm{TE}_{lm'}}_{1,\mathrm{TX}_{np}}(r) &= Y^{i}_{lm'\mathrm{TE}} V^{\mathrm{TE}_{lm'}}_{1,\mathrm{TX}_{np}}(r) \\ &= Y^{i}_{lm'\mathrm{TE}} \sum_{m} D^{l}_{m'm}(\xi_0,\theta_0,\phi_0) V^{\mathrm{TE}_{lm}}_{1,\mathrm{TX}_{np}}(r) \\ &= Y^{i}_{lm'\mathrm{TE}} D^{l}_{m'n}(\xi_0,\theta_0,\phi_0) V^{\mathrm{TE}_{ln}}_{1,\mathrm{TX}_{np}}(r) \end{aligned}$$

and likewise:

$$I'^{\mathrm{TM}_{lm'}}_{1,\mathrm{TX}_{np}}(r) = Y^{i}_{lm'\mathrm{TM}} D^{l}_{m'n}(\xi_0,\theta_0,\phi_0) V^{\mathrm{TM}_{ln}}_{1,\mathrm{TX}_{np}}(r)$$

With regard to figure 4.3, the complexity of the rotation has been reduced with one dimension. Each block D_l now has a single column that can be calculated with one single recurrence over m'. However, if we are not interested in reconstruction of the entire rotated magnetic field and just want to obtain the mutual admittances, we can further reduce the complexity by once again making use of the orthogonality between spherical and circular modes of different order. As we are looking for the mutual admittance from a TX_{np}-mode in aperture one to a TX_{qw}-mode in aperture two, we know that we are ultimately only interested in the spherical modes of order $m' = q$. The blocks D_l can therefore be further collapsed into single elements, that are calculated using (4.5) and (4.6):

$$\begin{aligned} D^{l}_{qn} &= e^{-jn\xi_0} d^{l}_{qn}(\theta_0) e^{-jq\phi_0} \\ d^{l}_{qn} &= \sqrt{\frac{(l+q)!(l-q)!}{(l+n)!(l-n)!}} \left(\cos\frac{\theta_0}{2}\right)^{q+n} \left(\sin\frac{\theta_0}{2}\right)^{q-n} P^{q-n,q+n}_{l-q}(\cos\theta_0) \end{aligned}$$

If the mutual admittance between circular modes of the same order is calculated, the rotation elements simplify to:

$$D^{l}_{nn} = e^{-jn(\xi_0+\phi_0)} d^{l}_{nn}(\theta_0) \tag{4.13a}$$

$$d^{l}_{nn} = \left(\frac{1}{2}(1+\cos\theta_0)\right)^{n} P^{0,2n}_{l-n}(\cos\theta_0) \tag{4.13b}$$

By substituting the expansions (4.11) and (4.12) into (4.9), and applying the orthogonality of the spherical modes, the mutual admittance becomes:

$$\begin{aligned} Y_{\mathrm{TX}_{np}\mathrm{TX}_{qw}} &= \frac{1}{V^{\mathrm{TX}}_{np} V^{\mathrm{TX}}_{qw}} \frac{1}{r^2} \sum_{l} \left(V'^{\mathrm{TE}_{lq}}_{2,\mathrm{TX}_{qw}} I'^{\mathrm{TE}_{lq}}_{1,\mathrm{TX}_{np}} + V'^{\mathrm{TM}_{lq}}_{2,\mathrm{TX}_{qw}} I'^{\mathrm{TM}_{lq}}_{1,\mathrm{TX}_{np}} \right) \\ Y_{\mathrm{TX}_{np}\mathrm{TX}_{qw}} &= \frac{1}{V^{\mathrm{TX}}_{np} V^{\mathrm{TX}}_{qw}} \frac{1}{r^2} \sum_{l=\max(|q|,|n|,1)}^{\infty} \left(V'^{\mathrm{TE}_{lq}}_{2,\mathrm{TX}_{qw}} D^{l}_{qn} Y^{i}_{lq\mathrm{TE}} V^{\mathrm{TE}_{ln}}_{1,\mathrm{TX}_{np}} + V'^{\mathrm{TM}_{lq}}_{2,\mathrm{TX}_{qw}} D^{l}_{qn} Y^{i}_{lq\mathrm{TM}} V^{\mathrm{TM}_{ln}}_{1,\mathrm{TX}_{np}} \right) \end{aligned} \tag{4.14}$$

If the spherical transmission line expansion coefficients $V'^{\mathrm{TE}_{lq}}_{2,\mathrm{TX}_{qw}}$ and $V'^{\mathrm{TM}_{lq}}_{2,\mathrm{TX}_{qw}}$ were obtained with the assumption of an incident circular mode carrying 1W of RMS power, then the voltages V^{TX}_{np}

and V_{qw}^{TX} can be obtained from equation (3.7). An important observation should be made at this point. The magnetic field in the expression for the mutual admittance is reconstructed from the modal voltages by using the input admittances of the spherical modes. This implicitly means that we have stepped over the discontinuity at the first aperture and the magnetic field used here is not the same as the one that was originally defined in the circular waveguide, but rather one that can exist over the surface of the sphere. This is the point where the numerical mode matching explained in chapter 2 could be applied to derive both the modal voltages and modal currents. This also means that care should be taken in the calculation of the self-admittances of each aperture. Consider for instance the self admittance of the TE_{11} mode in the primary aperture. Suppose that the circular mode carries 1W of RMS power, so that the integral in equation (4.9) equals $2 \times P_{av} = 2$. The denominator of the same equation becomes:

$$\frac{1}{(V_{11}^{\mathrm{TE}})^2} = \frac{\sqrt{1-(\frac{f_c}{f})^2}}{2\eta} \tag{4.15}$$

and the total equation equals the self-admittance for TE modes, that can be derived from equation (3.7):

$$Y_{\mathrm{TE}}^{i} = \frac{\sqrt{1-(\frac{f_c}{f})^2}}{\eta} \tag{4.16}$$

However, if equation (4.14) is used to calculate the self-admittance of the circular modes, a different value will be found. This apparent contradiction is due to the fact that in the latter equation we have stepped out of the circular waveguide already and inside the circular waveguide the equation is not valid. Therefore we will remember to calculate the self-admittances of the circular modes explicitly, when they are needed in the next chapter.

If the two apertures are identical, the mutual admittance between their RHCP fundamental TE_{11}-modes becomes:

$$Y_{\mathrm{TE}_{11}\mathrm{TE}_{11}} = \frac{1}{(V_{11}^{\mathrm{TE}})^2}\frac{1}{r^2}\sum_{l} D_{11}^{l}\left(Y_{l1\mathrm{TE}}^{i}\left(V_{1,\mathrm{TE}_{11}}^{\mathrm{TE}_{l1}}\right)^2 + Y_{l1\mathrm{TM}}^{i}\left(V_{1,\mathrm{TE}_{11}}^{\mathrm{TM}_{l1}}\right)^2\right) \tag{4.17}$$

$$\text{with} \quad D_{11}^{l} = e^{-j(\xi_0+\phi_0)}\frac{1+\cos\theta_0}{2}P_{l-1}^{0,2}(\cos\theta_0)$$

A circular waveguide that supports a TE_{11} RHCP mode will also support the LHCP mode, so that also the coupling into the TE_{-11}-mode must be taken into account. As we are outside the validity region for the Jacobi polynomial, the symmetry relations in (4.6) have to be used on this occasion.

$$Y_{\mathrm{TE}_{11}\mathrm{TE}_{-11}} = \frac{1}{V_{11}^{\mathrm{TE}}V_{-11}^{\mathrm{TE}}}\frac{1}{r^2}\sum_{l}\left(V_{2,\mathrm{TX}_{-11}}^{'\,\mathrm{TE}_{l(-1)}} D_{-11}^{l} Y_{l1\mathrm{TE}}^{i} V_{1,\mathrm{TE}_{11}}^{\mathrm{TE}_{l1}} + V_{2,\mathrm{TX}_{-11}}^{'\,\mathrm{TM}_{l(-1)}} D_{-11}^{l} Y_{l1\mathrm{TM}}^{i} V_{1,\mathrm{TE}_{11}}^{\mathrm{TM}_{l1}}\right) \tag{4.18}$$

$$\text{with} \quad D_{-11}^{l} = e^{-j(\xi_0+\phi_0)}\frac{1-\cos\theta_0}{2}P_{l-1}^{2,0}(\cos\theta_0)$$

With circular polarisation in the aperture, the magnitude of the mutual admittance depends

solely on the angle θ_0. We placed the primary element in NP, but the results for the mutual admittance are perfectly valid for any two arbitrary elements on the sphere, and θ_0 the angle spanned between the two elements, along the great arc connecting them. The angles ϕ_0 and ξ_0 only act on the phase of the obtained mutual admittances between the modes of two circularly polarised waveguides. There is some arbitrariness in the choice of the phase reference for the elements, and the angles chosen for ϕ_0 and ξ_0 must follow the adopted phase reference. We will come back to this issue in section 5.3.4.

For the calculation of the Jacobi polynomials a recurrence over its degree can be used[106]. This is especially attractive as the polynomial in any case needs to be evaluated over all degrees in the summation of (4.14):

$$\begin{aligned}
P_0^{(\alpha,\beta)}(x) &= 1 \\
P_1^{(\alpha,\beta)}(x) &= \frac{1}{2}\left(2(\alpha+1)+(\alpha+\beta+2)(x-1)\right) \\
P_{l+1}^{(\alpha,\beta)}(x) &= \frac{C1}{C3}xP_l^{(\alpha,\beta)}(x)+\frac{C2}{C3}P_{l-1}^{(\alpha,\beta)}(x) \\
C1 &= (2l+\alpha+\beta-1)\left((\alpha^2-\beta^2)(2l+\alpha+\beta)(2l+\alpha+\beta-2)\right) \\
C2 &= -2(l-1+\alpha)(l-1+\beta)(2l+\alpha+\beta) \\
C3 &= 2l(l+\alpha+\beta)(2l-2+\alpha+\beta)
\end{aligned}$$

4.4 Results for Mutual Coupling Coefficients

Having obtained the mutual admittances, we can convert these into scattering coefficients S_{ij}, also called mutual coupling coefficients. The interpretation of the radiation environment of the array antenna as a scattering network will be given more attention in chapter 5. Here we will just consider isolated mutual couplings, the scattering from one mode in one aperture to any other mode in any other aperture, with no other modes or apertures present than the considered pair. In chapter 5 this information will be embedded in a larger model that describes the complete array antenna. With only two modes in the problem, whether or not in the same aperture, a (2×2) admittance matrix can be calculated. The admittance matrix $\mathbf{Y}$ is converted to a scattering matrix $\mathbf{S}$ with the formula:

$$\mathbf{S} = (\mathbf{I_n} - \mathbf{Y})(\mathbf{I_n} + \mathbf{Y})^{-1} \tag{4.19}$$

In this matrix equation, S_{12} is the scattering coefficient between the two modes. Figures 4.7 and 4.8 show the obtained S_{12} scattering coefficient, plotted as a function of the distance between the apertures.

Figure 4.7 shows the effect of different curvatures of the spherical surface. The magnitude of the scattering coefficient between the TE_{11}-mode in one aperture and the TE_{11}-mode in another aperture is plotted against the geodesic distance between the two apertures. The first aperture was placed at NP and the second was subject to a variable Euler rotation around the y-axis over an angle θ. The other two Euler rotations would not change the magnitude of the

scattering between circularly polarised modes, however they do act on its phase. Three curves are plotted, for three different radii of the sphere. To see the effect of the curvature, and to compare apples with apples, the geodesic distance $gd = R\theta$ was used rather than the angle θ to plot the scattering coefficients against. From the figure it is apparent that smaller radii, or else a higher curvature of the sphere, decreases the mutual coupling as would be expected. The apertures had a diameter of $6cm$ and the frequency was 3 GHz. Spherical Transmission Line modes were used up to degree 500, which is more than sufficient in this case as no difference is observed when 2000 modes are used. There is an interesting ripple when the apertures are nearly in diametrical positions. This is a physical effect (it has nothing to do with simulation accuracies) and is caused by interferences between waves that travel around the sphere and meet again at the opposite side. A curved groundplane reduces the mutual coupling significantly for apertures that are wider apart. For apertures that are near to each other (within one wavelength), the effect is negligible.

Figure 4.8 shows how a TE_{11} mode in one apertures scatters to different circular modes in another aperture. Also in this figure the scattering coefficient between two apertures is plotted, only this time for a TE_{11} mode in the first aperture to several different circular modes in the other aperture. The radius of the sphere is kept constant at $0.4m$. Note that the admittances from which the scattering coefficient was calculated were not normalised to the self-admittance of the modes in the circular waveguide. Stated otherwise, the propagation in the circular waveguide is not taken into account here. In chapter 5 the effects on the overall system, including propagation into the feeding waveguide will be taken into account. Important here is to see that the coupling into other circular modes is comparable to the coupling into the corresponding mode in the other aperture. This is important because, while the effect on the free excitation will be small, the coupling into these other modes will have an effect on the realised gain pattern of an aperture. Indeed, evanescence of higher order modes in the feeding waveguide does not necessarily mean that this mode cannot re-radiate into the free space surrounding the antenna. Hence, it is not justified to consider only the modes that can propagate in the feeding waveguides (as in [100]). For definitions of the free excitation and the realised gain pattern, please refer to chapter 5.

Also noteworthy is the much stronger coupling in the TM_{01} mode. The reason for this was already given in chapter 3: TM_{01} modes radiate laterally and therefore excite surface waves over the spherical surface. In this case of a Perfectly Electric Conductor (PEC) metal sphere without dielectric coating, the surface wave takes the form of a creeping wave. Although creeping waves do attenuate with increasing curvature[1], they are responsible for the increased mutual coupling shown in this plot.

The results have not been verified with numerical simulations, as it is difficult to build an accurate and representative numerical model and solve it with an electromagnetic solver. It would take a very long time to solve a large and complex structure with an FDTD solver and in addition programs such as CST for instance require the waveguide ports to be parallel with one of three orthogonal planes. However, all the constituent parts of the algorithm have been thoroughly tested and verified. The rotation of the field has been demonstrated to work and the calculation of the expansion has been numerically verified in the previous chapter. As the algorithm as a

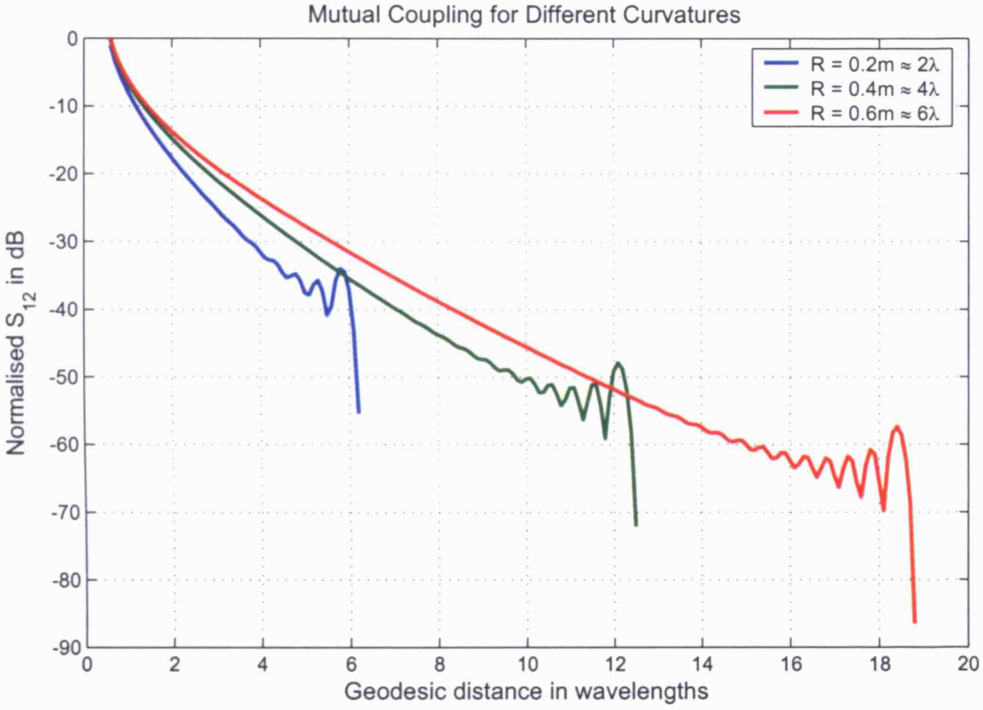


Figure 4.7: The magnitude of the scattering coefficient (in dB, normalised to the maximum value) between the TE_{11}-mode in one aperture and the TE_{11}-mode in another aperture is plotted against the geodesic distance between the two apertures, for different curvatures of the sphere.

whole has not been numerically verified, it would be reassuring to do this empirically, which will be further discussed in chapter 9.

4.5 Summary and Conclusions

In this chapter expressions were obtained for the mutual coupling between a pair of antenna elements on a conducting sphere, in the form of the mutual admittance, defined in equation (4.9). The use of SMT shows it full potential in the implementation of the field rotations in section 4.2. This concept of field rotation will repeatedly be used in other parts of this thesis.

The expressions derived in this chapter for the mutual coupling are not sufficient to analyse a spherical array antenna with more than two radiating elements, where there are multiple interactions between the elements. To accept the second point of the hypothesis in section 1.3.1, a model is necessary that can capture spherical array antennas of any complexity. The derivations in this chapter will be used in chapter 5 to take on spherical array antennas with an arbitrary number of elements, placed in arbitrary positions.

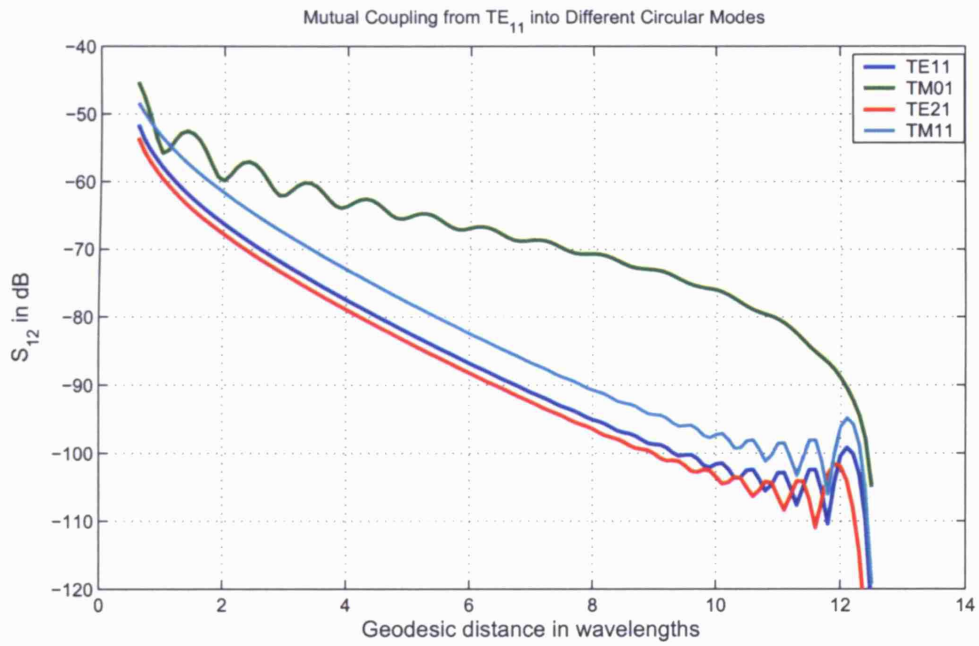


Figure 4.8: The scattering coefficient between two apertures is plotted, for a TE_{11} mode in the first aperture to several different circular modes in the other aperture (see legend).

Chapter 5

Scattering Matrix Description of the Spherical Array Antenna

5.1 Introduction

In chapter 3 the radiation pattern was obtained for a single array element on a large conducting sphere. A spherical array antenna combines a large number of these elementary radiators and they all contribute to the radiation pattern. Part of the radiated energy from each element is however coupled into its neighbouring elements, which affects the radiation pattern of the array antenna. In chapter 4 the isolated external coupling between two radiating elements on the sphere was quantified. Isolated, because only one pair of elements was considered at a time, and external, because only the scattering in the radiation environment was considered.

The complete array antenna is a much more complex environment than the simple cases of one single or one pair of radiating elements. For instance, the portion of the radiated energy from one element that couples into some other element will partially be re-radiated, and can couple in turn into other elements. Another portion of the coupled energy will be guided into the receiving channel, where further scattering between the receiving channels of different elements may occur. Some of the energy can even loop back and be re-radiated by the first element. The analysis of this complex system of internal and external scattering and multiple reflections is the subject of this chapter.

In this chapter, the results obtained in chapters 3 and 4 will be incorporated in a powerful model, called the scattering matrix description. The scattering matrix description is based on the work of Borgiotti for the analysis of conformal array antennas[2][107]. The scattering matrix description describes the array antenna as two cascaded networks, one that describes the feed network and one that describes the radiation environment. The methodology of microwave networks can be applied to describe the interactions between the networks with matrix operations.

In accordance with the general approach taken throughout this thesis, the model obtained completely describes the array antenna and its radiation characteristics based on rigorous modal methods. By implementing the model in software, fast and accurate simulations can be run that predict the behaviour of an electrically large and complex spherical array antenna.

5.2 Top-Level description

The description of a conformal array antenna naturally divides into two parts. The first part is the determination of the fields or currents on the surface of the array, as a function of the applied excitation signals. The second part is the determination of the radiated field, as a function of the fields or currents on the array surface.

The second part of the problem is a common problem in electromagnetics and several methods exist to compute a solution[77][1]. For the spherical array antenna however, the problem is essentially the same as the problem solved in chapter 3 for a single element. The modal method can be applied to a superposition of known fields or currents on the surface of the sphere. The remainder of this chapter will concentrate on the first part of the problem.

Throughout this chapter, the example of an array antenna of circular apertures is used, and the field on the surface of the sphere is given by a set of modal aperture voltages. For other types of element, the formalism proceeds in exactly the same way. As explained in the introduction to this chapter, scattering occurs outside the sphere and inside the feeding network. To tackle the overall scattering problem, it is further broken down in two easier-to-handle subproblems:

- Internal scattering in the feed network (FN)
- External scattering between the apertures outside the body of the array, described by a radiation network (RN).

With reference to figure 5.2, the two different networks can be identified, represented by the blue vertical blocks. The block on the left is the feed network (FN) that accounts for all internal coupling effects. The network on the right is the radiation network (RN). Each network is mathematically represented by a scattering matrix and described below.

5.2.1 The Feed Network (FN)

The feeding structure for the array antenna is represented as an RF network, with a scattering matrix describing the reflections and transmissions between the input and output ports. The meaning of input and output ports is arbitrary - here we refer to the ports that connect to the array elements as output ports and those that connect to the upstream feed chain as input ports.

The network that describes the internal scattering is obviously dependent on the type of RF feed chain, and it can be made to cover as much detail as required. In our example of a waveguide-fed array of apertures, the input ports could be the terminals to a circular waveguide and the output ports could each represent a propagating circular mode in a circular waveguide. In a more elaborate example, the network could be broken down into a series of cascaded networks, each describing a part of the feed chain. Extra networks can be bolted on as required. For instance, a network with a single input representing a power divider, could connect a number of output ports to the input ports of the feeding waveguides. To avoid an excessive level of detail though, we will only make the assumption of a set of circular apertures fed by circular waveguides. Rather than taking any assumptions on the upstream implementation of the feed chain, we remove the input ports from the network and add an excitation vector in parallel to

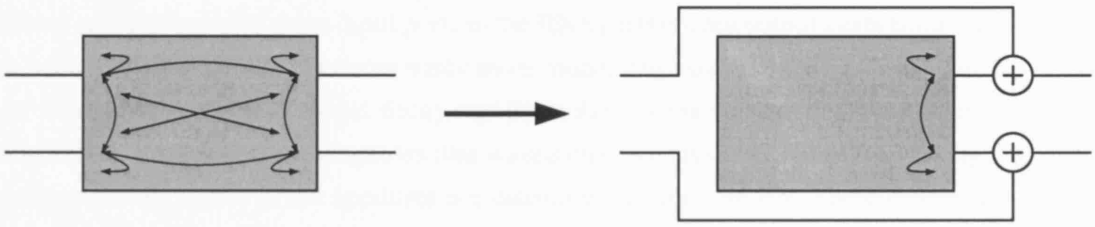


Figure 5.1: The picture shows schematically how the inputs to the Feed Network are replaced with an equivalent parallel excitation vector, while the scattering between the outputs is retained. This makes an implementation of the model simpler, but also more flexible to use with different RF feed chains.

the outputs of the network. This makes an implementation of the model simpler, but also more flexible to use with different RF feed chains. The principle is schematically illustrated in figure 5.1.

In the circular waveguide fed architecture, each waveguide has typically one propagating fundamental mode, whereas the higher order modes are evanescent. There are in fact two degenerate propagating modes, one for each polarisation. The two polarisations may each have a separate terminal to connect to the feed chain. The excitation vector is therefore of the same length as the number of waveguides times the number of fundamental modes times 2.

The FN represents all the internal scattering that occurs in the feed system. Often it is possible to make some simplifying assumptions and complete the missing information from the FN scattering matrix with measurements. For the case-study of the spherical array fed by circular waveguides, it could for instance be assumed that scattering only occurs between modes that exist within the same waveguide.

5.2.2 The Radiation Network (RN)

Also the radiation environment can be represented by a similar network, with one port for each circular mode that exists in any of the apertures. For absolute gain calculations, an extra port can be added for radiation to the far-field.

In the RN shown in figure 5.2, the black diamonds represent the ports. The arrows represent the scattering that occurs between the ports. Each arrow corresponds to one complex scattering coefficient in the RN scattering matrix.

5.2.3 Complete Description with Network Interactions

The complete system of feeding waveguides, array surface, apertures and radiation environment is included in a scattering matrix description, which is schematically depicted in figure 5.2. In the figure a separate transmission line is shown for each circular mode that is included in the analysis. The horizontal yellow blocks group the modal transmission lines according to the feeding waveguide in which they exist. As we have excluded scattering between modes of different waveguides, the FN contains a number of separate scattering environments. The scattering matrix of the FN takes a block-diagonal form, and each of these blocks is represented by a white square in the schematic of figure 5.2. In our implementation, each of these squares is further simplified as depicted in figure 5.1.

There are clearly many more input ports to the RN than there are output ports from the FN. That is because in the circular apertures many more modes can exist that are evanescent in the circular waveguides. As these modes decay rapidly in the feeding waveguides, they cannot transmit any power through these waveguides (the waveguides are presumed to be long enough) and the corresponding modes in the apertures are disconnected from the FN. These evanescent modes however do exist in the apertures and can contribute to the scattering in the RN. Therefore they should be included in the analysis of the array antenna. In the model, the modal transmission lines that correspond with non-propagating modes are terminated with their characteristic impedance, which automatically takes care of their evanescent character.

The system is excited by a set of signals, contained in the excitation vector $\mathbf{F}$. The apertures are fed by a set of signals $\mathbf{a}$, which is called the free excitation of the array[2]. Consider the hypothetical situation where all but one of the input ports to the RN are disconnected from the FN and match terminated. The remaining radiating element is driven with a signal of 1V. Stated otherwise, the free excitation $\mathbf{a}$ is set to a vector of zeros, except for one entry which is equal to 1. The obtained radiation pattern is called the element Realised Gain Pattern (RGP) (if more than one mode or polarisation is used, there will be a realised gain pattern for each of them). The Realised Gain Pattern (RGP) is different from the isolated pattern (obtained if only one radiating element were present) due to the external scattering in the RN. By its definition, the interactions between the FN and the RN cannot have any influence on the realised gain pattern. The array RGP, is then found as the linear superposition of the element RGPs, according to the coefficients in the free excitation.

From the scattering matrices of the RN and the FN together, it is possible to work out the relation between the free excitation $\mathbf{a}$ and the applied excitation $\mathbf{F}$. Hence, the knowledge of the realised gain patterns, and the scattering matrices for the RN and the FN is sufficient to completely characterise the array antenna.

To work out the free excitation from the excitation vector, it is necessary to recognise that it will be affected by the simultaneous action of:

- Scattering between the array elements in the radiation environment
- Scattering between the feedlines and terminals inside the feed network
- Multiple reflections that occur between the array elements and the feed network.

The three processes are a necessary condition for the free excitation to be affected by mutual coupling effects. Without reflections from the feed network back into the radiation environment for instance, no mutual coupling effects exist. Note that the scattering in the radiation network will still have an impact on the realised gain patterns.

The presented model is very generic and can in principle be adapted to facilitate any conformal array configuration. The focus here lies on spherical array antennas. In section 5.3 the details of the scattering matrix description will be developed for spherical array antennas. In sections 5.4 and 5.5 the formalism will be used more specifically to analyse and synthesise a spherical array of circular apertures, fed by circular waveguides.

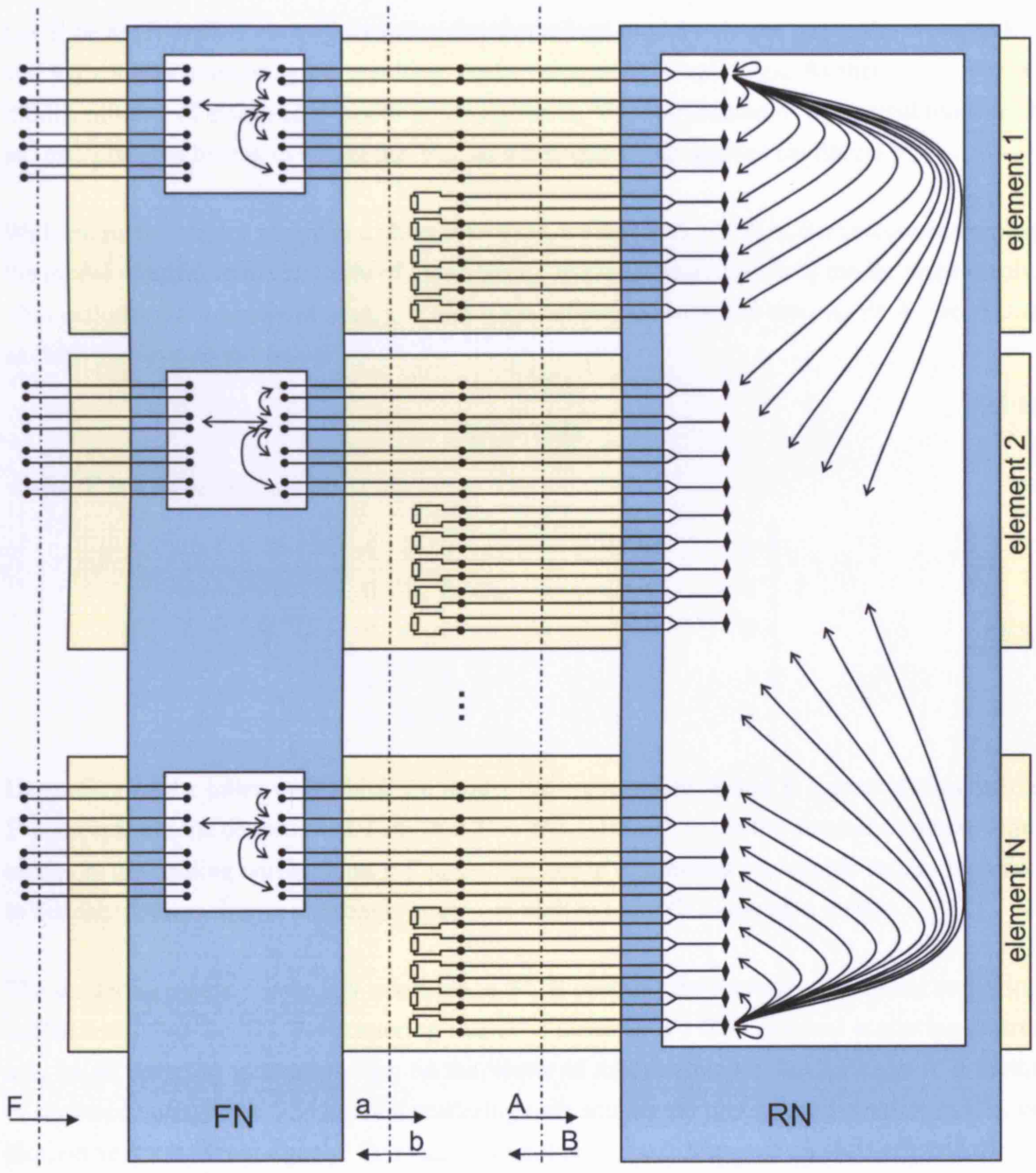


Figure 5.2: The scattering matrix description of an array antenna is shown.The array is represented by two scattering networks FN and RN. RN embodies the scattering outside the spherical surface of the array, FN the internal coupling effects. The internal coupling effects consist here solely of effects at the terminal side of the feeding waveguides, but it is easy to include other possible couplings in the scattering matrix.

5.3 The Scattering Matrix Description Unravelled

5.3.1 Aperture Voltages

Let $\mathbf{F}$ be an excitation vector containing the normalised modal voltages fed in the waveguides, and V_{ap} a vector containing the resulting modal voltages in all apertures. As there may exist an infinite number of evanescent modes in the apertures, $\mathbf{V_{ap}}$ is truncated to a practical number of modes. The goal here is to obtain the $\mathbf{V_{ap}}$, as a function of the applied excitation $\mathbf{F}$.

With the surface of the sphere as a phase reference, we define $\mathbf{A}$ and $\mathbf{B}$ as the vectors containing the modal voltages in the aperture of the outgoing and incoming (reflected) modes respectively. This includes the evanescent modes. $\mathbf{a}$ and $\mathbf{b}$ are defined as truncated versions of $\mathbf{A}$ and $\mathbf{B}$ that exclude the evanescent modes:

$$\begin{aligned} \mathbf{a} &= \mathbf{TA} \\ \mathbf{b} &= \mathbf{TB} \end{aligned} \tag{5.1}$$

where $\mathbf{T}$ is a truncation matrix:

$$\mathbf{T} = \left[\begin{array}{ccccc|ccc} 1 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & \dots & 0 \end{array}\right] \tag{5.2}$$

Depending on the order with which the modes are organised in $\mathbf{A}$ and $\mathbf{B}$ a permutation matrix $\mathbf{P}$ is included in the definition of $\mathbf{T}$, so that $\mathbf{T} = \mathbf{P}[\mathbf{I}_{m1}|\mathbf{O}_{m2}]$. $m1$ is the number of propagating modes in the feeding waveguides, $m2$ is the number of evanescent waveguide modes included in the calculations, $\mathbf{I}_{m1}$ is an identity matrix of rank $m1$ and $\mathbf{O}_{m2}$ the zero matrix.

The scattering matrix for the RN is written as $\mathbf{S^E}$. It contains the normalised external scattering coefficients. For the FN, the scattering matrix is noted as $\mathbf{S^I}$. The external scattering matrix acts on all outgoing modes, or else, on the vector of modal voltages $\mathbf{A}$. It results in a vector of scattered voltages $\mathbf{B}$. The internal scattering only acts on the propagating circular modes, or else, on vector $\mathbf{b}$. From figure 5.2, it can be seen that:

$$\begin{aligned} \mathbf{B} &= \mathbf{S^E A} \\ \mathbf{a} &= \mathbf{F} + \mathbf{S^I b} \end{aligned} \tag{5.3}$$

Using (5.1) and (5.3) we can derive an expression for $\mathbf{A}$ as a function of $\mathbf{F}$:

$$\begin{aligned} \mathbf{TA} &= \mathbf{F} + \mathbf{S^I T S^E A} \\ \mathbf{A} &= (\mathbf{T} - \mathbf{S^I T S^E})^{-1}\mathbf{F} \end{aligned} \tag{5.4}$$

Note that if $\mathbf{S^I} = 0$ we have $\mathbf{TA} = \mathbf{a} = \mathbf{F}$.

Finally, The aperture voltages are obtained from the normalised voltages in $\mathbf{A}$ and $\mathbf{B}$ as:

$$\begin{aligned} \mathbf{V}_{ap} &= \sqrt{\mathbf{Z}}(\mathbf{A}+\mathbf{B}) \\ \mathbf{V}_{ap} &= \sqrt{\mathbf{Z}}(\mathbf{A}+\mathbf{S}^{E}\mathbf{A}) \end{aligned} \tag{5.5}$$

where $\mathbf{Z}$ is a diagonal matrix containing the modal impedances in the feeding waveguides.

5.3.2 Array Element Patterns vs Element RGPs

Having obtained the aperture voltages $\mathbf{V}_{ap}$ as a function of the applied excitation $\mathbf{F}$, the array RGP can be obtained. Attention must be drawn to the fact that nowhere we have used the element RGP as an intermediate step towards the array RGP. Instead the aperture voltages were obtained immediately from the excitation vector, and the array RGP is obtained from the aperture voltages. This approach is entirely equivalent and is in this case more straightforward. A situation where realised gain patterns would be useful, is when these can be obtained by measuring the radiation pattern of one active array element and match terminating all other elements. The behaviour of the completely active array antenna can then be inferred from the realised gain pattern. This is especially useful if the RGPs of each element have enough similarity to be modelled as identical.

It is still useful to see how the radiation pattern of a single element is affected by the proximity of and the mutual coupling effects with the other array elements. To see this, we will use the Array Element Pattern (AEP). We define this as the radiation pattern of an element that is fed with a signal of 1V, while all other elements are fed with 0V. There is a subtle difference between the definition of the element RGP and the AEP. The element RGP uses the free excitation, whereas the AEP uses the excitation vector. Hence, the AEP includes the full effects of mutual coupling, whereas the element RGP only includes the external scattering.

The AEP is obtained in the same way as the array RGP, with an excitation vector $\mathbf{F}$ which has one element equal to 1 and all others equal to 0. The first step is to establish the aperture voltages $\mathbf{V}_{ap}$. In equation (2.56), the isolated element patterns were obtained for the case of a circular aperture in a conducting sphere. Combining these according to the element position and the voltages in (5.5), the array RGP or the AEP can be obtained. This will be illustrated in section 5.4.

5.3.3 Scattering Coefficients

In the previous section, it was assumed that the scattering matrices for the RN and FN are known. The problem of obtaining a complete scattering description of the array has been narrowed down to the determination of the scattering matrices $\mathbf{S}^{I}$ and $\mathbf{S}^{E}$.

The internal scattering matrix depends on the exact structure of the feeding waveguides and the probes at the terminal end, and can best be obtained by measurements. Often it is reasonable to assume all feeding waveguides are identical, and no coupling exists between them (as shown in figure 5.2). $\mathbf{S}^{I}$ then takes the form of a block-matrix with all blocks identical.

The external scattering coefficients in $\mathbf{S}^{E}$ are obtained from the normalised admittance

matrix[108]:

$$\mathbf{S}^{\mathrm{E}} = \frac{\mathbf{I_m} - \mathbf{y}}{\mathbf{I_m} + \mathbf{y}} \tag{5.6}$$

where $\mathbf{I_m}$ is the identity matrix of rank $m = m_1 + m_2$ and $\mathbf{y}$ is the normalised admittance matrix. The elements y_{ij} in y relate to the un-normalised elements Y_{ij} as:

$$y_{ij} = \frac{Y_{ij}}{Y_i} \tag{5.7}$$

where the Y_i are the characteristic modal admittances. This is also true for the evanescent modes that are terminated with their imaginary characteristic admittance (as illustrated in figure 5.2). On page 120 a remark was made regarding the calculation of the self-admittances. These should be calculated directly from equation (4.16).

All that is missing now to complete the scattering matrix description are the mutual admittances between the apertures in the sphere. These were obtained in chapter 4 and given in equation (4.14).

5.3.4 Polarisation and Phase Reference

In section 4.3 an expression was derived for the mutual admittances between any two modes that may exist in any two circular apertures on the sphere. The solution for the mutual admittances was a function of the Euler angles θ, ϕ and ξ between the apertures in which the modes exist. With circularly polarised modes, the magnitude of the mutual admittances is only a function of θ, while the angles ϕ and ξ act on their phase. It is not immediately clear what value needs to be used for the difference in tilt angle ξ between any two apertures, as this would depend some phase reference. In this section we will remove the arbitrariness that exists on the phase of the mutual admittances by properly defining and adopting a phase reference for all elements on the sphere.

Choosing a phase reference for all circularly polarised apertures on the sphere is the same as associating a reference vector with each aperture that corresponds to the field orientation in the aperture at time zero. This is equivalent to the definition of a linear polarisation over the surface of the sphere. Once that each aperture has been associated with such a reference vector, the Euler angles that are needed in the calculation of the mutual admittance between aperture 1 and aperture 2, define the rotation that would align the reference vector of aperture 2 with the reference vector of aperture 1.

The determination of the rotation between the reference vectors of any two apertures will be the subject of section 5.3.5. What is left is the determination of a reference vector for each aperture. This problem is equivalent to choosing a polarisation definition for a vector-field that is tangential to a spherical surface. The latter problem will be dealt with in more detail in section 6.1. In that section some more attention is given to the definition of circular polarisation we have adopted in section 2.4.3 for the polarisation of the Far-Field. The equivalent linear polarisation is Ludwig's third definition for polarisation, for which the co-polar basis vectors are given as:

$$\vec{r} = \hat{e}_\theta \cos\phi - \hat{e}_\phi \sin\phi \tag{5.8}$$

for a linear polarisation aligned with the x-axis. Also in section 2.4.3 it is explained why that particular definition for polarisation is considered optimal to use with spherical array antennas. Adopting the same polarisation on the surface of the array antenna to determine the phase references for the circularly polarised elements has the following additional advantages:

- The polarisation definition for the Far-Field was used to define circularly polarised basis functions in which the Far-Field was expanded. In section 7.1 it will be shown how the Far-Field expansion can be backprojected to an expansion for the tangential Near-Field with the same polarisation. Having the array elements aligned with this polarisation simplifies the description of the sampling process, as will be seen in section 7.2.2.

- If the radiation pattern from the array antenna is to achieve maximum directivity in the direction of the positive z-axis, then the radiation from the array elements needs to add up coherently in that direction. The chosen polarisation aligns the reference vectors along the fieldlines depicted in figure 2.4. After aligning the reference phasors of the circularly polarised elements with these fieldlines, the phases and amplitudes are further corrected for curvature and path delay by the scalar modal functions.

- If the maximum directivity is to be obtained in any other direction, this can be achieved by a linear operation on the modal expansion coefficients. The phase of the radiating elements will automatically be adjusted so that they align in the direction of maximum directivity. There is no need to change the reference phase. This can only be done with circularly polarised elements. If the elements were linearly polarised, it would be necessary to have dual polarised elements, with double the amount of controls, so that the tilt angle of the polarisation axis can be changed.

From the discussion above and the arguments made in section 6.1 we decide to phase the array elements as follows:

- Define a right-hand coordinate system in the origin of the sphere.

- Place the circularly polarised elements on the sphere, according to a convenient distribution. this will be dealt with in chapter 8.

- With the spherical coordinates defined as usual in the coordinate system from above, apply a correction to the phase of all elements so that they initially line up with the vectors defined in equation (5.8). These vectors are the phasors of the tangential field in the immediate vicinity of each radiating element at time zero, with an element excitation of zero phase and unit amplitude.

- The element excitation is superimposed on the array elements and bears the phase and amplitudes to synthesise some array pattern.

5.3.5 Obtaining the Rotation Angles

With one aperture at NP, the Euler angles to be used in the equation for the mutual admittance with a second element are simply the position angles θ, ϕ and ξ that would align the reference vector at NP with the reference vector of the second aperture. To get the Euler angles between any two apertures, given by a distribution as a set of (θ, ϕ) pairs, two more derivations are necessary:

- The third Euler angle $\xi(\theta, \phi)$ that corresponds with the phase reference obtained in section 5.3.4
- The Euler angles of the rotation that would align the reference vector of an aperture with Euler angles $(\xi_1, \theta_1, \phi_1)$ to the reference vector of an aperture with Euler angles $(\xi_2, \theta_2, \phi_2)$

The first point is easily solved. If we start with the reference vector at NP and rotate it around the y-axis in the direction $\phi = 0$ over some angle θ, the tangential reference vector is still pointing straight to the southpole. If we rotate the reference vector around the z-axis over $\phi = \frac{\pi}{2}$ radians, it should become horizontal to the equatorial plane by rotating it over an angle $\xi = \frac{-\pi}{2}$. If we rotate it further to $\phi = \pi$, it should point straight up to NP, which is done by rotating it over $\xi = \frac{-\pi}{2}$. With reference to figure 5.3 We can more formally derive the expression for ξ from:

$$\begin{aligned} \hat{e}_\theta &= \cos\phi \\ \hat{e}_\phi &= -\sin\phi \\ \tan\xi &= -\frac{\sin\phi}{\cos\phi} \\ \xi &= -\phi \end{aligned} \tag{5.9}$$

The second point is solved with the aid of rotation matrices. For the two apertures, the Euler angles are converted to rotation matrices R_1 and R_2 with the formula:

$$R_i = R_z(\phi_i) R_y(\theta_i) R_z(\xi_i) \tag{5.10}$$

Where the rotation matrices $R_y(\alpha)$ and $R_z(\beta)$ are defined as:

$$\begin{aligned} R_y(\alpha) &= \begin{bmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{bmatrix} \\ R_z(\beta) &= \begin{bmatrix} \cos\beta & -\sin\beta & 0 \\ \sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned} \tag{5.11}$$

These rotation matrices would align the reference vector at NP with the reference vectors of the apertures, as shown in figure 5.4. Let P_1 be the reference vector in aperture 1, P_2 be the

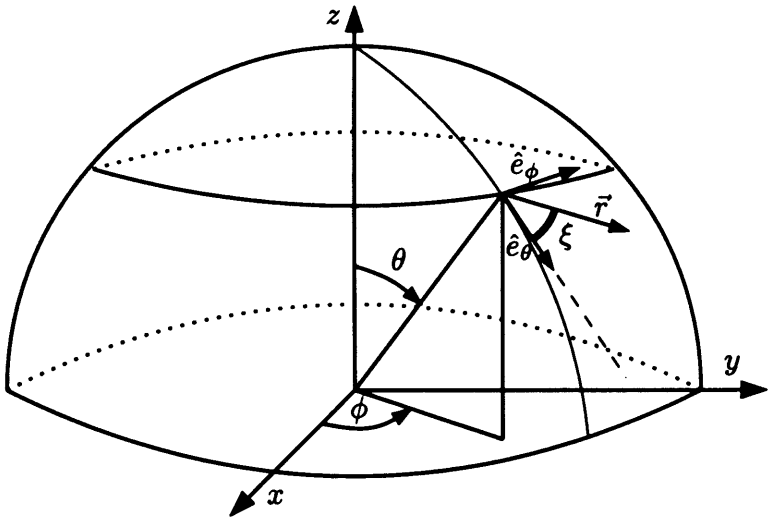


Figure 5.3: The Euler angle ξ that sets the orientation of the reference vector in the tangential plane for each aperture is obtained from the Euler angle ϕ in accordance with the definition of $\vec{r}$.

reference vector in aperture 2 and NP the reference vector at $z = 1$. We have:

$$\begin{aligned} P_1 &= R_1 \cdot \mathrm{NP} \\ P_2 &= R_2 \cdot \mathrm{NP} \end{aligned} \tag{5.12}$$

The coordinate system is rotated so that P_1 becomes aligned with NP:

$$\begin{aligned} P_1' = \mathrm{NP} &= R_1^{\mathrm{T}} P_1 \\ P_2' &= R_1^{\mathrm{T}} P_2 \end{aligned} \tag{5.13}$$

Combining the previous equations we get:

$$P_2' = \underbrace{R_1^{\mathrm{T}} R_2}_{R_{12}} P_1' \tag{5.14}$$

So it can be concluded that $R_{12} = R_1^{\mathrm{T}} R_2$ is the rotation matrix from NP to the second aperture, in a rotated coordinate system in which the first aperture has been aligned with NP. Note that in the original coordinate system the rotation matrix from P_1 to P_2 is given by $R_{21} = R_2 R_1^{\mathrm{T}}$ (and note that this is not the inverse of R_{12}).

Finally, the solution for the Euler angles between the two apertures is obtained by retrieving the Euler angles from the rotation matrix R_{12}:

$$R_{12} = \begin{bmatrix} \cos\theta\cos\phi\cos\xi - \sin\phi\sin\xi & -\cos\theta\cos\phi\sin\xi - \sin\phi\cos\xi & \sin\theta\cos\phi \\ \cos\theta\sin\phi\cos\xi + \cos\phi\sin\xi & -\cos\theta\sin\phi\sin\xi + \cos\phi\cos\xi & \sin\theta\sin\phi \\ -\sin\theta\cos\xi & \sin\theta\sin\xi & \cos\theta \end{bmatrix} \tag{5.15}$$

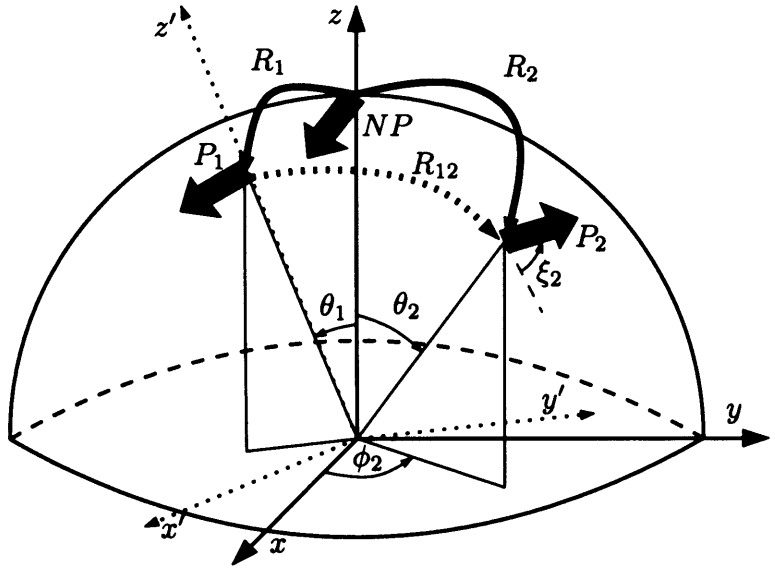


Figure 5.4: The figure illustrates the geometrical problem of obtaining the rotation from P_1 to P_2, after rotating the coordinate system so that P_1 is aligned with NP. The rotation R_{12} is defined relative to the coordinate system $x'y'z'$. In this coordinate system P_1 is at NP position. The positions of P_1 and P_2 are known in the coordinate system xyz.

5.4 Spherical Array Analysis

In this section we will use the formalism developed in the previous sections to analyse a spherical array of waveguide-fed circular aperture antennas. The objective of the analysis is to obtain the Far-Field and its properties of the spherical array, for a given excitation $\mathbf{F}$.

The array consists of a conducting spherical shell of 80 cm diameter. The circular waveguides have a diameter of 3 cm. Operating at a 3 GHz, only the two fundamental TE_{11} modes (one for each polarisation) propagate through the waveguides. We will assume that the fields in the aperture are sufficiently well represented by the first $M_2 = 10$ circular modes. These are the TE_{11},TE_{-11},TM_{01},TE_{21}, TE_{-21},TE_{01},TM_{11},TM_{-11}, TE_{31} and TE_{-31} modes, where negative values for n indicate LHCP waves. The length of the vectors A, B and V_{ap} becomes $N \times M_2 = 10N$. The length of the a, b and F vectors becomes $N \times M_1 = 2N$. Let the operator $\mathbf{B}_x\{B\}$ be a constructor for a block-diagonal matrix, with x identical blocks B placed on its diagonal and let $O_{n \times m}$ be the $(n \times m)$ zero-matrix. The truncation matrix T can then be written as:

$$\mathbf{K} = [\mathbf{I}_{\mathbf{M}_1} \mid \mathbf{O}_{M_1 \times \mathbf{M}_2 - \mathbf{M}_1}]$$
$$\mathbf{T} = \mathbf{B}_N\{\mathbf{K}\}$$

We will now demonstrate how $\mathbf{S}^{\mathrm{I}}$ is defined. Looking into the waveguide from the aperture, we estimate an order-of-magnitude figure of -20dB for reflected power of the RHCP TE_{11} mode, and the same is true for the LHCP wave. Furthermore, we estimate -40dB of the power is reflected as a cross-polarised wave, again for both polarisations. We will further assume that the lengths of the waveguides are an integer multiple of the wavelength. Hence we can construct

$\mathbf{S}^{\mathrm{I}}$ as follows:

$$
\begin{aligned}
s_{11} &= s_{22} = 0.1 \\
s_{21} &= s_{12} = 0.01 \\
K &= \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} \\
\mathbf{S}^{\mathrm{I}} &= \mathbf{B}_N\{\mathbf{K}\}
\end{aligned}
$$

A suitable excitation has to be defined in the vector $\mathbf{F}$. This can for instance be a vector that is zero everywhere, except for a 1 at the position of a fundamental mode for one element. Another possible excitation is the one that optimises the directivity in some direction. Retrieving the necessary excitation $\mathbf{F}$ to obtain a certain far-field pattern is the reverse problem from what is discussed in this section and will be discussed in section 5.5.

To build up the scattering matrix model, all elements of the admittance matrix need to be calculated, from which $\mathbf{S}^{\mathrm{E}}$ can be obtained. The circular modes are expanded in spherical modes up to a sufficiently high degree, for which we choose LMAX $= 2000$. The result of the scattering matrix calculations is a set of modal aperture voltages V_{ap}. The modal aperture voltages are subsequently used to obtain the far-field pattern of the array antenna. This is achieved by using the spherical expansion of each circular mode at NP and rotating that expansion to the correct position and orientation of the corresponding aperture. All such obtained expansions are then linearly combined according to their aperture voltages. To rotate the spherical expansions, the rotation matrices are calculated based on the Euler angles of each element. Previously the algorithm has calculated only those elements of the rotation matrices that were necessary to calculate the mutual admittances between circular modes in different apertures. It still makes sense not to calculate the complete rotation matrices in that step because in the calculations for the mutual admittances the transmission line version of the spherical expansions were used up to a degree LMAX $= 2000$. The expansions to obtain the far-field are only calculated and rotated up to the cut-off degree $l_m = kR + 10 \approx 34$.

Figure 5.5 illustrates the radiation pattern of a pure TE_{11} mode in an aperture at NP, without any internal or external scattering. The configuration is schematically shown, with a yellow patch showing the position and size of the aperture on the sphere. The co-polar and cross-polar radiation patterns are shown on the second row, and cuts through the patterns are shown on the third row of plots on a logarithmic scale. The plot in the top-right corner shows the axial ratio. Figure 5.6 contains plots that show the same information, but for a single waveguide where two propagating modes are considered, one for each polarisation of the TE_{11} mode. External and internal scattering occurs between the two modes, based on the assumptions in the previous paragraphs. A slight perturbation of the axial ratio is visible in the boresight direction. Figures 5.7 to 5.9 show again the same kind of plots, but for an increasing number of waveguide-fed apertures. It is clear that although there is only one active aperture for each of the simulated configurations, the mutual coupling effects have a severe effect on the radiated pattern.

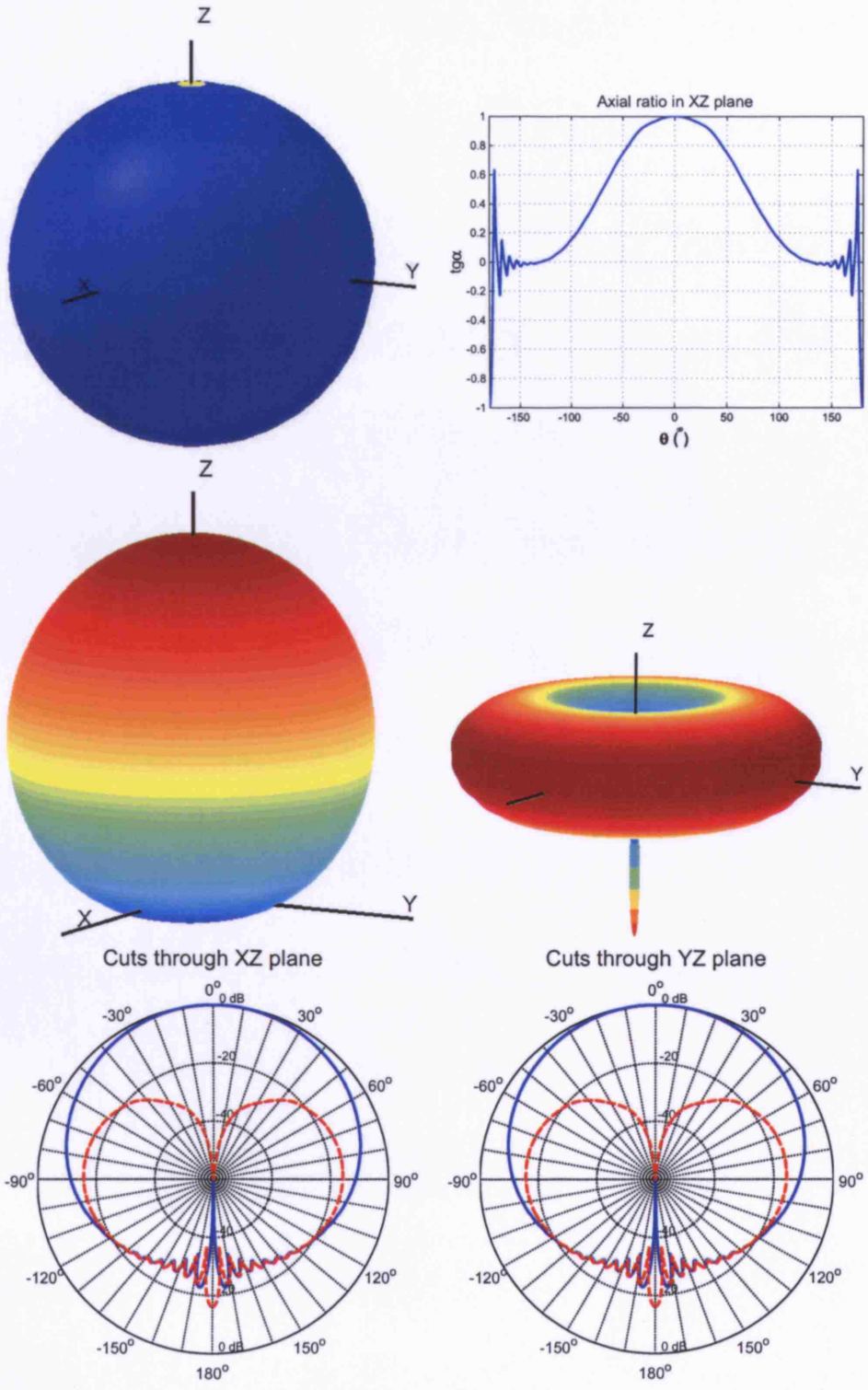


Figure 5.5: The co- and cross-polarised far-field element patterns are shown for a single radiating element at NP. Only the fundamental RHCP TE_{11} mode exists in the waveguide with no coupling into the LHCP mode.

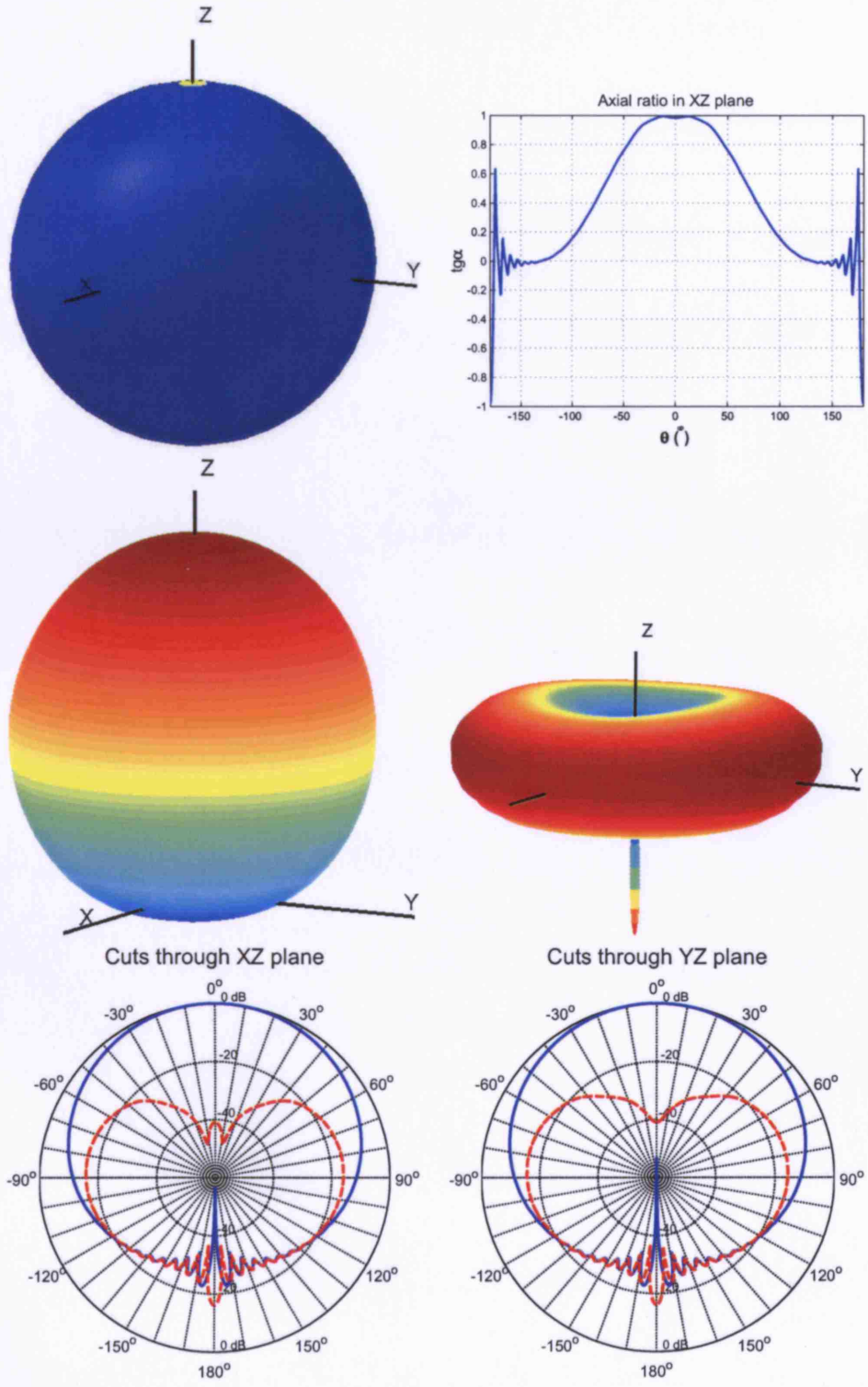


Figure 5.6: The co- and cross-polarised far-field element patterns are shown for a single radiating element at NP. Only the fundamental TE_{11} mode exists in the waveguide. There is some coupling into the LHCP polarised mode and both modes are seeing some reflections from the feedpoint.

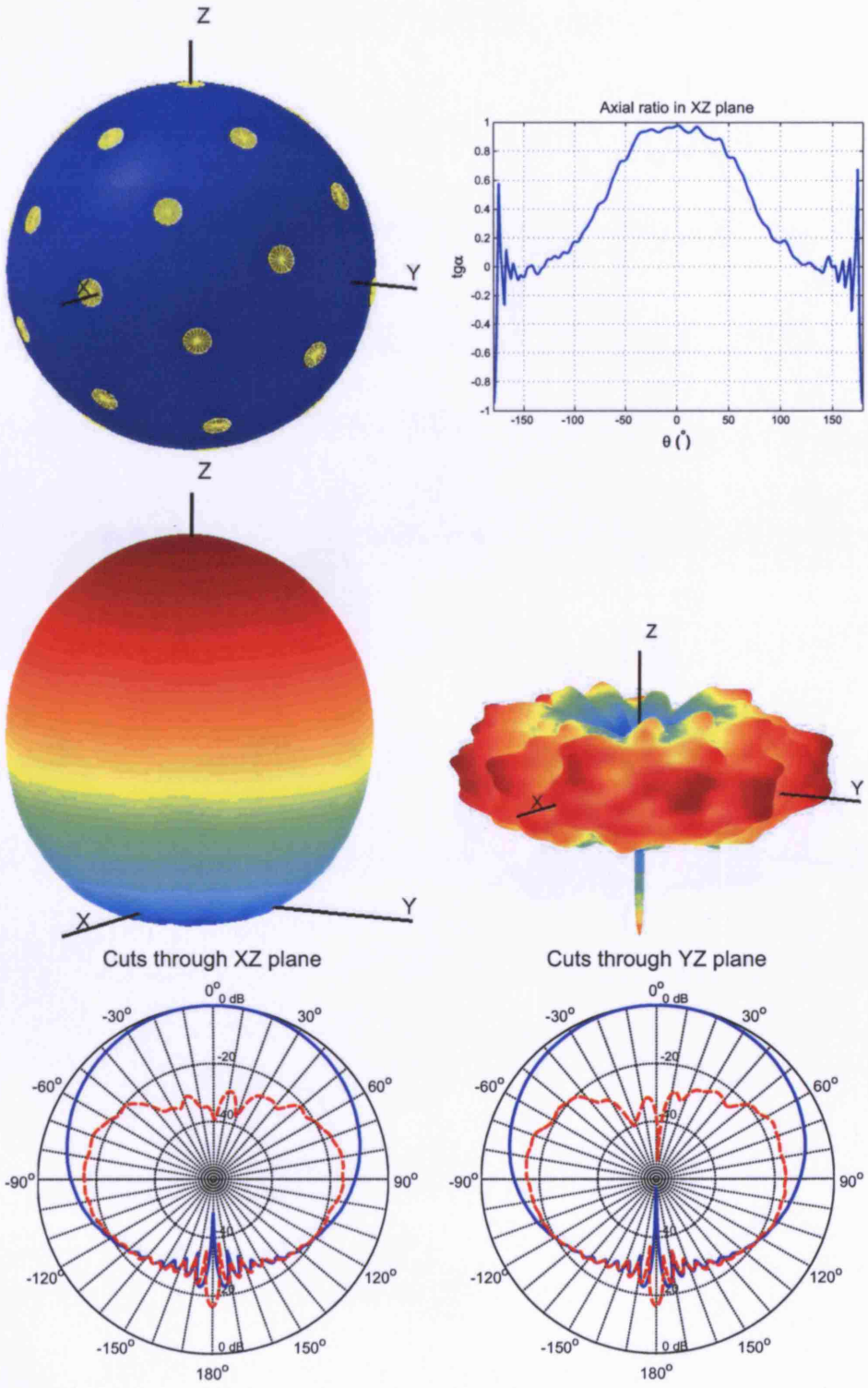


Figure 5.7: The co- and cross-polarised far-field element patterns are shown for one radiating element at NP that is part of a spherical array of 32 elements. The element distribution is shown in the top-left plot, the axial ratio in the top-right plot.

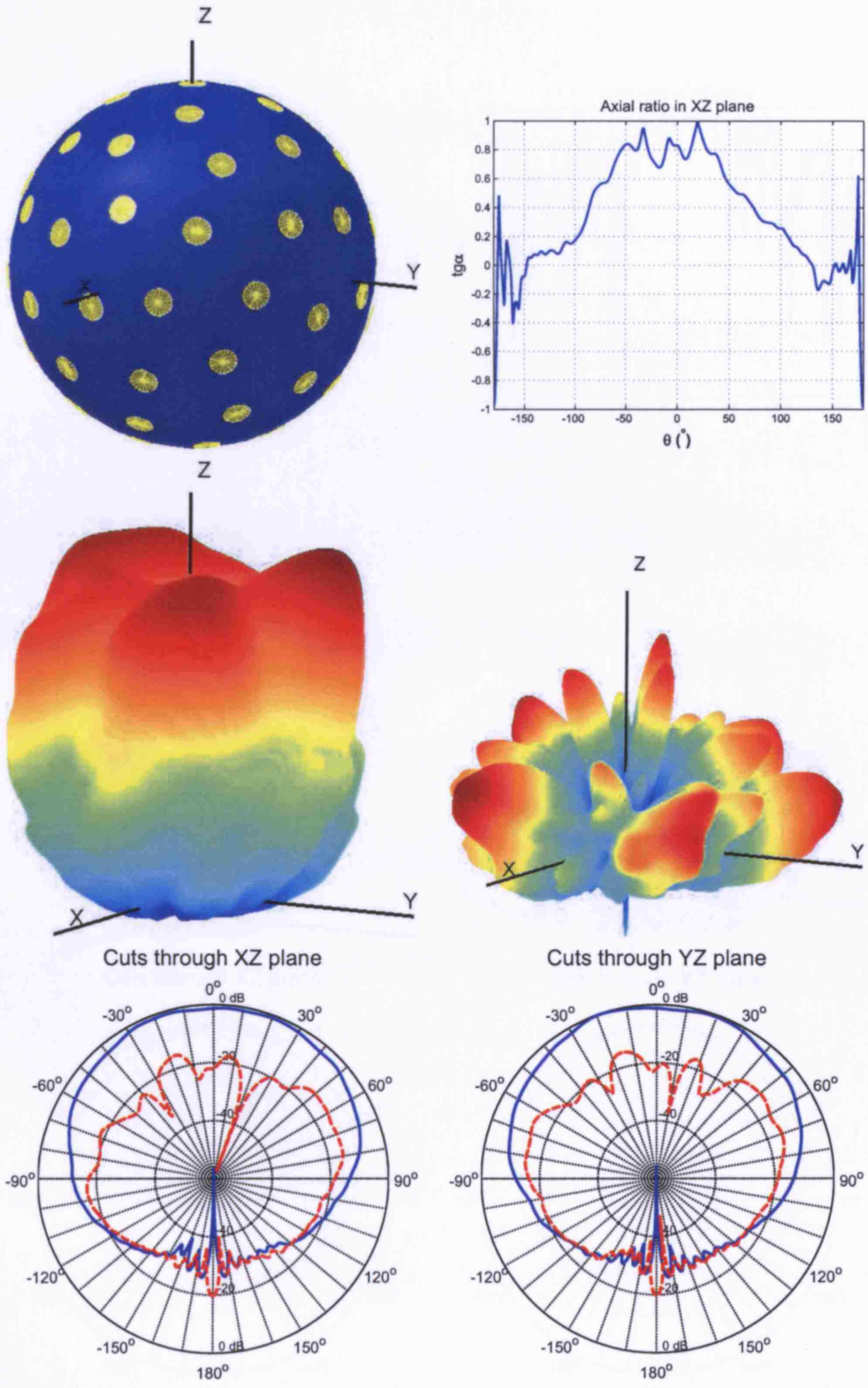


Figure 5.8: The co- and cross-polarised far-field element patterns are shown for one radiating element at NP that is part of a spherical array of 64 elements. The element distribution is shown in the top-left plot, the axial ratio in the top-right plot.

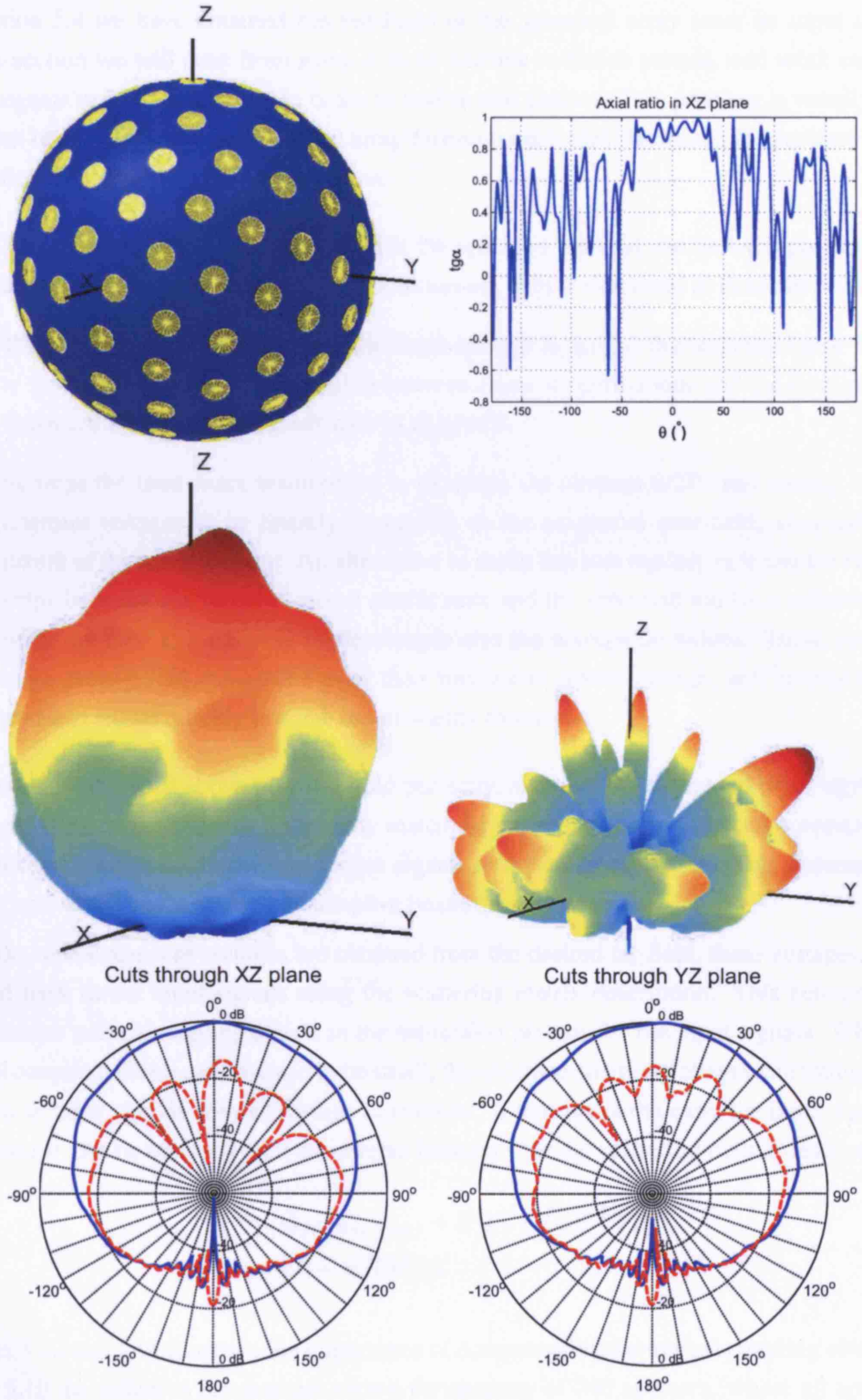


Figure 5.9: The co- and cross-polarised far-field element patterns are shown for one radiating element at NP that is part of a spherical array of 120 elements. The element distribution is shown in the top-left plot, the axial ratio in the top-right plot.

5.5 Spherical Array Synthesis

In section 5.4 we have obtained the Far-Field of the spherical array from its input signals. In this section we will start from some desired antenna radiation pattern, and work out what input signals need to be applied in order to realise that pattern. This problem is usually more complex to deal with, and the spherical array forms no exception. However, the problem can be simplified by accepting a few assumptions:

- The sphere is sufficiently large so that the specified far-field can be back-projected to a tangential near-field with the same polarisation. This is described in detail in section 7.1.
- The distribution of array elements is dense enough to resolve the tangential near-field up to the required degree. The relation between element distributions and the level of detail in the radiation pattern is dealt with in chapter 8.
- Perhaps the least exact assumption, is to ignore the element RGPs and simply take the aperture voltages to be linearly dependent on the tangential near-field, sampled in the centre of the array element. An alternative to make this assumption, is to use the relationships between the feedline modal coefficients and the spherical modal coefficients and work out how the spherical modes couple into the waveguide modes. There are many more propagating spherical modes than waveguide modes though, and in practise this problem would quickly become too unwieldy to solve.

The assumptions listed above may not hold perfectly, and with the estimated input signals the obtained radiation pattern may not exactly match the required pattern. This can however easily be checked by feeding the estimated input signals to the analysis routine. This process could also be automated in the form of an adaptive beamforming technique.

Once the modal aperture voltages are obtained from the desired far-field, these voltages can be worked back to the input signals using the scattering matrix description. This automatically incorporates mutual coupling effects in the estimation process for the input signals. When the mutual coupling effects are assumed to be small, the obtained aperture voltages can immediately be used as input signals with no further adjustment. If that is not the case, the input signals in the vector $\mathbf{F}$ can be obtained with the reverse formula's for the scattering matrix description:

$$\mathbf{A} = (\mathbf{I}_{N\times(m_1+m_2)} + \mathbf{S}^{\mathrm{E}})^{-1}(\sqrt{\mathbf{Z}})^{-1}\mathbf{V}_{\mathbf{ap}} \tag{5.16}$$

$$\mathbf{F} = (\mathbf{T} - \mathbf{S}^{\mathrm{I}}\mathbf{T}\mathbf{S}^{\mathrm{E}})\mathbf{A} \tag{5.17}$$

Figures 5.10 and 5.11 illustrate the importance of compensating for mutual coupling effects. In figure 5.10 the radiation patterns are shown for an array of 240 aperures, where all apertures were excited according to a synthesised excitation that maximises directivity in the direction of the z-axis. However, the mutual coupling effects were assumed to be compensated for, by bypassing the mutual coupling calculations in the algorithm and applying the synthesised voltages directly to the apertures. In figure 5.11 the mutual coupling effects were fully calculated as before and taken into account in the analysis of the same maximum-directivity excitation. The

pattern is thereby destroyed, especially the polarisation purity seems to suffer from the mutual coupling effects.

5.6 Summary and Conclusions

In this chapter a model has been presented, the scattering matrix description, in which the results obtained in previous chapters were combined into one single description of an entire spherical array antenna. This model is schematically illustrated in figure 5.2. The powerful model can be adapted to suit spherical array antennas with different kinds of radiating elements and different feeding structures. The position of the elements, their electromagnetic radiation properties and mutual coupling between the radiating elements are all automatically taken into account in the calculations of the model. Once calculated, the array antenna is fully specified, and the radiation pattern, input impedances and so forth are readily obtained from any given array excitation.

The model has been implemented into MATLAB, at the moment only for waveguide elements, but the code should readily be upgradable to cover the type 2 elements described in chapter 3. For a simple imaginary spherical array antenna of circular waveguides, the model was used to show the effect of mutual coupling on the array radiation pattern in figures 5.5 to 5.9.

It has been shown how the scattering matrix description can be used to analyse the performance of a spherical array antenna, starting from a given excitation. To solve the inverse problem, synthesising an excitation from a desired far-field pattern, other methods are required that will be given attention in chapter 7. Nevertheless the scattering matrix description can be used to predict and compensate for mutual coupling effects, the effect of which has been illustrated in figures 5.10 and 5.11. The powerful analysis tool goes a long way to describe spherical array antennas of arbitrary complexity, as it can be used with any number of elements, any kind of elements and any element distribution. To accept the second point of the hypothesis in section 1.3.1 it is also necessary to develop fast and accurate procedures to synthesise an array excitation and test the simulations empirically with a real-world prototype. These activities are the subject of chapters 7 and 9. In the following chapter, we take a step back and approach the design of spherical array antennas without relying on the elaborate analysis tools that have been developed over the previous chapters.

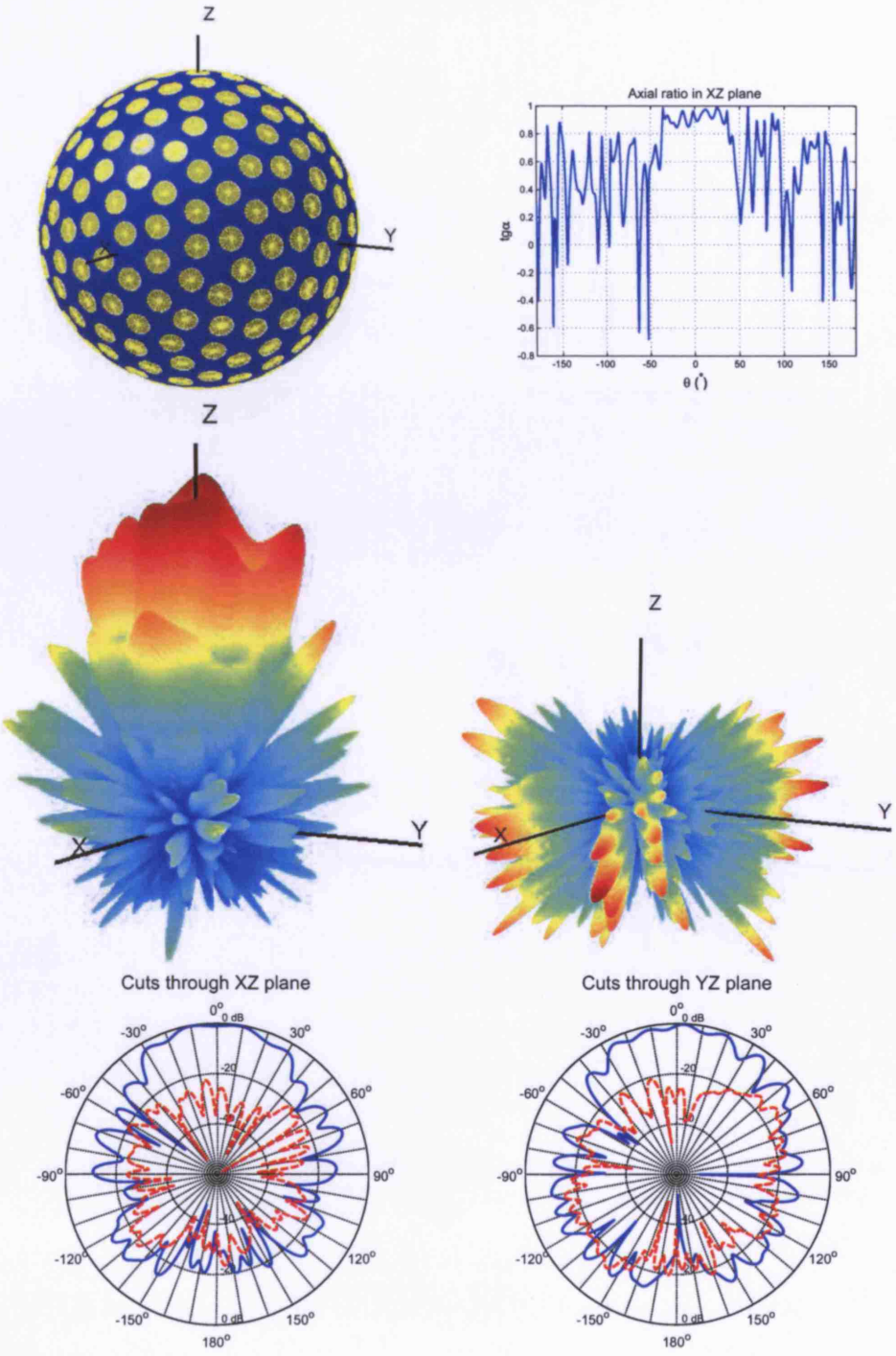


Figure 5.10: The far-field is shown for a 240-element array excitation that was optimised to achieve a maximum directivity in the direction of the z-axis. In the analysis procedure, the mutual coupling was shortcircuited and the excitation directly applied to the aperture fields. This is equivalent to anticipate and compensate the mutual coupling effects.

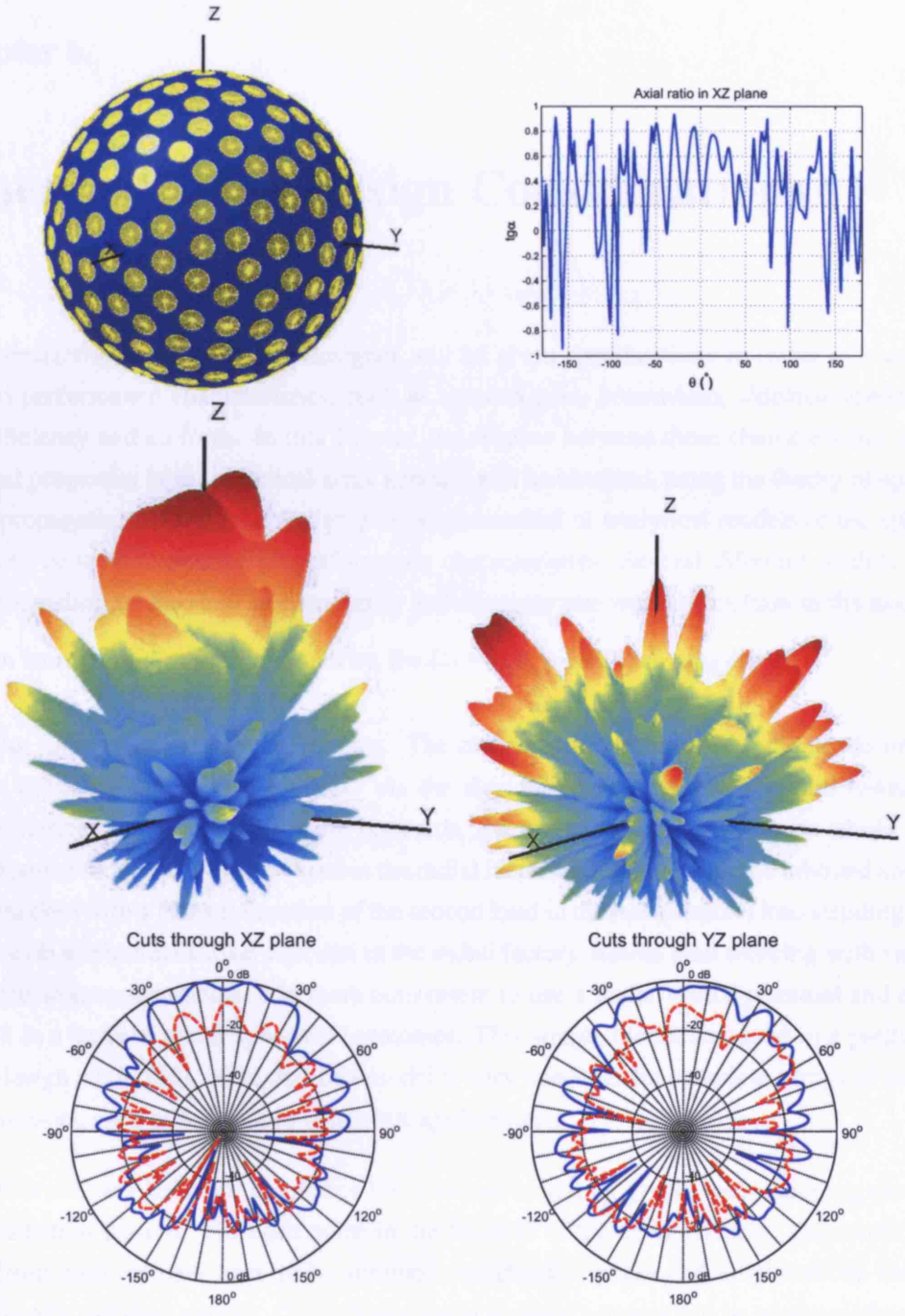


Figure 5.11: The far-field is shown for the same 240-element array excitation that was optimised to achieve a maximum directivity in the direction of the z-axis. This time the excitation was applied at the feed terminals, without compensation for the mutual coupling effects that disturb the pattern.

Chapter 6

Spherical Array Design Considerations

When designing an antenna, the designer will be given specifications in terms of some well defined performance characteristics, such as antenna gain, beamwidth, sidelobe levels, aperture efficiency and so forth. In this chapter, the relation between these characteristics and the physical properties of the spherical array antenna will be obtained, using the theory of spherical mode propagation. During the design process, numerical or analytical models of the spherical array can be used to analyse its performance characteristics. Several different models can be used, depending on the level of complexity and accuracy one wants to include in the analysis.

With an increasing degree of complexity, the following models are suggested:

- An array of point source elements. The elements are omnidirectional and do not have a radiation pattern of their own. As the elements radiate as much inward towards the spherical centre as they radiate outwards, the outbound spherical modes (those with a Hankel function of the first kind in the radial factor) recombine with the inbound spherical modes (with a Hankel function of the second kind in the radial factor) into standing waves (with a spherical Bessel function in the radial factor). Rather than working with vectorial electromagnetic fields, it is more convenient to use a scalar vector potential and expand it in a basis of scalar spherical harmonics. This simple model was used in a preliminary design study [28]. Note that this model is very adequate for spherical arrays of pressure sensors, such as those used in SONAR applications.

- It is not very difficult to enhance the previous model by giving each radiating element a radiation pattern. For each point in the far-field of the array antenna, the contributions from each element have to be summed, weighted in phase and amplitude by each element's radiation pattern. Only simple scalar element patterns can be accommodated, and they are usually approximated with a simple analytic function such as a raised-cosine pattern. This approach was used in the author's MSc dissertation, of which some conclusions relevant to this thesis are highlighted in appendix C.

- The radiation pattern of each array element can be represented by a full vectorial spherical wave expansion. These expansions can be obtained accurately, by measurement, simulation or if possible, an analytic approach. The latter can be simplified by breaking up the

field of the array element at NP position in orthogonal electromagnetic modes and deriving the spherical wave expansion of each mode. This approach was followed in chapter 3. The array elements are usually identical and it is easy to rotate a field that is expanded in spherical modes to obtain the field of any array element.

- Ultimately, in order to accurately model the array environment, it is necessary to take the interactions between the elements into account. The mutual coupling effects influence the excitation applied to the elements and may disturb the obtained array pattern. A powerful model for this kind of detailed analysis was developed in chapter 5.

This chapter aims to develop some simple design formula's for the spherical array, so that a preliminary design can be made, that can then be taken through a number of design cycles with varying level of analysis, using the models described above.

6.1 Choice of Polarisation

Although our discussions are applicable to any particular polarisation, often we assume a circular polarised far-field and circularly polarised array elements. This section explains why circular polarisation is the best suited polarisation to use in conjunction with spherical array antennas.

Assume an x-oriented linearly polarised plane wave, travelling in the direction of the z-axis, impinges on a sphere. The fieldlines are given by the unit vectors:

$$\begin{aligned} \vec{v} &= \hat{e}_x \\ &= \hat{e}_r \sin\theta\cos\phi + \hat{e}_\theta \cos\theta\cos\phi - \hat{e}_\phi \sin\phi \end{aligned} \tag{6.1}$$

The field $\vec{v}$ travels southwards and hits the sphere first at NP. As the field continues to travel southwards towards the other pole SP, it is sensed by the elements distributed over the spherical surface. Assuming the elements are tangential to the sphere, they can only sense the tangential component of the field $\vec{v}_t$:

$$\vec{v}_t = \hat{e}_\theta \cos\theta\cos\phi - \hat{e}_\phi \sin\phi \tag{6.2}$$

The field vectors are illustrated in figure 6.1. Small dipoles on the sphere, aligned with this tangential vector field, will also maximise the radiation in the direction of the z-axis, provided they have the proper phase delays according to a wave travelling in the z-direction. Path delay phase compensation is dealt with by the phase of the excitation signals to the dipoles. The tangential field is curved with respect to the incident planar field, which still causes a non-uniformity in amplitude and the blind spots around the positive and negative x-axis. It would make sense to include this effect in the amplitudes of the excitation signals, and have all dipoles on the sphere of the same size. It is indeed easily verified that the amplitudes of the field $\vec{v}$ are not uniformly equal to one, making it unsuitable as a definition for polarisation. Dropping the dependence on θ however does make it have this property. We therefore drop the dependence on θ from the definition of the polarisation, and transfer the responsibility over path delay compensation and curvature compensation to the setting of correct phases and amplitudes in

the excitation signals to the radiating (or receiving) elements. Using spherical modes to build up radiation patterns simplifies and automates this task.

Dropping the θ dependence from the polarisation definition is equivalent to opening the sphere at SP, peeling it open in all directions, and flattening it to a disk in the xy-plane that contains NP. Let $\vec{r}$ be the tangential vector field of unity vectors, corresponding to an incident linearly polarised planar wave as before, but with the effect of curvature removed. The field $\vec{v}$ has parallel field lines over the disk (as shown in figure 6.2). By projecting these lines back to the sphere, we obtain a definition for $\vec{r}$:

$$\begin{aligned} \vec{r} &= \vec{v}_t(\theta = 0, \phi) \\ &= \hat{e}_\theta \cos\phi - \hat{e}_\phi \sin\phi \end{aligned} \tag{6.3}$$

This definition for $\vec{r}$ is identical to the definition of the co-polar component of the far-field, as given in equation (2.55) and shown in figures 2.4, 6.2 and 6.3. Therefore, we can use the spherical basis functions that were derived in accordance with that polarisation definition to ensure that the fieldvectors are optimally aligned for a specific purpose, such as maximising the directivity in some direction.

There is another issue that needs to be resolved, that is immediately apparent from figures 6.1 and 6.3. By adjusting the phases and amplitudes of the excitation signals to the dipoles on the sphere, the effects of curvature and path delay can be compensated, so that the radiation pattern has a maximum in the direction of the z-axis. However, without rotating the dipoles, this can not be achieved in any other direction than the z-axis. Instead of dipoles, we need some kind of dual polarised elements, so that the tilt angle of the polarisation axis can be changed in the tangential plane, which requires a doubling of the number of controls to the array antenna. Alternatively, and much more simply, is to use circularly polarised elements. The observations made before for linear polarisation remain valid, but now pertain to the reference phase, or the orientation of the phasors at time zero. Aligning the circularly polarised elements so that their radiation combines coherently, is now a matter of setting the phases correctly. By analogy with orienting the linearly polarised elements, we give each circularly polarised element a fixed phase offset, after which the amplitudes and phases of its excitation signal will take care of the path delay and curvature compensation.

Another reason to use circular polarisation with spherical array antennas may come from the applications in which they are used. Indeed, typical applications that can benefit from spherical array antennas are space communications, where it is customary to use circular polarisation to minimise polarisation mismatch between two antennas with unknown relative position.

6.2 Directivity, Beamwidth and Sidelobe Levels

In this section some observations will be made on the directivity, beamwidth and the sidelobes of radiation patterns consisting of a finite spectrum of spherical modes. The observations stand apart from the physical shape of the antenna that creates the pattern, so that the obtained results remain perfectly valid for any antenna. Of course the results will only be useful as long as the

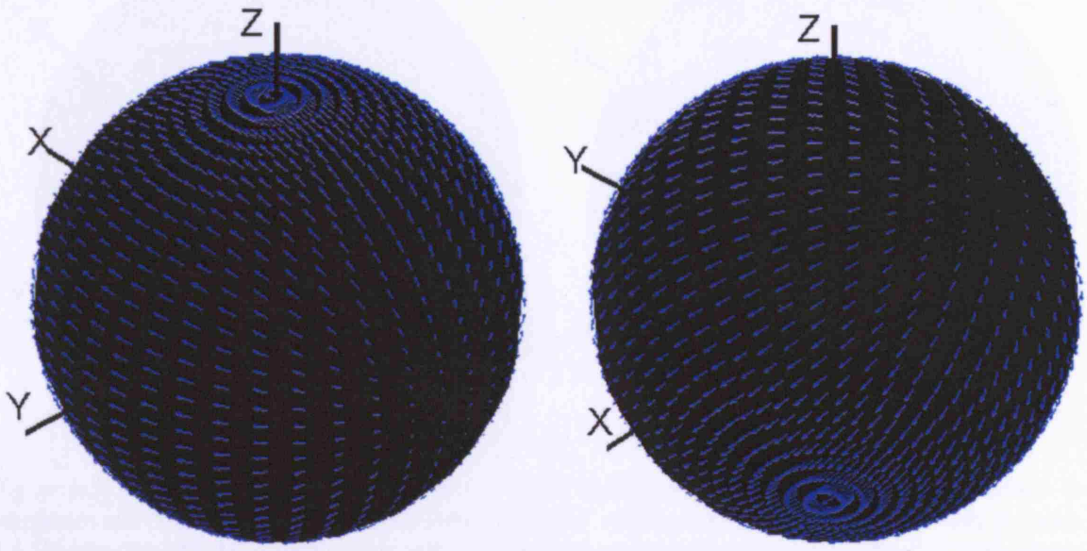


Figure 6.1: The tangential vectors correspond with the fieldlines of a linearly polarised plane wave that travels in the direction of the z-axis. The effect of curvature causes the blind spots in the vector field. Propagation delays associated with the wave characteristic are not taken into account. The vector field is unsuitable as a definition for polarisation because the amplitudes are not uniform.

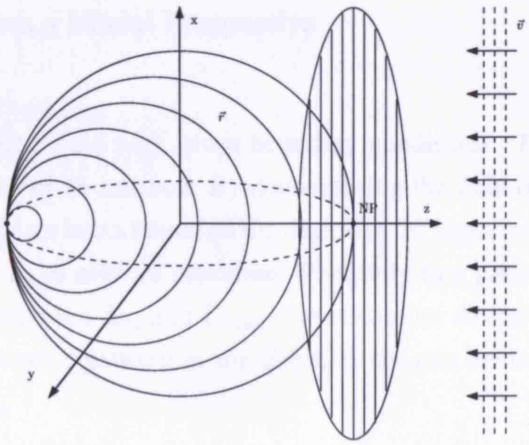


Figure 6.2: Ludwig's third definition for co- and cross-polarisation of the far-field is equivalent with the tangential vector field that is obtained when a plane wave hits the surface of a flattened sphere, peeled open from the southpole SP. This interpretation adds an intuitive feel to Ludwig's third polarisation definition.

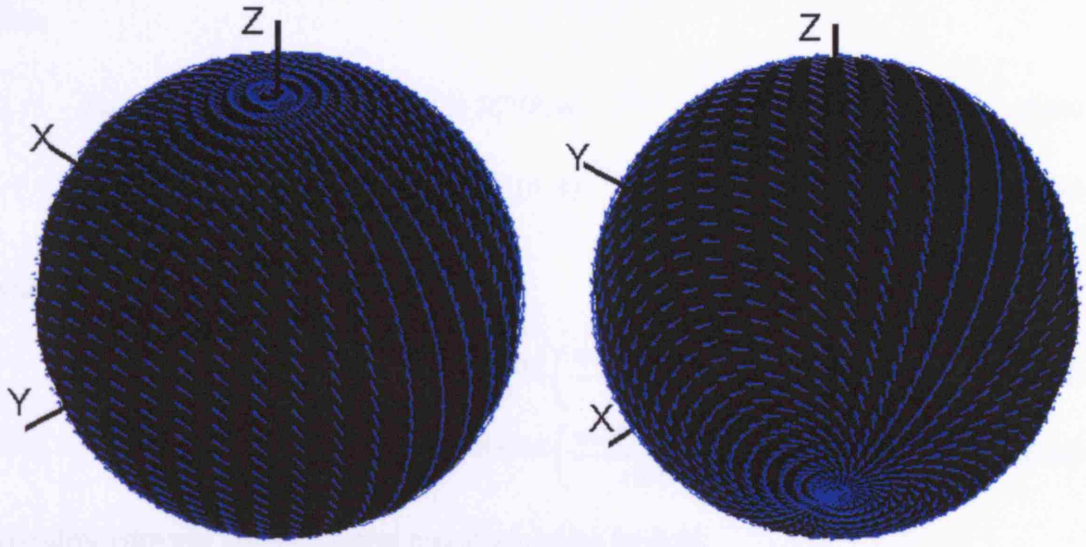


Figure 6.3: By removing the effects of the curvature, the tangential vector field becomes uniform in amplitude and coincides with Ludwig's third definition of polarisation, that was also illustrated in figure 2.4. Neither the effects of path delays and curvature are taken into account. The effects of path delays and curvature need therefore to be included in the scalar functions that describe the co- and cross-polar components of the field. Using the spherical modes to construct the radiated field will automatically take care of that.

spherical decomposition of the antenna radiation pattern can be obtained relatively easily, as is the case with spherical array antennas.

6.2.1 Directivity, from a Modal Perspective

Assume an electromagnetic field with given boundary conditions. The boundary conditions typically relate to the size of an antenna. By decomposing the field into spherical modes, the boundary conditions translate into a bound on the degree of the highest spherical mode. Suppose the field is optimised so as to achieve maximum directivity in a particular direction. We will establish a relationship between L_m and $D_{\max}$, the maximum directivity of this field. As we can rotate the spherical wave expansion in any direction, the maximum achievable directivity is obtained in any direction.

In section 3.4.2 we defined the far-field in spherical modal terms, and used a particular definition for the co-polar and cross-polar components to split the vectorial modes in two sets of scalar modes. Provided that the coefficients C_{lm}^{TE} and C_{lm}^{TM} of the spherical wave expansion are known, the far-field was obtained (relative to the excitation level) as:

$$\vec{\mathcal{F}}(\theta,\phi) \;=\; \mathcal{F}_{rhcp}(\theta,\phi)\,\hat{e}_{rhcp} + \mathcal{F}_{lhcp}(\theta,\phi)\,\hat{e}_{lhcp} \tag{6.4a}$$

with:

$$\mathcal{F}_{rhcp}(\theta,\phi) = \sum_{lm} (C_{lm}^{\mathsf{TE}} + C_{lm}^{\mathsf{TM}})\, R_l^m(\theta,\phi) \tag{6.4b}$$

$$\mathcal{F}_{lhcp}(\theta,\phi) = \sum_{lm} (C_{lm}^{\mathsf{TE}} - C_{lm}^{\mathsf{TM}})\, L_l^m(\theta,\phi) \tag{6.4c}$$

where the scalar functions were defined as:

$$R_l^m = \sqrt{2\pi}(-j)^l N_{lm} e^{j(m-1)\phi} \left(\frac{m P_l^m(\cos\theta)}{\sin\theta} + \frac{dP_l^m(\cos\theta)}{d\theta} \right) \tag{6.4d}$$

$$L_l^m = \sqrt{2\pi}(-j)^l N_{lm} e^{j(m+1)\phi} \left(\frac{m P_l^m(\cos\theta)}{\sin\theta} - \frac{dP_l^m(\cos\theta)}{d\theta} \right) \tag{6.4e}$$

The directivity was just the squared magnitude of the far-field:

$$D(\theta,\phi) = |\vec{\mathcal{F}}(\theta,\phi)|^2 \tag{6.5}$$

When maximum directivity is required in the direction (θ',ϕ'), it can be achieved by acknowledging that the spherical modes offer an orthogonal basis in the spherical system, so that the coefficients can be chosen to maximise (6.4b) or (6.4c). Suppose we want to maximise the right hand circularly polarised far-field, then we choose:

$$C_{lm}^{\mathsf{TE}} = R_l^m(\theta',\phi')^* \tag{6.6a}$$

$$C_{lm}^{\mathsf{TM}} = C_{lm}^{\mathsf{TE}} \tag{6.6b}$$

With this choice of coefficients, the RHCP component will achieve a maximum in the direction (θ',ϕ'), but also, the left hand polarised component will be zero in all directions. It is therefore theoretically possible to achieve zero cross-polarisation from a spherical array antenna, regardless of the direction of observation. This is not the same as to say that a spherical wavefront with a single polarisation can be created, because that is untrue. The multipole expansion for vectorial spherical waves starts from degree one. There is no uniform mode of degree zero, and any pattern of a single polarisation will have crests and bumps.

We choose $(\theta',\phi') = (0,0)$. By using the following equalities (see appendix A1 of [63]):

$$\left.\frac{dP_l^m(\cos\theta)}{d\theta}\right|_{\theta=0} = \begin{cases} \dfrac{-l(l+1)}{2} & \text{for } m = 1 \\ \dfrac{1}{2} & \text{for } m = -1 \\ 0 & \text{for } |m| \neq 1 \end{cases} \tag{6.7a}$$

$$\left.\frac{m P_l^m(\cos\theta)}{\sin\theta}\right|_{\theta=0} = \begin{cases} \dfrac{-l(l+1)}{2} & \text{for } m = 1 \\ \dfrac{-1}{2} & \text{for } m = -1 \\ 0 & \text{for } |m| \neq 1 \end{cases} \tag{6.7b}$$

in equation (6.4d), we obtain a solution for $R_l^m(0,0)$:

$$R_l^m(0,0) = \begin{cases} \frac{1}{\sqrt{2}}(-j)^l\sqrt{2l+1} & \text{for } m = 1 \\ 0 & \text{for } m \neq 1 \end{cases} \tag{6.8}$$

so that the expansion coefficients that maximise the directivity in the direction of the z-axis become:

$$C_{lm}^{\text{TE}} = C_{lm}^{\text{TM}} = \frac{1}{\sqrt{2}}(j)^l\sqrt{2l+1} \tag{6.9}$$

The obtained expansion that maximises the directivity in the direction of the z-axis can be rotated into any desired direction. The expansion coefficients will be altered, however the obtained pattern and the maximum directivity remain the same. We obtain the maximum directivity as:

$$\begin{aligned} D_m &= \left|\vec{\mathcal{F}}(0,0)\right|^2 \\ &= |\mathcal{F}_{rhcp}(0,0)|^2 \\ &= \sum_{l,m} 2\left|R_l^{m^*}(0,0)\right|^2 \\ &= \sum_{l=1}^{L_m}(2l+1) \\ D_m &= L_m(L_m+2) \end{aligned} \tag{6.10}$$

The resulting patterns, optimised for maximum directivity, can be obtained by using the coefficients in (6.9) into the expressions for the far field (6.4). The resulting patterns are shown in figure 6.4.

6.2.2 Beamwidth and Sidelobe Levels

Having obtained an expression for the maximum directivity of the spherical array, we can now relate this quantity to the corresponding beamwidth of the radiation pattern. We can expect that we can trade off some of this optimal directivity and beamwidth with lower sidelobe levels.

Let $\mathcal{P}_m$ be the power density radiated in the direction of maximum directivity and $\mathcal{P}_i$ be the power density of an isotropic radiator, transmitting with the same absolute power level P_t. Then by definition we have:

$$D_m = \frac{\mathcal{P}_m}{\mathcal{P}_i} = \frac{\mathcal{P}_m}{\frac{P_t}{4\pi R^2}} \tag{6.11}$$

The beamwidth θ_{3dB} is defined as the angle between the half-power (3dB) points of the main lobe in the radiation pattern. We define α as half this angle, from the point of maximum directivity to the 3dB point. Suppose that all the power was radiated uniformly in the region between the 3dB points, and none of it outside this region. At any distance R, the area of the illuminated spherical cap is $A = 2\pi RH = 2\pi R^2(1-\cos\alpha)$. With the uniform illumination of the spherical cap, the power density inside the spherical cap would be $\frac{P_t}{A}$. With a non-uniform illumination but no sidelobes, the disregarded portions outside the spherical cap would more or less

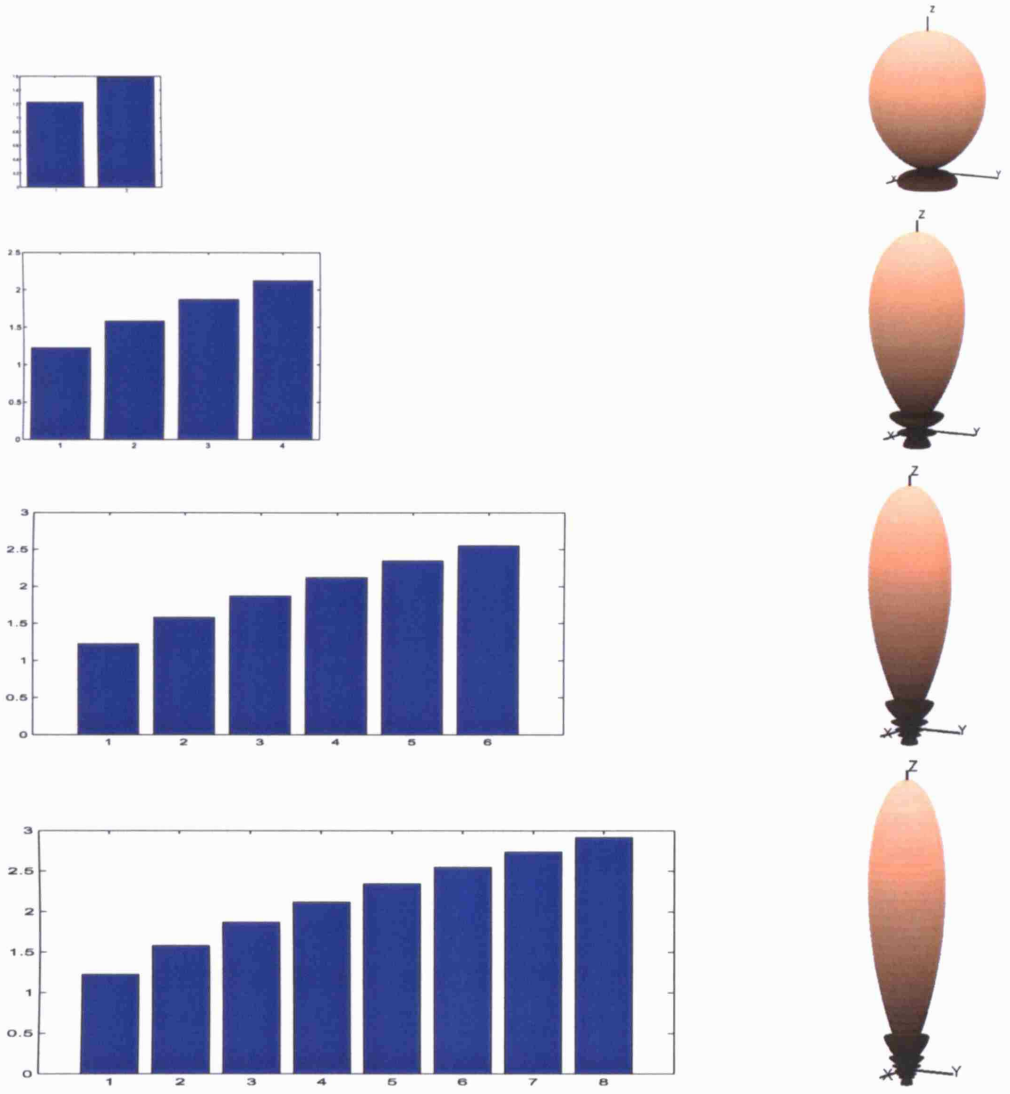


Figure 6.4: Optimal directivity patterns are shown for a maximum degree of 2,4,6 and 8. The patterns are shown on a linear scale. The left column shows the magnitude of the expansion coefficients as a function of the degree l. The spectral domain is 1-dimensional as $m = 1$ for axially symmetric fields around the z-axis.

compensate the less illuminated portions of the spherical cap, so that the found power density is a good approximation for $\mathcal{P}_m$. In reality, there is quite some power put into the sidelobes, so that we can expect $\mathcal{P}_m$ to be somewhat lower than $\frac{P_t}{A}$:

$$\mathcal{P}_m = \sigma \frac{P_t}{A} \tag{6.12}$$

We introduced the factor σ to reflect roughly how much power is contained in the sidelobes. Substituting (6.12) into (6.11), we obtain:

$$D_m = \frac{2\sigma}{1 - \cos\alpha} \tag{6.13}$$

so that

$$\theta_{3dB} = 2\alpha = 2\arccos\left(1 - \frac{2\sigma}{D_m}\right) \tag{6.14}$$

$$= 2\arctan\left(\frac{\sqrt{1 - \left(1 - \frac{2\sigma}{D_m}\right)^2}}{1 - \frac{2\sigma}{D_m}}\right) \tag{6.15}$$

$$\approx 4\sqrt{\frac{\sigma}{D_m}} \tag{6.16}$$

where the latter approximation holds for harmonic degrees above 1 and improves for larger D_m. The equation was compared to the true values, found by constructing the far-field with the expressions given in section 6.2.1 and searching for the angle where the power has fallen to -3dB. It was found that a value of $\sigma = 70\%$ gives an excellent match between the rule of thumb and the true beamwidths. Hence, we can conclude that a pattern optimised for maximum directivity will contain roughly 30% of its power in the sidelobes. This suggests that large sidelobes will be present in the patterns that are optimised for directivity, just as with other types of array antennas.

The formula above, including the empirically found value for σ, is entirely in line with an existing rule of thumb to calculate the directivity of high-gain antennas. It was first presented in [109] and repeated in many reference books on antennas, including [82]. The formula expresses the directivity of an antenna D_0 as:

$$\frac{1}{D_0} = \frac{1}{2}\left(\frac{1}{D_1}\frac{1}{D_2}\right) \tag{6.17a}$$

$$D_1 \approx \frac{16\ln 2}{\Theta_1} \tag{6.17b}$$

$$D_2 \approx \frac{16\ln 2}{\Theta_2} \tag{6.17c}$$

where Θ_1 and Θ_1 are the half-power beamwidths (in radians) in the E- and H-planes, and $\ln 2 \approx 70\%$.

For the directivity-optimised patterns from above, figure 6.6 shows the position and height of the

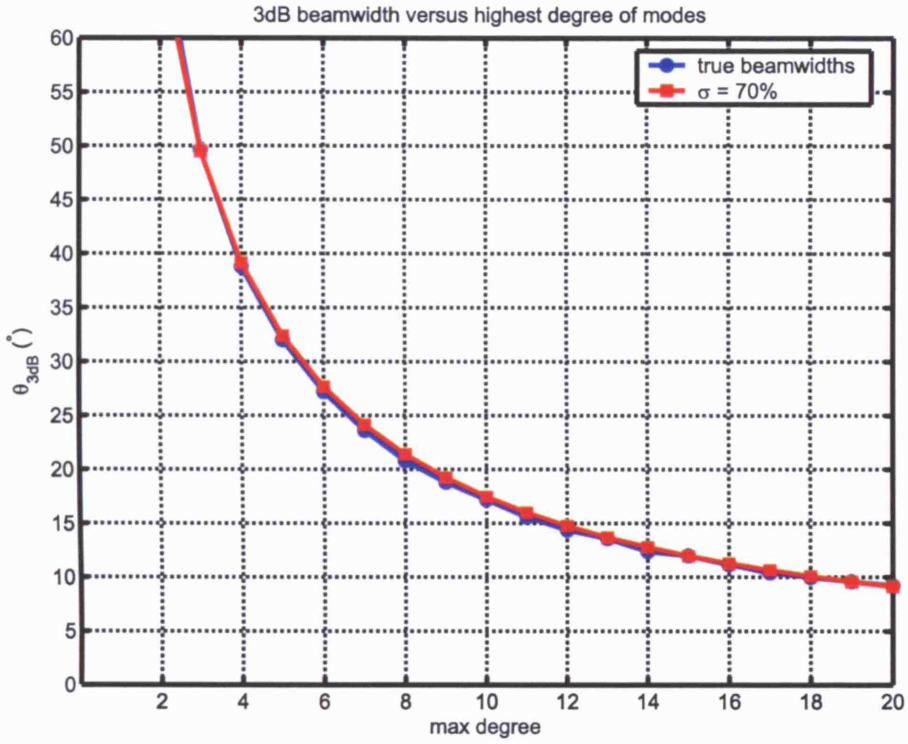


Figure 6.5: 3dB beamwidths are shown as a function of the highest degree mode in the optimal directivity pattern. The rule of thumb in equation (6.16) was compared with the true beamwidths of the patterns in figure 6.4. (The given beamwidths are full beamwidths, not half beamwidths.)

nearest sidelobe, parameterised by the maximum degree of the spherical modes. As expected, the first sidelobe is large and may be unacceptable for many applications.

Also from figure 6.6, it can be seen that the more spherical modes are used to increase the directivity, the more the sidelobes close in, without lowering significantly. This could be expected from the modal spectrum of the patterns, defined in equation (6.9). The magnitude of the coefficients of the spherical wave expansion increases with the square root of the highest degree. with no smooth transition back to zero. This is equivalent to multiplying a pattern of infinite directivity (a Dirac pulse at zero) with a rectangular window function up to degree L_m. In the spatial domain, this amounts to a spherical convolution of the Dirac pulse with the inverse spherical Fourier transform of the rectangular window. The sidelobes of the latter are what causes the sidelobes to appear in the radiation patterns. As with any Fourier transform, a smoother window would result in lower sidelobes. This observation suggests how the sidelobe structure can be improved by smoothing the transition back to zero, for the highest degree modes. Necessarily, this will come at the price of compromising a little on directivity, as the original patterns were shown to be optimal from that perspective.

Let $L_{rolloff}$ be the degree from where the spherical wave coefficients will smoothly fall back to zero, which is achieved at L_m. As an example on how the sidelobe structure can be improved, a window function is defined as being equal to 1 below $l = L_{rolloff}$ and falls back to zero as a cosine function $\cos(x)/2 + 0.5$ between $L_{\text{roll-off}}$ and L_m. This is illustrated in figure 6.7. With

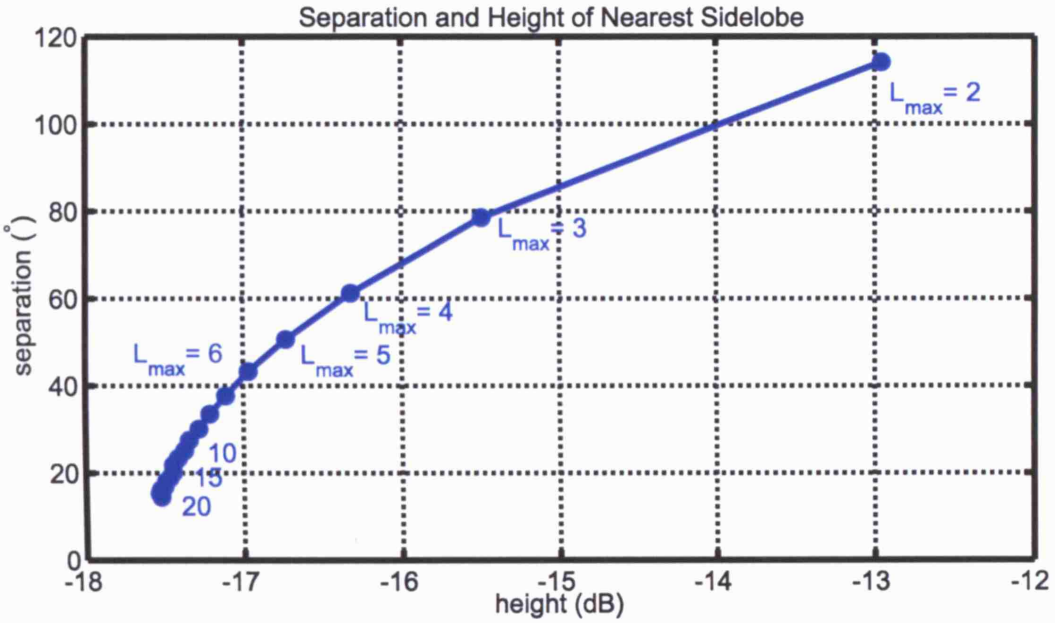


Figure 6.6: Position and magnitude of first sidelobe as a function of the maximum degree in the optimal directivity pattern.

a smooth transition from $l = L_{\text{roll-off}} = 0.8 \times L_m$, the effect on the nearest sidelobe is shown in figure 6.9. The improvement is much more dramatic for lower maximum degrees. This should not come as a surprise, as for large degrees (and hence a very broad window function) $l = L_{\text{roll-off}}$ should be lowered in order to keep a smooth window function. The jumps in figure 6.9 are there as a consequence of how $L_{\text{roll-off}}$ was defined. While these discontinuities vanish for a high maximum degree, it can be seen by comparison of figures 6.9 and 6.8 that the jumps occur whenever $L_{\text{roll-off}}$ is increased.

The decrease in directivity as a consequence of tapering the higher degree modes is illustrated in figure 6.10, which shows the increased beamwidth of the main lobe in the radiation pattern. This effect can be quantised by adjusting the parameter σ, so that more power is transferred from the

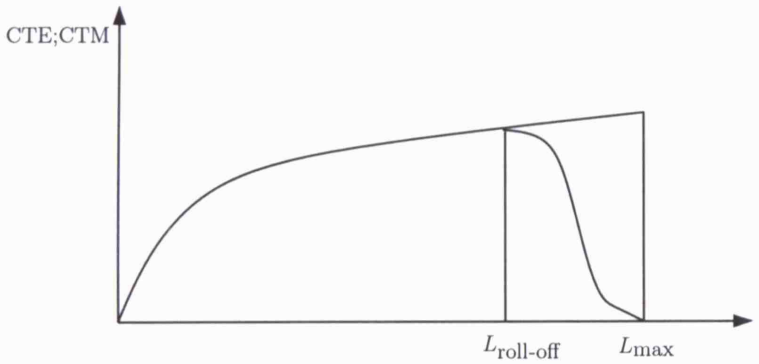


Figure 6.7: The maximum directivity expansions shown in figure 6.4 can be adjusted with a smoothing taper that gently rolls off to zero at the highest degrees. This enhances (reduces) the sidelobe levels in the radiation pattern as shown in figure 6.9.

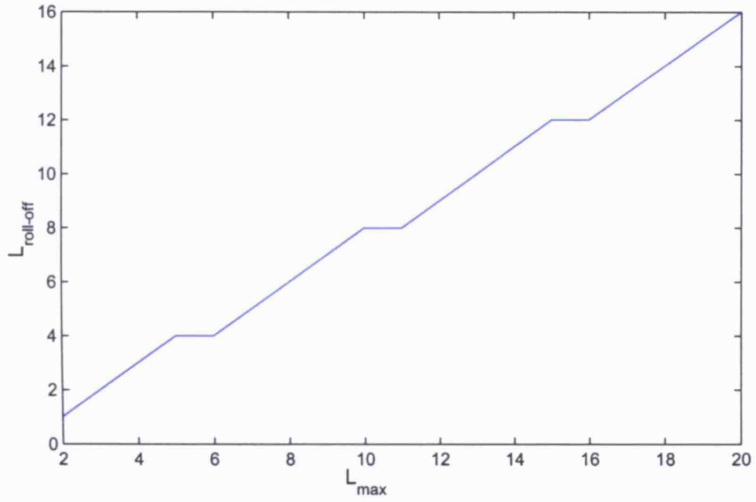


Figure 6.8: With reference to figure 6.7 $L_{\text{roll-off}}$ was chosen close to $0.8 \times L_m$. But as the spectrum is discrete (the sphere is a periodic domain) there are only discrete values for $L_{\text{roll-off}}$ to choose from. The exact values that were used for $L_{\text{roll-off}}$ are shown as a function of L_m. The discontinuities explain the jumps in figure 6.9

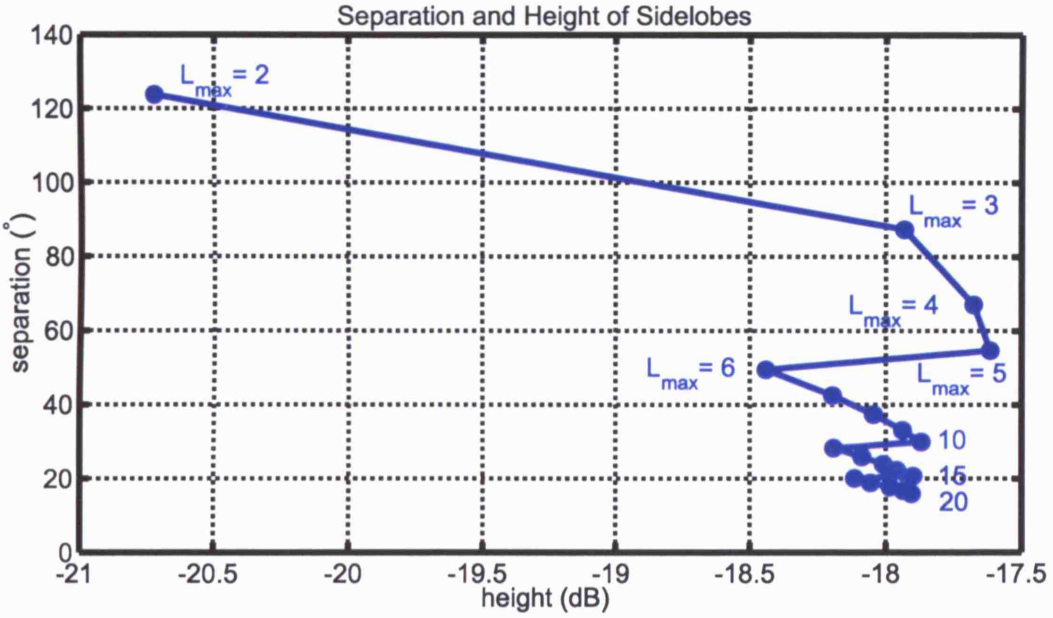


Figure 6.9: Position and magnitude of first sidelobe as a function of the maximum degree in the optimal directivity pattern for the amplitude tapered optimal directivity expansions.

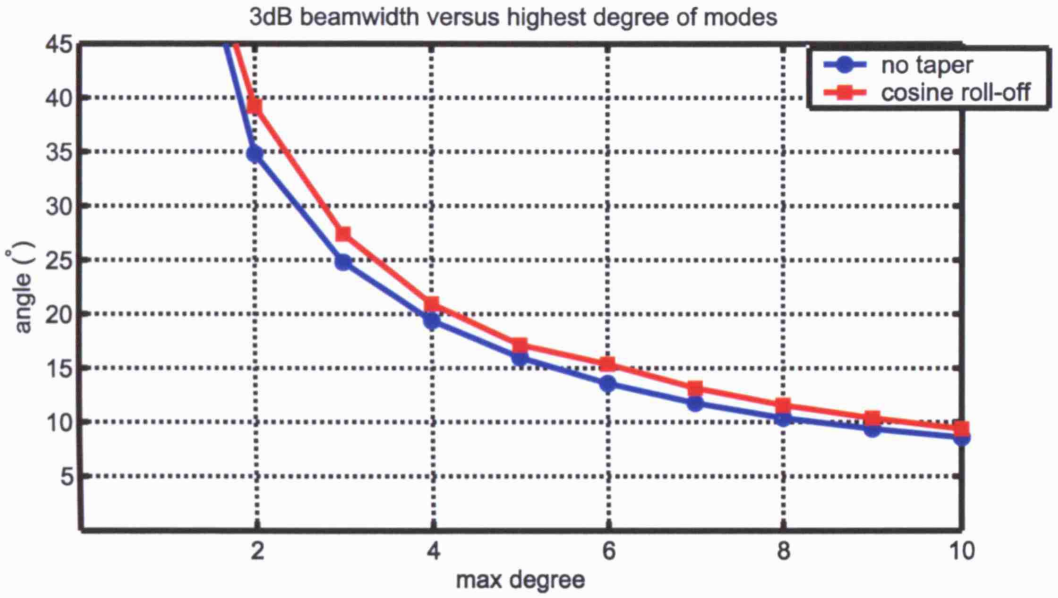


Figure 6.10: The amplitude tapering of the higher degree modes in the optimal directivity patterns to decrease the sidelobe levels slightly increases the beamwidth and decreases the directivity of the pattern.

sidelobes into the main beam. With the definition for the tapering of the coefficients used here, an adjustment of $\sigma = 80\%$ is appropriate. That means that 10% of the power was transferred from the sidelobes into the main beam. The plots in figure 6.11 illustrate the radiation patterns for both tapered and untapered spherical spectra.

6.3 Gain, Area and Effective Aperture of Spherical Antennas

The gain of an antenna is related to its effective aperture, projected in the direction of interest. For a spherical antenna, the projected shape in any direction is that of a circular disk. The gain of any reciprocal antenna relates to its effective aperture as follows:

$$G = \frac{4\pi}{\lambda^2} A_e \tag{6.18}$$

The physical area of sphere projected in a particular direction, is the area of a circular disk:

$$A = \pi R^2 \tag{6.19}$$

The question is how this physical area relates to the effective aperture or else, what is the **aperture efficiency** of a radiating spherical body?

$$\begin{aligned} A_e &= \eta_a A \\ \eta_a &= ? \end{aligned} \tag{6.20}$$

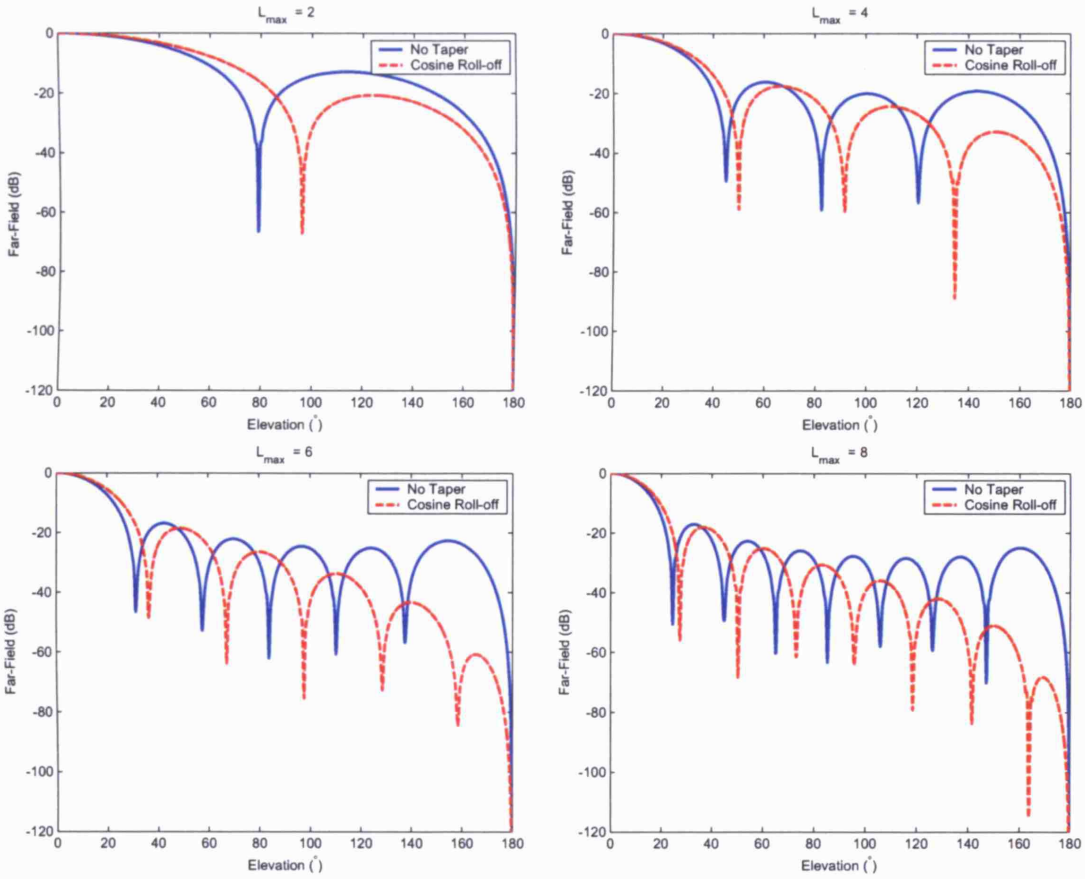


Figure 6.11: Radiation patterns are shown for both tapered and untapered optimal directivity patterns. The four plots differ in the maximum degree that is used in the modal expansion of the field.

To address this question, the aperture efficiency is split into parts that can be investigated separately:

$$A_e = \eta \quad \eta_f \quad \eta_s \quad \underbrace{\eta_m \quad A}_{A_m=\frac{\lambda^2}{4\pi}D_m} \quad \underbrace{\underbrace{\underbrace{}_{A_s=\frac{\lambda^2}{4\pi}D_s}}_{A_f=\frac{\lambda^2}{4\pi}D_f}}_{A_e=\frac{\lambda^2}{4\pi}G} \tag{6.21}$$

In equation (6.21), η is the radiation efficiency of the antenna, due to ohmic losses and relating the gain with the directivity of the antenna. If the antenna would have a continuous excitation, meaning that the sphere is excited over its full body, than the projected physical area A has to be corrected with an efficiency factor η_m in order to get the directivity of this type of antenna. However, a spherical array antenna has a finite set of radiating elements placed in a spherical configuration, so that the effective aperture needs to be adjusted with η_s to account for that effect. The elements are considered to be point sources that sample the continuous excitation function. η_s is unrelated to the reduction of the antenna's physical area, because the cumulative area of the point sources is zero. Finally, η_f takes into account the effect of having finite sized elements that each have a radiation pattern of their own. This factor depends on the design of the radiating elements.

6.3.1 Continuous Excitation Function

Without any restrictions on the excitation function that is defined over a spherical conducting body, the maximum directivity that can be obtained is still bounded by the size of the sphere. This can be seen from equation (6.10) that relates the maximum directivity to the highest degree spherical mode:

$$D_m = L_m(L_m+2) = \frac{4\pi}{\lambda^2}A_m \tag{6.22}$$

and acknowledging the fact that the metal sphere acts as a low-pass filter on spherical modes. The highest degree L_m of spherical modes that can appear in the far-field pattern is dictated by the electrical size of the sphere:

$$L_m \approx kR \tag{6.23}$$

where k is the propagation constant and R the radius of the sphere.

By substituting (6.23) into (6.22), an expression for the effective aperture of a continuous excited sphere is obtained:

$$\begin{aligned} A_m &= \underbrace{\left(1+\frac{2}{kR}\right)}_{\eta_m} A \\ &\approx A \qquad \text{for} \quad \lambda \ll R \\ &= 3A \qquad \text{for} \quad \lambda = 2\pi R \end{aligned} \tag{6.24}$$

It is evident that for an electrically large sphere, its aperture is just the projected area of a circular disk. However, as the sphere becomes electrically smaller, creeping waves from portions of the

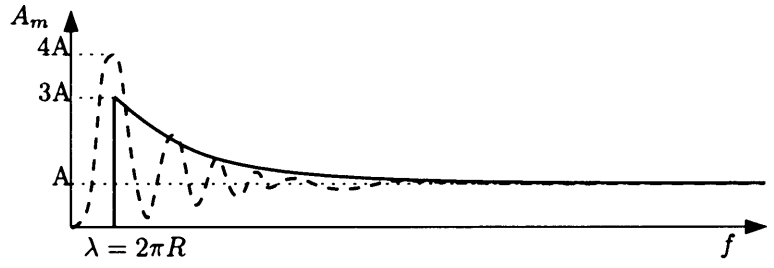


Figure 6.12: The sketch compares the typical shapes of the effective aperture A_m of a continuously excited sphere (solid line) and its RCS (dashed line), when plotted against frequency.

sphere that are beyond the horizon can contribute and the maximum achievable directivity is greater than one would expect from a sphere of that size. Normalised to the projected area, the maximum perceived aperture occurs when the wavelength equals the circumference, the condition with maximum constructive interference between direct and creeping waves. When this condition of $kR = 1$ occurs, the aperture will seem to be three times as large as the projected area. This can be compared with the well known Radar Cross Section (RCS) of sphere, as shown in figure 6.12. In the optical region the same result is obtained: $\sigma_{\text{RCS}} = \pi R^2$. In the Mie range of frequencies however the RCS behaves oscillatory as a result of resonances between incoming and reflected waves. In the case under study here no incoming waves are present, and the resonances do not show. In the Rayleigh zone, the sphere behaves as a point scatterer and its RCS varies with the 4^{th} power of frequency. This behaviour is beyond the validity region of our analysis, as there is not enough phase variation available over the spherical surface to define spherical modes. One must not be misled by figure 6.12 and conclude that the condition $\lambda = 2\pi R$ is optimal, as the projected area is still a quadratic function of R, so that larger spheres result in higher directivity, as they should. Any practical array size will have $R \gg \lambda$.

6.3.2 An Array of Point Sources

A spherical array consists of a number of radiators (array elements) placed in a spherical configuration. The continuous excitation of the sphere is then approximated by sampling it at the element positions and exciting those elements accordingly. Assuming the elements are dimensionless point sources, the spherical equivalent of the Nyquist sampling theorem can be applied to establish a relationship between the density of the sampling grid and the available spatial bandwidth, the highest degree B_s of spherical modes available to synthesise a radiation pattern. Again, we can use (6.10) to relate this number to the maximum achievable directivity. Hence, we define the sampling efficiency as:

$$\begin{aligned} \eta_s &= \frac{B_s(B_s+2)}{L_m(L_m+2)} \\ \eta_s &\leq 1 \end{aligned} \tag{6.25}$$

and the corresponding "virtual" effective area as

$$A_s = \eta_s A_m \tag{6.26}$$

At this point questions can be raised as to what would happen when the number of elements is increased so that $B_s > L_m$. With such a hyperdense distribution of elements the excitation function would be oversampled and a larger number of modes can be used than can actually travel to the far-field. Although these higher degree modes could theoretically be used to control the reactive near-field, they do not contribute to the far-field and therefore to the directivity. That is why we have limited η_s to values not exceeding 1.

To establish a relation between the distribution of the elements and the highest degree B_s of spherical modes that can be used in the radiation pattern is not as simple as with circular geometry where the Nyquist theorem applies. The relationship we seek is dependent on the definition of the element distribution, which is in itself a lot more involved than for circular arrays. We will therefore take a shortcut here and define a sensible relationship, with an extra parameter that will allow us to make corrections afterwards. We will then further analyse this relationship in a separate chapter, chapter 8.

As there are $2B_s(B_s + 2)$ modes of degree less or equal to B_s (where the factor 2 comes from the fact that there are both TE and TM modes), and we want to control all of these orthogonal modes separately, an equal number of controls, C, need to be present in the system:

$$2B_s(B_s + 2) = C \tag{6.27}$$

However, when designing an array antenna (and not, for instance, a spherical near-field antenna measurement system) one is usually only interested in controlling the far-field. The far-field is tangential to its wavefront, and usually decomposed into two orthogonal components, the co- and cross-polarisation. If it is reasonable to assume that one of these polarisations can be excited without over-spilling too much into the other polarisation, than the number of controls is reduced by half. But doing that requires in turn doubling the amount of controls. When it is also possible to relate each of two polarisations in the far-field to one of two orthogonal excitations of an array antenna, than the number of elements is effectively reduced by half, as the other set of elements is simply left out. While this is straightforward in a planar array of dipoles for instance, it is more complicated for a spherical array. This will be dealt with in more detail in section 7.1. As the definition of the polarisations are always dependent on some reference direction, they change as the array antenna is steered into a different direction. Therefore, it is not straightforward to leave out half of the controls (as with an array of dipoles) and maintain good steering capability. We will assume the number of array elements, N to be half the number of spherical modes. The question of whether this number needs to be doubled for polarisation control is answered in section 7.1.

$$N = \frac{C}{2} \tag{6.28}$$

We establish a generic sampling theorem as:

$$B_s(B_s + 2) = \eta_d N \tag{6.29}$$

where N is the number of elements and the distribution efficiency η_d, an extra parameter that we can use to adjust this relationship later on. From the number of elements we can define the average element density as

$$d = \frac{N\lambda^2}{4\pi R^2} \tag{6.30}$$

Finally, we can rewrite the reduction in directivity as a consequence of (sub-)sampling as

$$\eta_s = \eta_d \frac{Rd}{R\pi + \lambda} \tag{6.31}$$

Now, suppose we do not want to reduce the directivity by the action of sampling, or else, we want to sample with an optimal density. That means η_s should be equal to 1. Furthermore, suppose we have defined the sampling grid in a way so that we achieve a distribution efficiency of $\eta_d = 1$. We then find the average element density to be:

$$d = \pi + \frac{\lambda}{R} \quad \left[\frac{\text{elements}}{\lambda^2}\right] \tag{6.32}$$

This means that for electrically large spheres, the optimal average element density is π elements per square wavelength, less than 4 elements per square wavelength for the planar array! The total number of elements is thus roughly equal to the number of wavelengths in the circumference, squared. As we have pointed out earlier, the sphere becomes less efficient as it decreases in electrical size. And in accordance to our findings on its effective aperture, we find that the optimal density increases to a maximum of 3π when $kR = 1$.

As a final remark in this section, it must be noted that no assumptions on the single element radiation patterns or modal propagation characteristics were taken into account. The term 'point source' was used to describe the sampling action of the array antenna and stands apart from the interpretation as a source with an isotropic radiation pattern.

6.3.3 Directive Elements

The effect of using directional elements with a finite size, rather than sampling points, deserves closer attention. In future work, rules of thumb could be established for the parameter η_f to take this effect into account.

The impact of using directional elements with a finite effective area on the overall gain of the array antenna is complicated. At first sight, it can be expected that η_f be greater than one. This is expected because the sampling distribution of array elements limits the highest degree of mode that can be used without aliasing. Finite sized elements integrate the near-field over a finite area and may act as anti-aliasing filters. The greater the effective area of the array elements, and thus the higher their directivity, the less aliasing effects will occur. On the other hand, if the overall antenna gain is to be maximised in some direction, the contributions from all array elements but those pointing somewhat in that direction will be less if the elements have greater directivity. This is true because of the curvature of the spherical array. Lastly, the gain will also be reduced if the effective areas of the array elements overlap. For all these reasons

together, one may expect that:

$$\eta_f > 1 \tag{6.33a}$$

$$\eta_f \eta_s \leqslant 1 \tag{6.33b}$$

To investigate this further, the analysis tools developed in the previous chapters can be used to compare arrays with varying element directivities. Isotropic radiators could also be analysed, but it was already mentioned in the introduction that these kind of arrays have unfavourable bandwidth characteristics (due to mode instability, just as with circular arrays[82]) and that in this thesis a metal spherical body is always assumed at the core of the antenna. Using the analysis tools to compare the maximum gain of the array for different element radiation patterns departs from an important assumption: that the array excitation that maximises the directivity be known. So far, only the modal expansion was obtained that maximises the directivity. Synthesising the array excitation from that is dealt with in section 6.3.4.

6.3.4 Excitation Synthesis for Maximum Directivity

In this chapter the far-field modal excitation was derived that maximises the directivity for a certain maximum degree in the radiation pattern, but nothing was established about the excitation of the array elements to maximise the directivity. To synthesise an excitation from a desired radiation pattern, it is necessary to obtain the electromagnetic reaction of each array element on the desired radiation pattern, which is obtained by calculating a spherical convolution. More on this topic will be covered in chapter 7, where it will be shown that in the spectral domain the convolution becomes a multiplication. The spherical modal expansions obtained for single elements in chapter 3 can be used, together with the expansion of the desired radiation pattern (e.g. the maximum directivity expansion obtained in equation (6.9) of this chapter) to synthesise the excitation. This procedure has also been used for circular array antennas and is called Fourier synthesis[82].

The focus of this chapter is to propose simple design formula's or methods for spherical array antennas, and application of the more involved concepts of chapters 7 and 3 for array synthesis is out of its scope. Instead, a simple alternative procedure can be used to get the unknown array excitation that maximises the directivity, provided the element RGP is known. If the array directivity is to be maximised in a direction (θ_0, ϕ_0), and the RGP of element i is known to be $F_i(\theta, \phi)$ (assuming a single scalar polarisation component), then the excitation coefficient C_i for element i is given as[61]:

$$C_i = F_i^*(\theta_0, \phi_0) \tag{6.34}$$

which follows from the simple signal theoretical consideration that each element must maximally contribute to the pattern in the direction (θ_0, ϕ_0). To apply this procedure, it is necessary to have the RGPs of each element, which can be rotated versions of the same pattern. The usual spherical modal expansion can be used to simplify the rotation of the element patterns using the equations from section 4.2. This simple approach will be followed in chapter 9 to obtain some array patterns from a measured RGP.

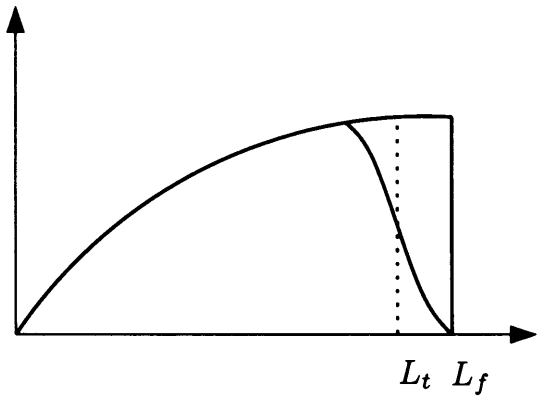


Figure 6.13: The equivalent maximum degree of the radiation pattern that is used to calculate the directivity can be adjusted to take into account a decrease in directivity as a consequence of anti-sidelobe tapering.

A more in-depth study of the effect of element gain on the overall gain of the spherical array, along the lines described in the previous paragraphs, is left for future work.

6.3.5 Coefficient Tapering Correction

A spherical expansion that is optimised to achieve maximum directivity in some direction has its expansion coefficients continuously rising as a square root function of the degree of the spherical harmonics. At the maximum achievable degree, there is a discontinuous drop to zero, as higher degree modes are not processed. This sudden drop causes high sidelobes to occur in the radiated pattern, which can be decreased by enforcing a smoother transition at the highest degrees of the expansion. This was explained in section 6.2.2. We have also seen that this slightly decreases the directivity of the antenna pattern. No separate efficiency factor was foreseen in equation (6.21) to account for this effect, but if necessary this effect can be incorporated by adjusting the equivalent degree L_f to L_t as shown in figure 6.13.

Near-field to far-field propagation from an antenna of finite size, acts as a low-pass filter on the spherical expansion of the antenna excitation. At the highest degrees, the transition from pass-band to stop-band is smooth. A rule of thumb is to take $L_m = kR + 10$ and $L_t = kR$.

6.4 Summary and Conclusions

In this chapter a number of parameters that are commonly used to describe the performance of an antenna were obtained for the spherical array antenna by placing them in the context of SMT. The focus lay on the establishment of simple design formula's, in the spirit of the third point of the hypothesis in section 1.3.1.

A detailed treatment of the polarisation definition already used in previous chapters was given and the choice for this particular polarisation was justified. The adoption of a suitable polarisation and the definition of the associated CPST in chapter 2 were of key importance to render the derivations in this chapter in their simplest form.

Attention was given to the maximum attainable directivity as well as to the beamwidth, sidelobe

levels and aperture efficiency of the spherical array antenna. Also, a preliminary sampling theorem was intuitively formulated, that will be given much more attention in chapters 7 and 8. In the margin, a simple and well known formula to synthesise a maximum directivity radiation pattern was restated.

There is scope for future work to complete the few presented formula's with more performance characteristics, such as for instance the quantification of various error terms.

Chapter 7

Spherical Array Signal Processing

In chapter 6 the coefficients of the spherical wave expansion were obtained that optimise the far-field pattern in some way, e.g. to obtain high directivity with some acceptable level of sidelobes. Once the expansion coefficients are known, they can be used to back-project the field to the surface of the spherical array. This procedure will be outlined in section 7.1.

Once the relationship between the near- and the far-field is established by backprojection, a continuous excitation is known over the array surface that corresponds to the desired far-field pattern. We have characterised this excitation as a superposition of spherical modes. The array antenna will approximate the continuous excitation with a number of separately controllable radiating elements on its surface. Each of these radiating elements can receive and/or transmit a signal carrier with a certain phase and amplitude. The combination of all these signals represents a sampled near-field around the antenna. Suppose now we have a system X with a number of input/output ports. The set of all phase and amplitude pairs for all array elements appear at the element ports of system X.

As the excitation function is defined as a superposition of spherical modes, it would be useful to have a set of input/output ports, the modal ports, at the other end of system X that each excite or listen to one specific mode. System X then transforms the signals at its modal ports to the appropriate signals at the element ports and vice versa.

Inside system X (shown in figure 7.1) the signals from the element ports are transformed into signals at the modal ports or vice versa. In this chapter we are concerned about this processing at system level (the implementation of the processing will be dealt with in chapter 9). The relation between the element ports and modal ports is governed by the spherical equivalent of a discrete Fourier transform. This allows a number of signal processing theorems and techniques to be applied on system X. For instance, the spherical equivalent of the sampling theorem can be used to relate the number of element ports to the number of modal ports. In this chapter we will apply the techniques of Fourier analysis to the signals from the spherical array.

7.1 Backprojecting the Far-field

In this section the known expansion that sythesises some desired far-field is used to obtain the near-field over a sphere that coincides with the array surface. The vectorial near-field is split into its radial component and two tangential components. As for the far-field, the vectorial

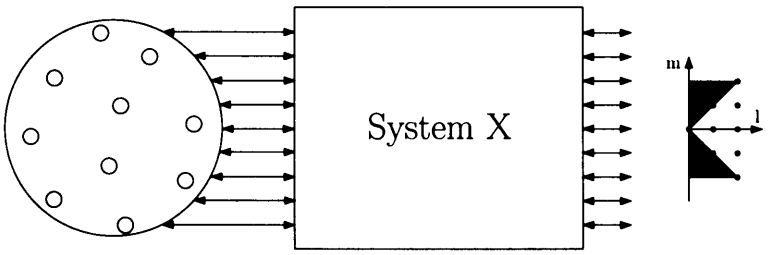


Figure 7.1: In this chapter we will study the transformation that happens inside system X. The ports on the left are the array element ports, the ports on the right are the modal ports. The graph on the right depicts the spectral domain of the spherical modes, parameterised by their degree and order.

tangential near-field can be projected on two orthogonal unity vector distributions, which for the sake of simplicity we will also refer to as polarisations. We will concentrate on the electric field, although we may as well do the same for the magnetic field.

Let $\vec{E}_{\mathrm{T}}^{\mathrm{NF}}$ be the tangential electric near-field at the radius of the spherical array and $\vec{E}^{\mathrm{FF}}$ the (tangential) electric far-field. By definition of the far-field we have:

$$\begin{aligned} \vec{E}^{\mathrm{FF}}(\theta,\phi) &= V \frac{k\sqrt{\eta}}{\sqrt{4\pi}} \frac{e^{jkr}}{kr} \vec{\mathcal{F}}(\theta,\phi) \\ &= \frac{k\sqrt{\eta}}{\sqrt{4\pi}} \frac{e^{jkr}}{kr} \sum_{lm} \left(C_{lm}^{\mathrm{TE}} \vec{\mathcal{M}}_l^m + C_{lm}^{\mathrm{TM}} \vec{\mathcal{N}}_l^m \right) \end{aligned} \tag{7.1}$$

From this and the definition for circular polarisation in (2.55) the circularly polarised components were derived in (2.56):

$$\begin{aligned} \vec{E}_{rhcp}^{\mathrm{FF}}(\theta,\phi) &= \frac{k\sqrt{\eta}}{\sqrt{4\pi}} \frac{e^{jkr}}{kr} \sum_{lm} r_{lm} R_l^m(\theta,\phi) \cdot \hat{e}_{rhcp} \qquad (7.2a) \\ &= \frac{k\sqrt{\eta}}{\sqrt{4\pi}} \frac{e^{jkr}}{kr} \cdot \mathcal{F}_{rhcp} \cdot \hat{e}_{rhcp} \\ \vec{E}_{lhcp}^{\mathrm{FF}}(\theta,\phi) &= \frac{k\sqrt{\eta}}{\sqrt{4\pi}} \frac{e^{jkr}}{kr} \sum_{lm} l_{lm} L_l^m(\theta,\phi) \cdot \hat{e}_{lhcp} \qquad (7.2b) \\ &= \frac{k\sqrt{\eta}}{\sqrt{4\pi}} \frac{e^{jkr}}{kr} \cdot \mathcal{F}_{lhcp} \cdot \hat{e}_{lhcp} \end{aligned}$$

Using equations (2.39), the far-field mapping of the spherical modes in (2.46) can be undone. Applied to the far-field in (7.1), the corresponding near-field is:

$$\vec{E}^{\mathrm{NF}}(r,\theta,\phi) = k\sqrt{\eta} \sum_{lm} \left(C_{lm}^{\mathrm{TE}} \vec{M}_l^m(r,\theta,\phi) + C_{lm}^{\mathrm{TM}} \vec{N}_l^m(r,\theta,\phi) \right) \tag{7.3}$$

This expression for the field can be used to reconstruct the field at any radius, including at the radius of the spherical array antenna, a process called backprojection. We can take the tangential part of this field:

$$\vec{E}_{\mathrm{T}}^{\mathrm{NF}}(r,\theta,\phi) = k\sqrt{\eta}\sum_{lm}\left(C_{lm}^{\mathrm{TE}}\vec{M}_l^m + C_{lm}^{\mathrm{TM}}\{\vec{N}_l^m\}_{\mathrm{T}}\right) \tag{7.4}$$

On page 53, we converted the tangential part of the SWG expansion into a form with basis functions that were independent of radius and the variation with radius was incorporated in the coefficients. The obtained expansion was called the STL expansion. We will do the same here with the expansion above, but we will work it into a form that reuses the far-field basisfunctions from equation (7.1) above, so that we can repeat the projection on each of the two polarisations, as done to obtain equation (7.2a). The STL expansion was given in equation (2.38):

$$\vec{E}_{\mathrm{T}}^{\mathrm{NF}}(r,\theta,\phi) = \sum_{lm} V_{lm}^{\mathrm{TE}}(r)\frac{\vec{e}_{lm}^{\mathrm{TE}}}{r} + \sum_{lm} V_{lm}^{\mathrm{TM}}(r)\frac{\vec{e}_{lm}^{\mathrm{TM}}}{r} \tag{7.5}$$

$$= \frac{-1}{\sqrt{4\pi}}\sum_{lm}\left(\frac{j^{l+1}V_{lm}^{\mathrm{TE}}}{r}\vec{\mathcal{M}}_l^m + \frac{j^l V_{lm}^{\mathrm{TM}}}{r}\vec{\mathcal{N}}_l^m\right) \tag{7.6}$$

where we used:

$$\vec{\mathcal{M}}_l^m = \sqrt{4\pi}(-j)^{l+1}(-\vec{e}_{lm}^{\mathrm{TE}}) \tag{7.7a}$$

$$\vec{\mathcal{N}}_l^m = \sqrt{4\pi}(-j)^{l}(-\vec{e}_{lm}^{\mathrm{TM}}) \tag{7.7b}$$

so that by comparison of (7.6) with (7.1), the circularly polarised components of the tangential near field are obtained:

$$\vec{E}_{\mathrm{T},rhcp}^{\mathrm{NF}}(r,\theta,\phi) = \frac{-1}{\sqrt{4\pi}}\left(-k\sqrt{\eta}\frac{e^{jkr}}{kr}\right)\sum_{lm} r_{lm}^{\mathrm{NF}}(r)R_l^m(\theta,\phi)\cdot\hat{e}_{rhcp} \tag{7.8a}$$

$$= \frac{-1}{\sqrt{4\pi}}\left(-k\sqrt{\eta}\frac{e^{jkr}}{kr}\right)\mathcal{F}_{rhcp}^{NF}\cdot\hat{e}_{rhcp}$$

$$\vec{E}_{\mathrm{T},lhcp}^{\mathrm{NF}}(r,\theta,\phi) = \frac{-1}{\sqrt{4\pi}}\left(-k\sqrt{\eta}\frac{e^{jkr}}{kr}\right)\sum_{lm} l_{lm}^{\mathrm{NF}}(r)L_l^m(\theta,\phi)\cdot\hat{e}_{lhcp} \tag{7.8b}$$

$$= \frac{-1}{\sqrt{4\pi}}\left(-k\sqrt{\eta}\frac{e^{jkr}}{kr}\right)\mathcal{F}_{lhcp}^{NF}\cdot\hat{e}_{lhcp}$$

with the radius-dependent near-field polarised spherical expansion coefficients defined as:

$$r_{lm}^{\mathrm{NF}} = \frac{\sqrt{2\pi}}{\left(-k\sqrt{\eta}\frac{e^{jkr}}{kr}\right)}\left(\frac{j^{l+1}V_{lm}^{\mathrm{TE}}}{r} + \frac{j^l V_{lm}^{\mathrm{TM}}}{r}\right) \tag{7.9a}$$

$$l_{lm}^{\mathrm{NF}} = \frac{\sqrt{2\pi}}{\left(-k\sqrt{\eta}\frac{e^{jkr}}{kr}\right)}\left(\frac{j^{l+1}V_{lm}^{\mathrm{TE}}}{r} - \frac{j^l V_{lm}^{\mathrm{TM}}}{r}\right) \tag{7.9b}$$

By plugging equations (2.39) into (7.9), we can relate these back to the SWG expansion coeffi-

cients:

$$r_{lm}^{\mathrm{NF}} = \sqrt{2\pi}\left(\frac{j^{l+1}\left(-\sqrt{\eta}krz_l(kr)C_{lm}^{\mathrm{TE}}\right)}{-kr\sqrt{\eta}\frac{e^{jkr}}{kr}} + \frac{j^{l}\left(-\sqrt{\eta}\frac{d}{d(kr)}\left[krz_l(kr)\right]C_{lm}^{\mathrm{TM}}\right)}{-kr\sqrt{\eta}\frac{e^{jkr}}{kr}}\right) \tag{7.10a}$$

$$= \sqrt{2\pi}\left(\frac{z_l(kr)}{(-j)^{l+1}\frac{e^{jkr}}{kr}}C_{lm}^{\mathrm{TE}} + \frac{\frac{1}{kr}\frac{d}{d(kr)}\left[krz_l(kr)\right]}{(-j)^{l}\frac{e^{jkr}}{kr}}C_{lm}^{\mathrm{TM}}\right) \tag{7.10b}$$

$$l_{lm}^{\mathrm{NF}} = \sqrt{2\pi}\left(\frac{z_l(kr)}{(-j)^{l+1}\frac{e^{jkr}}{kr}}C_{lm}^{\mathrm{TE}} - \frac{\frac{1}{kr}\frac{d}{d(kr)}\left[krz_l(kr)\right]}{(-j)^{l}\frac{e^{jkr}}{kr}}C_{lm}^{\mathrm{TM}}\right) \tag{7.10c}$$

Which in turn can be related to the far-field CPST coefficients by utilising:

$$C_{lm}^{\mathrm{TE}} = \frac{1}{2\sqrt{2\pi}}(r_{lm} + l_{lm}) \tag{7.11a}$$

$$C_{lm}^{\mathrm{TM}} = \frac{1}{2\sqrt{2\pi}}(r_{lm} - l_{lm}) \tag{7.11b}$$

We have seen that one of the polarisations can be made to vanish in the far-field by choosing the expansion coefficients to obey $C_{lm}^{\mathrm{TE}} = \pm C_{lm}^{\mathrm{TM}}$. To maximise the RHCP component for instance, the SWG expansion coefficients would be $C_{lm}^{\mathrm{TE}} = C_{lm}^{\mathrm{TM}} = \frac{r_{lm}}{2\sqrt{2\pi}}$ and the near-field polarised spherical expansion coefficients become:

$$r_{lm}^{\mathrm{NF}} = F_1(kr)\cdot r_{lm} \tag{7.12a}$$

$$l_{lm}^{\mathrm{NF}} = F_2(kr)\cdot r_{lm} \tag{7.12b}$$

with:

$$F_1(kr) = \frac{1}{2}\left(\frac{z_l(kr)}{(-j)^{l+1}\frac{e^{jkr}}{kr}} + \frac{\frac{1}{kr}\frac{d}{d(kr)}\left[krz_l(kr)\right]}{(-j)^{l}\frac{e^{jkr}}{kr}}\right) \tag{7.12c}$$

$$F_2(kr) = \frac{1}{2}\left(\frac{z_l(kr)}{(-j)^{l+1}\frac{e^{jkr}}{kr}} - \frac{\frac{1}{kr}\frac{d}{d(kr)}\left[krz_l(kr)\right]}{(-j)^{l}\frac{e^{jkr}}{kr}}\right) \tag{7.12d}$$

This equation relates the backprojected near-field circularly polarised expansion coefficients to their far-field counterparts. From the large-argument approximations for the spherical Hankel functions given in equation (2.46) it is immediately clear that the expansion coefficients do not change as long as the backprojection is on a sphere of very large radius (as would indeed be expected):

$$r_{lm}^{\mathrm{NF}} = r_{lm} \tag{7.13a}$$

$$l_{lm}^{\mathrm{NF}} = l_{lm} = 0 \tag{7.13b}$$

In order to gain some knowledge of the electromagnetic fields on the surface of the spherical antenna, we need to backproject the far-field expansion to a sphere of a practical radius, e.g. when $kR \approx 100$. The functions F_1 and F_2 from above are plotted in figure 7.2, as a function

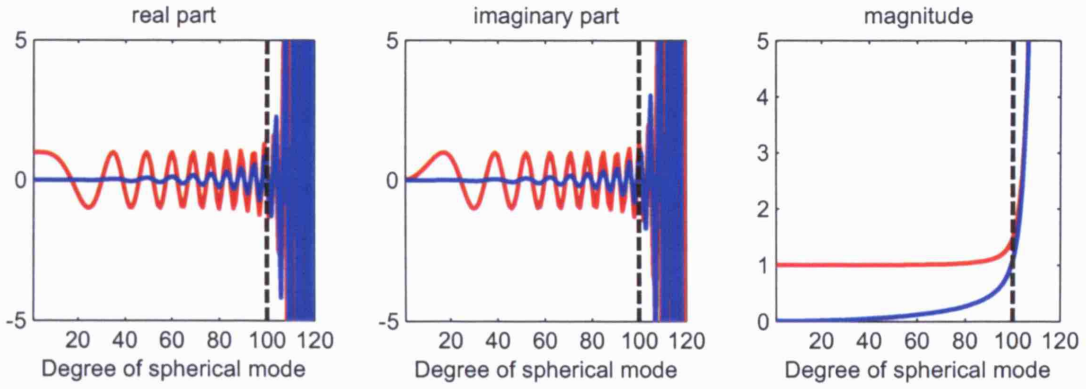


Figure 7.2: The functions $F_1(kr)$ and $F_2(kr)$ are shown, in red and blue respectively, as a function of the degree l of the backprojected component. kr is set to 100, so that l can be seen as the percentage thereof. The vertical dashed line indicates where $l = kr$. For not too high degrees, the cross-polar spherical spectrum of the near-field is negligible, whereas the co-polar spectrum is the same as that of the far-field.

of degree l. We find that the near-field expansion will be in close agreement with the far-field expansion and that the cross-polar component is negligible, provided that the electrical size of the sphere is larger than the highest degree in the expansion: $L_{max} < kR$. This is usually true for large array antennas, as L_{max} is limited by the number of array elements (this was covered in chapter 6). Figure 7.3 shows an example of a backprojected field. For a maximum degree of 10, the far-field was obtained with maximum directivity and zero cross-polarisation, as explained in section 6.2.1. The backprojected tangential near-field shows that polarisation is preserved, although one should realise that the obtained field is back-projected from the far-field, and defined only up to an arbitrary superposition of evanescent spherical modes.

In this section we obtained the electric tangential near-field from a backprojection of the desired far-field on a sphere of radius R. We have seen that the tangential field has the same polarisation as the far-field, provided that R is not too small for the required directivity or detail of the directivity function. Furthermore, an arbitrary set of higher degree evanescent modes may be superimposed on the near-field without altering the far-field. The boundary between these higher degree modes and lower degree modes is set by the electrical size of the conducting sphere. In the next section we will address how to produce the required near-field, that was given as a superposition of circularly polarised spherical modes, with a finite number of small radiators.

7.2 Spherical Sampling Theory

In chapter 2 the spherical harmonics were identified as the basisfunctions for spherical Fourier analysis. A key principle used with Fourier analysis on periodic signals is formulated by the Nyquist theorem that states that a function can perfectly be reconstructed from its samples, provided the sample interval is smaller than $\frac{1}{2B}$, where B is the Fourier bandwidth of the function. It will come at no surprise that a similar theorem can be formulated that relates the sample density on the sphere with the bandwidth in the domain of the spherical harmonics. We will

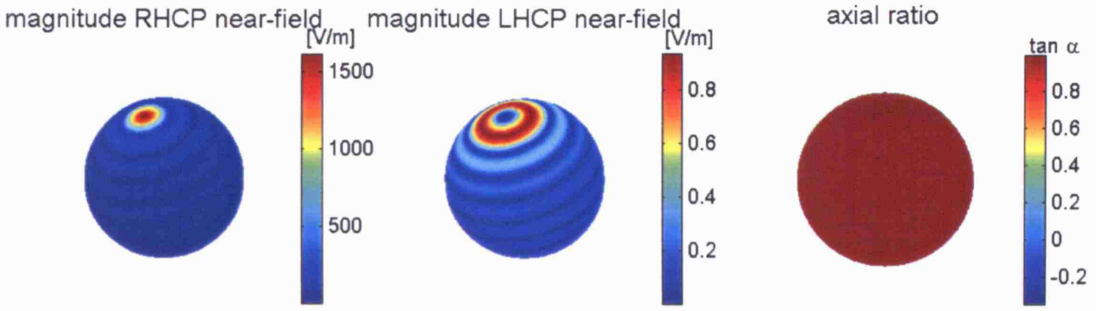


Figure 7.3: An example of a backprojected field on a sphere of radius $0.4m$ and with $f = 13GHz$, so that $kR \approx 100$. The directivity of the RHCP far-field was maximised in the direction $(\theta, \phi) = (\frac{\pi}{6}, 0)$ with zero LHCP. For these low orders, the tangential near-field replicates the far-field.

refer to this spectral bandwidth by the term spatial bandwidth, to avoid confusion with the time-frequency related Nyquist rate and bandwidth of electromagnetic signals.

7.2.1 Scalar Sampling Theorem

In this section we will formulate the scalar sampling theorem. It follows from this theorem that a bandlimited function f can be perfectly reconstructed after a sampling operation with a sample distribution of sufficiently high bandwidth. A function f is bandlimited to bandwidth B if its spherical Fourier transform $\hat{f}$ has no non-zero components of degree $l > B$. The bandwidth of a distribution will be defined further on, after we have rigorously defined what a distribution is. This is necessary because on the sphere the familiar concepts of sampling density and uniform sampling can no longer unambiguously be defined and can therefore not be used as such in a formulation of the sampling theorem. In fact, establishing the spatial bandwidth of a distribution is a research topic in its own right and we will defer an in-depth treatment to chapter 8.

A Definition of a Distribution

Before we formulate the spherical sampling theorem, we need a mathematical description for the sampling operation. A tool commonly used for this operation on periodic signals is a multiplication with a periodic Dirac pulse-train. We will use a slightly more general mathematical tool for the same purpose, that can for instance describe a set of Dirac pulses in some known points on the sphere. This tool is called a distribution or generalised function[110][111].

A distribution T_s acts on the integrable function f over S^2 as:

$$T_s(f) = \int_{S^2} s \cdot f \, d\omega \tag{7.14}$$

The function s defines a sampling measure. For instance, if $s = \delta_{\theta'\phi'}$, a Dirac pulse in (θ', ϕ'), than the distribution $T_s(f)$ will evaluate the function f in one sample point (θ', ϕ'). We can extend the concept of the Fourier expansion of a function to the Fourier expansion of a distribution. Let $\hat{s}$ be the Fourier transform of s and $\hat{T}_s$ the Fourier transform of T_s. It is easy to see

that:

$$T_s(Y_l^{m^*}) = \int_{S^2} s \cdot Y_l^{m^*} \, d\omega \tag{7.15a}$$

$$= \hat{s}(l,m) \tag{7.15b}$$

$$= \hat{T}_s(l,m) \tag{7.15c}$$

The concept of the Fourier transform of a distribution allows us to associate a spherical bandwidth with the distribution. We define a distribution of bandwidth B as a distribution with the following properties:

$$\hat{T}_{s_B}(0,0) = 1 \tag{7.16a}$$

$$\hat{T}_{s_B}(l,m) = 0 \quad \text{for} \quad 0 < l < 2B \tag{7.16b}$$

More details on distributions will be given in chapter 8.

The action of sampling a smooth continuous function f can mathematically be described as taking the product of f with a distribution T_s. The result is a distribution in itself:

$$f \cdot T_s(h) = T_s(f \cdot h) \tag{7.17}$$

B Fourier Transform of a Sampled Function

The objective is to obtain the spherical Fourier transform $\hat{f}$ of a function f, defined over S^2:

$$\hat{f}(l,m) = \int_{S^2} f(\theta,\phi) \cdot Y_l^{m^*} \, d\omega \tag{7.18}$$

The problem is that we do not know the function f, but we know a sampled version of it, which we will call f_s. Mathematically we describe the action of sampling as a product with a sampling distribution. For the sampling distribution we will assume a distribution of spherical bandwidth B:

$$f_s = f \cdot T_{s_B} \tag{7.19}$$

The sampling theorem states that $\hat{f}(l,m) = \hat{f}_s(l,m)$ for all degrees $l \leqslant B$. From (7.17) we know that f_s is a distribution:

$$f_s(h) = f \cdot T_{s_B}(h) = T_{s_B}(f \cdot h) \tag{7.20}$$

And from (7.15) we know its spherical Fourier transform:

$$\hat{f}_s(l,m) = f_s(Y_l^{m^*}) \tag{7.21}$$

so that for $l \leqslant B$ the spherical Fourier transform of f is obtained as:

$$\hat{f}(l,m) = \hat{f}_s(l,m) = T_{s_B}(f \cdot Y_l^{m^*}) = \int_{S^2} f \cdot s_B \cdot Y_l^{m^*} \, d\omega \tag{7.22}$$

C Sampling Theorem

We yet need to proof the sampling theorem itself, $\hat{f}(l,m) = \hat{f}_s(l,m)$ where the sampling distribution has bandwidth B and function f is bandlimited to B. To see this, we write the sampling distribution of bandwidth B as:

$$T_{s_B} = 1 + \sum_{\substack{l>2B \\ |m|<l}} \hat{T}_{s_B}(l,m) \cdot Y_l^m \tag{7.23}$$

Which is a direct consequence of (7.16). The sampled function f_s can than be written as:

$$f_s = f \cdot T_{s_B} \tag{7.24a}$$

$$= f\left(1 + \sum_{\substack{l'>2B \\ |m'|<l'}} \hat{T}_{s_B}(l',m') \cdot Y_{l'}^{m'}\right) \tag{7.24b}$$

$$= f + f \cdot \sum_{\substack{l'>2B \\ |m'|<l'}} \hat{T}_{s_B}(l',m') \cdot Y_{l'}^{m'} \tag{7.24c}$$

$$= f + \cdot \sum_{\substack{l\leqslant B \\ |m|<l}} \sum_{\substack{l'>2B \\ |m'|<l'}} \hat{f}(l,m)\hat{T}_{s_B}(l',m') \cdot Y_l^m Y_{l'}^{m'} \tag{7.24d}$$

$$= f + A \tag{7.24e}$$

where the term A represents the aliasing from sampling. The spherical Fourier transform becomes:

$$\hat{f}_s(l,m) = \hat{f}(l,m) + \hat{A}(l,m) \tag{7.25}$$

The aliasing term is the sum of products of spherical harmonics. We can use the Clebsch-Gordon decomposition to expand these products in a spherical harmonic decomposition. It is known that the degrees L of the harmonics in such a decomposition follow the triangle rule:

$$|l - l'| \leqslant L \leqslant l + l' \tag{7.26}$$

Therefore, the aliasing term has no spectral content of degree $l = B$ or lower so that the spectra of the sampled and the original function are the same up to degree B. it follows that the function f can be reconstructed from the sampled function f_s by bandlimiting the sampled function to B. On the sphere, this will amount to a convolution, but we will return to that in section 7.4.

7.2.2 Polarised Modes

So far we have formulated a scalar sampling theorem for the scalar spherical harmonic transform. However, an electromagnetic wave is represented by a vector field, and needs to be sampled with a distribution of unit vectors in order to obtain scalar data that can be processed further. We could sample the radial component of the near-field with radially oriented dipoles. This would amount to taking the scalar product of the radial near-field with a set of radially oriented vectors. The result of that operation is a sampled scalar function that can be expanded in

scalar harmonics. Although this approach may be feasible, the relationship between the radial near-field and the polarisation of the tangential far-field does not seem straightforward and we have not investigated this further. To capture a field with a certain polarisation, it makes more sense to sample the near-field with a distribution of tangential vectors. If nothing is known a priori about the polarisation of the electromagnetic field, than two orthogonal vectors will sense the field in each sampling point. Without loss of generality, we will assume we have projected the far-field (and as we have demonstrated also the near-field) on the polarisation of interest, say the right-hand circular polarisation. The near-field can then be sampled with a distribution of vectors of the same polarity. We define such a distribution as:

$$\vec{S}_{rhcp} = T_s \cdot \hat{e}_{rhcp} \tag{7.27}$$

where T_s is a scalar distribution of sample points as before. In section 7.1 we have written both the far-field (equation (7.2a)) and the tangential near-field (equation (7.8)) as an expansion of the polarised modes R_l^m. We will now use this expanded near-field and sample it with the vectorial distribution from above. If we define the continuous near-field $\vec{\mathcal{F}}^{NF}$ as:

$$\vec{\mathcal{F}}^{NF} = \frac{\vec{E}^{\text{NF}}(r,\theta,\phi)}{\frac{-1}{\sqrt{4\pi}}\left(-k\sqrt{\eta}\frac{e^{jkr}}{kr}\right)} \tag{7.28}$$

we get for the sampled near-field:

$$\begin{aligned}
\vec{\mathcal{F}}^{NF} \cdot \vec{S}_{rhcp} &= \mathcal{F}^{NF}_{rhcp} \cdot T_{s_B} \\
&= \mathcal{F}^{NF}_{rhcp}\left(1 + \sum_{\substack{l'>2B \\ |m'|<l'}} \hat{T}_{s_B} Y_{l'}^{m'}\right) \\
&= \mathcal{F}^{NF}_{rhcp} + \mathcal{F}^{NF}_{rhcp} \sum_{\substack{l'>2B \\ |m'|<l'}} \hat{T}_{s_B} Y_{l'}^{m'} \\
&= \mathcal{F}^{NF}_{rhcp} + \sum_{\substack{l\leqslant B \\ |m|<l}} r^{NF}_{lm} R_l^m \sum_{\substack{l'>2B \\ |m'|<l'}} \hat{T}_{s_B} Y_{l'}^{m'} \\
&= \mathcal{F}^{NF}_{rhcp} + \sum_{\substack{l\leqslant B \\ |m|<l}} \sum_{\substack{l'>2B \\ |m'|<l'}} r^{NF}_{lm} \hat{T}_{s_B} \cdot R_l^m Y_{l'}^{m'}
\end{aligned} \tag{7.29}$$

Again we obtain that the sampled function is the sum of the function itself and an aliasing term. As for the scalar modal functions, we have $\left\langle Y_{l'}^{m'} R_{l''}^{m''}, R_l^m \right\rangle \neq 0$ only for $|l''-l'| \leqslant l \leqslant |l''+l'|$. Hence, the aliasing term has only modal functions with a degree $l > B$, so that by filtering out those higher degree modes, the original near-field can be perfectly reconstructed from its samples.

In section 2.4.4 the polarised expansion was rewritten so that it could be obtained as a recombination of a scalar spherical expansion. The highest degree of the scalar modes needed in the computation was the highest degree in the polarised expansion plus one. This intermediate cal-

culations however have no implications on the required spherical bandwidth of the expansion, that remains unaffected by this observation.

7.3 The Sampling Action of an Array Antenna

In chapter 6 an available spatial bandwidth B_s was defined as the highest degree that can be used in an aliasing-free far-field pattern of an array antenna.

In section 7.2 it was shown that the sampling operation replicates the spectral coefficients of the sampled function at higher degrees, depending on the distribution bandwidth. We define the sampling degree L_s as twice the bandwidth of the sampling distribution, or else, the period with which the spectrum is repeated due to the sampling action. This definition takes an information theoretical point of view, irrespective of electromagnetic propagation properties and L_s is by definition independent of electromagnetic properties. B_s on the other hand, must be dependent of the wavelength because as the wavelength gets shorter, the radiating elements will need to be put closer together to avoid aliasing effects in the radiation pattern.

The relation between L_s and B_s needs to be established in order to relate the required number of array elements via the sampling theorem to the electrical size of the array. With a sufficient number of radiating elements, the array is usually operated to achieve maximum directivity, so that

$$B_s = \frac{L_s}{2} \tag{7.30}$$

Which was also the assumption taken in section 6.3.2 to obtain the required element density for spherical array antennas. This need not always be the case however, and a more general relation between B_s and L_s must be put in place.

As it is L_s that is governed by the sampling distribution, it is more appropriate to reformulate the generic sampling theorem in equation (6.29) using equation (7.30) as:

$$\frac{L_s}{4}(L_s + 4) \;=\; \eta_d N \tag{7.31}$$

If the spherical array antenna is to achieve maximum directivity for a given size, than $\eta_s = 1$ in equation (6.25) and $B_s = L_m \approx kR$. In a receiving regime, this is equivalent to the correct reconstruction of the field that results from a point source at infinity: the superposition of an incident plane wave and a wave reflected of the conducting sphere. As an example, consider a spherical array antenna of 80 cm diameter operating at 3 GHz. We have that $kR = 25$ so that $B_s = 25$. We therefore need a distribution of the same bandwidth, and the sampling degree is $L_s = 2B_s = 50$. The number of radiating elements can be estimated from the generic sampling theorem in equation (7.31) with a distribution efficiency $\eta_d = 1$. We get $\frac{L_s}{4}(L_s + 4) = 675$. Hence, the minimal number of radiating elements required to get the most out of an antenna of that size is 675. It should be born in mind that in reality $\eta_d < 1$ so that even more elements are needed. But this will be the subject of chapter 8. The maximum directivity that can be achieved with an array of that size is obtained from equation (6.10) and given by $D_s = 20\log_{10}(675) \approx 57dBi$.

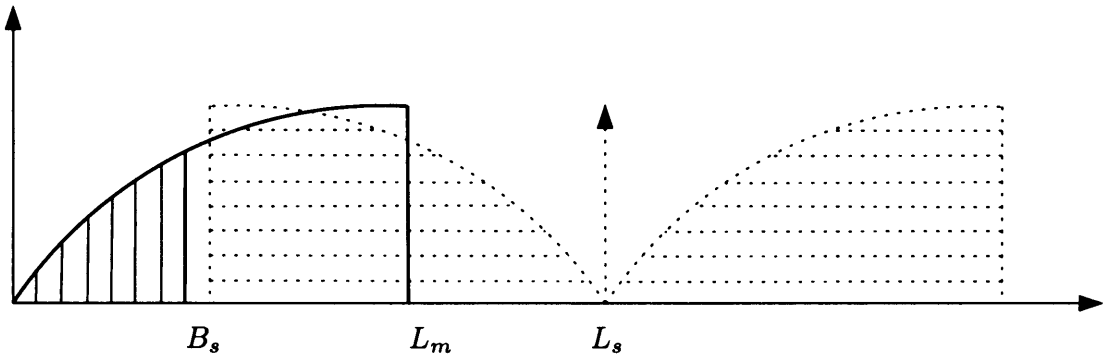


Figure 7.4: Aliasing in the higher degree modes above B_s is tolerated and filtered out after sampling. By allowing the aliasing to occur and being content with a decrease in directivity, the array can be constructed with less radiating elements. The 1-dimensional spectral domain is for an axially symmetric function around the z-axis, but the principle holds true for rotated versions of the pattern.

It is possible however to use less radiating elements in the array and allow some aliasing to occur in the higher degree modes. If these modes are not used or filtered out, a lower degree radiation pattern of less directivity can still be formed without aliasing effects. This is illustrated in figure 7.4. In that case, the spatial bandwidth of the radiation pattern B_s is smaller than L_m. The higher degree modes of the wave existing outside the array antenna are aliased in the radiation pattern for modes of a degree above $L_s - L_m$ so that the requirement for B_s becomes:

$$B_s \leqslant L_s - L_m \tag{7.32}$$

The reconstruction is equivalent to low-pass filtering after the sampling operation and removes the aliased higher degree modes. The low-pass filtering needs to remove all degrees above B_s which reduces the directivity of the pattern. For instance, for an array of the same size as before, 240 array elements could be used instead. Again using equation (7.31) and with the same note regarding the distribution efficiency, we obtain $\frac{L_s}{4}(L_s + 4) = 240$ from which $L_s = 29$. Therefore, by equation (7.32), we obtain for the available bandwidth: $B_s = 29 - 25 = 4$ and the maximum achievable directivity becomes $D_s = 20\log_{10}(4 \times 6) \approx 28dB$.

Although spherical array antennas are effective in terms of the required element density, in absolute number they typically will have a large number of elements. As the cost of an array antenna is mainly governed by the number of channels to the active elements, it may be necessary to keep the costs down by reducing the number of element channels, while size of the antenna may be dictated by other considerations. Allowing aliasing to occur in the higher degree modes and filtering these out as described above is one approach to achieve this. Another approach that can be taken is to filter out the high degree modes from the field that exists outside the antenna prior to sampling it. A spatial anti-aliasing filter is convoluted with the field prior to the sampling operation. This filter acts as a low-pass filter on the modes that can be radiated from or received by the array antenna. The array elements can achieve this by spatially integrating the field. To describe this mathematically the mechanisms in section 7.4 are required, but the principle is illustrated in figure 7.5.

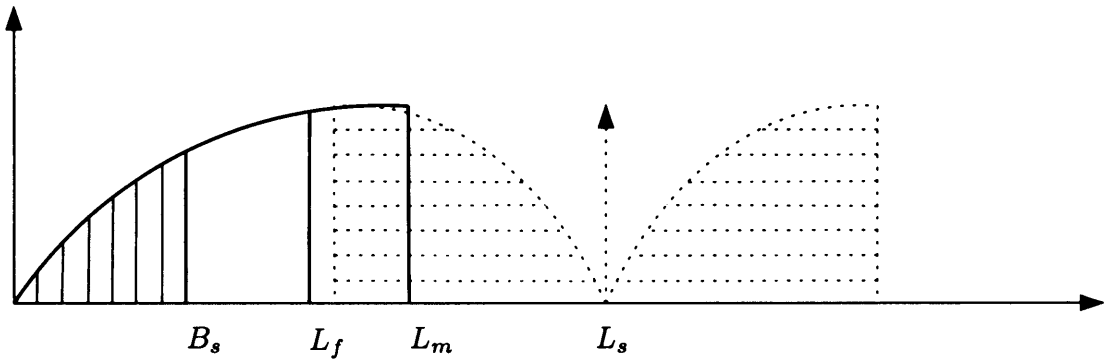


Figure 7.5: Filtering by spatial integration before sampling reduces the bandwidth of the function to be sampled and can further reduce the number of array elements.

In practise the radiating elements will always have some directivity, and therefore an associated effective area on the surface of the sphere. The spatial filtering will therefore always be present, intentionally or not. In section 6.3.3 this effect was incorporated in the efficiency factor η_f. As already mentioned there, quantification of this factor is left for future work and an accurate prediction of the directivity of a complex spherical array antenna of high-gain elements still requires the analysis procedures that were presented in chapter 5.

7.4 Spherical Convolution and Filtering

In this section the concept of spatial filtering on the sphere is described, starting from the formulation of spherical convolution. In chapter 8 we will use the concepts developed here to look into anti-aliasing filtering for the spherical array antenna. A standard reference for (scalar) convolution on the sphere is [54]. Other work on spherical convolution and signal processing in general can be found in the following papers on spherical microphone arrays: [29][112][40][41][113][114].

Before continuing, the reader's attention must be drawn to the following point: The scalar spherical convolution will be presented and it will be shown how it can be efficiently computed using the SFT. Throughout this text we have used a multipole SWE definition for the left-hand and right-hand circularly polarised far-field functions R_l^m and L_l^m in the CPST, that should not be substituted for the scalar functions used in the SFT. The reason for this is that the basis functions we used to define the scalar far-fields in the equations (2.56) are not compatible with the standard spherical harmonics used for scalar spherical functions, and the obtained coefficients would not be the r_{lm} and l_{lm} defined there. Even the bandwidth of the projection is not preserved if for instance R_l^m were to be projected on the scalar harmonics Y_l^m, hence it is really not advisable to do so, even to roughly estimate the spatial bandwidth of a given radiation pattern. The formalism proposed in section 2.4.4 can be used to obtain the multipole coefficients r_{lm} and l_{lm} with the standard SFT. Spherical convolutions of vector-fields and efficient computation thereof using spherical multipole expansions are presented in 7.4.4.

7.4.1 Convolution and Cross-Correlation of Complex Functions on $\mathbb{R}$

The mathematical concepts that are needed for spatial filtering on the sphere are the spherical equivalents of convolution and cross-correlation. Before making the extension to the spherical domain, we will briefly review these concepts for complex functions defined on the field of real numbers $\mathbb{R}$.

The convolution of two functions f and h is defined as:

$$(f * h)(t) = \int f(\tau)h(t-\tau)d\tau \tag{7.33}$$

Before taking the integral of the products, one function is reversed and shifted. If these functions are defined on a periodic domain, the convolution is cyclic and the shifting operation becomes a rotation on a circle. This already gives some insight in how we can generalise the concept of convolution to spherical domains.

Closely related with the convolution is the concept of cross-correlation between two functions. It is defined as:

$$(f \star h)(t) = \int f^*(\tau)h(t+\tau)d\tau \tag{7.34}$$

Cross-correlation and convolution are indeed very similar and the relation is given by:

$$f(t) \star h(t) = f^*(-t) * h(t) \tag{7.35}$$

7.4.2 Convolution and Cross-Correlation on the Sphere

In section 2.2 we defined the rotation operator on square integrable functions in $L^2(S^2)$ as:

$$\Lambda(g)f(\omega) = f\left(g^{-1}\omega\right) \tag{7.36}$$

The definition of spherical convolution is obtained by replacing the shifting operation in equation (7.33) with the rotation operator above. Let η be the north pole, $\eta = (0,0,1)$ and dg an elementary rotation in $SO(3)$, $dg = \sin\theta d\theta d\phi d\psi$. The right convolution of f with h on the sphere is often defined as:

$$(f * h)(\omega) = \int_{g \in SO(3)} f(g\eta)\left(\Lambda(g)h(\omega)\right) dg \tag{7.37}$$

which simplifies to:

$$(f * h)(\omega) = \int_{g \in SO(3)} f(g\eta)h(g^{-1}\omega)dg \tag{7.38}$$

Note the asymmetry in the theorem - only the zero-order coefficients of $\hat{h}$ are used for the right convolution. This can be explained with the following argument: g is a rotation in $SO(3)$ and is itself a function of 3 independent parameters, the 3 Euler angles (ξ, θ, ϕ). Indeed, the position or rotation of a point on the sphere is sufficiently described by $\omega(\theta, \phi)$, which is a function of two Euler angles only, but the position or rotation of a function or vector on the sphere needs an extra parameter to define it uniquely. We can see that this definition can be problematic.

The output of the convolution is defined as a function of ω, or on the surface of the sphere. The input function $f(g\eta)$ however is clearly defined as a function of g. The variable $g\eta$ maps all rotations g to a point on the sphere, and the integration over the third Euler angle averages all positions $h(g^{-1}\omega)$. Stated otherwise, the convolution kernel h is forced into a shape with azimuthal symmetry. If h would not be axially symmetric around the z-axis and we would not want to average that out, than we need a third Euler angle to describe its position on the sphere. This is typically necessary when the functions represent one polarisation of an electromagnetic field. In that case, the definition of the polarisation can be used to write the third Euler angle in function of the first two and the convolution would be defined on a 2-dimensional manifold in a 3-dimensional space of Euler angles. If h is taken to be axially symmetric, the third Euler angle is irrelevant while h is taken round the sphere by the convolution. This results in the left and right convolutions to be different.

An alternative definition for the convolution is:

$$(f * h)(g) = \int_{S^2} f(\omega)\,(\Lambda(g)h(\omega))\,d\omega \tag{7.39}$$

This definition does not suffer from the previous limitation.

The Cross-correlation on the sphere can similarly be defined as:

$$(f \star h)(g) = \int_{S^2} f(\omega)\,(\Lambda(g)h(\omega))^*\,d\omega \tag{7.40}$$

7.4.3 The Fourier Transform of the Spherical Convolution

As with circular Fourier transforms, a convolution theorem can be formulated that states that the spherical Fourier transform of a spherical convolution can be written as a pointwise product of the transforms. Again, care must be taken in the interpretation of this statement. With the first definition of convolution in equation (7.37), the result is defined over the sphere and the SFT can be calculated using the definition for the SFT in (2.2). The second definition in equation (7.39) is defined over the rotations in $SO(3)$. Hence, the SFT as we defined it in this text cannot be used on this function. Luckily a Fourier transform exists on $SO(3)$ that can be used here. Its basisfunctions are the Wigner-D functions[56]. We will simplify the result afterwards to the case of an axially symmetric filter function h.

We can use the SFT in (2.2) to decompose the function h and use equation (2.6) to rotate all the harmonic components individually. Recognising the SFT on f, the cross-correlation can be

rewritten as:

$$(f \star h)(g) = \int_{S^2} f(\omega) \left(\Lambda(g) h(\omega)\right)^* d\omega \tag{7.41a}$$

$$(f \star h)(g) = \int_{S^2} f(\omega) \left(\Lambda(g) \sum_{l \in \mathbf{N}} \sum_{|m| \leqslant l} \hat{h}(l,m) Y_l^m(\omega) \right)^* d\omega \tag{7.41b}$$

$$(f \star h)(g) = \sum_{l \in \mathbf{N}} \sum_{|m| \leqslant l} \sum_{|k| \leqslant l} \hat{h}^*(l,m) D_{k,m}^{(l)^*}(g) \int_{S^2} f(\omega) Y_l^{k^*}(\omega) d\omega \tag{7.41c}$$

$$(f \star h)(g) = \sum_{l \in \mathbf{N}} \sum_{|m| \leqslant l} \sum_{|k| \leqslant l} \hat{f}(l,k)\, \hat{h}^*(l,m)\, D_{k,m}^{(l)^*}(g) \tag{7.41d}$$

In the last line we recognise the pointwise product of the spherical Fourier transform of f and h, with the Wigner-D Fourier basis functions for $SO(3)$.

Now if the kernel function h has axial symmetry, it will only have spherical harmonic components for which $m = 0$. In that case the output can be further simplified, based on the following expression ([63], appendix A2):

$$d_{m0}^l(\theta) = \sqrt{\frac{4\pi}{2l+1}} N_l^m P_l^m(\theta) \tag{7.42}$$

where the normalisation constant N_l^m was defined in (2.13). It follows from the definition of the Wigner-3D function in (4.5) that:

$$\begin{aligned} D_{m,0}^{(l)}(g) &= \sqrt{\frac{4\pi}{2l+1}} N_l^m P_l^m(\theta) e^{-jm\phi} \\ D_{m,0}^{(l)^*}(g) &= \sqrt{\frac{4\pi}{2l+1}} Y_l^m(\theta,\phi) \end{aligned} \tag{7.43}$$

With this simplification the cross-correlation becomes:

$$(f \star h)(\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \sqrt{\frac{4\pi}{2l+1}} \hat{f}(l,m) \hat{h}^*(l,0) Y_l^m \tag{7.44}$$

Figure 7.7 illustrates the spherical correlation of two scalar hat-functions defined on the sphere.

The derivation for the spherical convolution proceeds along the same lines (it also makes use of the fact that $N_l^m P_l^m = (-1)^m N_l^{-m} P_l^{-m}$) and results in:

$$(f * h)(\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \sqrt{\frac{4\pi}{2l+1}} \hat{f}(l,m) \hat{h}(l,0) Y_l^m \tag{7.45}$$

In the domain of the spherical coefficients, the convolution is a simple pointwise multiplication:

$$(f * h)_{lm} = y_{lm} = \sqrt{\frac{4\pi}{2l+1}}\, \hat{f}(l,m) \hat{h}(l,0) \tag{7.46}$$

In [54] a proof is directly given for a convolution theorem based on the definition in (7.37).

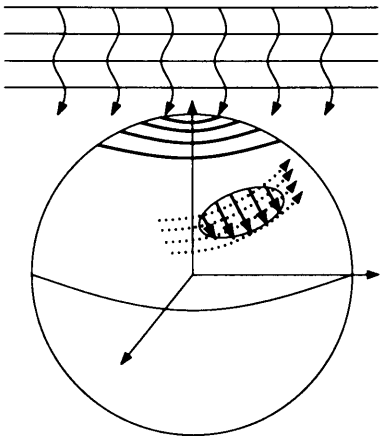


Figure 7.6: The figure illustrates the following problem formulation: A vectorial field, consisting of one or two polarisations is impinging on the sphere. On the surface of the sphere, the tangential part of the impinging field can be decomposed in two tangential polarisations, that are defined over the surface of the sphere. This could for instance be the polarisation as defined in equation (2.56) and illustrated in figure 2.4. In this figure, the polarisation definition is illustrated with the dotted lines. Now we can define a second vectorial function of the sphere, for instance in the shown elliptical patch, that senses the tangential part of the impinging field. We will call this the filter-field. In order to calculate the linear projection of the impinging field on the fieldlines of the filter-field, the position of the filter-field needs to be known with its 3 Euler angles. In section 7.4.4 it is demonstrated how the concept of spherical convolution can be extended to the vectorial case described here.

There an additional factor of 2π is found, because the integration is taken over $SO(3)$ and the integration over the third Euler angle results in an averaged $\int_0^{2\pi} d\psi = 2\pi$ factor. Therefore, in the literature one often encounters the following convolution theorem:

$$\widehat{f * h}(l,m) = 2\pi\sqrt{\frac{4\pi}{2l+1}}\hat{f}(l,m)\hat{h}(l,0) \tag{7.47}$$

7.4.4 Convolution of Tangential Vectorfields on the Sphere

In order to deal with the situation where an electromagnetic field impinges on a sphere and is subsequently sensed by a number of polarised antennas or subarrays of antennas, it is necessary to have a mathematical tool that is equivalent to the spherical convolution, but able to cope with vectorial fields. Figure 7.6 illustrates the problem of vectorial convolution on the sphere. Two tangential vectorial functions are considered, with fieldlines on the surface of the sphere. These fieldlines do not necessarily line up with the fieldlines of the chosen polarisation definition. Even if one of the vector functions is defined in some preferred reference system and with some preferred polarisation definition, the other vector function is rotated with all rotations g before the inner product of the two functions is taken. The rotations g contain three Euler angles, and the rotated vector function, which we will call the filter-field, is generally not aligned with the chosen polarisation definition. This raises the question of whether it is possible to work with two polarisations separately, without mixing them up in the vectorial convolution. With circularly polarised expansions this is indeed possible, as will be shown hereafter.

We will assume the same polarisation definition as defined in equation (2.56) and illustrated in

figure 2.4. We will further assume one vectorfield $\vec{F} = F_{rhcp}(\theta,\phi)\cdot\hat{e}_{rhcp}$ that is fully right-hand circularly polarised. The extension to a (partially) left-hand polarised field is straightforward. Firstly, we define the vectorial inner product of $\vec{F}$ with $\vec{H}$ as:

$$\langle\vec{F},\vec{H}\rangle \;=\; \int_{S^2}\vec{F}(\theta,\phi)\cdot\vec{H}^*(\theta,\phi)\,dS \tag{7.48}$$

With $\vec{F}$ purely right-hand circularly polarised, this becomes:

$$\langle\vec{F},\vec{H}\rangle \;=\; \int_{S^2}(F_{rhcp}(\theta,\phi)\hat{e}_{rhcp})\cdot(H_{rhcp}(\theta,\phi)\hat{e}_{rhcp}+H_{lhcp}(\theta,\phi)\hat{e}_{lhcp})^*\,dS \tag{7.49a}$$

$$\langle\vec{F},\vec{H}\rangle \;=\; \int_{S^2}F_{rhcp}(\theta,\phi)\cdot H^*_{rhcp}(\theta,\phi)\,dS \;=\; \langle F_{rhcp},H_{rhcp}\rangle \tag{7.49b}$$

The vectorial inner product can thus readily be written in scalar form, provided that the two functions are defined in the same reference system and with the same polarisation definition. As the polarisations are orthogonal, the inner product contains no cross-polarisation between the vector functions. In order to generalise the definition of convolution on the sphere to vector functions, we will therefore need to be able to refer a vector function in a rotated reference system to the reference system of the other function. To do that, we will again make use of an expansion in a rotationally invariant basis. We have done this already extensively in chapter 4 where we used the so-called STL expansion for this purpose. Here however we would like to make use of the farfield expansions defined for the circularly polarised spherical transform in equations (2.56). With equation (7.49a) we only need to consider the scalar expansions of the F_{rhcp} and H_{rhcp} components, given by:

$$F_{rhcp} \;=\; \sum_{lm} f_{lm}R_l^m(\theta,\phi) \tag{7.50a}$$

$$H_{rhcp} \;=\; \sum_{lm} h_{lm}R_l^m(\theta,\phi) \tag{7.50b}$$

with the R_l^m defined in equation (2.56g). As for the STL expansion in (4.4) or for the scalar basis functions in (2.6), we can use the Wigner 3D functions to rotate the circularly polarised expansion of H:

$$\Lambda(g)H_{rhcp}(\theta,\phi) \;=\; \sum_l\sum_{|m|\leqslant l}\sum_{|k|\leqslant l}\left(D^l_{km}h_{lk}\right)R_l^k(\theta,\phi) \tag{7.51}$$

The derivation of the scalar spherical correlation on the sphere as outlined in equations (7.41a) can be repeated now to correlate two tangential vector fields on the sphere. The only difference are the basis functions used, and the definition for the vectorial correlation is obtained by replacing the Y_l^m with the R_l^m. This is necessary as otherwise a varying kernel $[\Lambda(g)H_{rhcp}]\hat{e}_{rhcp}$ is rotated instead of $\Lambda(g)[H_{rhcp}\hat{e}_{rhcp}]$. Still assuming $\vec{F}$ to be right-hand circularly polarised,

we obtain:

$$(\vec{F} \star \vec{H})(g) = \int_{S^2} F_{rhcp}(\omega) \left(\Lambda(g) H_{rhcp}(\omega)\right)^* d\omega \tag{7.52}$$

$$(\vec{F} \star \vec{H})(g) = \sum_{l \in \mathbf{N}} \sum_{|m| \leqslant l} \sum_{|k| \leqslant l} f_{lk}\, h^*_{lm}\, D^{(l)^*}_{k,m}(g) \tag{7.53}$$

Note that this definition is defined over all rotations g in $SO(3)$ as indeed it should. These rotations are a function of 3 Euler angles. In the scalar case, we could simplify the convolution and correlation further by assuming an axially symmetric filter function. Here we will also make the following simplifying assumptions:

- H_{rhcp} is axially symmetric around the z-axis.

- The 3-dimensional domain of rotations $g(\xi_0, \theta_0, \phi_0)$ is reduced to a 2-dimensional manifold by writing one of the 3 Euler angles (ξ_0) in function of the two others (θ_0 and ϕ_0). This is done by a parametrisation of the polarisation definition, so that the vector in the centre of the rotated filter-field is locally aligned with the polarisation vector.

The parametrisation for Ludwig's third polarisation, defined in (2.55), is simply:

$$\begin{aligned} \theta_0 &= \theta \\ \phi_0 &= \phi \\ \xi_0 &= -\phi \end{aligned} \tag{7.54}$$

where ξ_0, θ_0 and ϕ_0 are the Euler angles that completely define the rotation g and θ and ϕ define a position on the sphere.

From the axial symmetry of H_{rhcp} it follows that only those basis functions for which $m = 1$ are included in its pattern. This follows from the definition of the R^m_l on page 58. Therefore, the cross-correlation in equation (7.53) simplifies to

$$(\vec{F} \star \vec{H})(g) = \sum_{l \in \mathbf{N}} \sum_{|m| \leqslant l} f_{lm}\, h^*_{l1}\, D^{(l)^*}_{m,1}(g) \tag{7.55}$$

Again we use the definition of the Wigner-3D function as defined in equation (4.5):

$$D^l_{m,1}(\xi_0, \theta_0, \phi_0) = e^{-j\xi_0} d^l_{m1}(\theta_0) e^{-jm\phi_0} \tag{7.56}$$

With the parametrisation in (7.54) this becomes:

$$D^l_{m,1}(\theta, \phi) = d^l_{m1}(\theta) e^{-j(m-1)\phi} \tag{7.57}$$

In appendix A2 of [63] the following equations can be found:

$$d^l_{1m}(\theta) + d^l_{-1m}(\theta) = -\frac{2}{\sqrt{l(l+1)}}\frac{m d^l_{0m}(\theta)}{\sin(\theta)} \tag{7.58a}$$

$$d^l_{1m}(\theta) - d^l_{-1m}(\theta) = -\frac{2}{\sqrt{l(l+1)}}\frac{d}{d\theta} d^l_{0m}(\theta) \tag{7.58b}$$

Summing, we obtain

$$d^l_{1m}(\theta) = -\frac{1}{\sqrt{l(l+1)}}\left(\frac{m d^l_{0m}(\theta)}{\sin(\theta)} + \frac{d}{d\theta} d^l_{0m}(\theta)\right) \tag{7.59}$$

Using $d^l_{m,1} = (-1)^{m+1} d_{1,m}$ and $d^l_{0,m} = (-1)^m d^l_{m,0}$ as well as equation (7.42) and the definition of N_{lm} in (2.35), we obtain:

$$D^l_{m,1}(\theta,\phi) = (-1)^m \frac{1}{\sqrt{l(l+1)}}\left(\frac{m d^l_{0m}(\theta)}{\sin(\theta)} + \frac{d}{d\theta} d^l_{0m}(\theta)\right) e^{-j(m-1)\phi} \tag{7.60a}$$

$$D^l_{m,1}(\theta,\phi) = \sqrt{\frac{4\pi}{2l+1}} N_{lm}\left(\frac{m P^m_l(\theta)}{\sin(\theta)} + \frac{d}{d\theta} P^m_l(\theta)\right) e^{-j(m-1)\phi} \tag{7.60b}$$

so it follows from the definition of R^m_l on page 58 that:

$$D^{(l)^*}_{m,1}(\theta,\phi) = \sqrt{\frac{4\pi}{2l+1}} j^l R^m_l(\theta,\phi) \tag{7.61}$$

which simplifies the cross-correlation in (7.55) to:

$$(\vec{F} \star \vec{H})(g) = \sum_{l\in\mathbb{N}} \sum_{|m|\leqslant l} \sqrt{\frac{4\pi}{2l+1}} j^l f_{lm}\, h^*_{l1} R^m_l(\theta,\phi) \tag{7.62}$$

and the convolution to:

$$(\vec{F} * \vec{H})(g) = \sum_{l\in\mathbb{N}} \sum_{|m|\leqslant l} \sqrt{\frac{4\pi}{2l+1}} j^l f_{lm}\, h_{l1} R^m_l(\theta,\phi) \tag{7.63}$$

In the domain of the spherical coefficients, the convolution is a simple pointwise multiplication:

$$(\vec{F} * \vec{H})_{lm} = r_{lm} = \sqrt{\frac{4\pi}{2l+1}} j^l\, f_{lm}\, h_{l1} \tag{7.64}$$

Figure 7.8 illustrates the spherical correlation of two vectorial tangential circularly polarised hat-functions, defined on the sphere.

7.5 Direction Finding with Matched Filtering

Having defined the spherical cross-correlation we can proceed to have a look at spatial filters, defined over the sphere. With the convolution theorem we have seen how the cross-correlation can be calculated with simple pointwise products in the spherical harmonic domain. The fast-

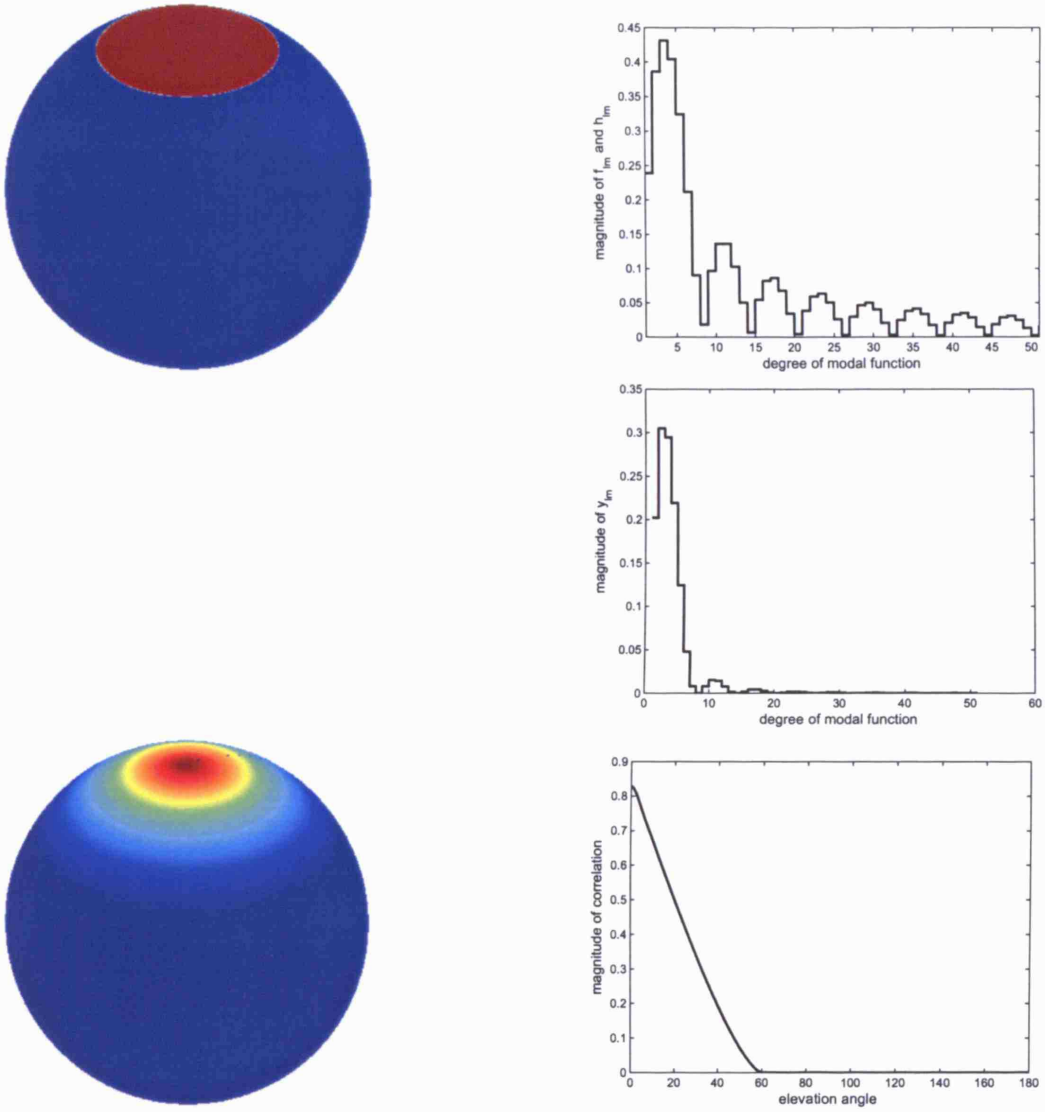


Figure 7.7: A spherical cap that extends from the NP to an elevation of $30°$ is correlated with itself by a pointwise multiplication of its spherical harmonic coefficients. The spherical cap is shown in the top left plot, its spherical harmonic coefficients on the right of it. The latter can be shown in a two-dimensional plot because the cap is axially symmetric and therefore has only non-zero coefficients for $m = 0$. The convolution is calculated in the domain of the spherical harmonics, the result is shown on the second row. By calculating the inverse spherical transform the result is obtained in the spatial domain and shown on the third row. The plot on the right plots the correlation as a function of elevation. As the cap is the equivalent of a hat-function on the sphere, we expect the correlation to be triangular in shape. The slight smoothing of the triangle is the result of cutting of the higher degree modes at $l = 50$. The results illustrate the correlation of two scalar functions, similar results are obtained for the circularly polarised far-field functions as illustrated in figure 7.8.

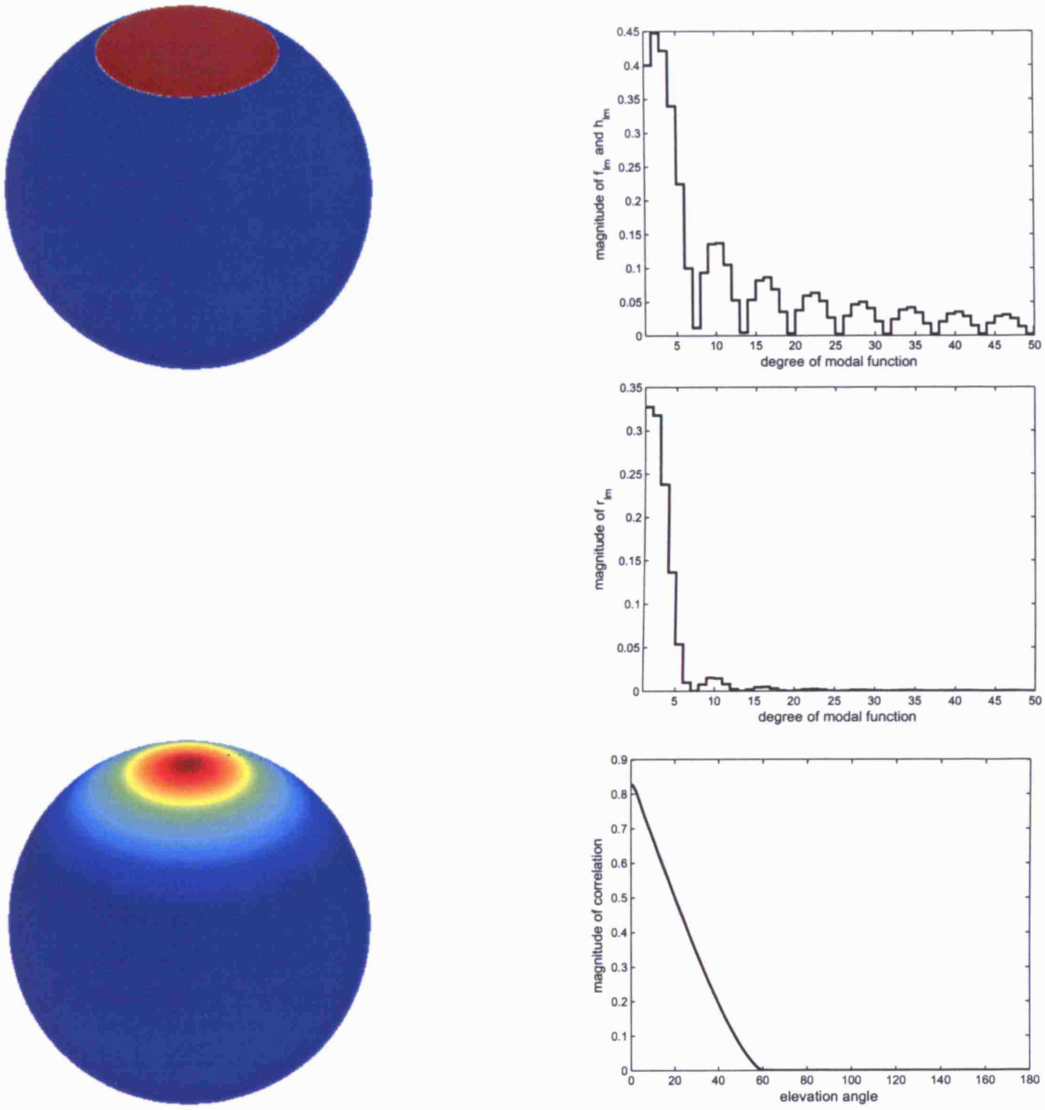


Figure 7.8: The results obtained for the scalar correlation have been generalised to cover the circularly polarised far-field functions. The results for the vectorial correlation are very similar to the results for the scalar case as illustrated in figure 7.7. Note that here the spherical modal coefficients are those for which $m = 1$.

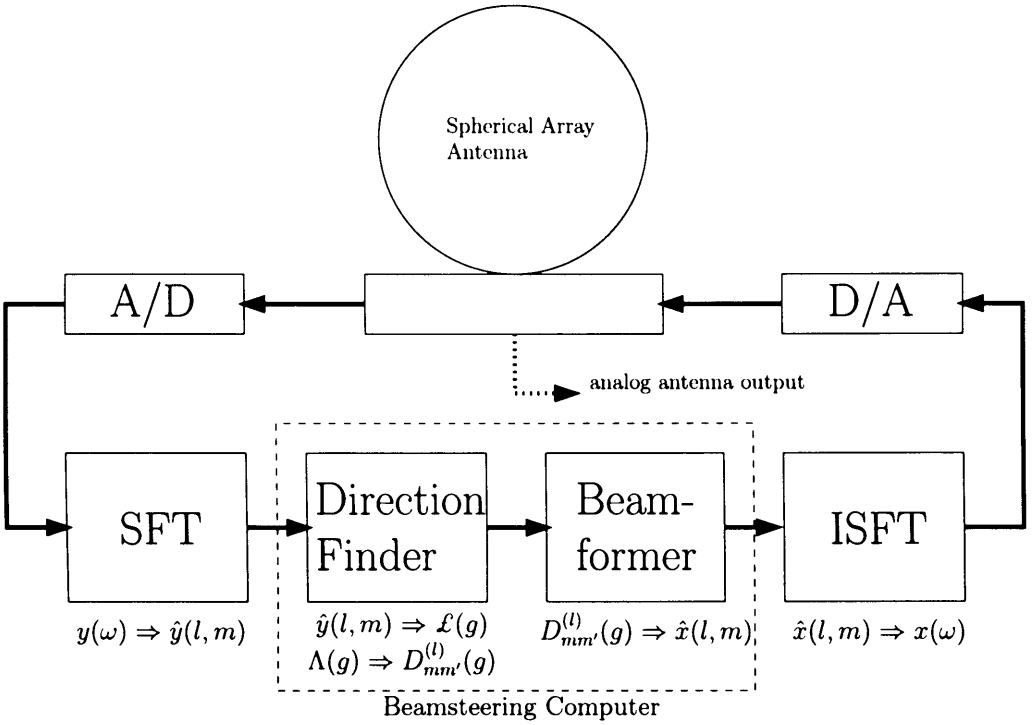


Figure 7.9: The figure shows schematically how the beamsteering computer (the dashed box) controls the radiation pattern of the spherical array antenna. It communicates with the array antenna by getting a vector of signals $\hat{y}(l,m)$ at its input that represents the SFT transformed received radiation pattern. From this information, the beamsteering computer calculates in a few simple steps the most likely direction of arrival $\pounds(g)$ and the required rotation $\Lambda(g)$ to steer its beam in the corresponding direction. Then the rotation matrices $D^{(l)}_{mm'}$ are calculated with which the radiation pattern is rotated in the domain of spherical harmonics. The resulting radiation pattern, represented by its harmonic coefficients $\hat{x}(l,m)$ is then fed back to the array antenna for further processing. The lower level processing in the antenna interface outside the dashed box will be dealt with in section 7.7

track route to spatially filter a function f with a filter h is to use the SFT to calculate $\hat{f}(l,m)$, execute the product with the typically stored $\hat{h}(l,m)$ and computing the inverse SFT. This can be used in all applications that seek to detect, match or estimate a known pattern, in data that is gathered on a sphere.

In this section we will concentrate on the signal processing that needs to be performed to steer the beam of the spherical array antenna to the estimated direction of arrival of the incident radiation. We assume the signals of the array elements have been digitised and transformed using the SFT. The vector $\hat{y}(l,m)$ of transformed signals are the inputs to the beamsteering computer. At the output of the beamsteering computer a vector of signals $\hat{x}(l,m)$ will be presented that are the spherical harmonic coefficients of the radiation pattern that is rotated in the desired direction. The lower-level processing that is necessary to obtain the signal vectors $\hat{y}(l,m)$ and $\hat{x}(l,m)$ and interface them so that they work together with the spherical array antenna is handled in what we will call System X and dealt with in section 7.7. The beamsteering computer that deals with the direction finding and beamsteering is put into context in figure 7.9, where it is shown in the dashed box.

A central idea to the spatial filtering process and the driver behind the SFT is the rotational

invariance of the eigenspaces formed by spherical harmonics of a certain degree l. This turns the rotation of a function, expanded into the spherical harmonic basis, into a multiplication of its SFT coefficients with a block-diagonal matrix, as given in equation (2.6) and repeated here:

$$\Lambda(g)Y_l^m(\omega) = \sum_{|k|\leqslant l} Y_l^k(\omega)D_{k,m}^{(l)}(g) \tag{7.65}$$

The semi-infinite block diagonal rotation matrix D can be computed from the required rotation g alone. It is independent from the positions of the array elements, with which the function x is sampled. The computation of the rotation matrix was the subject of section 4.2 in chapter 4.

The previous describes how to perform a rotation on the antenna radiation pattern, or the translation from a known rotation $\lambda(g)$ to a rotated set of expansion coefficients $\hat{x}(l,m)$, which is the task of the beamformer in the beamsteering computer (see figure 7.9). Prior to that, the desired rotation g needs to be retrieved from the received radiation pattern $y(\omega)$. This is the task of the direction finder.

The direction finder bases its operation on the convolution theorem discussed in section 7.4. It is not unreasonable to expect a certain form for the incident wave. If the incident wave has the typical planar wavefront for instance, we know its expansion in spherical harmonics as a function of the electrical size of the sphere. Without lack of generality, we assume the incident wave originated from the direction of positive z-axis. The expected illumination function $f(\omega)$ corresponds with a radiation pattern that maximises the directivity in the same direction. This will be discussed in more detail in section 7.6. The received radiation pattern will be a rotated version of the assumed pattern, in addition to noise. The task here is to find the unknown rotation on the received radiation pattern. The following discussion is based on an example of SFT processing in [55].

Let $\mathbf{y}(\omega)$ be the vector with measured data on the sphere. g is the unknown rotation between the expected radiation pattern $f(\omega)$ and the true direction of incidence. $\mathbf{n}(\omega)$ is a random process on the sphere representing additive noise. We can write:

$$\mathbf{y}(\omega) = \Lambda(g)f(\omega) + n(\omega)$$

We want to find the rotation g that is *the most likely* to have caused this received $y(\omega)$. From statistics we know that this can be found by trying all possible rotations $\Lambda(g)$ on the known illumination f and correlating that with the received y. Mathematically this translates into the following definition of the most-likelihood function $\mathcal{L}(g)$:

$$\mathcal{L}(g) = \arg\max_g \left(\int_{S^2} y(\omega)\,[\Lambda(g)f(\omega)]^*\,d\omega \right)$$

In which we recognise the spherical correlation of equation (7.40). Thus, we can calculate the maximum likelihood function by multiplying the received and transformed $\hat{y}(l,m)$ with a precomputed mask of $\hat{f}(l,m)$ as described in section 7.4.3.

Putting the direction finder and beamformer together, the array antenna is capable of "listen-

ing" in all directions simultaneously for an expected source and adjusting its antenna pattern so that its gain is maximised in the direction of the source. As the source moves around, the antenna could electronically track the movement and follow with its beam so that it is continuously pointed at the source (or target in the context of RADAR antennas). Of course more complex schemes of signal processing are possible. For instance, in the meantime a zero in the radiation pattern could be held pointing in the direction of a jammer. All this processing can be implemented efficiently in the spherical harmonic domain.

7.6 Plane Wave Decomposition and Reciprocity

It has been mentioned on several occasions throughout this text that an antenna radiation pattern that maximises its directivity in a certain direction will have a near-field that is equivalent to a plane wave incident on the sphere from a point source at infinity. This statement may require some clarification.

First of all, in the statement above, "equivalent" should not be misinterpreted as "equal". The spherical wave expansion incorporates a radial factor, the most general form of which is given by equation (2.19). For inbound waves towards the centre of the sphere, the radial function is different than for outbound waves from the centre. When an antenna is radiating, the field travels towards infinity, and can be imagined to originate from a phase centre in the centre of the sphere. When the array is receiving, it will be hit by a plane wave, originating from a transmitter at infinity. If the receiving antenna is not there, the waves will close in to the centre, pass through, and emerge again as outbound waves. The inbound and outbound waves, described with a radial function that is a Hankel function of the second kind and first kind respectively, form a standing wave that is described by a spherical Bessel function. With the receiving antenna in place, part of the incident wave will be absorbed (received by the antenna) but another part of the wave that impinges on the metal surface of the sphere is scattered back into the radiation environment. The overall effect is that the field outside the surface of the antenna can only be described as a superposition of spherical Hankel and spherical Bessel functions (or an imbalanced superposition of spherical Hankel functions of both kinds). Clearly, the field outside the antenna is different upon transmission and receive, and not just by the direction of propagation. This is of course not true for the gain function that - by the reciprocity theorem for antennas - should be the same upon transmit and receive.

The remarks made in the previous paragraph have their importance in direction finding applications. When a transmitter is to be located based on the wave incident on the array, then the expected wave is a planar wave coming from an unknown direction. The correlation of the received signals should therefore be made with the expansion of a planar wave front, and not with the expansion of maximum directivity. This is known as plane wave analysis, and has been proposed as a signal processing technique for spherical SONAR arrays[29][112][40][41][113]. With the extension of the spherical convolution to multipole fields from section 7.4.4, the concept of plane wave analysis is also readily extendible to electromagnetic spherical antenna arrays.

7.7 A Digital Signal Processing Architecture

In this chapter we have related the polarised spherical expansion in the far-field to the tangential polarised near-field expansion. We have used the sampling theorem to obtain this expansion from a set of phase and amplitude samples collected over the surface of the spherical array. In section 2.4.4 we have seen how the transform could be altered so that the standard scalar spherical transform can be built in as the main engine of the transform process. In this section we will combine all these techniques in one system that takes in phase and amplitude samples collected at the spherical surface and outputs the coefficients of the polarised expansion. Although we will not yet go into the details of hardware implementation (that is done in chapter 9), we will outline the processing inside a system based on a digital architecture. A digital architecture takes preference for a first proof-of-concept array design, which will be developed in chapter 9. Nevertheless, some initial efforts have been made to develop an analogue processing stage, that could perform the spherical harmonic transform on the RF signals. A description of the latter will be given in section 7.8.

In figure 7.11 an overview is given of the processing going on inside the digital system that transforms between a sampled near-field and the spectral components of the far-field. As established in section 7.2.2, the set of sample values F_s contains the result of a sampling operation:

$$F_s = \vec{\mathcal{F}}^{\mathrm{NF}} \cdot \vec{S}_{rhcp} \tag{7.66}$$

Physically, the sampling operation is of course performed by the array elements. The conversion of the element signals to complex baseband is of no concern in this section, suffice it to say that this can (and will) be implemented with IQ demodulators. The complex samples may need to be corrected for phase and amplitude, depending on how they are placed on the sphere. If they are non-uniformly distributed, they will need to be weighted to correct for this effect (this will be dealt with in more detail in chapter 8).

The processing stage assumes that the near-field is sampled with a vectorial distribution, e.g. $\vec{S}_{rhcp}$. The sample vectors in that distribution have an orientation or, more appropriate in the case of circular polarisation, a phase reference. Due to different path lengths for the RF signals, a phase distortion $e^{j\beta_n}$ can be present, that is different for every element. Also, by physically mounting the array elements on the spherical surface, the phase reference of each element is set. Unless the elements are carefully oriented so that their phase reference is aligned with the assumed sampling distribution, element-dependent phase errors $e^{j\alpha_n}$ will be present. Assuming that we have taken out the previous phase errors, so that for instance the reference vector for an element at position (θ_j, ϕ_k) is $\hat{e} = \frac{1}{\sqrt{2}}(\hat{e}_{\theta_j} + j\,\hat{e}_{\phi_k})$, we can re-align the phase reference to whatever distribution we like, for instance that of $\vec{S}_{rhcp}$. The latter distribution has its reference vectors aligned with the fieldlines shown in figure 2.4 and requires a multiplication with a phase factor $e^{-j\phi_k}$. However, we will see that the first step in the processing chain consists of a multiplication with $\frac{e^{-j\phi}}{\sin\theta}$, which will remove that phase adjustment. The latter phase shifts are indicated for illustrative purposes, in reality of course we would not implement two opposite phase shifts. On the contrary, in reality the whole chain of phase corrections and settings would

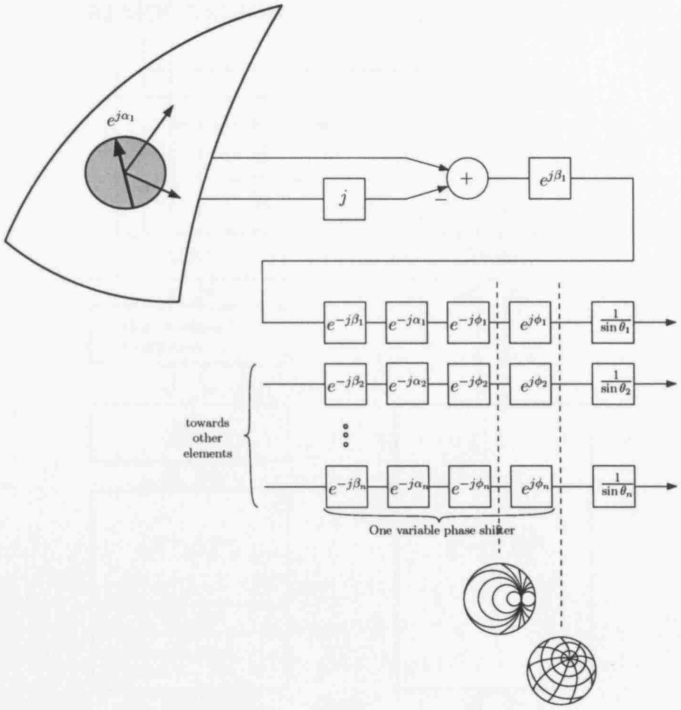


Figure 7.10: A correct setting for the phase of the circularly polarised elements needs to take into account their position on the sphere, and the chosen phase reference.

be combined in one variable phase shifter, so that the samples have the correct phase that is required by the processing stage. This is illustrated in figure 7.10. The correction of phases and amplitudes will be made as the first step in the digital system. As this step will need to correct phase errors that result from an unknown misalignment of the array elements, we need to establish some calibration procedure. We can further argue that we do not need to bother to align the array elements according to some particular distribution, as we set their alignment through the process of phase calibration anyway. (We are assuming circular polarised elements are used throughout this discussion. It is sensible to assume circular polarised elements from the discussion in 6.1.) We will come back to the phase calibration procedure in section 9.3.3.

After amplitude weighting (gain calibration) and phase calibration, the sample vector F_s is ready for further processing. The next four stages of processing are just an implementation of the theory laid down in section 2.4.4. The central part of the processing is the scalar spherical harmonic transformation of the sample vector. There are several algorithms available in the literature that perform the spherical harmonic transform. Therefore it is advantageous to split the overall CPST in a preprocessing stage, a scalar transform and a post-processing stage. The algorithms to solve the spherical harmonic transform are often based on some assumption of the sample distribution, and therefore will therefore be covered in chapter 8.

The overall polarised transform has so far been performed on the samples collected in the near-field of the antenna. It might be desirable to work the obtained expansion coefficients back to some other radius, often to the far-field at infinite radius. This has been theoretically dealt with in section 7.1. It can be noted that for beamsteering purposes with a radiator at finite distance

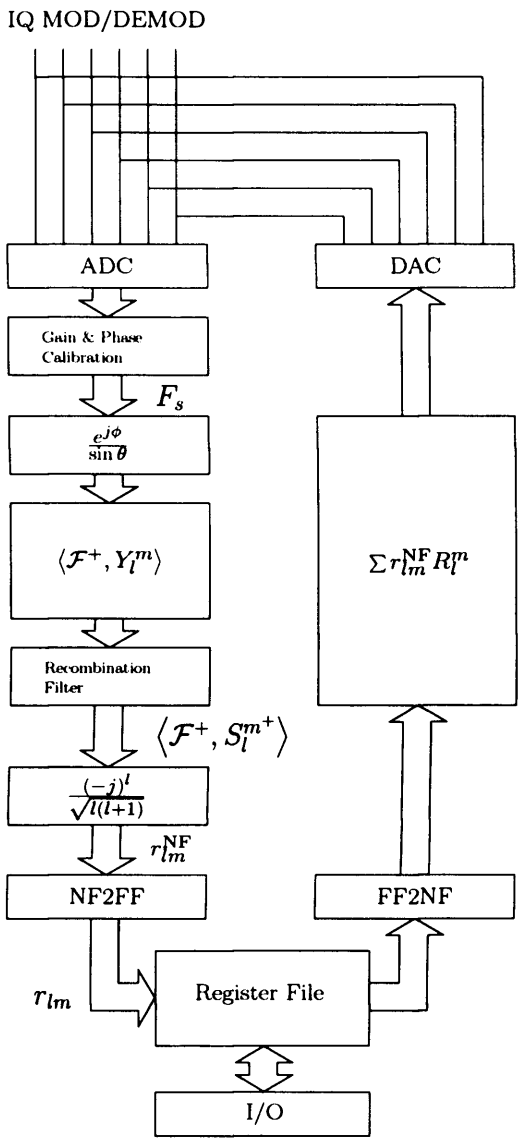


Figure 7.11: The processing path inside a digital system that performs the CPST.

such as in a near-field scanner, the expansion could be focused on any intermediate distance instead of infinite distance by adjusting equation (7.12) accordingly.

The reverse operations are straightforward. When an harmonic expansion is given, the spherical modes just need to be superimposed, weighted with the expansion coefficients, to construct the total field. Again, the hardware implementation of this operation is of no concern in this chapter.

Finally, it must be understood that the way we have broken down the processing in different stages was for illustrative purposes mainly. In reality, the hardware implementation may have a different breakdown in functional units, but that is outside the scope of this chapter.

7.8 Towards an Analogue Architecture

Some early work has been done on an analogue implementation for the spherical harmonic transform. This work has not been developed beyond a preliminary study reported in more

detail in appendix D, perhaps to be picked up in future work.

An analogue implementation will not be free from any assumptions on the sampling distribution, as was the case for the discussion in section 7.7, but has so far been developed around an equiangular distribution.

A possible advantage of an analogue architecture could be to offer a cheaper solution for the signal processing stage, while retaining a large number of elements. One could for instance design a system with a dense spherical sampling grid, where the cost of downconverting and digitising a large number of signals from all radiators would be prohibitively high. If we could reduce the number of signals to a smaller number of signal ports that allow independent control over some spherical modes (sufficient to attain some directivity), this would result in a saving in the number of channels that need to be downconverted and digitised. Of course the complexity of an analogue processing stage might also prohibit a design with a large number of elements.

Like the Butler matrix is an analogue implementation of the butterfly algorithm to calculate a Fast Fourier Transform (FFT), we will base our system on a fast algorithm to compute the spherical harmonic transform. Such algorithms benefit from symmetries in their input data, and for the spherical transform that amounts to the assumption that the data are collected on an equiangular grid. We will further assume this is the case. Even if the samples are not collected on the inefficient equiangular grid, one can imagine a preprocessing stage that interpolates the data of a thinned equiangular grid in order to obtain the required dataset. As the development of such an analogue implementation is entirely beyond the scope of this programme, we have not pursued a detailed solution for this problem.

The interested reader can find the details of the work that has been performed on an analogue system in appendix D.

7.9 Summary and Conclusions

In this chapter the basic toolset was established on which advanced signal processing techniques for spherical array antennas can be built. Formula's were presented to backproject the CPST to any finite radius. The abstract machinery of distributions was used to formally obtain a sampling theorem for both scalar functions and vector fields tangential to the sphere. The sampling theorem relates the spatial bandwidth of the radiation pattern to the distribution bandwidth (or to the sampling degree). The determination of the distribution bandwidth for some specific distributions will be dealt with in chapter 8.

The spherical convolution for tangential vector fields has been developed, again using the CPST. Also the convolution theorem was generalised for vector fields tangential to the sphere, which allows an efficient implementation of filters in the spectral domain of spherical modal functions. This was illustrated with a correlation filter that finds the likelihood of orientation for a function or tangential vector field on the sphere. Such filter can be used with spherical array antennas in a direction finding algorithm for the signal of arrival.

Finally a system architecture was presented to implement a processing chain for digital spherical array signal processing. On top of this architecture, complex signal processing schemes can be

implemented directly in the spectral domain of the far-field radiation pattern of the antenna, with abstraction of the low-level hardware implementation. The possibility of implementing an analog architecture for preprocessing the array signals was mentioned and is given some more attention in appendix D.

With regards to section 1.3.1, this chapter focused on the fourth point of the hypothesis, that stated that based on SMT, a framework can be constructed in analogy with phase mode theory for circular array antennas, that facilitates the development of advanced signal processing techniques for spherical array antennas. Also here it must be concluded that a solid basis has been established and that continued future work is necessary to take this further with the development of high-level signal processing schemes.

Chapter 8

Array Element Distributions

In chapter 7 we have seen how the signals from the array elements are transformed into a spherical modal decomposition of the far-field. We have seen that the processing depends on the distribution of the array elements. Also, the formulation of the sampling theorem depends on the distribution of the elements. The importance of this observation has no equivalence when dealing with circular geometry or "normal" Fourier analysis. Indeed, for periodic signals or circular arrays one has no difficulty distributing the samples or array elements uniformly over a periodic interval or circle and there is no reason to give any alternative distribution much thought. This all changes when dealing with spherical geometry, where it is no longer possible to perfectly meet requirements on uniformity, sample density and symmetry, except for a few sample points or array elements. The difficulties and implications associated with choosing a spherical sample or array element distribution will be covered in this chapter.

In section 7.3, the sampling degree L_s, defined as twice the distribution bandwidth, was related to the available spatial bandwidth for the radiation pattern of a spherical array with the equation:

$$B_s \leqslant L_s - L_m \tag{8.1}$$

The determination of the distribution bandwidth of sampling distributions of different kinds and sizes is the subject of section 8.1.

In sections 6.3.3 and 7.3 it was indicated that for a given electrical size of an array antenna the number of field samples can be decreased, together with an anti-aliasing action of the array element patterns. In practise this can be achieved with subarrays of array elements, but this puts other requirements on the overall distribution of array elements. The technique of subarraying with spherical array antennas will be discussed in section 8.2.

8.1 Distributions of Isotropic Elements

When placing radiating elements on the sphere, one wants to maximise the attainable directivity of the array antenna. Modal theory predicts that the maximum directivity equals the number of spherical modes in the radiation pattern. The sampling distribution in turn determines the highest degree of modes that can be controlled, as was established in the previous chapters. When the elements are phased to form a narrower pattern, with modes of a higher degree than

can be controlled, grating lobes will occur. Although true grating lobes do not exist as with linear arrays, the effect will be that sidelobes increase and the higher directivity is not realised. It is therefore important to place and control the array elements so that the highest possible degree of modes can be controlled. To keep the number of sampling points as low as possible the elements could be placed in the configuration of one of the t-designs discussed in section 8.1.2. An extra benefit of using a t-design is that the element signals need not to be weighted to compensate for an uneven spread.

A multitude of methods exist, to distribute a number of points "evenly" over a spherical surface. An overview of methods is given in [115]. An example is for instance an algorithm that would search for the minimum-energy distribution on a sphere of a set of point charges. One such algorithm has a useful meaning for our purposes: an algorithm that maximises the minimal distance between any element and its closest neighbour. This would result in a pointset that is optimised with respect to the reduction of mutual coupling effects. However, the distributions that will be used here are optimal with respect to integration on the sphere and will automatically have a quasi-uniform nature.

Symmetry is another criterion that is often used to make spherical distributions. However it is important to realise that the usefulness of this criterion is a myth if the application is a system that uses spherical modes. Spherical array patterns synthesised with a number of point sources on the sphere will obey the principle of uniform resolution, provided the sampling theorem is not violated. This means that the synthesised pattern is invariant, regardless of its rotational position relative to the pointset. The argument that in order to generate a radiation pattern that does not vary with scan angle, one should see a "similar" arrangement of points sources on the sphere when looking from different directions, is invalid. An exception to this observation is when directional element patterns are taken into account. However, for large arrays with a large number of elements this may not make much difference. Another more important exception is when a subarraying technique is employed. An in-depth treatment of this is given in section 8.2.

8.1.1 Minimal Spherical Designs and Quadrature Rules

A generic sampling theorem is given in section 7.2. The sampling theorem is generic in the sense that no decision is made on the sampling measure, except for the constraint in equation (7.16). Enforcing this constraint amounts to solving a new formulation of the same problem, that will recast it as a quadrature rule. As a sampling measure we choose a weighted sum of Dirac pulses in the points $\{\eta_1, \eta_2, \ldots, \eta_N\}$ on the surface of the sphere: $s = \sum_j^N w_j \delta_{\eta\eta_j}$ Let $P(\eta)$ be a polynomial on S^2 of degree $L \leqslant 2B$, so that:

$$P = \sum_{l,m}^{L} p_{lm} Y_l^m \tag{8.2}$$

Let P be a test function for the distribution T_s:

$$T_s(P) = T_s\left(\sum_{l,m}^{L} p_{lm} Y_l^m\right) \tag{8.3}$$

It follows from equation (7.16) that[1]:

$$T_s(P) = p_{00} = \frac{1}{\sqrt{4\pi}} \int P ds \tag{8.4}$$

but we also have:

$$T_s(P) = \int_{S^2} s \cdot P \, ds = \sum_j w_j P(w_j) \tag{8.5}$$

We conclude that the distribution T_s is bandlimited with bandwidth B if and only if, for all test polynomials P of degree $L \leqslant 2B$:

$$\frac{1}{\sqrt{4\pi}} \int P ds = \sum_j w_j P(w_j) \tag{8.6}$$

Or else, the definition of the bandlimited sampling distribution is equivalent to an interpolation grid that facilitates a quadrature rule that is exact for polynomials up to degree $2B$.

8.1.2 On Minimal Designs and T-designs

It is shown in [116] that if a quadrature rule is exact for all polynomials of degree $l \leqslant 2B = L_s$, then the number of quadrature points satisfies

$$N \geqslant \sum_{l=0}^{B} d_l = (B+1)^2 \tag{8.7a}$$

where d_l is defined as the dimension of the harmonic eigenspace of degree l:

$$d_l = 2l + 1 \tag{8.7b}$$

An exact quadrature rule that attains this bound is a minimal design. Unfortunately, for the sphere in 3 dimensions it is impossible to attain this bound for degrees higher than 2 and minimal designs do not exist. It can also be shown that minimal designs have equal weights[116]. Although minimal designs do not exist, there are known exact quadrature rules with equal weights, the so-called t-designs[117]. With a t-design, the integration of a polynomial of degree t or less can be exactly calculated as an average of the samples over the sphere. Hence, using a t-design of double bandwidth $t = 2B$, a function of bandwidth B can be reconstructed from its samples without aliasing. From all possible choices of sampling distributions, the 2B-design is the optimum. Note that by definition, the distribution bandwidth of a 2B-design is B. For instance, with a 10-design aliasing-free sampling can be achieved up to degree $B = 5$. A minimal 10-design would have $(5+1)^2 = 36$ points. However, the smallest 10-design known

[1] The sampling measure s should not be confused with the integration measure $ds = \sin\theta d\theta d\phi$.

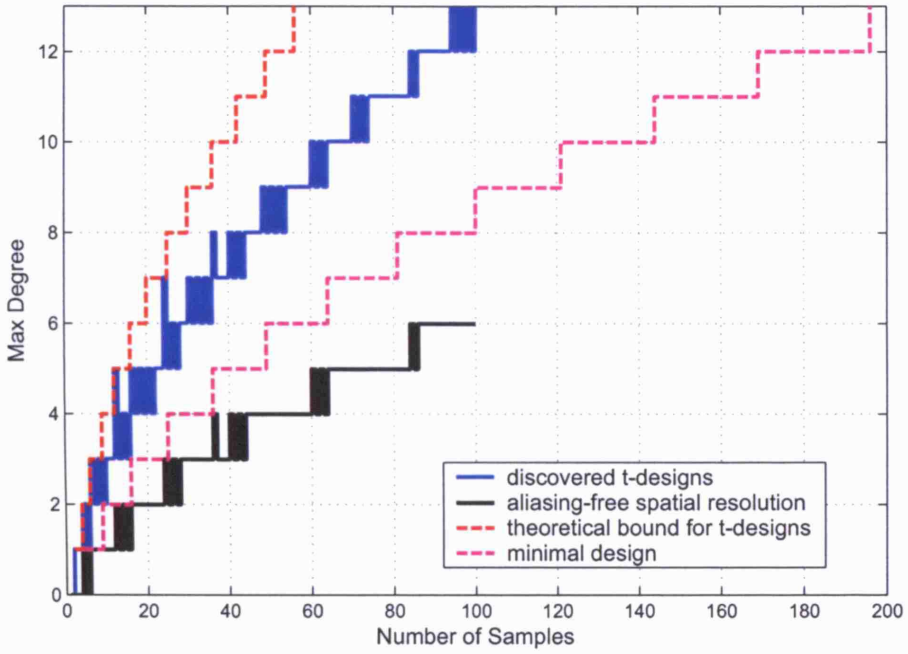


Figure 8.1: The known t-designs up to 100 sampling points. The resulting aliasing-free spectral resolution gives the minimum number of sampling points for a given maximum harmonic degree. The curve can never be pushed to the limit of a minimal design.

today has 60 points and a smaller 10-design may not exist[117].

T-designs are only known for not too many sampling points. Figure 8.1 shows the maximum degree of the quadrature rule as a function of the number of elements for the known t-designs up to 100 sample points. Also shown is a proven upper bound for the t-designs that can exist for a certain number of points. Taking into account that t-designs need to be exact up to $t = 2B$ in order to avoid aliasing, one obtains the aliasing-free spatial resolution, which represents the minimum number of samples to avoid aliasing. A theoretical lower bound for this is the hypothetical minimal system, also shown in the figure.

For lower degrees, the t-designs typically show a high degree of symmetry. Although strictly speaking the spherical harmonic transform obeys the law of uniform resolution, and it is therefore unnecessary to worry about symmetry as would be the case when taking a switched element approach, it is still elegant to maximise the symmetry in the distribution. Especially when considering subarray implementations, this becomes more important. For 20 elements or less, there are 5 platonic solids to choose from. This distributions are the only perfectly uniform distributions that can exist on the surface of a sphere.

Equation 8.7 was obtained for the scalar spherical harmonics. The spectrum of the SFT includes an harmonic of degree $l = 0$. The multipole expansions do not however include modes of degree zero, and the summation in equation should start from one instead of zero. In that case we obtain:

$$N \geqslant \sum_{l=1}^{B} d_l = B(B+2) \tag{8.8}$$

In which we recognise the form of the generic sampling theorem in equation 6.29. We therefore conclude that for minimal designs, the distribution efficiency becomes indeed $\eta_d = 1$ which justifies the formulation of the generic sampling theorem. It also means that we always have $\eta_d < 1$.

8.1.3 Hypersampling on the Sphere

The minimal number of sample points to calculate an aliasing-free Fourier transform equals the number of harmonics to be resolved, but this bound is shown to be unattainable. In practise, the absolute minimal number of points that can be used is given by the requirement that a known t-design must exist for twice the required bandwidth. T-designs are only known for low degrees, and the search for new designs is still ongoing among mathematicians. The rest of us can move away a little from the absolute minimum of sample points while relaxing the equal-weight rule for minimal designs. Both the quadrature weights and the sample positions can then be optimised according to some criterion. When sampling a function of bandwidth B, hyperinterpolation is defined as interpolation using a distribution with distribution bandwidth B or a sampling degree of double bandwidth $2B$[116]. The result is a reconstructed function that is free from aliasing effects. The distributions used for hyperinterpolation are those for which the $2B$ sampling degree can be proven. As we are focussing on sampling methods, we will use the term hypersampling when it can be mathematically proven that the sampling distribution has a $2B$ sampling degree. With Nyquist sampling we will refer to those distributions that allow aliasing-free sampling. Distributions that allow Nyquist sampling are accepted to have a $2B$ sampling degree, but it is not necessarily possible to prove it. We keep the term oversampling for when we are sampling on a denser grid than required by Nyquist sampling. All known hypersampling methods are oversampling the sphere.

In general it is not easy to prove the condition of double-bandwidth precision. An exception is formed by the tensor product rules, where the Fourier integral is separated in two one-dimensional Riemann sums. Examples of these quadrature rules are the Gauss-Legendre, Clenshaw-Curtis and Fejér rules[118]. The latter two correspond in fact with equiangular distributions, where the resolution in elevation is twice that in azimuth. These methods have roughly four times as many points as a minimal system. The Gauss-Legendre[2] distribution is equiangular in azimuth, but has its sample points on rings that are placed in the zeros of the Legendre polynomial of degree $B + 1$[119]. The rule translates roughly to twice as many points as a minimal system.

The FFT implementation samples in an equiangular fashion in elevation, with half the resolution of the azimuth sampling, and synthesises data for negative elevation angles. After this step a Fast Fourier Transform can be used over the elevation angle. This method also reduces the number of samples to twice the number of harmonics. The sampling efficiency remains low as the number of samples on each latitude circle remains constant[120].

The efficiency can be further increased by thinning the number of samples around the poles with a $\sin\theta$ factor. Although no sampling theorem has been formulated, the cut-off behaviour of high

[2]This is in fact Neumann's second method, but best known as the Gauss-Legendre method.

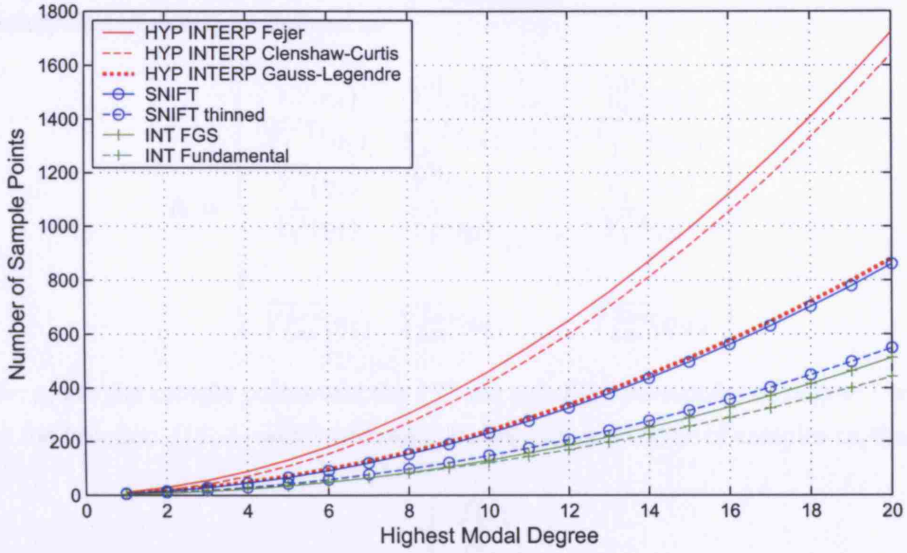


Figure 8.2: Several distributions are compared in terms of the number of sampling points that are used as a function of the highest degree spherical harmonic that is present in the function to be sampled. Among the hypersampling grids (shown in red), the Fejer and Clenshaw-Curtis quadrature rules are based on equiangular distributions and have roughly four times more sampling points than would be required by a fundamental system. The best known hypersampling grid is that from the Gauss-Legendre quadrature rule, which has twice as many sampling points than required by a fundamental system. As many points are needed for the FFT implementations of the spherical harmonic transform as implemented in the NIST computer code described in [63]. This grid could further be thinned around the poles, but this would destroy the guaranteed aliasing free character of hypersampling so that it belongs to the realm of interpolation grids. The interpolation point sets (shown in green) typically have the number of sampling points required by the fundamental system, or slightly more.

order Legendre functions near the poles makes the polar data irrelevant in the computation of harmonics of high order m so that the spectral resolution remains unaffected[120].

It is important to understand that hypersampling is not necessary to prevent aliasing. The equiangular distributions on which hypersampling occurs are enormously inefficient in terms of the number of samples that need to be collected. The weights of the quadrature rules are used to counteract the dense packing of samples around the poles. It is not true that these configurations are needed to avoid aliasing effects, but they are the only distributions for which the bandwith has been mathematically proven. A much more important rational for using an equiangular sampling grid is the efficiency of the computer algorithm that calculates the spherical Fourier transform. The algorithms that have been developed for optimum efficiency typically rely on the angular symmetries in the sampling grid[120][54].

8.1.4 Distributions and Quadratures Based on Interpolation

While hypersampling offers some benefits, mainly in the processing of the samples, it is found empirically that the true number of samples needed to prevent aliasing is closer to the minimal bound. With the relaxed requirement of choosing the weights, interpolating sampling grids can be tailored to minimise errors. Generally speaking, it pays to distribute the sampling points as uniformly as possible, although there is no single way to optimise this criterion.

Let the interpolation matrix be defined as

$$\mathbf{A} = \begin{bmatrix} \overline{Y_0^0}(\eta_1) & \overline{Y_0^0}(\eta_2) & \dots & \overline{Y_0^0}(\eta_N) \\ \overline{Y_1^{-1}}(\eta_1) & \overline{Y_1^{-1}}(\eta_2) & \dots & \overline{Y_1^{-1}}(\eta_N) \\ \overline{Y_1^0}(\eta_1) & \overline{Y_1^0}(\eta_2) & \dots & \overline{Y_1^0}(\eta_N) \\ \overline{Y_1^1}(\eta_1) & \overline{Y_1^1}(\eta_2) & \dots & \overline{Y_1^1}(\eta_N) \\ \vdots & \vdots & \ddots & \vdots \\ \overline{Y_{Lm}^{Lm}}(\eta_1) & \overline{Y_{Lm}^{Lm}}(\eta_2) & \dots & \overline{Y_{Lm}^{Lm}}(\eta_N) \end{bmatrix} \tag{8.9}$$

where the η_i are the sample points and the Y_l^m the spherical harmonics of degree l and order m. From the function $f(\theta, \phi)$ with bandwidth B, we collect a vector of samples on the sphere:

$$\mathbf{F_s} = \begin{bmatrix} f(\eta_1) \\ f(\eta_2) \\ \vdots \\ f(\eta_N) \end{bmatrix} \tag{8.10}$$

We want to find a vector of spherical harmonic coefficients so that the original function can be reconstructed:

$$\mathbf{F_s} = \mathbf{A}^* \cdot \mathbf{C} \tag{8.11}$$

If $\mathbf{A}$ is unitary, than $\mathbf{C}$ can be found as $\mathbf{C} = \mathbf{A}\mathbf{F_s}$, which is a linear projection of the sample vector on the sampled spherical harmonic base. Unfortunately, although the spherical harmonic functions are orthonormal, the vectors obtained by sampling them generally are not. This is very different from circular Fourier basis functions. Therefore, $\mathbf{A}$ is not unitary and equation (8.11) does not hold. $\mathbf{C}$ could be found as $\mathbf{C} = (\mathbf{A}^*)^{-1}\mathbf{F_s}$, as long as $\mathbf{A}$ is a regular square matrix, which is not very likely because generally, more samples are collected than there are harmonic coefficients to be calculated, and $\mathbf{A}$ has more columns than rows. Its rank equals the number of independent vectors of sampled harmonics. The latter are not orthonormal, and they may not even be independent if the sampling distribution is not well chosen, so that its rank decreases below the number of harmonics we are trying to resolve (the number of rows). Furthermore, the vector of samples $\mathbf{F_s}$ may incorporate higher degree harmonics that are beyond the Nyquist capability of the sampling distribution. For all these reasons together, it is unlikely that $\mathbf{F_s}$ will lie in the space spanned by the rows of $\mathbf{A}$ and we can only calculate an estimate for the spherical harmonic coefficients by minimising the error vector $\mathbf{E}$ in the equation:

$$\tilde{\mathbf{C}} = \mathbf{A}\mathbf{F_s} + \mathbf{E} \tag{8.12}$$

The least squares solution for $\tilde{\mathbf{C}}$ is found as:

$$\tilde{\mathbf{C}} = (\mathbf{A}\mathbf{A}^*)^{-1}\mathbf{A}\mathbf{F_s} \tag{8.13}$$

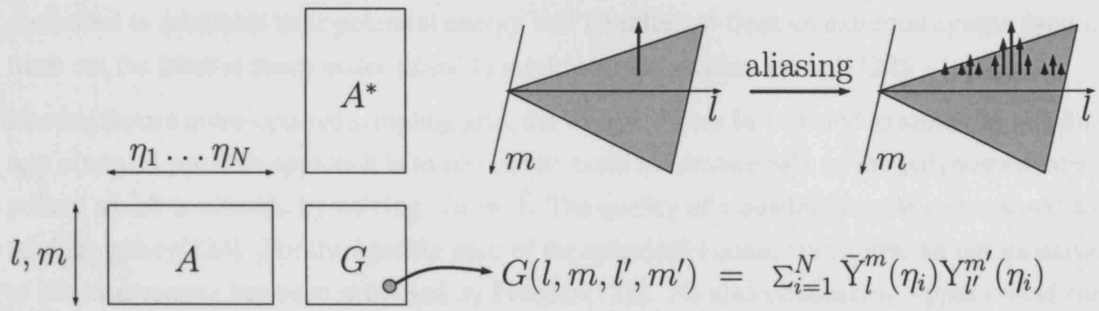


Figure 8.3: Each element in the Gram matrix is the image of four variables l, m, l' and m' under the aliasing function. For each spherical harmonic impulse, it gives the spectral response of other harmonics as a consequence of aliasing for the sampling distribution under consideration. This is schematically depicted in the figure.

Of course, this requires the Gram matrix to be invertible:

$$\mathbf{G} = (\mathbf{A}\mathbf{A}^*) \tag{8.14}$$

$$\det \mathbf{G} \neq 0 \tag{8.15}$$

In the aforementioned case when the sampling grid is trying to resolve modes that it cannot handle, not all rows of $\mathbf{A}$ are independent and $\mathbf{G}$ becomes singular. The redundant dimensions have to be removed from the solution space, before the least square projection on this space can be sought. To achieve this, a Gram-Schmidt orthogonalisation procedure can be followed, as presented in [121]. Starting with the lowest degree harmonic, and gradually moving to higher degrees, a vector of samples is added to $\mathbf{A}$, after orthogonalisation to all previously added vectors. When a vector of samples almost vanishes in the orthogonalisation process, the corresponding harmonic is removed from the solution space and the corresponding coefficient set to zero. This procedure is particularly helpful for badly conditioned sampling grids such as the equiangular one. For quasi-uniform distributions, the Gram matrix is usually invertible. When that is no longer the case, the aliasing is probably becoming too strong and it may not make much sense to try to add more higher degree modes to the problem.

An interesting interpretation of the Gram matrix is obtained by acknowledging that each element in the matrix is in fact the image of four variables l, m, l' and m' under a function $G(l, m, l', m')$. This function is called the aliasing function G and tells us how much of $Y_{l'}^{m'}$ gets aliased into Y_l^m for the given sampling distribution. A closer examination of the aliasing function is performed in [112].

It is possible to define a grid with $d_B = (B+1)^2$ sampling points and enforce the condition that G is regular. Such a pointsystem is called a fundamental system[122]. A good fundamental system is optimised according to some uniformity criterion. For instance, when $|\det(G)|$ is maximised the pointsystem is called extremal[122][123]. All criteria try to get to the optimum $G = \lambda_{\text{avg}} I$, where I is the identity matrix, along different paths. As it is impossible to get to this optimum (because then a minimal design would be obtained), the solutions for different optimising strategies are different[123]. That is why a system of pointcharges that has been

optimised to minimise their potential energy will be different from an extremal system (and it turns out the latter is much better suited to minimise interpolation errors)[122].

Starting from a quasi-optimal sampling grid, the weights $\vec{w}$ can be adjusted to minimise quadrature errors. A possible approach is to enforce an exact quadrature rule on the polynomial interpolant, which is effected by solving $G\vec{w} = \vec{1}$. The quality of a quadrature rule is measured by its discrepancy[124]. For the specific case of the spherical Fourier transform, an optimisation of the discrepancy has been proposed by Freeden[125]. He also obtained an upper bound for the error. We have found his method performs well with a quasi-uniform sampling grid. The sampling weights are different for each calculated harmonic, and therefore this method is only suitable in problems where the sampling grid is static. Figures 8.4 to 8.7 illustrate the process of sampling and reconstruction of scalar functions on the sphere, using Freeden's method.

8.1.5 Distribution for a Medium Size Array

For medium sized spherical array antennas, the 64-element 10-design is a good candidate distribution. The distribution is illustrated in figure 8.9. It is capable of resolving spherical modes up to degree 5, without aliasing errors. Thanks to the graceful degradation property of quasi-uniform distributions, the aliasing errors may be tolerable up to degree 7. For spherical arrays of a large electrical size ($kR > 7$), this is insufficient to avoid aliasing in the radiation pattern of the array antenna. This was explained in section 7.3. The size limit for a spherical array that employs this distribution is illustrated in figure 8.8 by the solid horizontal red line. On the right vertical axis of the same figure, it can be seen how this translates to a size restriction for each element. For the latter, we define the average available dimension *Aed* for the array elements as:

$$Aed = \frac{4R}{\sqrt{N}} \tag{8.16}$$

where R is the radius of the sphere and N the number of elements. This ratio is equivalent to the diameter of a circle that has the same area as the total area of the sphere, divided by the number of elements.

It should be mentioned that array elements with a higher gain can lift the aliasing limit, but will generally take up more space. Therefore the ratio of element gain to element size could be maximised, but this may adversely affect the gain of the array antenna.

In figures 8.11 and 8.12 the aliasing errors are shown, for a realistic Near-Field test-function. The test-function was obtained as a maximum-directivity expansion with cosine roll-off at the highest degrees to reduce sidelobe levels. The pattern was steered $\frac{\pi}{6}$ radians away from pole position. It is apparent that the errors increase gradually, but the distortions may be tolerable while directivity is improved beyond 30 dB. This is an important observation, typical for spherical sampling. The point where the sampling theorem is violated may be not as clearly defined, but when the sphere is undersampled, aliasing effects appear only gradually. The spherical modes are defined in the same physical space as the sphere itself, so that there is no equivalent to suddenly appearing grating lobes as is the case with planar arrays for instance. Practical designs may therefore willingly allow a limited amount of aliasing errors in the radiation pattern and thereby reduce the number of radiating elements that need to be placed on the sphere.

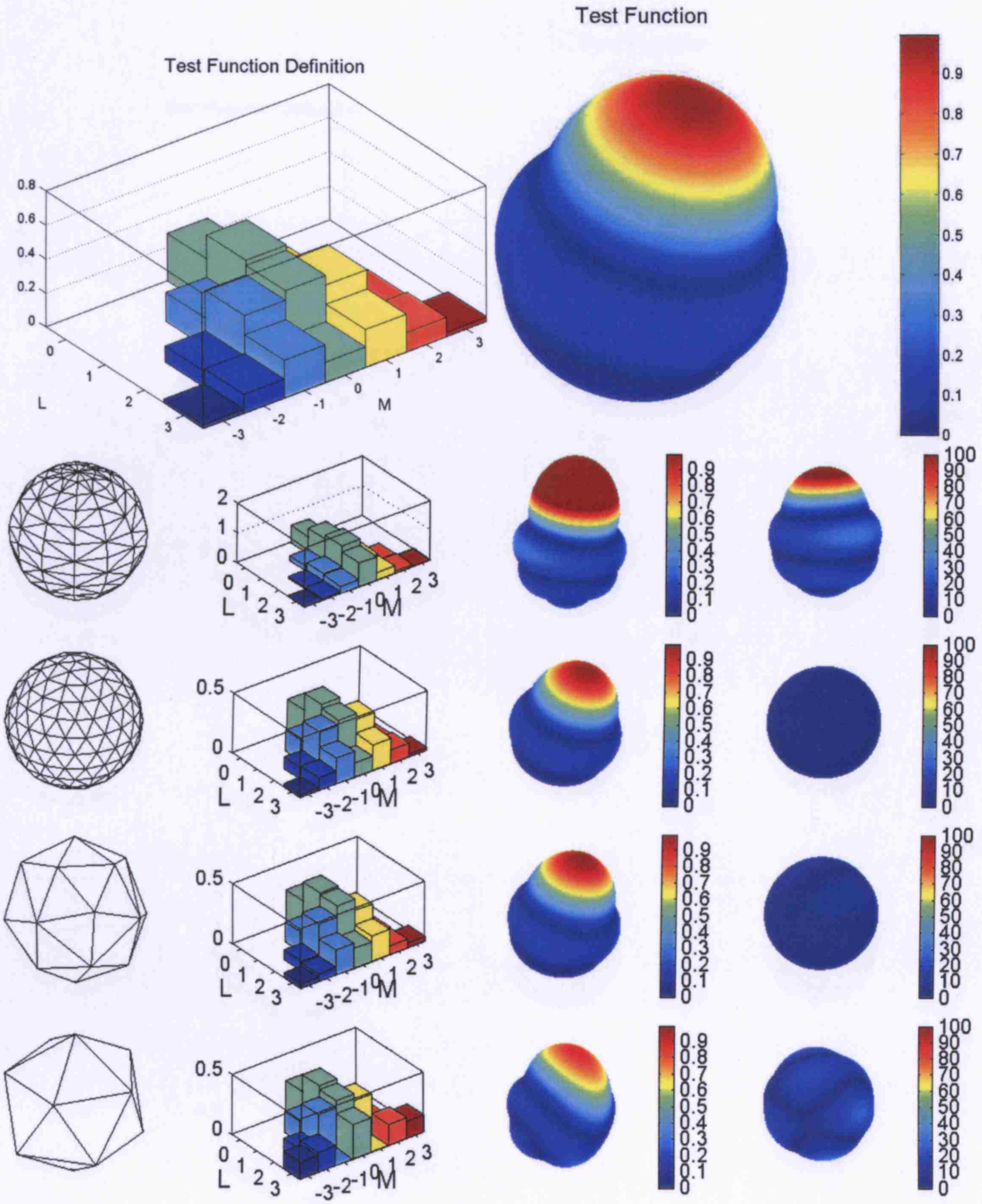


Figure 8.4: The test pattern shown in the top row is created with a maximum degree of 3. It is subsequently sampled with the sample point distributions shown in the left column, from top to bottom: a 169-element equiangular grid, a 162-element quasi-uniform grid, a 20-element t-design and finally, an icosahedron. The two columns in the middle show the function reconstructed from the estimated harmonic coefficients. The right column shows the absolute error magnitude, relative to the maximum magnitude of the original pattern. With the 162 and 20 element distributions, the reconstructions are good. With the icosahedral vertex distribution, aliasing errors are clearly visible. The equiangular distribution with 169 elements suffers from distortions that are attributable to the uneven packing of the samples.

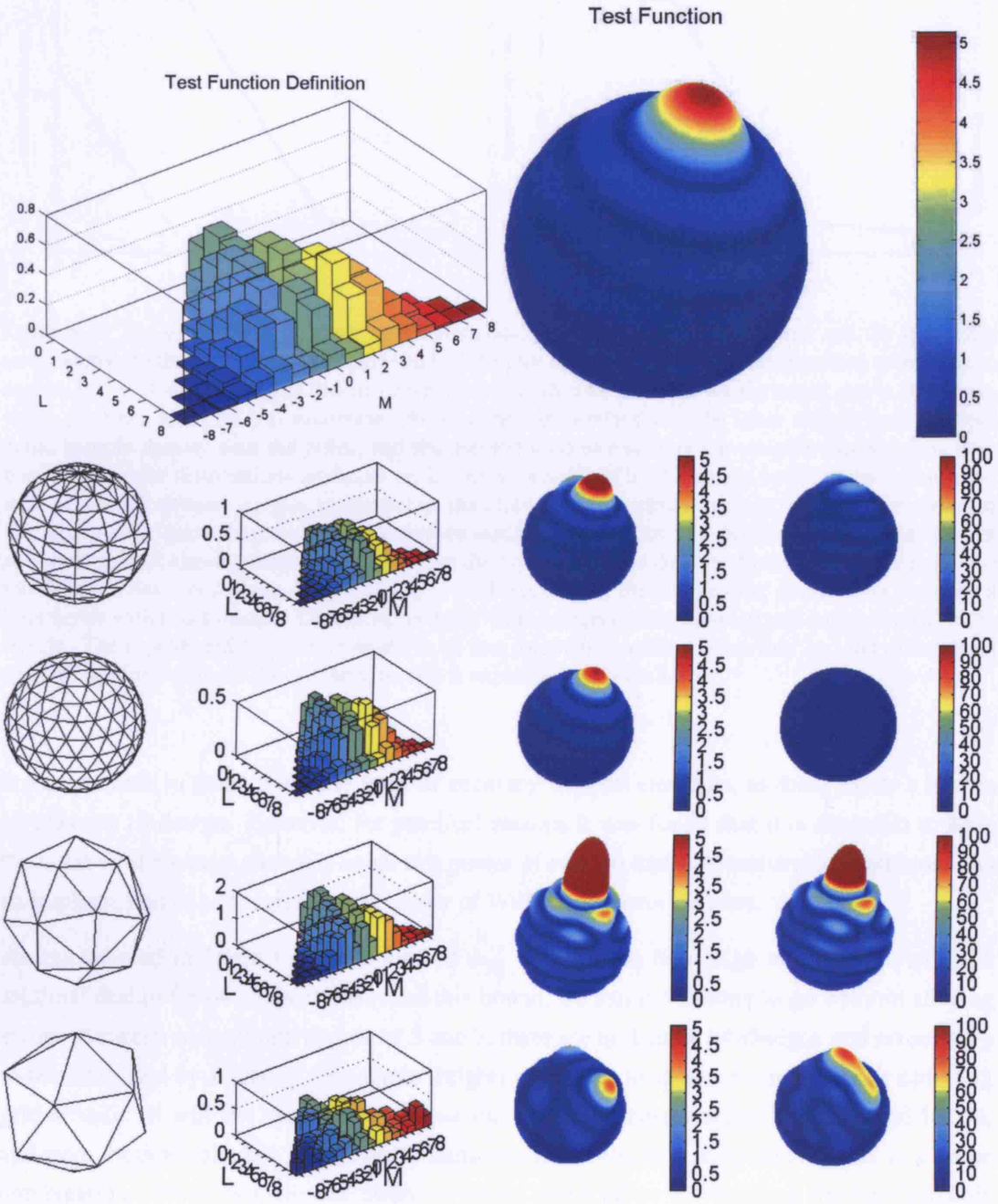


Figure 8.5: The test pattern shown in the top row is created with a maximum degree of 8. It is subsequently sampled with the sample point distributions shown in the left column, from top to bottom: a 169-element equiangular grid, a 162-element quasi-uniform grid, a 20-element t-design and finally, an icosahedron. The two columns in the middle show the function reconstructed from the estimated harmonic coefficients. The right column shows the absolute error magnitude, relative to the maximum magnitude of the original pattern. This function can not be reconstructed with the 20-element distribution. The 162-element distribution is still OK. Also here some distortion is visible from the uneven packing around NP, but as the function values are smaller in that region, distortions are also smaller.

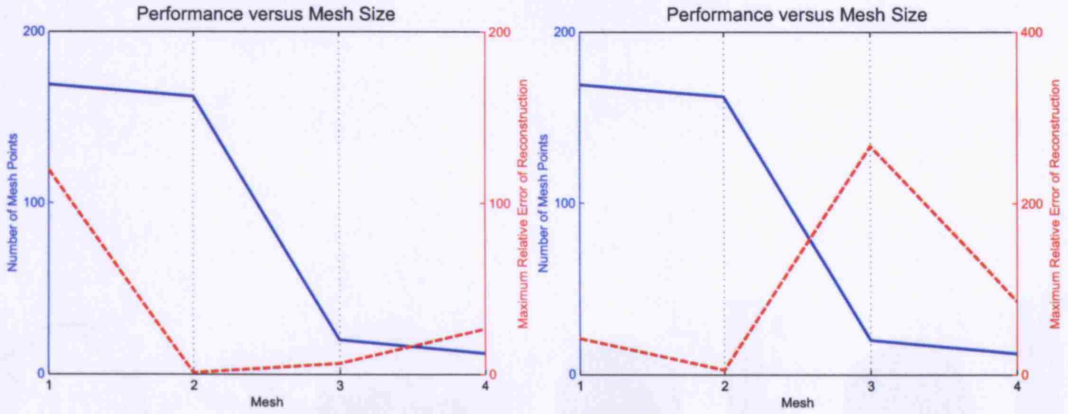


Figure 8.6: The plots show, for the four evaluated distributions, both the mesh size and the maximum error relative to the maximum field magnitude. The plot on the left is for the test-function of maximum degree $L_m = 3$ and aliasing is low or non-existent for all distributions. As the mesh size is increased, errors go down apart for one exception: the equiangular distribution. The latter suffers from the very dense sample density near the poles, and the method used to calculate the weights can not cope with this. Equiangular distributions are better dealt with as prescribed by the known hyperinterpolation rules, with accurately defined weights to counteract the close sample packing in polar regions. It is observed that equiangular grids should only be used when sampling on or above the Nyquist density, they are not so robust against aliasing errors when violating the Nyquist sample density. In the plot on the right, the test-function has a maximum degree of $L_m = 8$. It seems that the equiangular distribution can cope a little better with this situation. Distributions 3 and 4 are severely undersampled and suffer from aliasing effects. The reproduced function is prone to ad hoc rotation-dependent behaviour and the principle of uniform resolution breaks down. More on this is explained in figure 8.7

It was possible to attain the same level of accuracy with 60 elements, as there exists a known 60-element 10-design. However, for practical reasons it was found that it is desirable to have the number of element channels equal to a power of two, so that a power divider/combiner can be implemented as a straightforward ladder of Wilkinson power dividers.

As can be seen in figure 8.1, the bound of $L_m = 7$ for the 64-design corresponds with the minimal design for 64 elements. Beyond this bound, we can never hope to go without aliasing errors. Between a maximum degree of 5 and 7, there are no known 64-designs and errors have to be minimised by adjusting the sample weights according to some criterion. As our sampling grid is static (it will not vary over time) we can calculate these weights for once and for all, and store them to calculate the spherical harmonic transform of a sampled function as a linear combination. Therefore, we can afford to spend much effort in obtaining the best possible weights. To achieve this, we calculate a separate Least Square Error solution for the weights, which depends on the spherical harmonic on which the sampled function is projected. We use Freeden's solution of iterated Green functions to obtain these weights. Details of the procedure can be found in [125].

8.1.6 Spherical Near-Field Antenna Measurements

Although this is not directly related to the distribution of array antenna elements on a spherical surface, we shortly describe another application for which spherical sampling is very relevant. The measurements described in later chapters will rely on a Spherical Near-Field Measurement

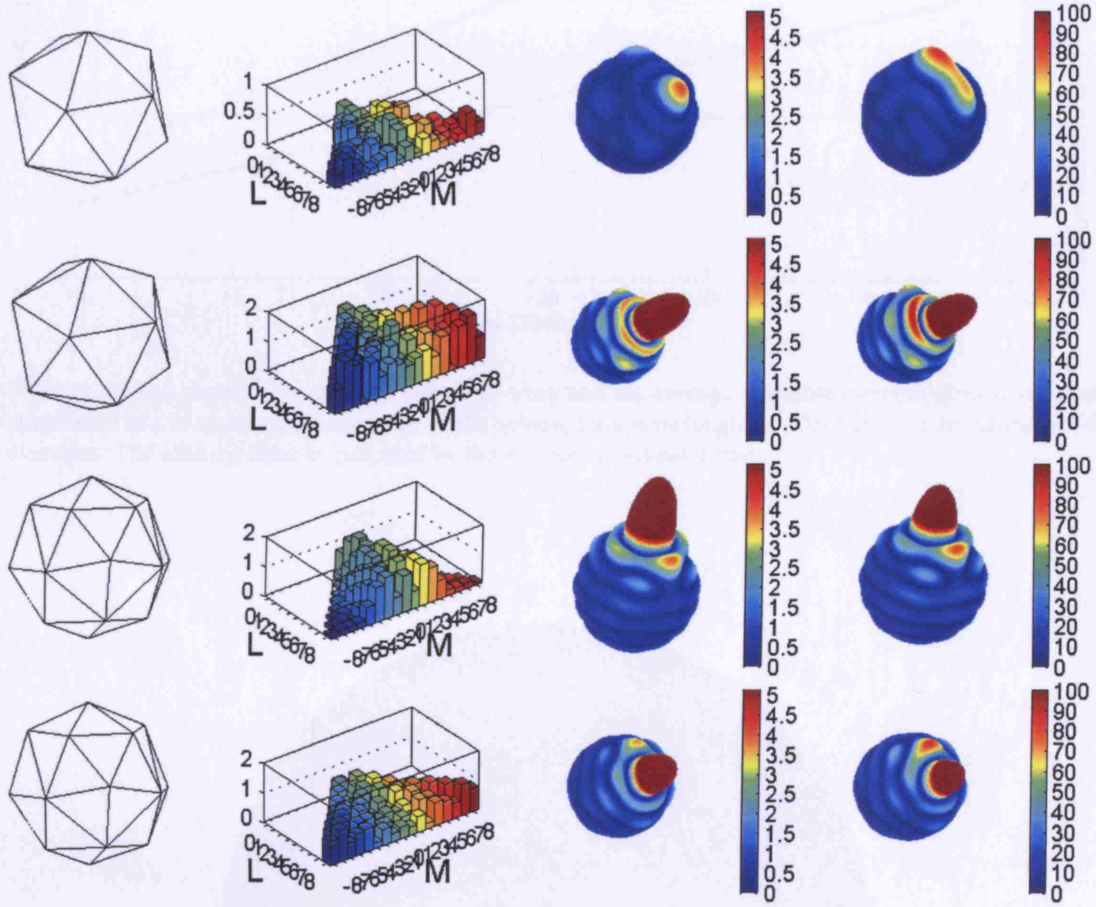


Figure 8.7: A spatial bandlimited function can be reconstructed from its samples, and this can be done with an arbitrary rotation applied to it. A rotation cannot shift energy from any harmonic subspace of one degree to another, and therefore, cannot increase the highest degree that is present in the modal expansion. However, when aliasing occurs, this principle of uniform resolution breaks down. That is because unrepresented higher degree modes manifest themselves at lower degrees, and may distort the reconstruction differently depending on its rotation. The top two rows of plots show this effect dramatically for $L_m = 8$ and the Icosahedral distribution. The bottom two rows show that the effect decreases for a 20-element 4-design. Note how well the 20-element 4-design is coping with the undersampling situation, as theoretically it can only guarantee aliasing free sampling when $L_m \leqslant 2$.

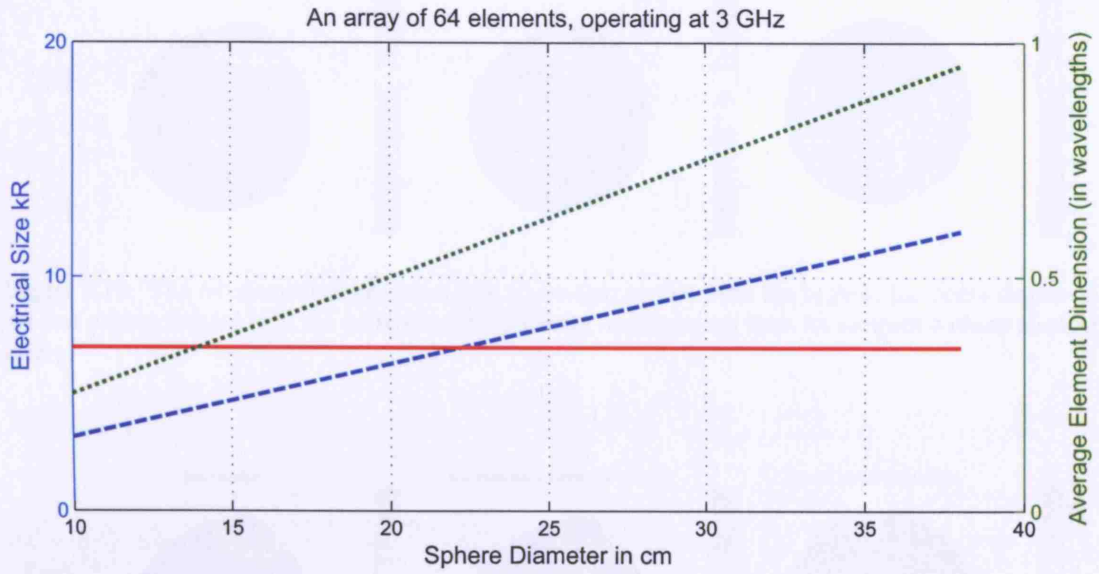


Figure 8.8: The electrical size of the spherical array and the average available element dimension Aed are plotted as a function of the diameter of the sphere, for a wavelength of $10cm$ and a distribution of 64 elements. The aliasing limit is indicated by the horizontal red solid line.

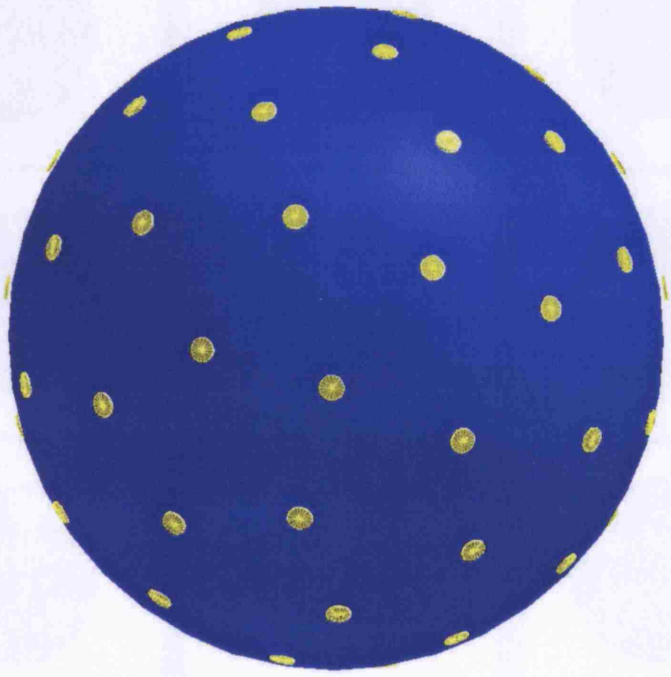


Figure 8.9: A very good 64-element candidate distribution for medium-size spherical array antennas. The distribution is a 10-design, which means it guarantees aliasing free sampling up to a maximum degree of 5. A 60-element 10-design exists, but unless each element has its own oscillator, it is desirable to have the number of elements equal to a power of two.

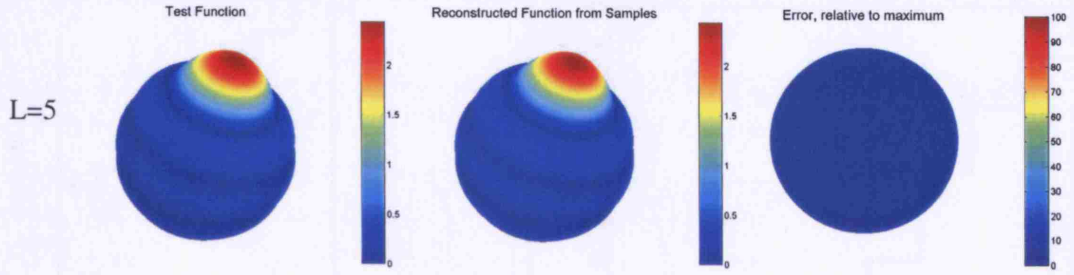


Figure 8.10: The 64-element distribution is a 10-design, so that with the highest harmonic degree in the test pattern limited to 5, the pattern can be perfectly reconstructed from its samples without aliasing errors.

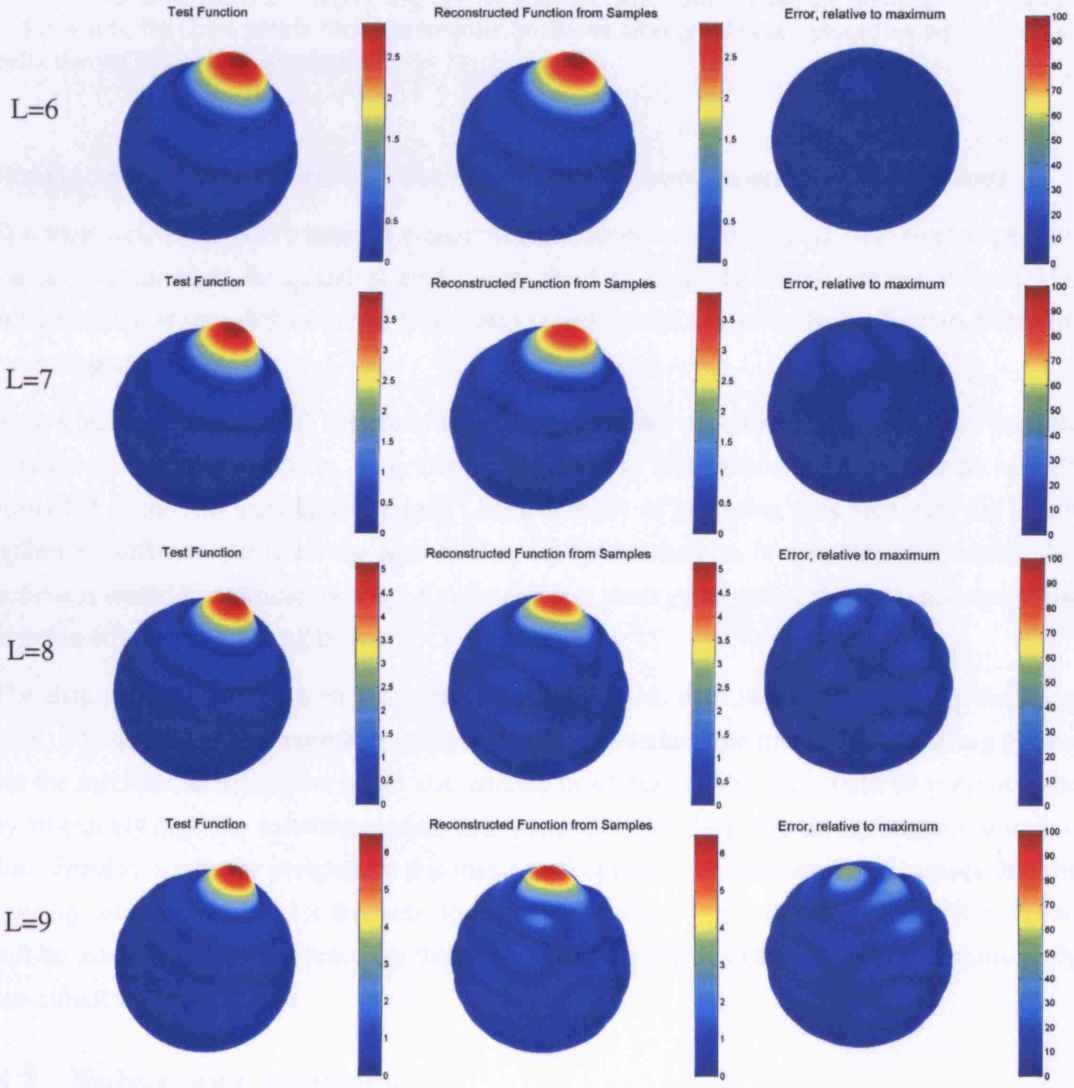


Figure 8.11: For the 64-element t-design, it is shown how the errors gradually increase as the maximum degree of the pattern rises above the aliasing-free limit of L=5. The left column shows the original patterns, the centre column shows the reconstructed patterns, the right column shows the error patterns, relative to the maximum pattern magnitude.

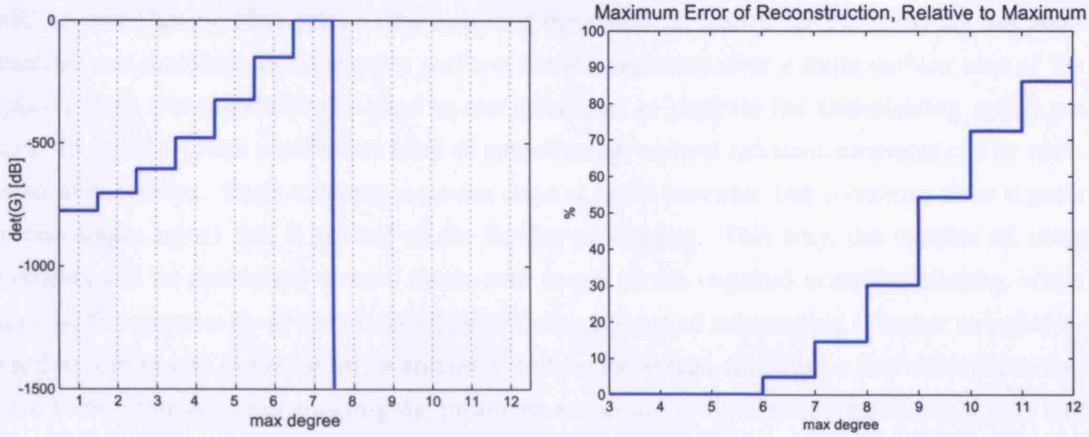


Figure 8.12: For the 64-element 10-design, aliasing-free sampling is possible up to degree 5. If the pattern contains harmonics of higher degrees, the aliasing errors start to jump up. Starting from degree 7 and upwards, the Gram matrix becomes singular, so that an orthogonalisation procedure with dimension reduction on $\mathbf{A}$ would be appropriate.

Facility, and therefore this could not be left out from a section on sampling distributions.

Spherical near-field (SNF) antenna measurement systems sample the near-field over a spherical surface and calculate the spherical modes from the data. Once the modal content of the field is determined, it is straightforward to propagate the orthogonal modes to the far-field to obtain the far-field pattern.

A disadvantage of the SNF system is the lengthy process of gathering enough data samples over the measurement sphere. This step takes typically several hours, which can be entirely attributed to the mechanical subsystem. After 6 hours of gathering data samples, the super-optimised software performs the near-field to far-field transform in a matter of minutes. We believe it would be sensible to adopt a more efficient strategy to gather the data, perhaps at the expense of some processing time.

The data could be gathered on the generalised spiral grid, with two simultaneously scanning axes. The quasi-uniform sampling grid would not only reduce the number of sampling points, but the mechanical subsystem could also operate much faster. The data could be preprocessed by an entirely separate software module that interpolates the data with an equiangular distribution. From an academic perspective this may be suboptimal, but it offers the advantage that the existing software module for the near- to far-field transform (typically SNIFT or TICRA) can still be used. Empirical evidence for this case can be collected with an SNF that supports only one simultaneous scan axis.

8.2 Subarray Clustering

Typically, a spherical array antenna is electrically large and a large number of spherical modes can propagate to the far-field. However, the number of modes that can be resolved with a practical number of elements is lower. In order to avoid aliasing, the higher degree modes need to be filtered out with a spatial anti-aliasing filter. This amounts to taking a spherical convolution

with an anti-aliasing filter prior to the sampling operation. In reality, the elements are not point sources, and therefore automatically perform some integration over a finite surface area of the sphere. This can potentially be used to our advantage to perform the anti-aliasing action we seek. In order to reach a sufficient level of anti-aliasing, several radiation elements can be clustered in subarrays. Each subarray contains several array elements but combines their signals in one single signal that is passed on for further processing. This way, the number of array elements can be multiplied several times over to get to the required sampling density, while keeping the complexity of the system down. The technique of subarraying is rather straightforward to implement in planar array antennas, but on the sphere there are a few difficulties that need to be addressed. Formulating the problems associated with subarraying on the sphere and providing the solutions for doing so are the subject of this section.

8.2.1 Problem Formulation

Let N_e be the total number of elements on the sphere and N_s the number of subarrays. The number of elements per subarray is $k = \frac{N_e}{N_s}$. It is instructive to think of the array antenna as an array with N_s elements, where the element radiation patterns are replaced with the pattern of a subarray of k elements. As the array with N_s elements is not sampling the sphere dense enough to avoid aliasing, the subarray radiation patterns must act as anti-aliasing filters.

Consider now the array of subarrays, with N_s elements. We like all the elements to be identical, or the behaviour of the array antenna becomes intractable. But each of these elements is in fact a subarray of k elements. Also, in order to benefit from the lesser complexity that subarraying can offer, it is necessary to combine k signals into one received signal. Therefore, each subarray will incorporate an RF circuit, that performs the coherent addition of the elements in the subarray. It is strongly desirable to have all these circuits and subarrays identical, rather than designing them individually. We conclude that a first requirement on the subarrays is that they are all identical. All subarrays are kept rigidly identical so that each subarray can be rotated to coincide with any other subarray. In other words, at all times we can rotate a point A in subarray X so that it coincides with a point B of subarray Y so that all other points of subarray X coincide with a point in subarray Y. This transitivity defines a symmetry group of rotations on the finite distribution of sampling points, which we will call the **global symmetry**. As far as the signal processing on the received signals is concerned, the subarraying process is transparent. The subarray patterns take the place of the single element patterns, and the number of sampling points is N_s . This means that the N_s phase centres of the subarrays now have to be distributed on the sphere in the optimal way. The optimal way maximises the number of spherical harmonics that can be resolved and is likely to be one of the t-designs, as was discussed in section 8.1. From an electromagnetic perspective, the array still has N_e elements distributed over a spherical surface. If no direction for the antenna beam is to be favoured over any other, it still makes sense to distribute the whole set of N_e elements evenly over the sphere, so that they are all allocated a fair share of the spherical surface. This will also minimise the mutual coupling between the elements. In this context, "evenly" can be interpreted as to maximise the minimal distance between any two radiating elements.

The subarrays are distributed over the sphere and will have their front faces pointing in all

directions. Throughout this thesis we have used array elements with an axially symmetric pattern around the radial line through their phase centre. This makes it easy to write such a pattern as an expanded sum of spherical harmonic terms and apply a rotation to it. Although the same can be achieved with element patterns that do not have this axially symmetric pattern, the analysis of the array antenna as a whole becomes a lot more complicated. From the moment that each radiation pattern is no longer axially symmetric, the array antenna as a whole may no longer be able to form an identical radiation pattern in all directions. It is therefore desirable to organise the elements within each subarray so that the pattern of the subarray is axially symmetric. We will call the symmetry of the distribution points around an axis through its phase centre **local symmetry**.

With the technique of subarraying we will divide the surface of the spherical array in smaller regions. The division will be based on the symmetrical arrangements of the faces of certain kinds of solids. It would be more practical to use these solids directly to place the radiating elements on, instead of the spherical surface. However, we want to keep the radiating surface as close to the sphere as possible, to not distort the spherical phase mode processing and prove the unique uniform resolution capability of the spherical array antenna. As a compromise, we will design the spherical array with a large number of planar tiles, that are each small compared to the radius of the sphere. All these tiles will be mounted tangentially on a sphere and meet each other at the edges. Provided the number of tiles is large enough, the radiating surface will be close to a sphere. We will have to synchronise that strategy with the concept of subarraying.

Finally, if the spherical array is to be approximated by a finite number of planar tiles, each of these tiles will have to be tangential to the same sphere and their phase centres will need to lie on a sphere as well. The first requirement is necessary to have the tile patterns all pointing outwards radially. Although this seems obvious, this does exclude the popular geodesic constructions to cover the sphere with planar tiles.

8.2.2 Enforcing Global Symmetry on the Subarray Distribution

In the previous section we explained that subarraying the spherical array requires the enforcement of symmetry on the subarray positions. The subarray positions are represented by a finite set of points on the sphere. These points permute transitively under the action of a finite set of rotations. This is trivial, because if there were an infinite number of different rotations acting on one of the points, this would result in an infinite number of points. All rotations are members of the special orthogonal group $SO(3)$. A rotation matrix $R \in SO(3)$ has the properties $R^T R = I$ and $\det A = 1$. There are five kinds of finite subgroups of $SO(3)$ (see appendix A for a more detailed overview of symmetrical polyhedra):

- C_k: the cyclic group
- D_k: the dihedral group
- T: the tetrahedral group
- O: the octahedral group

- I: the icosahedral group

The largest subgroup is the icosahedral rotation subgroup, which contains 60 different rotations. It is the symmetry group of rotations under which the set of vertices of the icosahedron is invariant. It is part of the larger full icosahedral symmetry group I_h of order 120, which also includes rotation-reflection combinations. Solids that have an invariant set of vertices under the rotations of I but not necessarily under the reflections of I_h are the chiral solids that come in two enantiomorphic forms. As we want to construct the spherical array with identical subarrays, all facing outward and excluding mirror images, we are only considering those solids that can be formed by pure rotations of a reference face. The solid can be chiral or achiral, but it cannot have more than 60 faces. depending on the manufacturing process, the cost impact of having two different kinds of faces may be small. In that case there are more solids to choose from, but none of them seem to offer an even distribution of face centres (the face centres will approximately become the phase centres of the subarrays). The reader can refer to appendix A for a more extensive list of illustrated solids.

The rotations in I come in five different flavours, its conjugacy classes are:

- E: identity element
- C_5 and C_5^2: Around 12 vertices of the icosahedron, or 6 vertex-pairs, 5 rotations can be made, one of which is the identity operation. Hence, $6 \times 4 = 24$ rotations can be retrieved from the 6 vertices in the upper half of the icosahedron. The rotation angles are 72° and 144° degrees, in both clock and counter clockwise directions.
- C_3: Around 20 faces of the icosahedron, or 10 vertex-pairs, 3 rotations can be made, one of which is the identity operation. Hence, $10 \times 2 = 20$ rotations can be retrieved from the 10 upward-looking faces of the icosahedron. The rotation angle is 120°, in both clock and counter clockwise directions.
- C_2: Around 30 edge-centres of the icosahedron, or 15 edge-pairs, 2 rotations can be made, one of which is the identity operation. Hence, $15 \times 1 = 15$ rotations can be retrieved from 15 not opposing edges of the icosahedron. The rotation angle is 180°.

By letting all rotations in a finite subgroup of rotations act on a point, a set of points is produced that form an orbit under the subgroup. We can choose an infinite number of startpoints, with an infinite number of associated orbits, that all together make up the surface of the sphere. To cover the whole sphere, it is sufficient to choose the starting points for the orbits in one particular area on the sphere, called the fundamental domain. The fundamental domain, when subject to the rotations in the subgroup, will cover the whole surface of the sphere. Several solids can be built in the same way, but instead of an area on the sphere a planar face can be rotated with its points rotating on spheres of different radii. Evidently the icosahedron itself is such an example. All the solids that can be built up with this construction share the same fundamental domain. This is illustrated in figure 8.13. In the figure the icosahedron is shown with blue edges. Each face of the icosahedron is divided into 3 white and 3 grey triangles. A white and grey triangle together

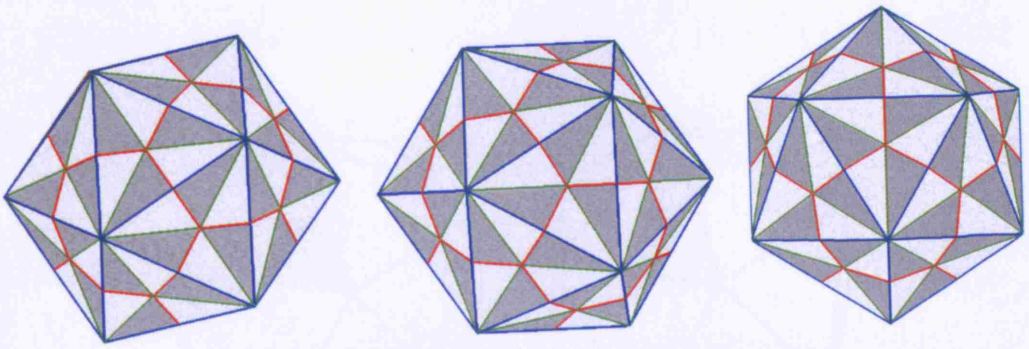


Figure 8.13: Three renditions of the icosahedron are shown that illustrate the group of icosahedral symmetries. The edges of the icosahedron are shown in blue. Icosahedral symmetry is important because it is the largest group of symmetries that exists on the sphere. Starting from any of the white coloured triangles, there are 60 rotations that will place the triangle in the position of any of the other white triangles. The same is true for the grey triangles. To go from a white triangle to a grey triangle, a rotation and reflection are required. As each white triangle is surrounded by grey triangles and vice versa, the points on the edges are always accessible with rotations only. The combination of a white and grey triangle forms a fundamental domain for the icosahedral rotation group, while a white or grey triangle is a fundamental domain for the full icosahedral group.

form a fundamental domain for the icosahedral rotation group, whereas a white or grey triangle on its own forms a fundamental domain for the full icosahedral group. A solid formed by applying the 60 icosahedral rotations to a point inside a white triangle (and not on the edge of it) has all its vertices in white triangles. The same is true for the grey triangles. Each grey triangle is surrounded by white triangles and vice versa. Hence, all starting points taken on one of the red, blue or green edges belongs to both the grey and the white subdomains. To go from inside a white triangle to its corresponding point inside a grey triangle, requires a symmetry operation that is in I_h but not in I. We will therefore restrict our attention to those solids that have their vertices in white triangles only, or on its edges. We can now construct solids by applying the rotations of the symmetry group to one or more points in the fundamental domain. An obvious example is the dodecahedron. The dodecahedron is found as the solid that has its vertices at the crossings of the red edges in figure 8.13. It is also the dual solid of the icosahedron - each vertex of the icosahedron corresponds to a face of the dodecahedron and vice versa. Many more solids can be formed in this way. However, of those the icosahedron and the dodecahedron are the only platonic solids. The process of selecting a point in the fundamental domain and creating a solid with icosahedral rotations is further illustrated in figure 8.14. The solid constructed in this way will represent the positions of the subarrays by letting its vertices coincide with the phase centres of the subarrays. As yet, there is no need to assume any particular shape or form for the subarrays themselves, so that one point suffices to locate their phase centre. As there are only 60 rotations in I and we start from one single point only, we necessarily end up with no more than 60 locations. That means that the largest possible number of identical subarrays in which we can group a set of points on the sphere is 60.

Of the infinite number of solids that can be constructed, some seemingly stand out as candidates for the subarray distribution, because they show a higher degree of uniformity. The two platonic ones, the icosahedron and the dodecahedron are entirely regular (see appendix A for a definition

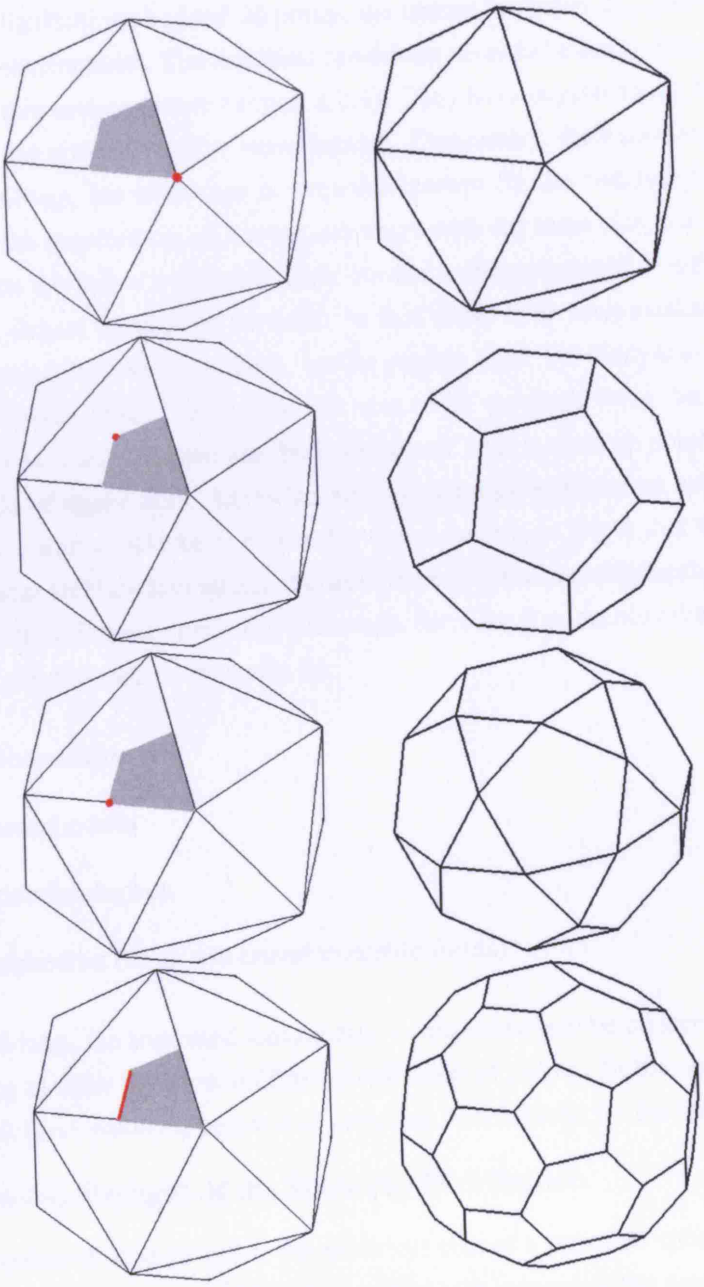


Figure 8.14: The figures illustrate the formation of different solids with icosahedral symmetry. The 60 rotations of the icosahedral group are applied to a point (shown in red) in the fundamental domain. Some of the rotations map the starting point to the same endpoint, so that we do not necessarily end up with 60 vertices. The figures in the top row show the construction of the icosahedron itself. The figures below show the construction of its dual dodecahedron. On the third row we recognise the icosidodecahedron and on the fourth row the truncated icosahedron. For the latter, the starting point can be chosen anywhere on the red edge of the fundamental domain. However, for one particular point we end up with the well known version that is also an archimedian solid. For that particular point the truncated icosahedron has regular pentagonal and hexagonal faces and is enjoying world fame as the shape of a football.

and listing of the platonic solids) but with their small number of vertices, 12 and 20 respectively, may not offer a sufficient number of subarrays. This is unfortunate, because as there are no perfectly uniform distributions beyond 20 points, we cannot hope that the subarrays will see an identical radiation environment. The next best candidates seem to be the archimedian solids (we will come back to this assumption in section 8.2.3). They have regular faces, but not all of the same type (again, see appendix A for more details). Conversely, their dual solids, the catalan solids, are face-uniform, but each face is irregular (except for the two mentioned hereafter), which establishes the requirement of having subarrays with the same size and shape. From the previous discussion we know we should look out for archimedian solid with 60 vertices, as this would be the largest number of vertices. In fact, there is an archimedian solid with 120 vertices, the truncated icosidodecahedron, but its catalan dual, the disdyakis triacontahedron, has two enantiomorphic faces. Of course this was to be expected from the aforementioned, as the truncated icosidodecahedron can be constructed with a starting point in both a white and a grey triangle of figure 8.13. Likewise, the truncated cuboctahedron and its dual catalan disdyakis dodecahedron should be excluded for the same reason. Note that there is no reason to exclude the chiral archimedian solids, the snub cube and the snub dodecahedron. Assuming we are interested in the largest number of subarrays, there are four archimedian solids that have 60 vertices (they are depicted in appendix A):

- truncated dodecahedron
- truncated icosahedron
- rhombicosidodecahedron
- snub dodecahedron (in its two enantiomorphic forms)

From these candidates, the truncated icosahedron could intuitively be chosen as the best candidate, for it seems to offer the most uniform distribution of points. In the next section we will address how well these pointsets perform as sampling distributions for the spherical harmonics.

8.2.3 Integration Strength of the Subarray Distribution

As we have established in section 8.1, the electrical size of a practical spherical array is such that 60 is a reasonable number of channels to control the pattern of the array and obtain good directivity. If the symmetry requirements were not an issue, then 64 channels would be a good choice, as it is a power of two and that makes life easier when designing the power divider for instance. We have also seen that the highest t-design that is currently known with that many elements is a 10-design. We have also concluded that using that number of elements on a sphere of $kR \approx 25$ would suffer from severe aliasing, which would destroy the radiation pattern. On the other hand, adding in a lot more radiating elements, each controlled with a separate channel, would be effective but unaffordable. We introduced subarraying as a compromise solution to this problem. In the previous sections we have seen how subarraying can be applied cost-effectively, provided that the array antenna has a sufficient level of symmetry to make all subarrays the same. It was shown that adding in the requirement of symmetry limits the number

of subarrays to 60 so that the selection of solids is narrowed down to the four archimedian solids with 60 vertices. The task of finding the perfect distribution for the subarrays is now confined to finding an archimedian solid with 60 vertices that is also a 10-design, the highest known t-design with 60 points. This is unfortunately not possible, so we will have to relax some of our requirements.

There is only one known 10-design with 60 points, and it does not exhibit icosahedral symmetry[117]. Hence, we cannot but relax this requirement and accept a t-design for which $t < 10$. On the other hand, one could ask himself whether there is also a minimum performance we can expect from the sampling distributions with icosahedral symmetry as obtained in section 8.2.2. We will call the pointsets found as an icosahedral orbit with an integration strength of t, by which we mean they are also a t-design, a t-orbit. It can be shown that all pointsets with icosahedral symmetry, starting from the minimum 12 vertices of the icosahedron, are 5-orbits [126][127]. Furthermore, there exists an infinite number of icosahedral 9-orbits. In [126] one of the 9-orbits is presented that is also a truncated icosahedron, but although they are almost indistinguishable, it is not exactly the same as the archimedian version. The archimedian truncated icosahedron, better known as the football, is only a 5-orbit. This shows that the acceptance of the vertices of the archimedian solids as the next best solution for the subarray distribution in section 8.2.2 was only partially justified. Partially, because their dual solids are face-uniform and form a good starting point to find other distributions with the same property but with better integration strength. Another example is given in [127], where one of the 9-orbits is shown to be optimal in respect to the integration error on 10 and 11-degree polynomials. In the same contribution, the authors prove that there is only one such optimal distribution and that no 10-orbits exist. The optimal distribution is shown in figure 8.16. From the figure it is clear that despite its superior integration strength, the distribution is less suitable for our purposes. This follows from the requirement of an even distribution, as formulated in the introduction. Indeed, for this particular distribution the minimal distance between the phase centres will be much smaller than for the distributions shown in figure 8.15 which makes it a bad compromise. That being said, we are dealing here with a distribution of subarrays and we cannot exclude the possibility of using this distribution for the phase centres of the subarrays, while pursuing a more even distribution for the complete set of radiating elements. The previous discussion suggests to further relax our requirements, and allow icosahedral orbits that are not archimedian to increase the integration strength. The best we can hope to achieve is a 9-orbit. The most suitable candidate seems to be the so-called improved football. It is a 9-orbit that offers a nearly semi-regular symmetry and is almost indistinguishable from the more familiar football shape, whose vertices offer only a 5-orbit.

8.2.4 Subarray Construction with Local Symmetry

In the previous sections we have focused on the distribution of the subarray phase centres. In this section we will address the organisation of radiating elements within the subarrays. As we found in section 8.2.3 the most suitable subarray distribution to be that of a truncated icosahedron that is very similar but different from the archimedian truncated icosahedron, we will develop a construction of the subarray geometry that is based on properties of the archime-

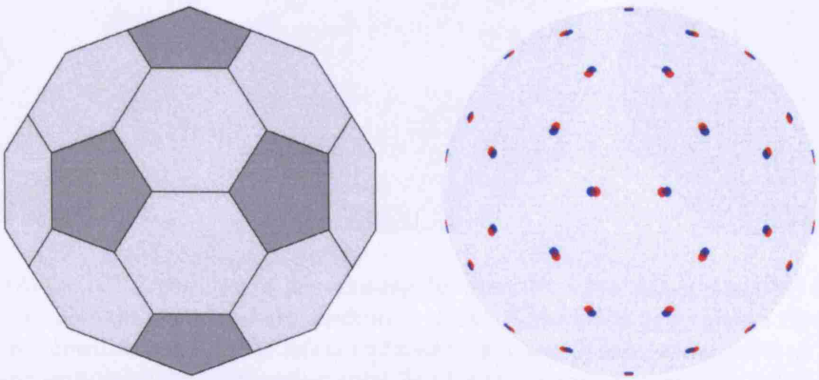


Figure 8.15: The blue dots are the vertices of the archimedian truncated icosahedron (football shape) and the red dots are the vertices of an irregular truncated icosahedron (the pentagons are slightly bigger) that preserves symmetry, improves integration strength, but is not archimedian. The dual solid of the perturbed design can still be face-uniform. Most importantly, the sampling points of the "improved football" are a 9-orbit, rather than just a 5-orbit for the archimedian version[126]. With reference to figure 8.14, the construction of the improved football can be obtained with the starting point of the orbit at 29.33% of the red edge length, counted from the point where the icosidodecahedron is formed. For the semi-regular football the starting point is chosen at $\frac{1}{3}$ of the same edge.

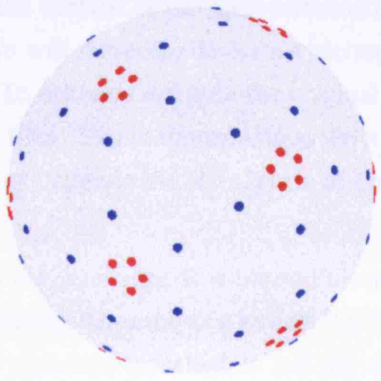


Figure 8.16: The blue dots are the vertices of the archimedian truncated icosahedron (football shape) and the red dots offer the optimal sampling distribution with respect to integration strength. In [127] it is shown that this distribution is the 9-orbit with the best approximation of integrals of spherical harmonics of tenth and eleventh degree and that there is only one such orbit. However optimal in terms of integration strength this distribution may be, as a subarray distribution it is not very useful as it is clear from the figure that sampling points are far from evenly distributed on the sphere, in the sense of maximising the minimal distance between the points.

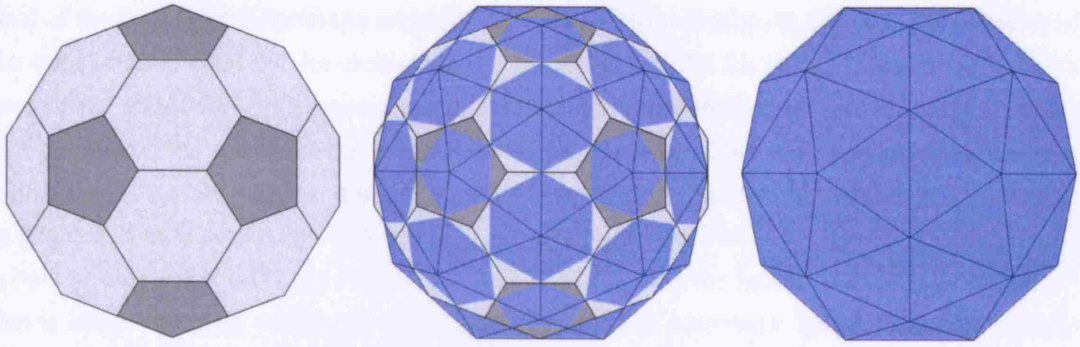


Figure 8.17: On the left a rendition of the semi-regular truncated icosahedron (football) is shown. On the right its dual solid, the pentakis dodecahedron, is shown. The middle figure shows clearly how each vertex of the archimedian solid on the left is replaced with a face in the catalan solid on the right and vice versa. The vertices of the archimedian solid lie on a circumsphere, while the faces of the catalan solid are tangent to an insphere. The edges of both solids are tangent to the midsphere. We will use the vertices of the archimedian solid as the phase centres of the subarrays and the faces of the catalan solid as a basis to construct the subarrays. Our completed construction will aim to approximate the midsphere.

dian solids and formulate the construction in a way that is general enough to also suit slight modifications of the archimedian solids.

The first question to be answered here is what to do when more than 60 subarrays are desired? For a large spherical array we still end up with fairly large subarray regions, which could perhaps be subarrayed further by a hierarchical succession of subarraying. For instance, the spherical array we are developing has a diameter of 800mm and a surface area of more than $2m^2$. That leaves us with a subarray area that is larger than a triangle of 30cm by 20cm. Although for our budget it is not desirable to further increase the number of signals from the array antenna, we would like the surface of the array antenna to be something that is close to a smooth spherical surface. We will therefore devise an hierarchical subarraying strategy that further divides the subarrays. To differentiate with the original subarrays, we will refer to the newly formed sub-subarrays as tiles. This is appropriate as we will not be processing the signals of the tiles separately, but rather combine the RF signals of the different tiles into one single signal before demodulation.

We start the construction from the semi-regular truncated icosahedron. Apart from its earlier described characteristics to be nearly the same as a known 9-orbit and having an apparent even distribution of its vertices, the truncated icosahedron has yet another desirable characteristic. Its dual solid, the pentakis dodecahedron, has isosceles triangles as faces. This makes it easier to devise a hierarchical tiling strategy, similar to a geodesic construction. The latter can not be applied without adaptation to the specific requirements of building a spherical array antenna, as was mentioned in the introduction to section 8.2. The truncated icosahedron and its dual pentakis dodecahedron are illustrated in figure 8.17.

As we will proceed with the dual pentakis dodecahedron, it is necessary to understand the concept of a dual solid. In the dual solid the vertices are replaced with faces and vice versa. Its construction is most easily defined for platonic or archimedian solids, and we will proceed defining the concept based on this assumption. Taking the dual solid works both ways, and the

dual of the dual solid is again the original solid. To avoid confusion, in this text we will reserve the term original solid for the archimedian solid and dual solid for its dual catalan solid. Each face of the dual solid is perpendicular to a radial line through the vertex it replaces. If there is a circumscribed sphere to the original solid that contains all its vertices (which is the case with the assumptions we have taken), than it follows that the faces of the dual solid are tangent to a sphere that is concentric with the circumscribed sphere. The latter sphere is the inscribed sphere to the dual solid. Let R be the circumradius and r be the inradius of the dual solid. All that is left to uniquely define the dual solid is the relation between r and R. The relationship between r and R is given by polar reciprocation about the midsphere of radius ρ:

$$\rho^2 = rR \tag{8.17}$$

The midsphere with midradius ρ is tangent to all edges of both the original archimedian solid and its dual catalan solid. Another construction that is valid for platonic and archimedian solids is the Dorman-Luke construction. In this construction, a circle is drawn through the midpoints of the edges that depart from a certain vertex. The polygon formed by drawing tangential lines where the circle crosses the edges is a face of the dual solid.

Using the construction of the dual solid defined above, we will outline a construction for the dual of the truncated icosahedron, the pentakis dodecahedron. The construction relies on the fact that the triangular faces of the pentakis dodecahedron have vertices on radial lines through the centres of the pentagons and hexagons of the truncated icosahedron. Because of the nature of the truncated icosahedron, a truncation of the icosahedron, these lines are also radial lines through the vertices of the icosahedron and the dodecahedron. We will therefore use these two other solids to make the construction of the pentakis dodecahedron easier. The construction will proceed as follows:

- In the fundamental domain of the icosahedral group, pick the orbital start point for the truncated icosahedron as a vertex of that solid. Call it T.
- In the same fundamental domain, pick the starting points (or vertices) for the icosahedron and the dodecahedron. Call them I and D respectively.
- Consider a radial line through T.
- Take a plane perpendicular to that line.
- Set the distance of the plane to the origin according to equation (8.17).
- Find two vertices in this plane as the intersection of the plane with radial lines through I and D.
- Using a suitable rotation out of the conjugacy class C_3, find the third vertex of the triangular face.
- Use 59 rotations of the icosahedral group (without the identity operator) to rotate the triangular face to obtain the 60 faces of the pentakis dodecahedron.

Let the vertex be given by $T = (t_x, t_y, t_z)$ and the origin by O. A plane perpendicular to the line OT has the equation:

$$t_x x + t_y y + t_z z + u = 0 \tag{8.18}$$

where u is a constant. Let this plane be tangent to the insphere in the point $K = (k_x, k_y, k_z)$. K is on the insphere with radius r and T is on the circumsphere with radius R. R is easily retrieved from T as $R = \sqrt{t_x^2 + t_y^2 + t_z^2}$, while the inradius of the pentakis dodecahedron can be obtained from it as:

$$r = R \cdot \frac{\frac{9}{2}\sqrt{\frac{1}{109}(17 + 6\sqrt{5})}}{\frac{1}{4}\sqrt{58 + 18\sqrt{5}}} \tag{8.19}$$

Assuming K and T to be on the same radial line, we can write:

$$K = \frac{r}{R} \cdot T \tag{8.20}$$

Evaluated in the point K, the equation for the plane becomes:

$$\frac{r}{R}(t_x^2 + t_y^2 + t_z^2) + u = 0 \tag{8.21}$$

$$r \cdot R + u = 0 \tag{8.22}$$

from which, using equation (8.17), we obtain u. The equation for the plane is therefore:

$$t_x x + t_y y + t_z z = \rho^2 \tag{8.23}$$

The lines through the vertices I and D are given by:

$$P_i \cdot I = P_i \cdot (i_x, i_y, i_z) \tag{8.24}$$

$$P_d \cdot D = P_d \cdot (d_x, d_y, d_z) \tag{8.25}$$

And the first two vertices of the pentakis dodecahedron are found as:

$$P_1 = P_i \cdot I \tag{8.26}$$

$$P_2 = P_d \cdot D \tag{8.27}$$

with:

$$P_i = \frac{\rho^2}{(t_x i_x + t_y i_y + t_z i_z)} \tag{8.28}$$

$$P_d = \frac{\rho^2}{(t_x d_x + t_y d_y + t_z d_z)} \tag{8.29}$$

The result of this construction is shown in figure 8.17. The construction as it was given above, relies on equation (8.19) for the inradius of the pentakis dodecahedron. If other catalan solids are to be obtained, this expression should be changed accordingly. Expressions for the radii of circumspheres and inspheres of archimedian and catalan solids (and of course platonic solids)

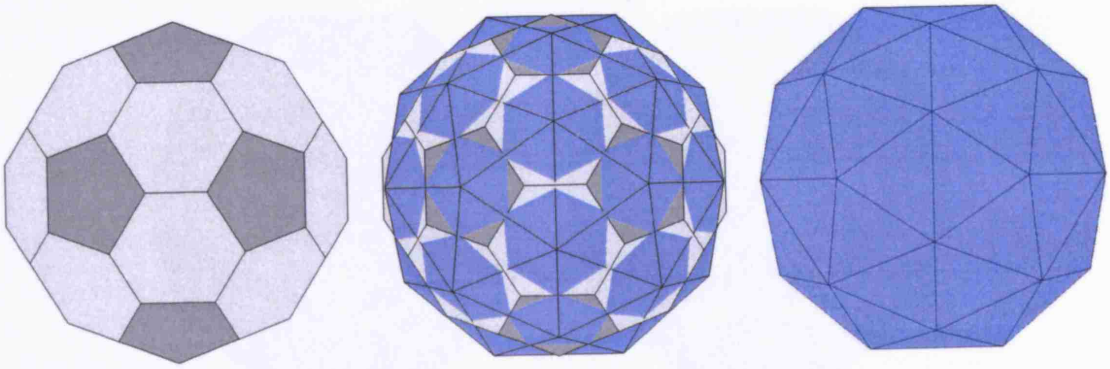


Figure 8.18: This figure shows the construction of the dual solid (on the right) of the improved football (on the left). The construction of the dual solid was carried out using properties of the archimedian version of the football as shown in figure 8.17, and as explained in the text. The obtained dual solid, which for the sake of simplicity we will also refer to as pentakis dodecahedron, is hardly any different from the original one. However, it turns out that the phase centres indicated by the vertices of the improved football are much more centred within their corresponding subarrays. This will be used in our advantage and is further illustrated in figure 8.19

are well documented in the literature and on internet. Some useful references are listed in appendix A. For solids that are not regular or semi-regular, it is not necessarily possible to find a circumscribed sphere and an inscribed sphere to its dual. However, in our construction we will use the expressions in (8.19) and (8.17) to obtain the radii, irrespective of the fact that we have a semi-regular truncated icosahedron or a slight perturbation thereof. Nothing else in the construction prevents us from obtaining another solid for solids that are not exactly archimedian solids (which for the sake of simplicity we will call dual solid as well). That is why the construction has been laid out here in so much detail. We now apply the construction to find the dual solid of the improved football, that was introduced in section 8.2.3 because it offers better integration strength. We use the orbital starting point of a non archimedian truncated icosahedron to start the construction, while keeping the rest of the construction unaltered. The result is shown in figure 8.18. Figure 8.19 illustrates, for both the semi-regular and the improved football, the structure of the original solid, whose vertices give the positions of the subarray phase centres (shown with red dots), as well as the obtained pentakis dodecahedron (shown with blue faces). The faces of the pentakis dodecahedron are not regular, but are isosceles triangles - the yellow and green vertices are different. Also shown, and that is apart from the coloured vertices the only difference with the centre figures in figures 8.17 and 8.18, are the midspheres. The black midspheres protrude through the faces of both the original and the dual solids. Only for the archimedian case on the left though, the midsphere is nicely tangent to the edges of both the original and the dual solid. Looking at a single subarray, any of the isosceles triangles with two yellow dots on the baseline and a green dot as the top of the triangle, it can be seen from the figures that in case of the improved football on the right, the red dots (representing the phase centres) are closer to the baseline than is the case in the figure on the left. Considering that these triangles are isosceles, the edge between the yellow dots being the longest, it is intuitively more difficult to design a subarray with its phase centres according to the red dots in the left figure. As a first order approximation this can also be understood from

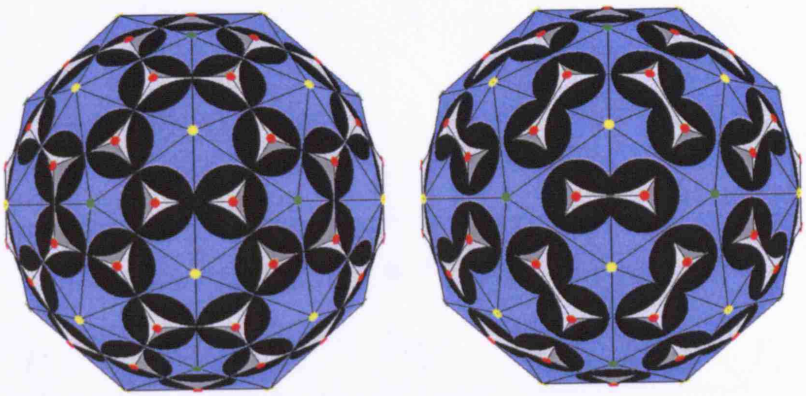


Figure 8.19: The centre figures of 8.17 and 8.18 are repeated, in addition to the midspheres (in black). On the left the construction is shown based on the semi-regular football, on the right it is shown for the improved football. The red dots, indicating the phase centres, are better positioned to represent the blue subarrays in the figure on the right. The construction can clearly be further improved to become a smoother approximation to the midspheres.

the fact that the centre of gravity of the triangles is closer to the long edge than the red dots in the left figure indicate. This seems to illustrate quite well why the improved football offers a better distribution for the phase centres. With each phase centre better positioned to represent the triangle around them, a better cover over the sphere is obtained. From figure 8.19 it is also clear that some parts of the subarray faces lie inside the midsphere, while some parts of it lie outside the midsphere. Furthermore, it seems a straightforward operation to lift the centre parts of the subarray outwards a little, so that the centre part becomes tangent to the midsphere and contains the phase centre. Rather than following a geodesic construction and just dividing the triangle in four smaller tiles and then radially projecting their vertices on the sphere, we want to be careful about keeping the centre tile tangent to the midsphere. We will therefore devise a more involved procedure, that allows us to keep control of the angular orientation of the tiles. As the tiles will each be radiating with a beam maximum in the normal direction, this is something that is important in the design of a spherical array antenna. When geodesic domes are constructed for construction purposes for instance, this may not be an issue.

The procedure starts in the same way as a geodesic construction. Let the three vertices of a triangular subarray be called P_1, P_2 and P_3. Let P_{12}, P_{23} and P_{31} be the midpoints of the edges. We can now divide the triangle into four smaller triangles as illustrated in figure 8.20. The centre tile contains the intersection point of the subarray with a radial line through the phase centre. The next step is to lift up the centre tile so that it becomes tangent to the midsphere, along a radial line through the desired phase centre. We do not just lift the centre tile, but rather find the intersecting points of an infinite tangent plane with radial lines through the points P_{12}, P_{23} and P_{31}. Let the new points be called P_{12L}, P_{23L} and P_{31L}. They are the final vertices of the centre tile. The centre tile contains the phase centre of the subarray (or at least a radial projection thereof) and is tangent to the midsphere in that point. Note that even if the original face would not have been perpendicular to the radial line through the phase centre, by construction we have forced to centre tile to be perpendicular to that line.

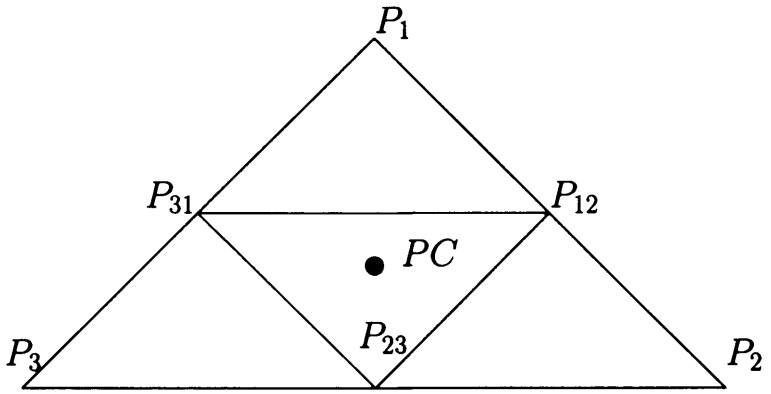


Figure 8.20: The isosceles triangle of a subarray is split into four smaller tiles. A construction is devised to reshape and replace the tiles so that we end up with a subarray that is likely to have its phase centre in the desired position PC and the maximum of its radiation pattern along a radial line through that point.

The next step is to construct the three other tiles surrounding the centre tile. The easiest solution at this point would be to consider the triangle $[P_1; P_{12L}; P_{31L}]$ and so on. However, this would result in a rather pointy construction, not very smooth on the sphere. The three peripheral tiles would also not be tangent to the sphere at all and the angles between the subarrays would be rather steep. Therefore, it makes sense to let the tips of the peripheral tiles dip down, so that they become tangent to the midsphere. However, note that the dihedral angle between the centre tile and the peripheral tiles would not be the same for the three tiles. As we started from an isosceles triangle, one angle is different, and smaller than, the two others. Ultimately we have to place radiating elements on these tiles. Ideally, the phase centre of their combined radiation pattern has to be in that one point on the centre tile that we are targetting. This task can be relatively easy if there is enough local symmetry in the subarray and almost impossible if there is not. It would therefore make sense to orchestrate the corrections on the tips of the peripheral tiles so that they all dip down an equal amount and their dihedral angles with the centre tile are all the same. To do this we first retrieve the three planes tangent to the midsphere and containing an edge of the centre tile. We then find the smallest dihedral angle of these planes with the centre plane and lift the other two planes so that they are no longer perfectly tangent, but at least the subarray as a whole becomes more symmetric by having three identical dihedral angles with the centre tile (it still is not rotationally symmetric as that would be impossible, but there is a definitive improvement that will become clearer when we place the radiating elements on the tiles).

As the outlined procedure solves a specific problem in the design of a spherical array antenna and is up to the authors knowledge not documented elsewhere in the literature, we will conveniently also include the design formula's for the two main steps. These are:

- Find the tangent point of a plane through the points $P_{12L} = (a_x, a_y, a_z)$ and $P_{31L} = (b_x, b_y, b_z)$ and a sphere of radius ρ.
- Find the position PP_1 of the tip of the peripheral tile $[P_{12L}; P_{31L}; PP_1]$, given that PP_1 lies on a radial line through P_1 and the dihedral angle between the tile and the centre tile

is α.

The first problem is solved by solving the following system for the tangent point $T = (t_x, t_y, t_z)$:

$$\begin{cases} t_x a_x + t_y a_y + t_z a_z &= \rho^2 \\ t_x b_x + t_y b_y + t_z b_z &= \rho^2 \\ t_x^2 + t_y^2 + t_z^2 &= \rho^2 \end{cases} \tag{8.30}$$

Two possible tangent points are found by solving the quadratic equation:

$$At_z^2 + Bt_z + C = 0 \tag{8.31}$$

with:

$$\begin{aligned} A &= \left(\frac{a_z b_y - b_z a_y}{a_x b_y - b_x a_y}\right)^2 + \left(\frac{a_z b_x - b_z a_x}{a_y b_x - b_y a_x}\right)^2 + 1 \\ B &= -2\left(\frac{(a_z b_y - b_z a_y)\rho^2(b_y - a_y)}{(a_x b_y - b_x a_y)^2} + \frac{(a_z b_x - b_z a_x)\rho^2(b_x - a_x)}{(a_y b_x - b_y a_x)^2}\right) \\ C &= \frac{\rho^4(b_y - a_y)^2}{(a_x b_y - b_x a_y)^2} + \frac{\rho^4(b_x - a_x)^2}{(a_y b_x - b_y a_x)^2} - \rho^2 \end{aligned}$$

and substituting the roots t_z in the equations:

$$t_x = \frac{-(a_z b_y - b_z a_y)}{(a_x b_y - b_x a_y)} t_z + \frac{\rho^2(b_y - a_y)}{(a_x b_y - b_x a_y)} \tag{8.32a}$$

$$t_y = \frac{-(a_z b_x - b_z a_x)}{(a_y b_x - b_y a_x)} t_z + \frac{\rho^2(b_x - a_x)}{(a_y b_x - b_y a_x)} \tag{8.32b}$$

from the two solution we choose the one that is closest to P_1. Having calculated the tangent point, the equation for the tangent plane is simply:

$$t_x x + t_y y + t_z z = \rho^2 \tag{8.33}$$

and the point PP_1 is the intersection of the radial line through P_1 with this plane.

The solution to the second problem can be reduced to a simple 2D geometrical problem by first finding the point H as the intersection of the line $[P_{12L}; P_{31L}]$ around which the tile rotates (the tile is rotated and deformed at the same time) with the plane of rotation of PP_1. As the latter slides on a radial line through P_1 (see figure 8.21), this is the plane containing that line and

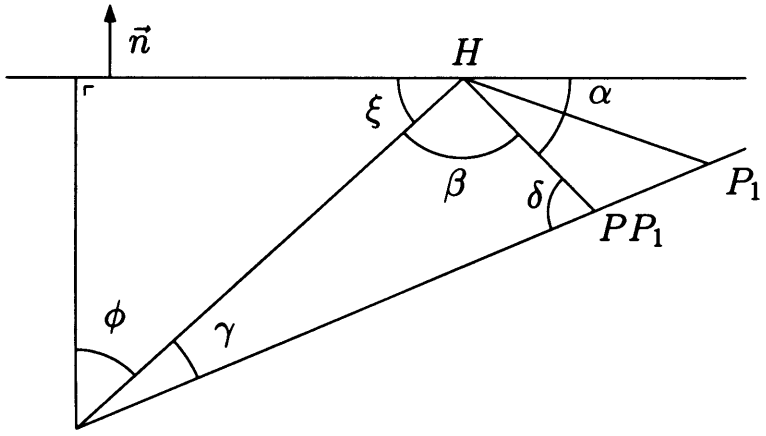


Figure 8.21: Geometry to retrieve the position of the tips of the peripheral tiles.

perpendicular to $[P_{12L}; P_{31L}]$. The point H is found as:

$$I_1 = \|P_1 - P_{12L}\| \tag{8.34a}$$

$$I_2 = \|P_1 - P_{31L}\| \tag{8.34b}$$

$$I_3 = \|P_{31L} - P_{12L}\| \tag{8.34c}$$

$$k = \frac{I_1^2 - I_2^2 + I_3^2}{2I_3^2} \tag{8.34d}$$

$$H = P_{12L} + k \cdot (P_{31L} - P_{12L}) \tag{8.34e}$$

With reference to figure 8.21, the solution is now found as:

$$\beta = \pi - \alpha - \xi \tag{8.35a}$$

$$\xi = \frac{\pi}{2} - \phi \tag{8.35b}$$

$$\phi = \widehat{\vec{n}, \vec{H}} \tag{8.35c}$$

$$\gamma = \widehat{\vec{P}_1, \vec{H}} \tag{8.35d}$$

$$\delta = \pi - \beta - \gamma \tag{8.35e}$$

$$\|OPP_1\| = \frac{\sin\beta}{\sin\delta}\|\vec{H}\| \tag{8.35f}$$

$$PP_1 = \frac{\|OPP_1\|}{\|P_1\|} \cdot P_1 \tag{8.35g}$$

Which concludes our construction for one subarray. As before, all subarrays remain identical and other subarrays are obtained with the rotations of the icosahedral group. The tiles within any one subarray are however not identical. As the differences are small, we can still manufacture all tiles to be identical and a little bit smaller than the sizes that come out of the construction above. That way we can leave some gaps between the tiles, so that the gaps are different rather than the tiles themselves. Figures 8.22 and 8.23 illustrate the finished construction with tiles covering the midsphere. In figure 8.22 the complete construction is shown for both the semi-regular

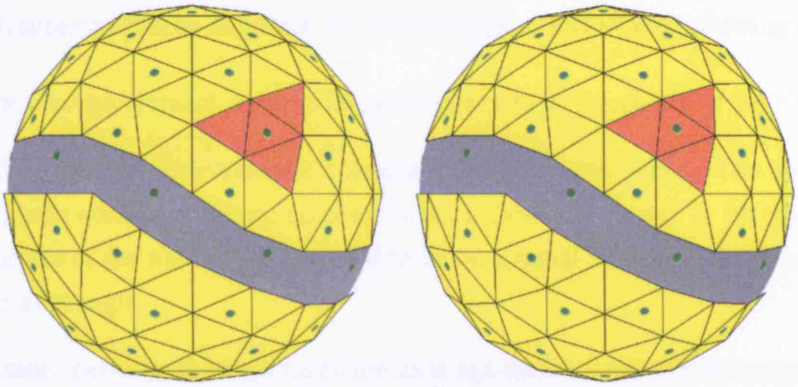


Figure 8.22: The finished construction with tiles covering the midsphere is shown for the semi-regular truncated icosahedron on the left and the improved football on the right. The green dots indicate the phase centres. One subarray is coloured differently to show the position of the phase centre within the subarray. Again we see a more favourable position for the construction based on the improved football. Two rows of tiles have been taken away to show how the tiles fit on the midsphere.

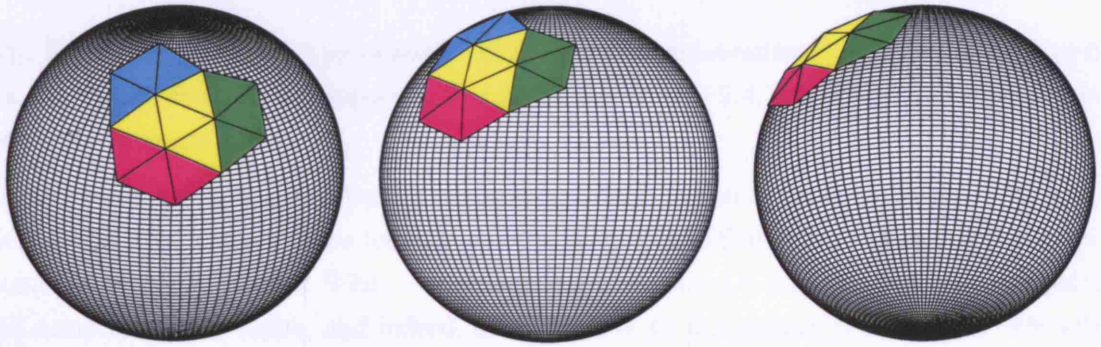


Figure 8.23: The four yellow tiles make up one single subarray. The other tiles belong to neighbouring subarrays. This setup will be used for our first measurements.

truncated icosahedron as the improved football. The green dots indicate the phase centres. One subarray is coloured differently to show the position of the phase centre within the subarray. Again we see a more favourable position for the construction based on the improved football. Two rows of tiles have been taken away to show how well the tiles cover the midsphere. In figure 8.23 one subarray is shown with a few tiles that belong to surrounding subarrays. We will come back to this setup later as it is the setup of one of our first measurements, but we show it here from a few viewpoints as another illustration of how the tiles fit on the sphere. In section 8.2.5 some attention will be given to the internal layout of radiating elements on each tile.

8.2.5 Single Tile Layout

On each triangular tile, three radiating elements are to be placed. The characteristics of the radiating elements were described in section 3.6.3. The development details of the tiles and elements will be covered in section 9.4.7. On the top layer, the radiating elements have only a simple ring. Here the layout of the tile, that is the positioning of the rings on the tile is considered. By locally changing the configuration of the antenna elements, the pattern of each tile and

therefore each subarray can be adjusted. We have chosen to follow the following strategy:

- For each antenna element, assume its phase centre to be the centre of the ring.
- Each tile is an isosceles triangle, containing 3 antenna elements. The triangle formed by the phase centres of the antenna elements is however chosen to be equilateral. The phase centre of the tile is then assumed to lie on a radial line through the centroid of the equilateral triangle.
- The distance between the antenna elements is optimised to achieve maximum gain, while keeping a circularly symmetric radiation pattern that appears to have a single phase centre.
- The position of the equilateral triangle of antenna elements on the tile is found so that the phase centre of a subarray of four tiles is in the right position, according to the calculations in the previous sections.

The first two points are straightforward. The third point is achieved by simulations of a tile with an EM solver. More on this aspect will be covered in section 9.4.7. Rests finding the position of the phase centre of each tile.

As the phase centre of each subarray has been specified to be in a specific position P, the positions of the phase centres of the four tiles need to be chosen so that their combined phase centre coincides with that point P. When combining several antennas it may be difficult to pinpoint the combined phase centre, and indeed, there may not be one unique phase centre. We will therefore make the simplifying assumption that the phase centre of a combination of antennas is found as the vectorial sum of their phase centres. That this assumption is generally incorrect but defendable in this particular case is illustrated in figure 8.24.

If all tiles are identical, and hence their phase centres in the same position on each tile, and the position of each tile is fixed within the subarray, then there are no degrees of freedom left to make sure that the four phase centres vectorially combine in a given point. An overall design aim was to have all tiles identical. In the manufacturing process of PCBs a mask for etching needs to be made and the manufacturing costs for 60 times 4 tiles is substantially higher than that of a single batch of 240 tiles. However, as the circuits are typically constructed on larger sheets of material and subsequently being cut, different cut-out shapes for each of the four tiles of a subarray may be acceptable. As explained in chapter 9, our approach has been to keep all tiles identical, and allow varying widths of spacing between the tiles.

8.3 Summary and Conclusions

This chapter focused on the placement of the elements of a spherical array antenna. Deciding how to distribute the elements evenly over the surface of the sphere is by and large not so straightforward as with planar or cylindrical array antennas.

A number of strategies were discussed to place the elements on the sphere, of which two were suggested as optimal, depending on whether the array has many or few radiating elements. If

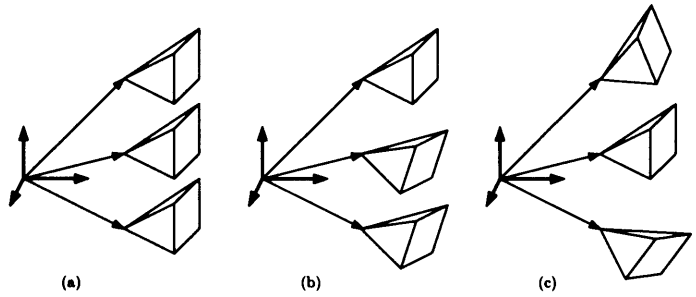


Figure 8.24: The estimation of the phase centre of an antenna array is not straightforward. A single and unique phase centre does generally not exist. Generally speaking, the estimation of a combined phase centre should take into account the positions of the constituent antennas, their relative orientation and their radiation pattern (including mutual coupling effects). Nevertheless we use a simple approximation by taking the vectorial mean of the phase centre positions of the array elements. This would incorrectly result in the same phase centre for the three cases (a), (b) and (c) shown in this figure, which would also be coincident with the phase centre of the antenna in the middle. However, by the construction explained in section 8.2.4, it was enforced that locally, within each subarray, the 4 tiles are oriented so that the situation is equivalent to (c). This was achieved by bending the three peripheral tiles away from the centre tile by an equal amount. For that case, taking the vectorial mean of the phase centres is correct. Note that because only two of the four tiles are tangent to the same spherical surface, the obtained phase centre is not exactly on the radial line to the phase centre of the centre tile. However, this error is negligible.

the number of elements is small enough to control each of them separately, then the optimal choice for the element positions is given by the t-designs. If several element signals are merged into a single channel, a technique called subarraying, then there are symmetry requirements that complicate the placement of the elements. Only for a very few number of elements (20 or less), a perfectly uniform distribution can be found in the form of the vertex set of one of the platonic solids.

The technique of subarraying was given an in-depth treatment, as it was decided to use this technique on LISA, a prototype spherical array the development of which is covered in chapter 9. Subarraying introduced extra complications, because of symmetry requirements on the overall distribution of array elements.

Chapter 9

Development of a Prototype Spherical Array Antenna (LISA)

For the sake of theory verification and empirical design, a prototype of a spherical array antenna, called LISA is being developed. The proceedings of this development are outlined in this chapter.

Upon completion of this thesis, LISA has been fully designed at system-level and for some parts down to component level. Development of the RF front-end is also in an advanced stage. Development of the digital back-end is left for future work, but is also described in this chapter.

As will be explained in section 9.1, LISA is an array with 720 radiating elements, organised on 240 flat triangular panels. The first batch of panels has been manufactured and tested, and been put together in a spherical subarray configuration for measurements. The results of the measurements were highly satisfactory and the manufacturing of the complete batch of panels has been ordered.

Construction of the full array antenna will be completed soon. This will be the start of a comprehensive testing programme in which the principles presented in this thesis will be demonstrated. After this proof-of-concept phase of experiments, the prototype will also help to study and test more advanced techniques that will be developed in the future. More on future work is covered in chapter 10. For the proof-of-concept phase that is to follow immediately, the measurements to be carried out have been designed already. LISA has been designed and developed towards these envisaged experiments, rather than to serve a particular application.

A key asset for the planned measurements is our new antenna test facility, described in section 9.2. The envisaged test plan is described in section 9.3. The development of LISAs subsystems is covered in section 9.4. The results from the preliminary measurements on a first batch of panels are reported in section 9.5. In the same section, the performance of the completed array antenna is anticipated, based on the results that are already available.

9.1 LISA Overview

Before getting into the details of the development of the antenna, the key facts about the design of LISA are resumed here. The design was obviously driven by the research reported in earlier chapters. Especially the findings in chapters 3, 6 and 8 have culminated in the design of LISA.

LISA is a prototype large spherical array antenna, designed to work at 3 GHz. The array was eventually designed as a multi-faceted body with planar faces, albeit with a very large number of faces. 240 panels fit together to form an approximately spherical surface of 84 cm diameter. Each of the panels features three radiating elements. The edge lengths of each triangular panel are only just over a wavelength in size, so that electromagnetically the structure is very smooth and not very different from a perfect sphere.

The radiating elements that are constructed on the tiles were already described in section 3.6.3. They are implemented on a multi-layer PCB, with a ring on the top layer in which a rotating TM_{11} mode is excited. The three radiating elements on each tile are rotated 120° relative to one another. Because the antennas are circularly polarised, this rotation can be compensated for by phasing the feeding signals so that the phasors align with one another. This not only irons out some asymmetries in the radiation pattern, it greatly improves the polarisation purity and the CP bandwidth of the panels, with respect to that of a single element. This principle is known as sequential rotation[128][129][130][131].

Each panel has an embedded feeding circuit on its backside. The circuit implements an equal 3-way power division and the relative phasing of the elements to achieve the sequential rotation. The elements cannot be controlled separately - there are a total of 240 connectors to the surface of the array antenna, for a total of 720 radiating elements. Each panel becomes in fact a single element with a high-gain element pattern, and the array can be modelled as an array of 240 high-gain elements. Refining the array model described in chapter 5 to accurately model the prototype array, and bring model and measurement results in correspondence, will be the next challenge.

The geometrical configuration of the array is the one proposed in section 8.2 on subarraying. The subarrays group 4 panels into one subarray, to reduce the number of channels further to 60. This is done chiefly to save on expensive IQ modulators/demodulators, of which one is needed per receiving channel.

To operate the array antenna, a variety of options are possible. Originally the array was designed to have 60 subarrays, each with a separate control. The first stage immediately behind the antenna panels will be a set of 60 4-way power dividers, followed by the 60 modulators/demodulators described in section 9.4.1. Fortunately, a 4-way power divider is straightforward to build with 3 Wilkinson power dividers.

For research purposes, it is a key advantage of the design that it has the flexibility to allow several alternative configurations. Possible configurations are:

- Standard configuration with 60 subarrays in a quasi-uniform distribution.
- Low-gain configuration, with 60 active panels (the centre panels of each subarray) instead of 240. The other panels are terminated.
- Maximum gain configuration, with an active side of the array and the other side switched off. The 60 available channels are used for 60 neighbouring tiles on one side of the

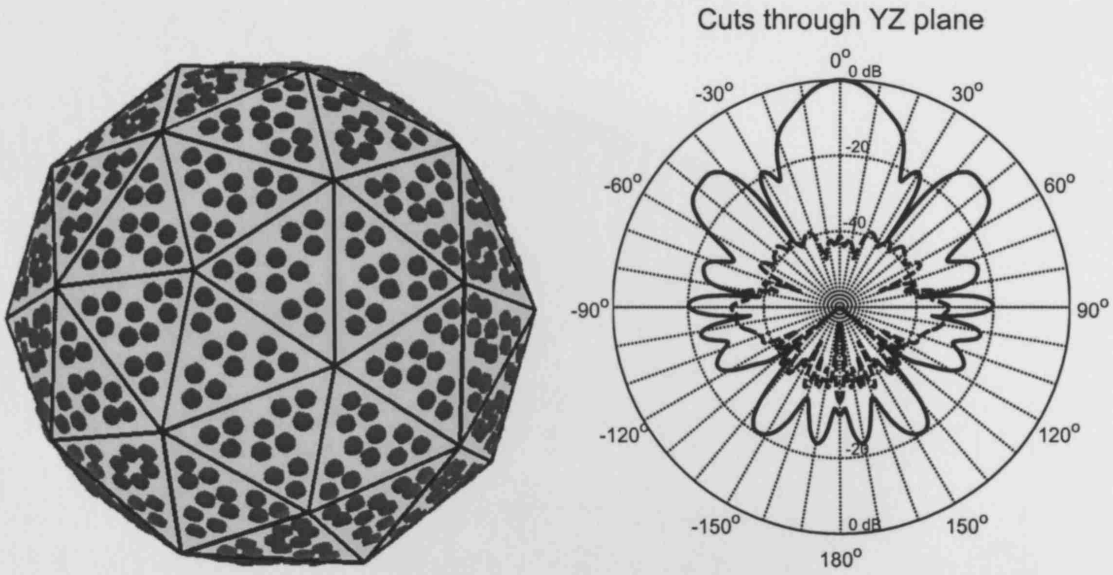
array. This does not allow to steer the beam in other directions, but makes it possible to determine the performance that can be achieved if 240 channels were available.

- Ultimately it is also possible to upgrade the array to have 240 separately controllable channels, so that a fully steerable pencil beam can be formed.

A possible reason to investigate the second configuration is that the subarrays have a very high gain and a narrow region of circular polarisation. It may be possible that holes in the radiation pattern are present between two subarrays, where the polarisation purity is poor.

As was explained in section 7.3, it is necessary to run a detailed analysis to predict the performance of a complex array structure as that of LISA. Simple design rules for the directivity and other antenna characteristics have not yet been developed. Also, the functionality of the code to run a detailed analysis does not yet include the newly developed radiating elements. A rough estimation of the array performance was therefore based on the analysis of a representative spherical array of waveguide fed apertures shown in figure 9.1a, with a similar amount of elements and subarrays, but a slightly different geometry. The feeding structure was assumed to be ideal so that no internal mutual coupling occurs. The results from the analysis explained in chapter 5 is shown in figure 9.1b. The array was excited to achieve a moderate directivity of 30 dB with the highest mode in its pattern $B_s = 5$. Polarisation purity was found to be good, but can be expected to be worsened in reality by increased mutual coupling effects. The directivity was estimated to be better than 20 dB. Aliasing effects are manifesting themselves as relatively high sidelobes in the radiation pattern. From the following argument, this also is to be expected.

The electrical size of the array is 25, and therefore its maximum achievable directivity (with its 720 elements separately controlled) is 56 dB. With only 60 controllable channels however, we know we can form a 9-design, so that $L_s = 9$. We could push the limit by assuming $L_s = 14$, as the corresponding minimal design would have 63 elements. But even then $L_s - L_m < 0$ and the errors from aliasing cannot be filtered out. From the other hand, if the spatial integration of the subarrays reduces the bandwidth of the sensed field enough, to $L_m = 13$ for instance, then there would be just enough spatial bandwidth left to form a directive pattern of $20\log_{10}((14-13)(14-13+2)) = 9.5$ dB (which is admittedly not more than a single element can achieve). Hence, it comes at no surprise that the pattern shown in figure 9.1 suffers from aliasing. We should from the other hand not be too pessimistic about the performance of the array with 60 subarrays. The pattern from a single subarray has been measured and already provides 25 dB directivity. In the worst case, the array performance can not be worse in at least 60 directions, as otherwise the other subarrays could be switched off. An educated guess for the attainable directivity of the array is therefore 20 dB. With the 240 tiles separately controlled, the situation looks better. In the example on page 178 it was already calculated that this results in a sampling degree of $L_s = 29$ so that $B_s = 4$ and the maximum achievable directivity becomes $D_s \approx 28$ dB.



(a) (b)

Figure 9.1: An approximate model of the prototype spherical array shown in (a) was analysed to predict the performance of LISA. In (b) the co- and cross-polarisation of the obtained radiation pattern are shown. The array consists of 60 planar subarrays, geometrically organised as the faces of a pentakis dodecahedron. Each face has 12 radiating elements, fed with an identical signal. The elements are assumed to be circular apertures, protruding from a sphere at the angular positions obtained as the radial projections of the element positions shown in figure (a).

9.2 Antenna Test Facility

In figure 9.2 a 3D model of the test facility is shown. The test facility consists of a shielded anechoic room and a spherical near-field measurement (SNFM) system.

The shield keeps unwanted radiation from sources outside the facility outside, and also keeps the radiation from the testing process inside, so that no interference occurs with the neighbours. The shield is designed and tested to provide 100 dB isolation between 2 and 40 GHz. This number compares a standard setup of two opposing horn antennas with and without the shield between them.

The chamber is anechoic, which means that despite its relatively small dimensions a free space environment is simulated. In a free space environment we can perform measurements without perturbations from reflections that occur at the chamber walls or other equipment inside the room. The anechoic behaviour of the room is obtained by covering the walls and RF equipment with an RF-absorbing material. The material is a carbon-impregnated foam and is shaped in panels with inward facing pyramids. The pyramids provide a smooth impedance transition from free-space to an impedance with a lossy characteristic. Also, the pyramidal shapes give the absorbing response a wideband characteristic.

The SNFM system collects near-field data that consists of phase and amplitude samples collected over a spherical surface (the measurement sphere) that completely surrounds the Antenna Under Test (AUT). Once the data is collected, the near-field (NF) is known up to a certain harmonic degree, dependent on the density of the sampling grid. This information is sufficient to

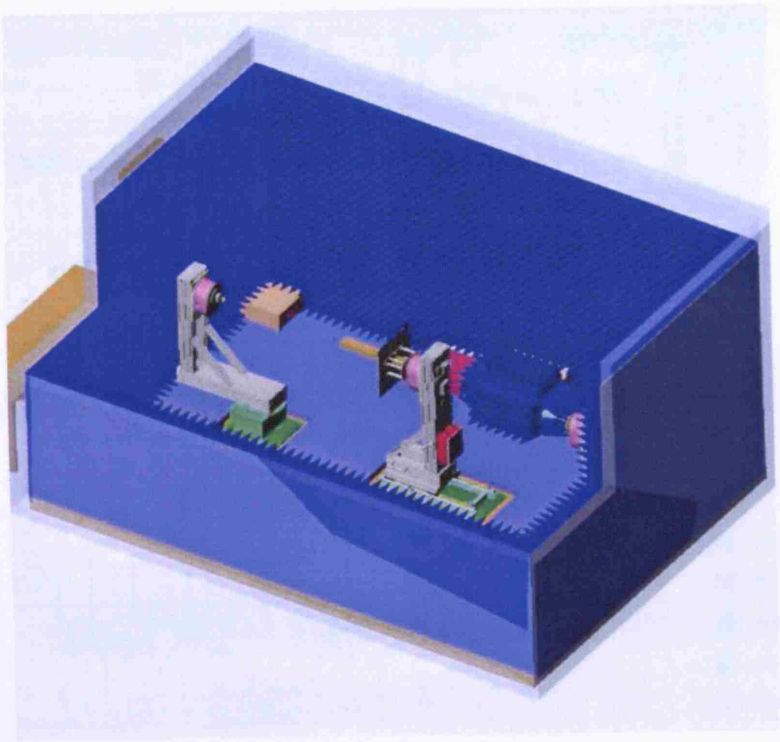


Figure 9.2: Layout of anechoic chamber

calculate the far-field, which is generally the desired information from antenna measurements. For processing of the NF data, it is irrelevant what is inside the sphere. Hence, the AUT can be mounted anywhere inside the measurement sphere. The minimal sphere that surrounds the AUT is called the minimal sphere. The size of the minimal sphere determines the highest degree of spherical harmonics that can be present in the AUT's radiation pattern, and therefore sets the minimal density of the grid on which the samples should be collected.

Processing of the data is based on the same theory of spherical wave expansions that was explained in chapter 2 and will only be briefly resumed here. The gathered data samples are projected on the spherical basis functions, so that the spherical wave expansion is obtained. The field can subsequently be reconstructed at any radius. The limit for infinite radius gives the far-field radiated by the antenna.

The complete SNFM system consists of the spherical near-field scanner, the RF subsystem, data capturing hardware, and processing hardware and software. The complete system is schematically depicted in figure 9.3.

The blue parts in figure 9.3 are part of the scanner or positioning system. The AUT is mounted on the leftmost tower and the probe that samples the near-field is mounted on the rightmost tower. The samples of the near-field are taken over a measurement sphere that surrounds the AUT. Rather than physically moving the probe around the AUT (which would be impractical in the limited available space) the AUT is rotated and the probe is mounted on a static tower. Therefore, the AUT tower has two axes of rotation in what is called a roll-over-azimuth configuration. The vertical axis is the azimuthal axis and rotates the AUT over 180 degrees. The

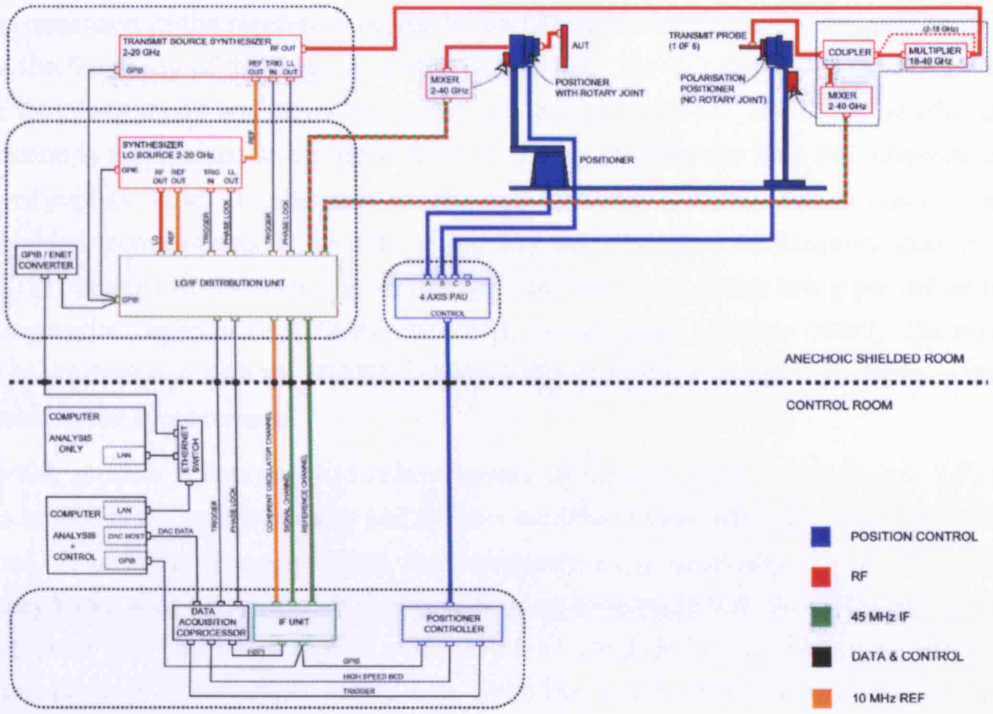


Figure 9.3: Anechoic chamber cable diagram

horizontal axis is the roll axis and rotates the AUT over 360 degrees. Hence, we can take measurements over a set of circles along the roll axis that together make up a full sphere. The angle convention deviates from what is used throughout this thesis, where the azimuthal angle ϕ corresponds with a rotation around what is here called the roll-axis and the elevation angle θ corresponds with a rotation around what is here called the azimuthal axis. We will only use this new convention when referring to the positioner hardware. The sampling grid on which the data is collected is defined by the governing software and programmed via a GPIB bus into the positioner controller. The controller steers the power amplifier unit (PAU) that drives the positioners. The positioners have synchro's, and feed an absolute position readout back to the controller. The controller is programmed to trigger the data acquisition every time that a position on the sampling grid is reached. An optional extra of the system is that the position data are made available in real-time to the acquisition system via a high speed BCD bus.

The near-field samples are taken by transmitting a continuous wave from the probe (whether the probe or AUT are transmitting should not make a difference in reciprocal antenna systems, but the sense could be reversed if necessary) and measuring phase and amplitude from the signal that is received at the AUT. This is repeated for every frequency of interest, and in every point of the sampling grid. The transmitted signal is synthesised in the transmit source synthesiser for frequencies between 2 and 18 GHz. Above 18 GHz, a multiplier will double the synthesised frequencies between 9 and 20 GHz to a maximum of 40 GHz.

Phase measurements are always relative to another phasor. In the SNFM system, the reference signal is taken from the transmitted signal by a directional coupler at the transmit tower. It is then downconverted in the reference mixer.

Downconversion in the receiver is driven by an LO source that runs at a frequency 45 MHz above the frequency of the transmit source synthesiser. The LO signal is split and amplified in the distribution unit and drives the signal and reference mixers. The IF signal after downconversion is multiplexed in the same cable with the LO signal for both the reference as the received signals. After amplification and filtering the received IF signals are brought outside the shielded room towards the IF unit, where they are digitised at 60 Mega-samples per second. The received digital samples are averaged and corrected before being passed on to the data acquisition coprocessor (DACP) over a high speed parallel interface (HSPI). The 60 MHz clock is synchronised with the 10 MHz reference signal to which the internal clocks of the two synthesisers are synchronised.

The whole process of data acquisition is controlled by the DACP, which triggers the RF instruments to step to the next frequency and triggers the IF unit to acquire data when phase lock is reached. Prior to the data acquisition, the instruments are programmed over the GPIB bus so that they know what to do when they receive triggers from the DACP. The RF instruments will also respond with phase lock signals when they are locked on the next frequency. This avoids that data needs to be collected over the slow GPIB bus, which allows the system to approach a speed of 10000 complex measurements per second.

While taking a new acquisition, a second computer system can access data that has been collected during previous acquisitions and can analyse the data.

9.3 Measurement and Operations

The first question to be asked when developing a piece of hardware is of course what function will it serve? In this section we will outline the experiments we are envisaging to perform once the array has been developed. These measurements are associated with modes of operation that may or may not need certain capabilities of the array antenna. Identifying the operational capabilities that will enable us to carry out the desired measurements on the array antenna is the objective of this section.

9.3.1 Measurements

Three classes of experiments have been identified to provide an empirical backbone for the work in this thesis.

1. Mutual coupling analysis experiments
   In this experiment we measure the near-field of a single radiating element on the surface of a conducting sphere. Mounting the sphere so that its centre exactly coincides with the centre of the measurement sphere allows to use the spherical wave expansion as it is obtained by the measurement system. This expansion can be compared to what is expected from the theory in chapter 4. The theoretically expected mutual coupling between two elements with a known spacing between them can be verified by connecting the transmitting end of the measurement system to one element and the receiving end to another element. This requires the design and manufacturing of two elements that are better suitable for this kind of measurement, perhaps the SAR element would be a good candidate.

Care would need to be taken to calibrate the measurement, as it is necessary to use a long cable from the transmitting end to the AUT. A high quality cable to do this is available in the facility.

2. Backprojecting experiments
   With the centre of the spherical array aligned with the centre of the measurement sphere, the spherical modes used in the software of the measurement system are defined in the same system of axes as the modes we have used to describe the field of the antenna array (if not, the tests described next can still be performed, but not so easily). The measured field can theoretically be backprojected to any radius greater than the minimal sphere, which in this case coincides with the surface of the antenna. By processing of the data, the near-field on the sphere can thus be obtained. However, the procedure is plagued by numerical difficulties (as shown in section 7.1). It would be interesting to see how accurate this can be done. What should be feasible without much difficulty is to work out which elements on the array are active and which ones are not. In the backprojected field, the active elements should "light up". It would be interesting to investigate the possibility to fine tune the array based on the backprojected near-field data.

3. Far-field synthesis experiments
   In this type of experiments the array antenna would be set so that it is capable of receiving with a theoretically predicted far-field pattern. Verifying this pattern should be a straightforward procedure in which the SNFM system is used in the way it is been designed for. This setup can also be used to obtain the Realised Gain Pattern (RGP) of one panel of the array, from which the performance of the complete array can be inferred as explained in section 5.3.2. This kinds of experiments have already been performed and the first results are available. They will be discussed in section 9.5. The setup required for this kind of experiments is illustrated by the topmost schematic in figure 9.4. We call this the receiving modus operandi of the array antenna. The shown data link would only be required between measurements, if the configuration of the array controller needs to be altered.

4. Beam steering experiments
   The ultimate proof-of-concept experiment would combine real-time beamsteering of the array antenna with a pattern measurement, so that the maximum of the antenna far-field pattern is kept pointing at the probe while the measurement is taken. The measured antenna pattern should be fairly isotropic but than with a positive gain everywhere, a condition previously unseen in any other type of antenna. In a first attempt to achieve this, the true position could be fed back from the measurement system to the array, so that it knows in which direction it has to steer the beam. In that case the same setup as before can be used, however this time the data port to the outside world will need to be used throughout the acquisition. In a more complicated experiment the array could work out the direction of incident radiation by analysing the received signals from its elements. A problem that arises is that the transmitting probe will most probably not be in the far-field of the array antenna. A combination of positioning the probe tower on the far-field

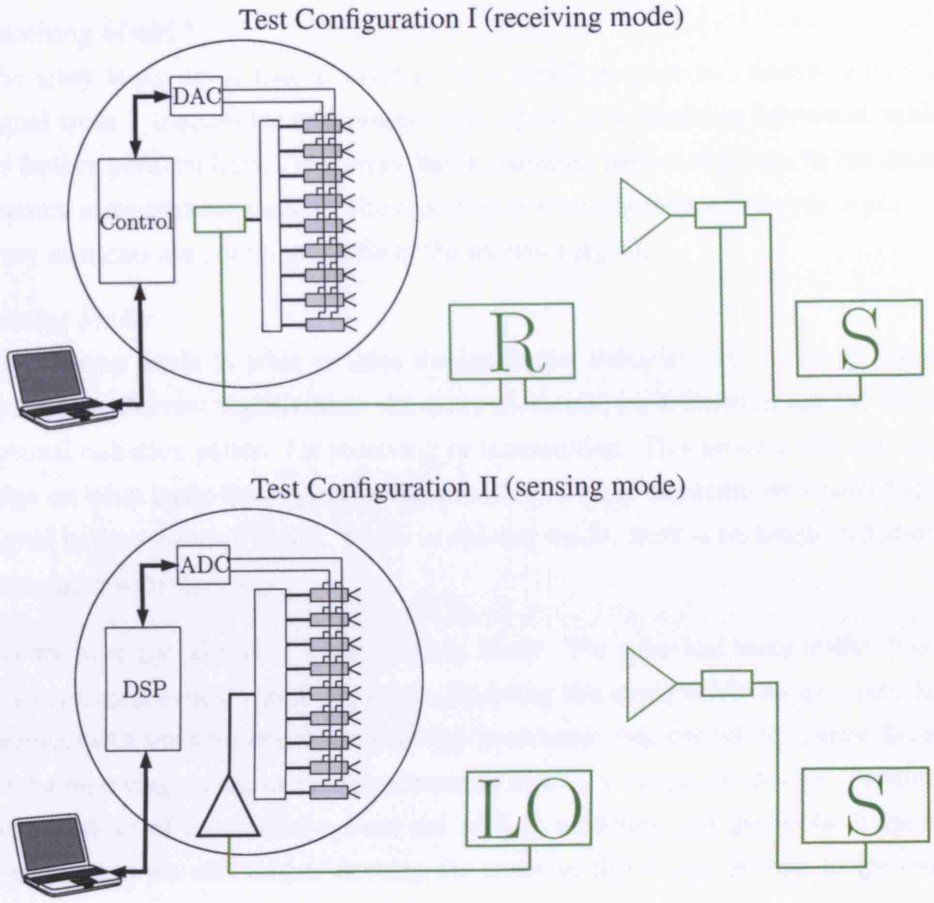


Figure 9.4: Two distinctly different setups are shown of the array antenna in the measurement facility. Both of these setups will be required to complete most of the measurements explained in this section. The components shown in green are part of the measurement facility. R is the receiver, containing an LO source. S is the source for the transmitted signal, LO denotes the receiver when used solely as an LO source. The thick cables carry RF signals. The parts behind the radiating elements are units that can either downconvert the RF signals to baseband in sensing mode or be used as phase and amplitude shifters of the received signals in receiving mode. The former requires an LO signal with enough power, which is derived from the (amplified) LO signal of the measurement system.

stand and focusing the array on a near-field radius rather than the far-field might offer the solution. During this experiment the array will need to be switched between two modi operandi, receiving and sensing, that are illustrated in the two schematics of figure 9.4.

The placement of the mixer at AUT side deserves closer attention. For sensing mode, an RF switch will need to switch the mixer out of the signal path. In 9.4 the mixer is shown to reside inside the AUT, so that switching between different configurations is performed inside the AUT. From figure 9.3 it can be seen that the standard placement of the mixer is on the AUT tower. It may be more practical to implement a bypass for the mixer with two switches on the AUT tower.

9.3.2 Modes of Operation

From the viewpoint of the array, the following two modes of operation were identified in the previous list of experiments:

1. *Receiving Mode*:
   The array is set up so that it could serve a useful purpose and receive a data carrying signal from a transmitter and present this signal to a receiving back-end, which is of no further concern here. The array has a radiation pattern that can be measured in an antenna measurement facility. The radiation pattern depends on how the signals from all array elements are combined to form the received signal.

2. *Sensing Mode*:
   The sensing mode is what enables the intelligent behaviour of the array antenna. By processing the raw signals from the array elements, an estimation can be made on the optimal radiation pattern for receiving or transmitting. This process delivers the knowledge on what basis the signals from the different array elements are combined into one signal in the receiving mode. While in sensing mode, there is no single radiation pattern associated with the array.

Note that we have not identified a transmitting Mode. The spherical array under development will not accommodate the transmitting mode. Enabling this mode would require extra hardware, but experiments in transmitting mode offer no knowledge that cannot be gained from experiments in the receiving mode, as the array antenna itself is a reciprocal device. Furthermore, in the standard set-up of the SNFM system the AUT is receiving and the probe is transmitting. In the next section we will further develop the array so that it can operate in the two modes described above.

9.3.3 Position Calibration

A procedure needs to be outlined to calibrate the phase references and amplitude levels for the array signal channels as explained in section 7.7. This calibration procedure could in fact retrieve position data for all array elements and feed it to the part of the processing unit that performs the spherical transform, while calibrating the phase references. The procedure could rely on the spherical geometry of the near-field scanner and the fact that all phase measurements will be relative to a fixed system delay. With the sphere of the spherical array made concentric with the measurement sphere, the following procedure could be followed:

- In case the panel positions would not be known already, find all element positions (θ_i, ϕ_i) by switching on one of them at a time and finding the maximum in the near-field pattern.
- Bring the first panel towards NP position and measure the phase of a pilot tone.
- From the position and phase of the first panel, a phase reference is calculated for the NP position.
- Bring one panel at a time towards NP position and adjust their phase to the NP reference plus what is required by the sample distribution and offset of the third Euler angle by the positioner.

Setting of the phase and amplitudes is done via the IQ modulator/demodulators described in section 9.4.1.

9.4 Development of LISA Subsystems

We can now identify at system level the functional blocks that can implement the two modes of operation from above. First we separate the hardware in the following categories: the RF front-end, the baseband circuitry, the digital signal processing unit (DSPU), the beam steering computer (BSC), the input & output subsystem (IOS), the power unit (PU), the array hardware and the radiating elements. Although merely part of the measurement system rather than the antenna, we will also discuss the development of a data-link between the antenna and the control room outside the anechoic chamber in the section on the IOS. Likewise, power needs to be provided to the antenna. The data and power links will remain part of the test facility.

9.4.1 RF front-end

The RF front-end has the same configuration in receiving and sensing modes, but is operated differently. In transmitting mode, the sense of amplification needs to be reversed, but as we will not implement this mode it will not be considered here any further.

In figure 9.5 the two configurations for the RF front-end are illustrated. The key elements in the RF front-end are IQ modulators/demodulators. The IQ mod/demods can be fed with the RF signals from the array elements and with an LO reference signal from the measurement system to downconvert the received RF signals to baseband. This is what is needed when the array is operating in sensing mode. In receiving mode, the IQ mod/demods can act as phase and amplitude shifters for the signals of the array elements, driven by baseband signals. The received signals are subsequently power combined to form the complete received signal. A picture of one of the IQ mod/demods is shown in figure 9.6. The manufacturers specifications are:

- RF/LO Freq: 3.0 ± 0.2 GHz
- IQ Freq: DC - 100 MHz
- LO Drive Level: +10 dBm nom
- Conversion Loss: 5.5/7.0 dB Typ/Max
- Image Rejection: 25/20 dB Typ/Min
- Phase Balance: 2 deg max
- Amplitude Match: 0.25 dB max
- L-R Isolation: 30 dB min
- L-IQ Isolation: 25 dB min

Apart from the IQ mod/demods, the RF front-end will incorporate a power combiner/splitter, an amplifier to boost the LO signal prior to being split over the elements in sensing mode, and possibly baseband filters on the baseband input/output lines. The latter are desirable when there is not sufficient isolation between the RF and baseband ports of the IQ mod/demods. RF

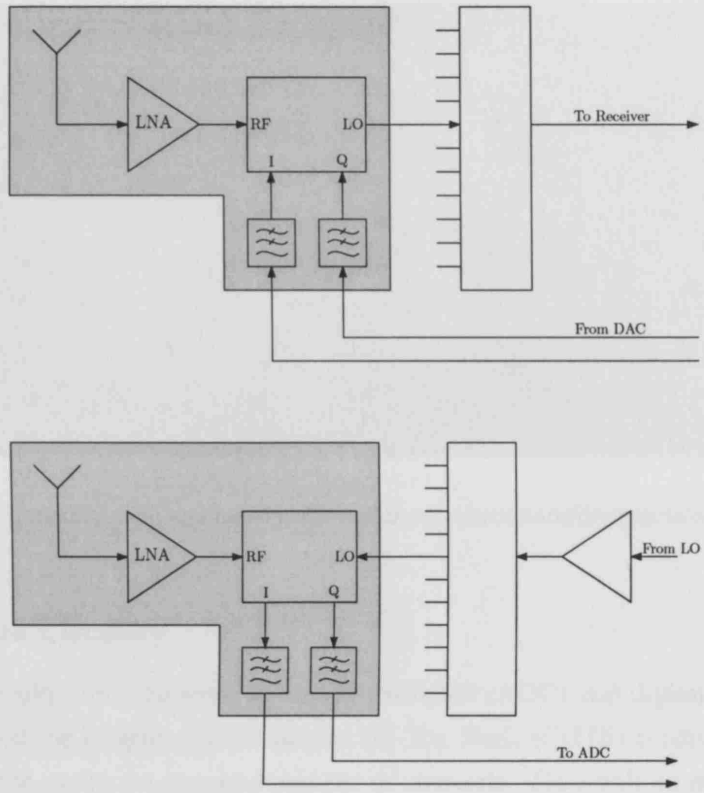


Figure 9.5: The RF-front is schematically drawn for both receiving and sensing modes.

filtering and low noise amplification should not be necessary if the array is operated only in the quite environment of the anechoic chamber. The power combiner/splitter will have to be enormous, with as many outputs as there are signal channels (60). This component has not yet been designed, and it is expected to be a challenge to avoid imbalances. It may be possible to correct the imbalances to some extent in the IQ mod/demods.

Although at a later stage, the DC-signals to the IQ mod/demods will be set digitally via a Digital to Analog Convertor (DAC), in a first instance, the I and Q DC-signals into the mod/demods are generated on a small circuit board that also holds the IQ mod/demods. The DC inputs are set manually with a variable voltage regulator. A board has been designed by another student that contains the regulators as well as some circuitry to protect the IQ mod/demods against damage from too high currents. Preliminary tests on the network analyser of the board with the IQ mod/demod have shown that the phasor of the S_{12} parameter can be accurately and repeatedly moved around within a complex circle. The next step will be to characterise the response of different IQ mod/demods and compare them with one another to check repeatability between the components. Also the reciprocity of the devices needs to be checked. If all goes well, the first RF subsystem is ready for mass production and the array antenna will be controllable. Until the digital back-end is developed, the procedure to set all the I and Q voltages will be tedious, and presumably only a few patterns will be measured this way.

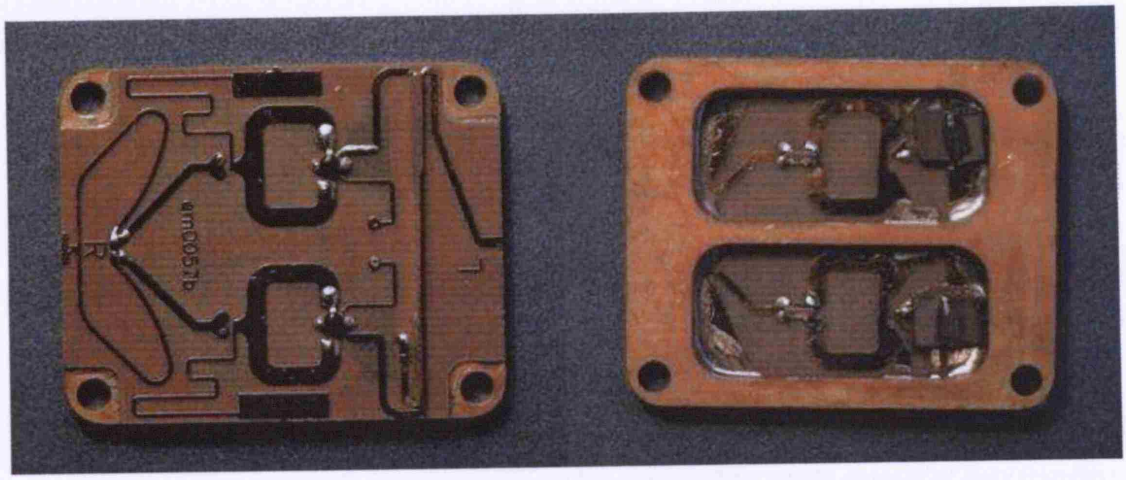


Figure 9.6: Top and bottom view of the IQ modulator/demodulator.

9.4.2 Baseband Circuitry

The baseband circuitry only consists of analog-to-digital (ADC) and digital-to-analog (DAC) converters, that will be bought as Commercial Of The Shelf (COTS) products. Several units will be combined to get to the required number of channels. They will be mounted inside the spherical array and connected to the IQ mod/demods. The units can be slow, as the required bandwidth is next-to-zero. With near DC operation, the units can operate sequentially over the baseband channels, which greatly reduces the cost. The total acquisition time T_a, possibly a multitude of the sampling time T_s determines the duration that the array needs to be put in sensing mode before switching back to receiving mode. The bandwidth of the baseband filters can simply taken to be the inverse of the total acquisition time:

$$BW_{BB} = \frac{1}{T_a}$$

9.4.3 DSP Unit

The design philosophy for LISA is that the antenna will be controlled via spherical modes. The beam steering computer works out the modal signals to excite the antenna with. The antenna can autonomously transform the signals, in forward and backward directions, between the IQ signals for the array signal channels and the modal signals for the communication with the beam steering computer.

This requires a hefty amount of data processing onboard the antenna. A simple controller is not sufficient to cope with this and the original idea was therefore to implement a controller with datapath on a Field Programmable Gate Array (FPGA) board, together with some other controllers for I/O with the beam steering computer, control of the DAC and Analog to Digital Convertor (ADC) units, overall system performance and so on. The high level data flow that would need to be implemented onboard the antenna is shown in figure 7.11 of chapter 7.

The FPGA implementation was abandoned however, in favour of a software implementation on a Compact-PCI platform. The reasons for taking this new approach were the following:

- There is no digital programming required with a software solution, which will significantly shorten development time.
- The application is not time-critical. The reason that it is not time-critical is that switching between the modes of operation explained in section 9.3.2 is bound to be too slow for the measurement system to run with continuous scans. When the operational mode of the antenna is changed, the measurement system will be paused and wait for a trigger from the antenna controller to resume the measurement. Hence, apart from keeping the overall measurement time within reasonable limits, there is no need to hurry.
- Relatively low-cost DAC and ADC convertors are available as c-PCI or PCI cards with drivers and interfaces for various program languages. The whole system of data transfers, processing and I/O with the external beamsteering computer can be implemented in one software-controlled unit.
- The RS-232 interface with the beam steering computer is readily available on a PC platform.
- development time could further be reduced by using MATLAB. In conjunction with the Real-Time Workshop and the xPC Target toolboxes, development and deployment on the antenna platform could potentially be very quick. Not in the least, because some of the DAC/ADC cards are delivered with drivers for MathWorks Real-Time Workshop. This is something to look out for when acquiring the ADC/DAC convertors.

9.4.4 BSC and I/O

Development of the the beam steering computer (BSC) is mainly a matter of writing software, that runs on a PC platform. It makes sense to place the BSC outside the array antenna, to allow for interaction with an operator. This is no problem when the interface is only being used to statically reconfigure the array between data acquisitions. However, when "live" beam steering is to be performed, connectivity must be retained during measurements. Even if no operator interaction is required, the BSC needs interaction with parts of the measurement system outside the anechoic chamber. More in particular, the BSC would either get position data made available via the high speed BCD bus and the DACP, or would interact with the DACP to pause acquisitions and command switching of the antenna to sensing mode to acquire position data from there. The BSC therefore needs to reside outside the antenna, even outside the anechoic chamber in the control room, and it will need to communicate with the I/O unit inside the antenna. A datalink needs to be in place between the two.

The input/output unit (I/O) provides a data interface to the DSPU from the BSC. The BSC is placed outside the anechoic chamber, whereas the DSPU sits inside the array that is rotated on the AUT tower. Hence a data interface between the rotating AUT and the static environment is inevitable.

An RS-232 data-link is desirable for its simplicity. Each PC platform has a programmable serial interface. An alternative option was considered to use the GPIB bus that is already in place to

connect the instruments of the measurement system with the computers outside the chamber, and is also faster than an RS-232 link. The complexity of interfacing with the GPIB makes it unsuitable for this link, also because future antenna systems may be able to benefit more from a permanent simple RS-232 link in the facility. The RS-232 link passes through the shield of the anechoic chamber to the control room over a fibre connection.

As the AUT is rotating on the roll-over-azimuth mechanical scanner, the data-link needs to carry across two points of rotation. The lowest point of rotation has limited movement and the cable can pass across without difficulty. The roll axis has continuous movement however, and a slipring assembly was necessary to complete the connection. An IR-link could have been used as well, but the more expensive slipring solution was chosen because it combines 5 signal channels with 3 power channels. The five signal lines allow an RS-232 protocol with flow-control and spoofed handshaking.

9.4.5 Power Unit

The PU provides power to the LO amplifier in the RF front-end, the ADC and DAC converters, the DSPU, and the I/O system. There are two options as to where to get the power from:

- Batteries
- External supply

A preliminary and roughly estimated power budget is as follows:

LNA	20	20 mA	15 V	6 W
FPGA	1	1 A	5 V	5 W
DAC	1	32 mA	10 V	320 mW
ADC	20	0.5 mA	5 V	50 mW
I/O	/			
LOA	1	200 mA	15 V	3 W

It is therefore a reasonable assumption to expect a provision of 15 W of power at 15 V will be necessary. That is 30 Ah for 30 hours of continuous operation. This could be provided with three rechargeable Li-ion external notebook batteries of 10 Ah, the cost of which exceeds the cost of a slipring three times ($\sim$£600 rather than $\sim$£200). Batteries thus offer no satisfactory alternative in this case, which justified the expense for the slipring assembly. The extra cost of a few data-channels on the slipring were marginal. The decision to revert to a PC based solution for the DSPU, rather than an FPGA will significantly increase the power demand and heat generation and more engineering may be required to overcome some problems in that area.

The power will be made available via a 3-way connector with 3 voltages: 0V, 5V and 15V. The current rating of the power channels is 20A on the 0V line and 10A on the other two lines. The PU is just a regulated DC power supply with extra fuses that resides inside the anechoic chamber and is fed by a the internal 250 V AC clean power circuit.

9.4.6 Hardware

As the front face of the array will ultimately consist of a large number of triangular panels, these could potentially be held together with a structure of rods. Despite the fact that this would be much lighter, the panels of LISA are mounted on a large metal spun sphere. This offers the following advantages:

- In the absence of a full panel cover, the presence of the metal sphere simulates the array environment when the pattern of one subarray or panel is being measured.
- Having a metal sphere offers a rigid mounting platform that is much more flexible and easy to use than a structure with rods. The latter, in order to be rigid enough, would probably need to be designed for a particular geometry. Any change in the geometry, as a consequence of inaccuracies for instance, could cause incompatibility with the mounting structure.
- The sphere can be reused for other interesting experiments, as it is after all the shape that is under investigation. More in particular, the computer models that predicted the radiation patterns for some radiating elements (patch and aperture types) on a conducting sphere could be verified with measurements.
- In order to test the theories, it is important to build a relevant demonstrator. As all the modelling departed from a spherical groundplane, it is important to approximate that shape as accurately as possible. It would not be easy to construct a mounting structure that is accurately spherical as well as rigid under the rotations of the positioner in the test facility.
- A practical reason to keep the mounting surface as accurately spherical as possible, is that otherwise the fabricated panels might not fit well together.
- Lastly, having a closed metal surface in the centre ensures that the elements are not coupled internally via standing waves in the interior volume of the spherical shell.

Originally the metal sphere was acquired because in an earlier stage of the design it was planned that LISA would utilise radiating apertures as array elements. It is possible that a future implementation of LISA will replace the metal sphere at its core with a lighter structure. Conversely, and perhaps more interesting, it is also possible that the design would become more advanced and use elements on flexible substrates that are flush mounted directly on the metal sphere.

The original design of the spun metal sphere and its mounting structure is illustrated in figure 9.7. Some pictures from the sphere are shown in figure 9.8. The tolerance on the sphericity of the sphere should be $2 - 3\%$. Rotation tests on the positioner of our antenna measurement facility show indeed that the amplitude of the wobble of the surface is around one centimeter, which would also include the effect of a slight misalignment with the axis of rotation. The average phase error due to non-sphericity can therefore be guessed to be less than 5% of a wavelength.

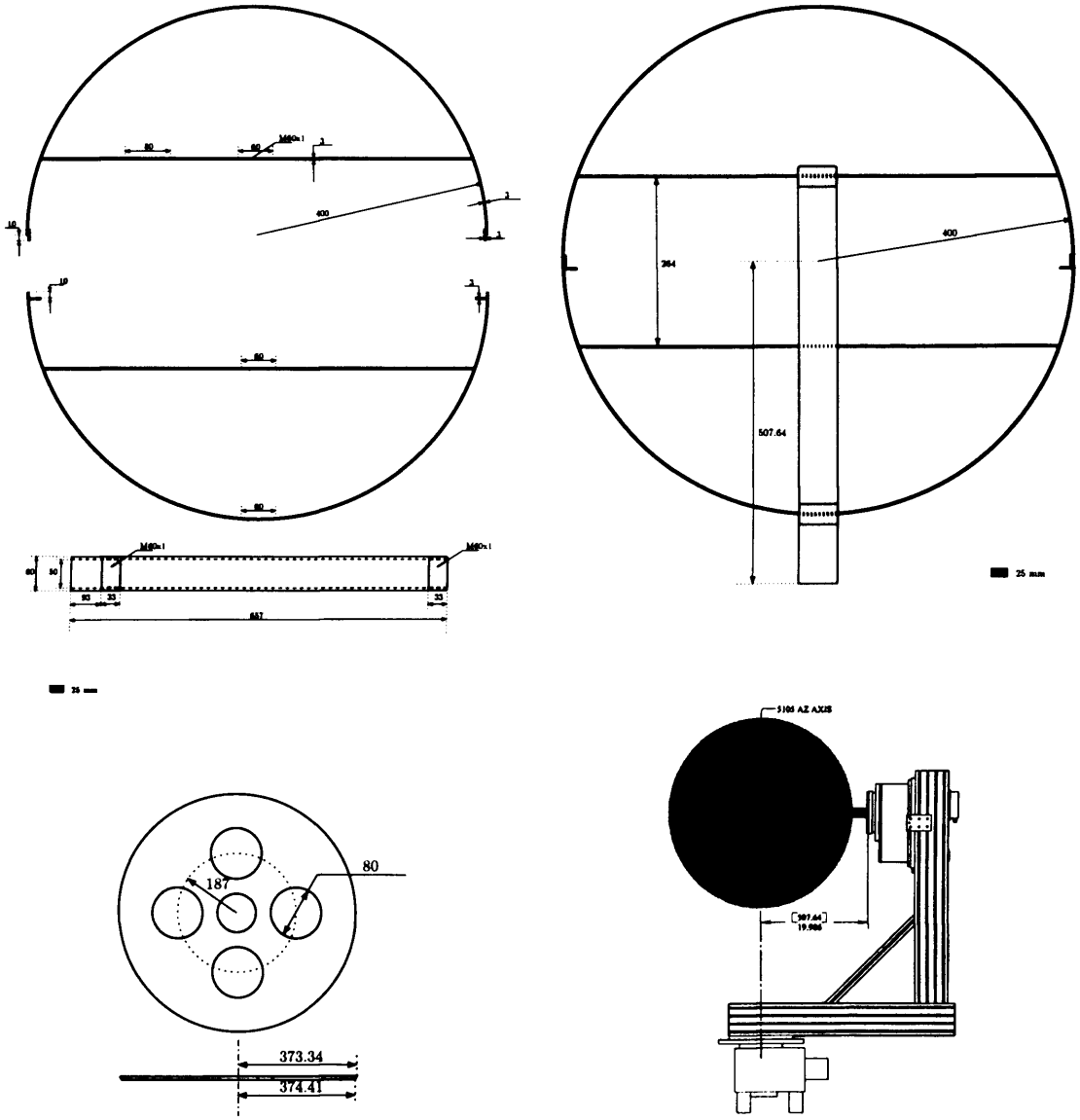


Figure 9.7: The dimensions for the spherical assembly, as it was originally designed. All dimensions are in mm. Threads are shown with a solid grey colour, rather then the usual symbol. The threads will allow to adjust the position of the centre of the sphere along the shaft, so that it can be positioned in the centre of the measurement sphere. If done accurately, the spherical mode decomposition used by the measurement system to calculate the far-field is referred to the same origin as the decompositions used in this text. Also, this alignment enables an easy calibration process (see section 9.3.3) and backprojection of the measured radiation pattern to the surface of the sphere.

Figure 9.8: Pictures from the spun metal sphere that is at the core of the spherical array antenna.

9.4.7 Antenna Panels

By far most of the design and development efforts have gone into the panels that contain the radiating elements. The radiating elements were already discussed in section 3.6.3. This section will add some design considerations, development issues and obtained results to the discussion.

Considering the large amount of array elements needed, it was of paramount importance to end up with a design whose performance does not vary too much with manufacturing inaccuracies. Therefore the feedline is designed to be fed with a plated through-hole via, rather than a soldered pin. A lot of attention went into the design of these viastructures and their connection with the feedline. Careful examination of figure 9.9 reveals that the vias are designed with a very fine printed collar to assure a good and reliable connection with the feedline.

To minimise development costs, the number of design cycles that include manufacturing were minimised. One prototype batch was made, and as of this writing the complete production batch is ordered with the manufacturer. Therefore, it was important to produce high-quality simulation models from which trustworthy performance predictions can be made. Therefore it was necessary to produce both coarser models for rapid simulations as well as models with a high level of detail. The simulations of models with high levels of detail posed major difficulties for the simulation software, especially for the FDTD method in CST. Nevertheless the design efforts proceeded in parallel using different EM solvers (ADS and CST). It was made sure that there was an agreement between the simulation results of the different solvers, as is illustrated in figure 9.10. In a previous design cycle, several prototypes were build of single radiating elements based on the J-feed design (see section 3.6.1). This design laid to the development of the NOSCIPRA antennas currently in use. Single elements of these antennas have however not

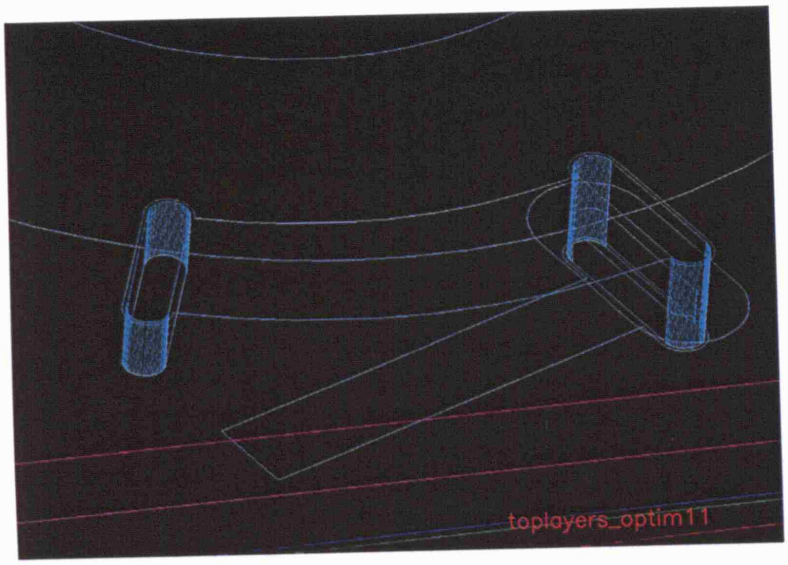


Figure 9.9: Detail of the via structures. Note the fine collar around the edges and the width of the straight edge that is equal to the width of the top feedline.

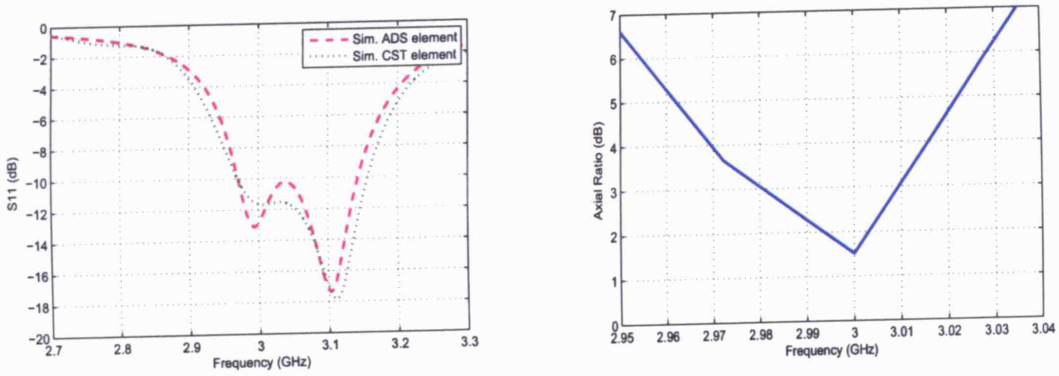


Figure 9.10: Return loss and axial ratio for a single simulated element. The two curves for the return loss are obtained from simulations with a MoM method (HPADS) and an FDTD method (CST).

been constructed and only measurement results from panels of three elements are available.

Figure 9.11 illustrates the different layers of the panel. The top layer was manufactured separately on a FoamClad substrate from Arlon. The substrate was chosen because of its extremely low dielectric constant and affordability. The substrate has a copper cladding on one side of a foam substrate and an adhesive peel-off layer on the other side. The triangles of the top and lower layers were delivered by different manufacturers and stuck together by hand. The top layer could not be processed in-house because milling was not possible on this soft substrate and the rings were made with an etching process. The lower 3 layers are made from RO4350 substrates from Rogers. These substrates offer a good trade-off between controlled dielectric properties and price. The boards consist of a double sided board and a single sided board, bond together with a very thin prepreg layer immediately below the groundplane.

The power divider on the fourth layer has a somewhat unusual design, shown in figure 9.13. It is based on a design published in [132]. A related design is also published in [133]. The design

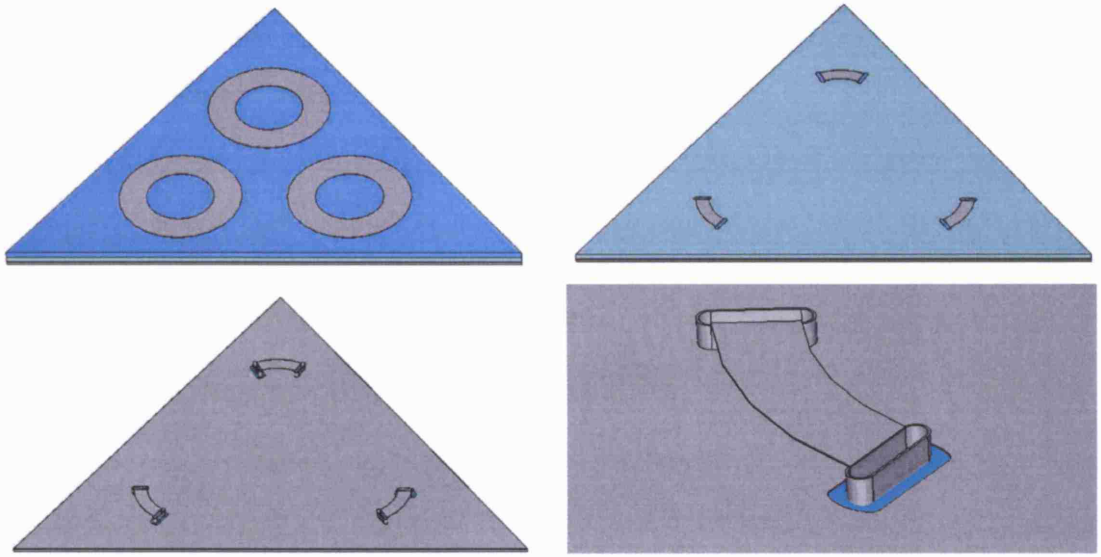


Figure 9.11: The panel consist of four layers. The top layer with the rings were manufactured separately from the other three layers that are bond together. The top layer is made on a foam substrate, the others layers are rigid. The fourth layer is not shown. It contains the power divider on the other side of the panel and is shown in figure 9.13. Pictures of the real tiles are shown in figure 9.12.

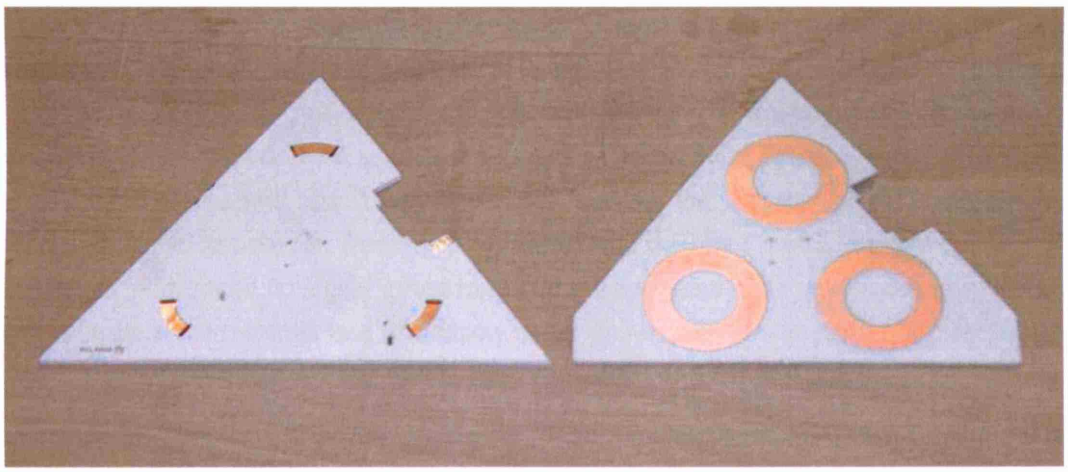


Figure 9.12: Pictures of the top layer and the lower three layers, before they were put together. The gaps on the edge is where the edge-mount SMA-connector will be attached. A small patch of the groundplane is visible because the legs of the connector need to connect to the groundplane and its pin to a microstrip line at the other side of the board.

was found to perform much better than any other of the designs we tried, including unequal Wilkinson dividers, Bagley dividers[134], circular sector dividers[135][136] and microstrip disc divider[137]. Ideally, we want to combine the following characteristics for the power divider:

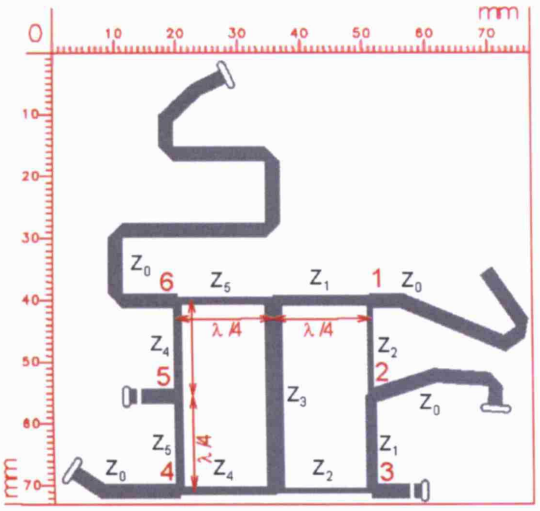
- Four port device: 1 input and 3 output ports
- Lossless
- All ports matched to 50Ω
- Equal power division over three output ports
- All the output ports isolated from one another
- Planar implementation with PCB technology possible
- Reciprocal, no anisotropic materials are used

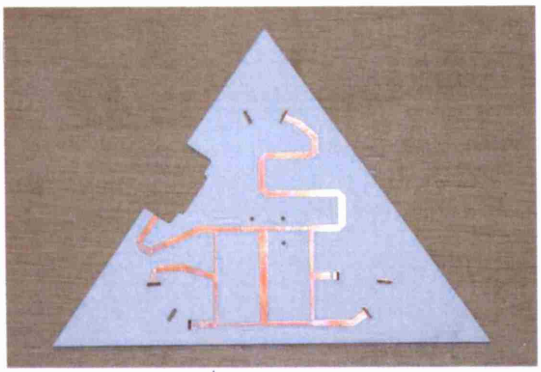
The latter is very important in this case, because the amount of mutual coupling between the 3 elements on one tile immediately depends on the isolation of the output ports of the power divider. The importance of internal mutual couplings was shown mathematically on page 5.3.1, where it was found that the free excitation equals the excitation vector if there is no internal coupling.

By formally imposing the requirements above on a (4×4) scattering matrix, no solution can be found - it is impossible to make such a device. In fact the three conditions for an ideal microwave component: matched, reciprocal and lossless, result in directional couplers, which have one port isolated from the input and a power division over the two other ports. By giving up one of the requirements, a solution can be found. The presented solution can be interpreted as either a lossless 6-port, of which two ports will need to be match terminated with their characteristic impedance, or as a lossy 4-port that includes these resistive terminations. The resistive terminations will absorb any unbalance between the 3 output ports, in other words they ensure isolation between these ports. A signal into the input port is divided equally over the 3 output ports, with no signal going to the terminated ports. The terminated ports take up the same role as the resistors in a Wilkinson power divider, but the latter do not offer a planar implementation. Simulated results for the power divider are given in figure 9.14.

The completed panels, with power divider and 3 radiating elements have also been simulated extensively with both HPADS and CST. A first batch of panels has been delivered and measured using a network analyser and our anechoic chamber. Results obtained for the radiation pattern will be shown in section 9.5. Figure 9.15 shows the results from simulations and measurements of the return loss and the axial ratio.

The sequential rotation explained in section 9.1 improves the CP bandwidth of a panel, compared to that of a single element. This is shown in figure 9.15 (b). Also the impedance bandwidth improves, as is shown in figure 9.15 (a).





Line	Z_1	Z_2	Z_3 Z_0	Z_4 Z_5
Impedance (Ω)	61.2	68.2	50	70.7
Width (mm)	1.54	0.77	2.19	1.18

Figure 9.13: The design layout and picture of the power divider are shown. The meandered lines ensure the correct phasing of the elements for the sequential rotation. The 2 small gaps in the layout are bridged with 2 parallel 100Ω resistors that connect the terminated ports of the divider with shorting vias to the groundplane. The value of Z_3 is not important as it has a length of half a wavelength. The values for the impedances achieve an equal 3-way power division, but other power ratios can also be achieved[132].

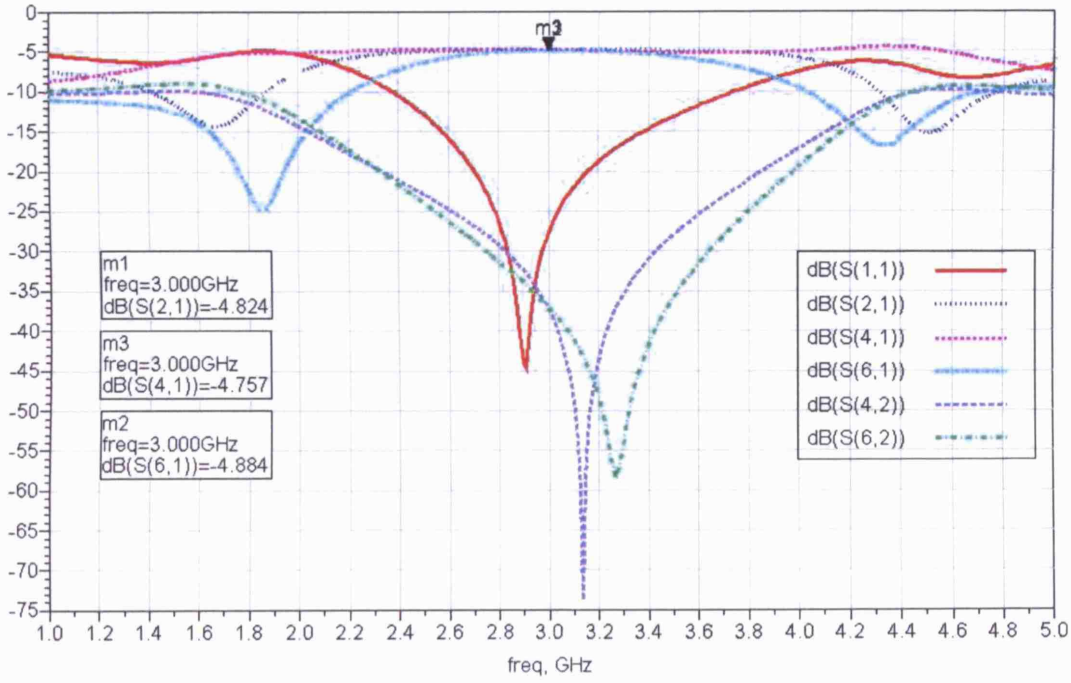
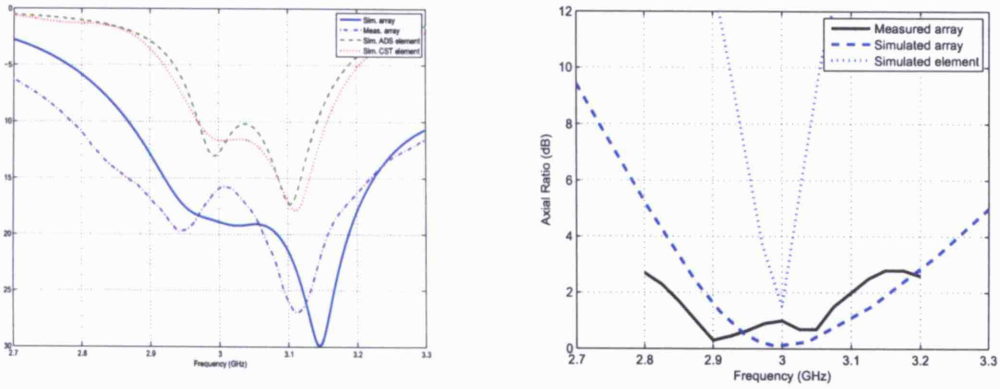


Figure 9.14: Some S-parameters are shown for the simulated power divider. From S_{21}, S_{41} and S_{61} it can be seen that at the design frequency of 3 GHz the power is indeed divided equally over the output ports 2, 4 and 6. The port numbers are indicated in red on figure 9.13. S_{42} and S_{62} indicate a very good isolation of port 2 from the other two output ports.



(a) (b)

Figure 9.15: The CP bandwidth of a panel of three sequentially rotated elements improves significantly compared to that of a single element. There is a good agreement between simulations and measurements for the panels. It is therefore likely that the simulation results for the single elements were also reasonably accurate.

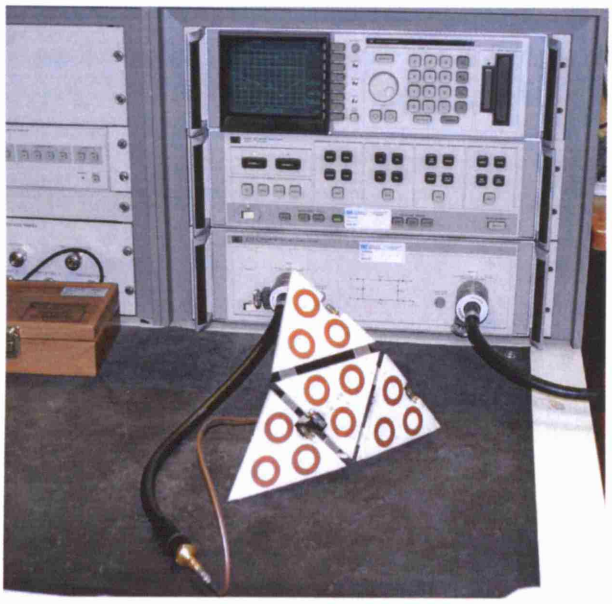


Figure 9.16: Measuring the return loss of a single panel with passively terminated neighbouring panels, using the network analyser.

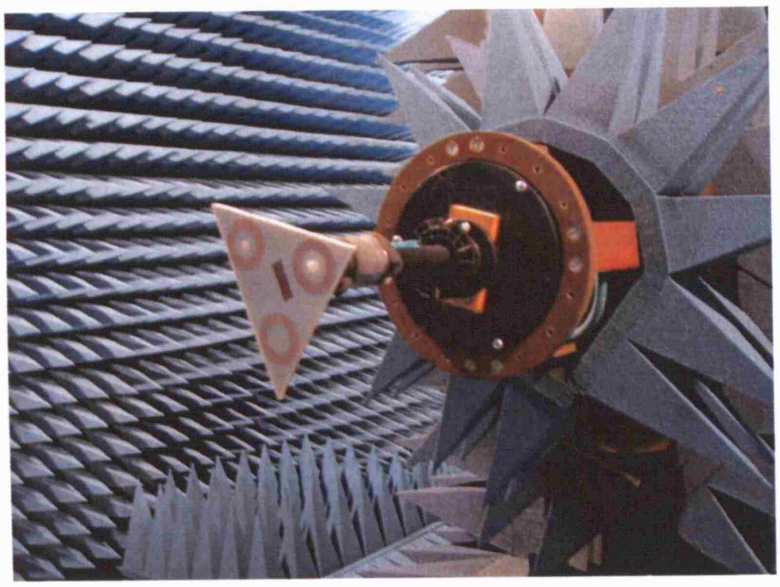


Figure 9.17: An early prototype for the array panels is mounted in the anechoic chamber. The measurement setup used to obtain the data from the current batch of panels was identical to the one shown in this picture.

9.5 Results from Measurements

In this section some preliminary test results are presented. The prime goal of these measurements was to verify the performance of the manufactured panels and how they would act together as a subarray. Overall the results are satisfactory, and sometimes even exceed expectations based on simulations. Apart from verification of the panel performance, some data were gathered to make some predictions on the performance of LISA once completed. This work is in an early stage, but some interesting conclusions can be drawn already.

Section 9.5.1 will describe in some detail the hardware and three different setups for the preliminary measurements. In section 9.5.2 the results of the three measurement setups are shown and discussed. In section 9.5.3 the obtained data will be used to synthesise radiation patterns of the completed array antenna.

9.5.1 Description of Preliminary Measurements

The first measurement setup is illustrated in figure 9.17, where a single tile is mounted in the anechoic chamber to verify its performance. Of particular interest are its polarisation quality and the symmetry of its radiation pattern. Although the reported data was obtained from the latest available panels, the panel shown in the picture is actually an early prototype that was build in-house with low precision. The design was still using a different kind of elements (referred to as J-feed elements in section 3.6) and changed later on to the a high-precision panel with NOSCIPRA elements.

In order to obtain a Realised Gain Pattern (RGP) for the panel, a second kind of measurements was performed with the tile mounted on the metal sphere shown in figure 9.8. The reasoning behind this setup is as follows: For the RGP only external scattering from the array environment

is of importance. Internal scattering should be absent for an RGP measurement. The external scattering in the radiation environment of the array is influenced by a most probable large reflection from the body of the array that is formed by the metal sphere and the groundplanes of the tiles mounted on it. The dimensions of the tiles are much smaller than the radius of curvature and in the same order of magnitude as the operational wavelength. The reflections from the body of the array antenna can therefore be assumed not be very different than that from the metal sphere alone. What is neglected is the absorption of energy by the other tiles of the array antenna, an effect that should be included in the RGP. Therefore a tile was measured with its three match terminated neighbours mounted next to it. Absorption by tiles that are further away is probably very small. Adding to the argument is that the gaps between the tiles act as obstacles for surface waves over the array body, whereas its curvature reduces the coupling via space waves. The dihedral angle between two neighbouring tiles ranges from 7.3° to 15.8°, and the edges of the triangles may help to shield the antennas on one tile from the antennas that are further away.

A third set of experiments for which results will be shown consider the RGP of a subarray of 4 tiles. Four tiles were combined with a 4-way hierarchical Wilkinson power combiner. Ideally, also here a set of neighbouring tiles would have been present to measure the RGP, as was done to measure the RGP of one panel. Due to time constraints this has not been done yet at the time of this writing.

Pictures of the subarrays of four tiles are illustrated in figure 9.18. The tiles are held together with brackets that are bend on a handpress with $\pm 0.5°$ precision. Over the large number of panels, the errors will hopefully average out to some extend. There is also some limited movement possible as the structure is not entirely rigid. A higher precision would not make much sense in respect to the tolerance on the sphericity of the sphere mentioned previously. The pictures also show the power divider in the background and the cables connected to the panels. The cables are not high quality cables as so many of them are needed. Errors attributable to the cables have not yet been quantified. All cables are kept to an equal minimum length of $50cm$. Figure 9.19 shows the subarray of four tiles mounted on the sphere, which is in turn mounted on the positioner of the measurement facility.

As a final remark to the followed approach, the measured RGP's of the subarray and the single tile are to be used as the RGP of an imaginary single element that acts as a substitute for the panel or subarray in an array with lesser elements. Of course each panel consists of multiple elements by itself, and mutual coupling can and will exist between the elements on the same tile. The effects of coupling within a tile are included in the RGP of the tile. To minimise the coupling between elements on the same tile, to actions have been taken: the elements were designed to radiate minimally in the lateral direction (see section 3.6.3) and the power dividers on the back of the tiles were designed to offer maximum isolation between the output ports. Quantification of the mutual coupling between elements on the same tile is not easily achieved in practise.

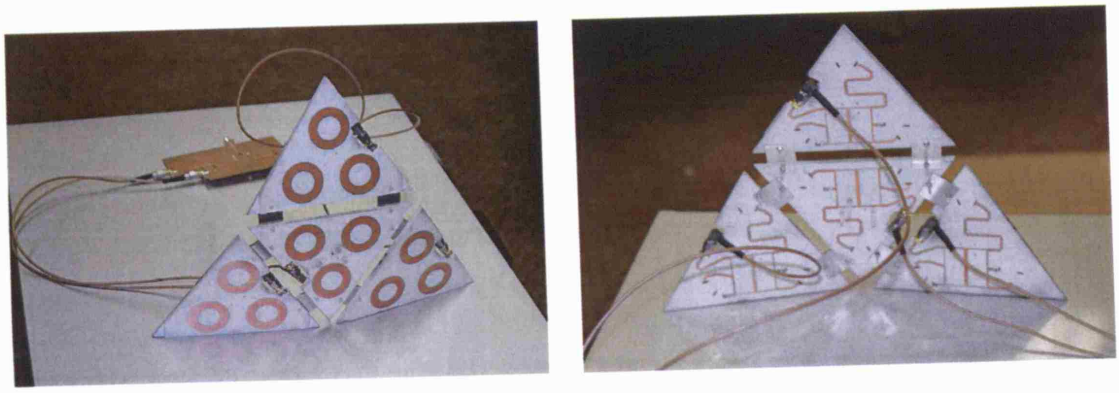


Figure 9.18: Front and back side of a complete subarray of four tiles. Also seen in the left picture is the power combiner that sums the signals from the four tiles together. Eventually 60 of these subarrays will cover the entire sphere.

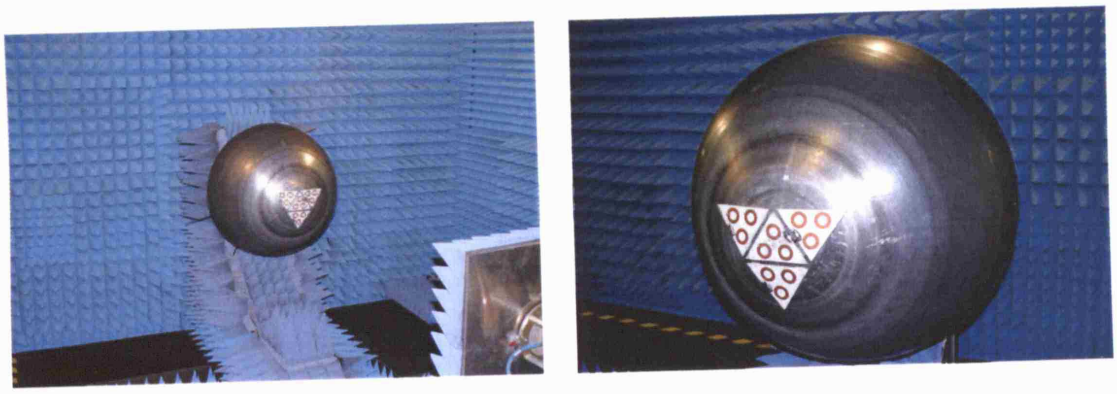


Figure 9.19: A subarray of four tiles can be seen mounted on the core metal sphere. The whole structure is ready to be measured in the antenna measurement facility, after aligning the approximate position of the phase centre with the axis of rotation with the aid of a LASER. The alignment is required by the roll-over-azimuth configuration of the spherical near-field test facility.

9.5.2 Data from Preliminary Measurements

In this section the measurement results are presented from the three measurement setups described in the previous section. The axial ratio of the subarray is optimal at 3.05 GHz. All directivity plots shown in this section are taken at a frequency of 3.05 GHz. For the preliminary checks on the measured data presented in this section, the raw data from the spherical near-field measurements is used without probe correction.

The data for the single tile, not mounted on a sphere, is visualised in the plots of figure 9.20. The top plot shows a 3-dimensional rendering over the sphere of the far-field, broken down in a right-hand circularly polarised component on the left and a left-hand circularly polarised component in the middle. No gain calibration has been performed on the measurement data, and the shown plots are directivity plots. The field strengths are plotted in dBi, relative to the total integrated power. The axial ratio is shown on the right. The triangular geometry of the tile can be seen lightly reflected by its radiation pattern. The plot below shows two-dimensional cuts through both polarisations of the far-field. The cuts are taken in the ZX and ZY planes. For these plots the directivity is normalised to its maximum. The plot below shows the boresight on-axis AR. It was obtained by taking a data acquisition during a complete revolution of the roll-axis of the antenna mount and fitting a sinusoidal curve to the obtained data. Indeed, with a linearly polarised probe and a circularly polarised antenna, a sinusoidal signal should be obtained, with the axial ratio determined by the amplitude of the sinusoid. As this method averages data from an entire revolution, the obtained result is more accurate than that can be obtained from a direct calculation with the complex values of the two measured polarisations. The latter can however be obtained over the full spherical surface.

The axial ratio is below 1 dB at the design frequency. It's 3-dB bandwidth is wider than the measurement interval. At 5% bandwidth, the axial ratio is 2.8 dB.

The data shown in figure 9.21 was gathered from measurements on a single active tile, mounted on the metal sphere, with three passive match-terminated tiles next to it. The optimal axial ratio has even further reduced, but has also shifted somewhat in frequency to 3.05 GHz. The axial ratio bandwidth has decreased significantly to about 3.6 %.

Figure 9.22 shows the results for an entire subarray, with the signals of four tiles combined in a power divider. Perhaps it is possible to further improve the results by phasing the three peripheral tiles relative to the centre tile, but this has not been tested yet. Apparent is the much narrower beam, as would be expected, and the improved circular polarisation. The triangular symmetry - of the subarray this time and not the triangle - is also clearly visible in the radiation pattern. The boresight axial ratio is below 1.8 dB for more than 5 % of the bandwidth. The 2-dimensional cuts look different because of the triangular symmetry: the YZ plane cuts through a sidelobe, whereas the XZ plane cuts between the sidelobes. A nicely formed beam exists with its maximum at boresight, so the angles betwen the tile sof a subarray are small enough to form a consistent pattern with a single phase centre, rather than breaking up into three overlapping beams.

In figure 9.23, the axial ratio is plotted, for the same three measurements, over the front hemi-

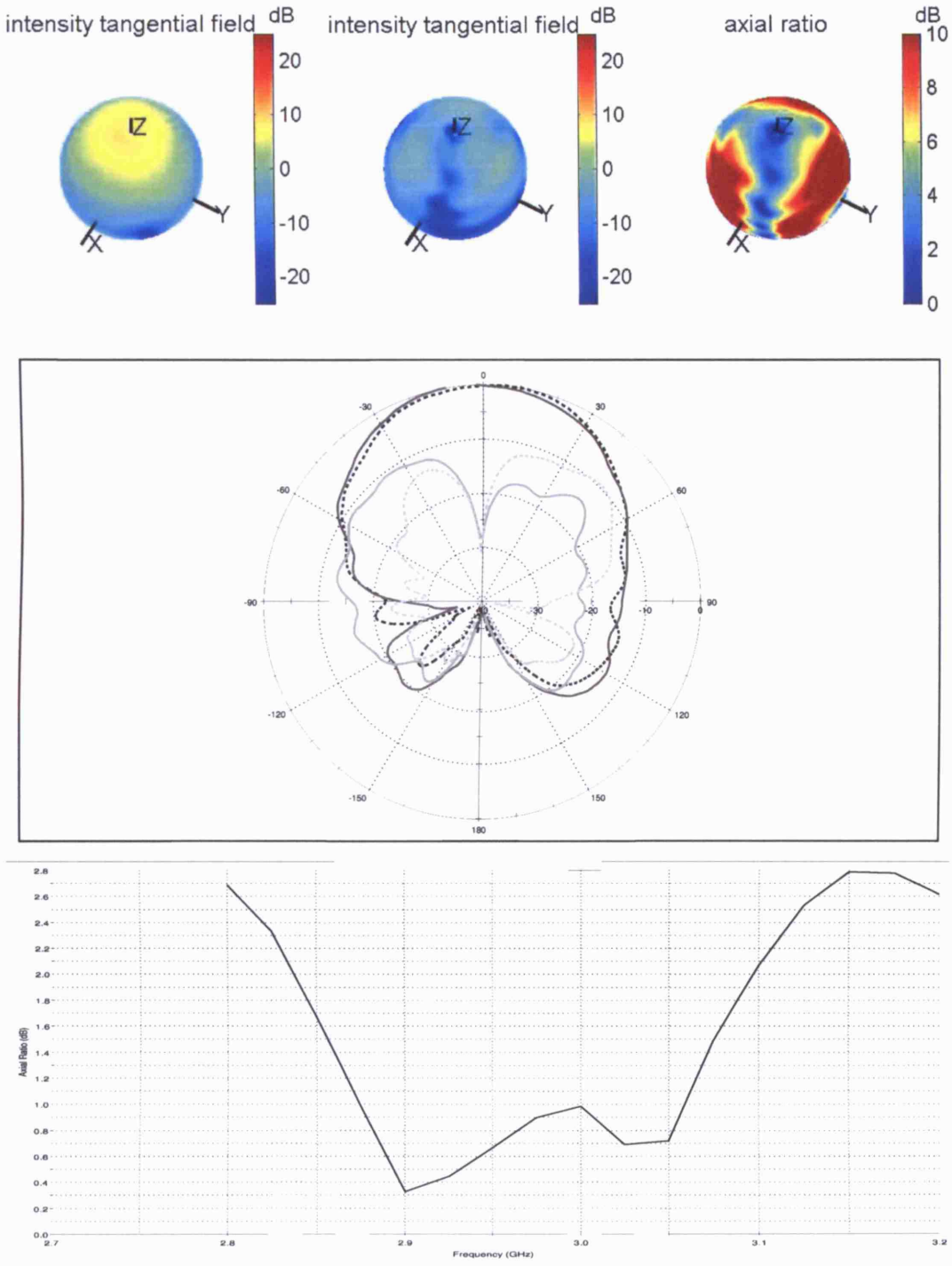


Figure 9.20: Data obtained from measurements on a single tile, without the sphere or other tiles present. The three spherical plots in the top row show the RHCP (right) and LHCP (middle) field strengths, as well as the axial ratio (right). The plot in the middle row shows two planar cuts through the RHCP and LHCP patterns. The plot on the bottom row shows the boresight on-axis AR.

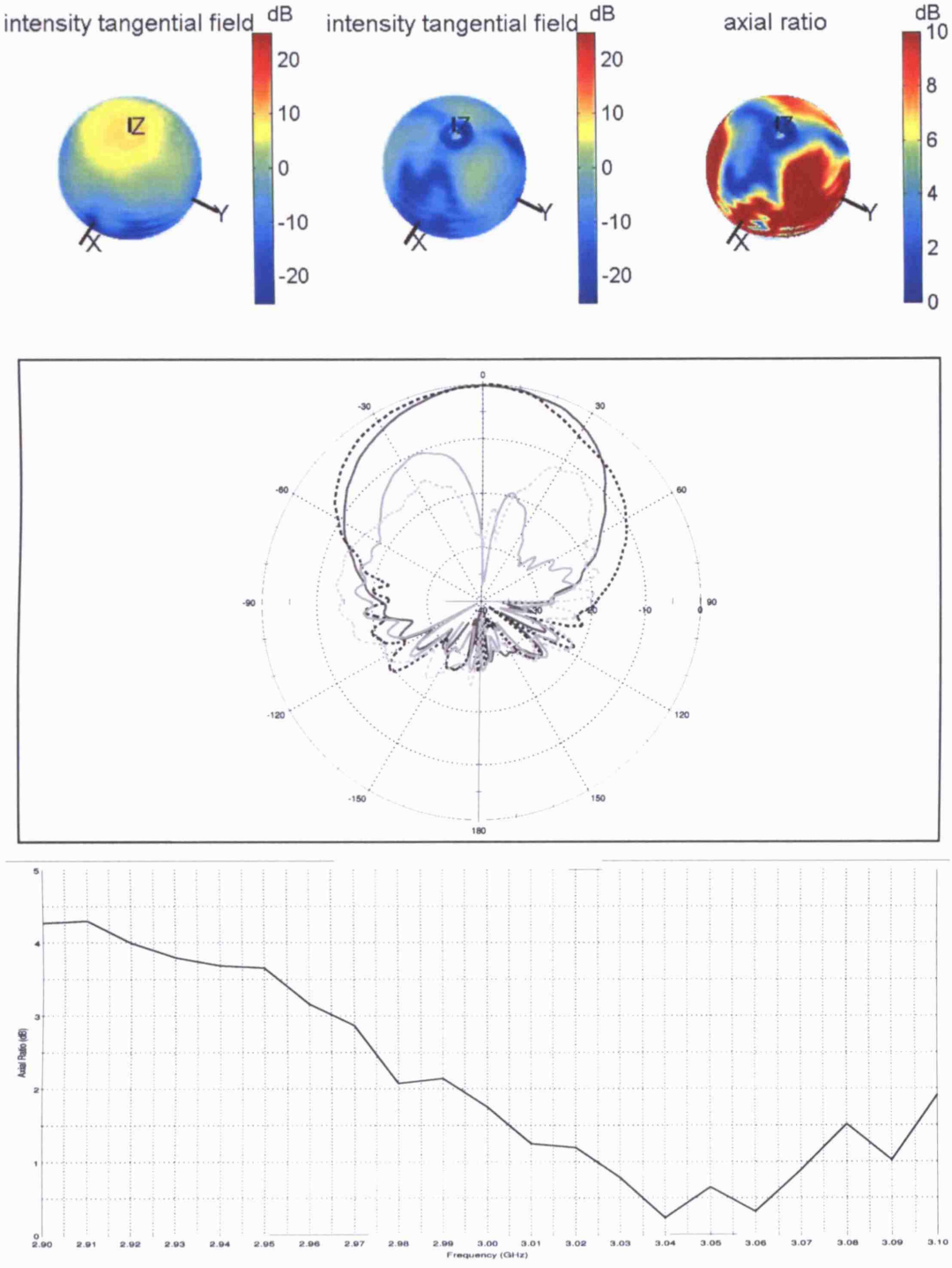


Figure 9.21: Data obtained from a single active tile mounted on the sphere, together with three passive neighbouring tiles. The data plots are equivalent to those in figure 9.20.

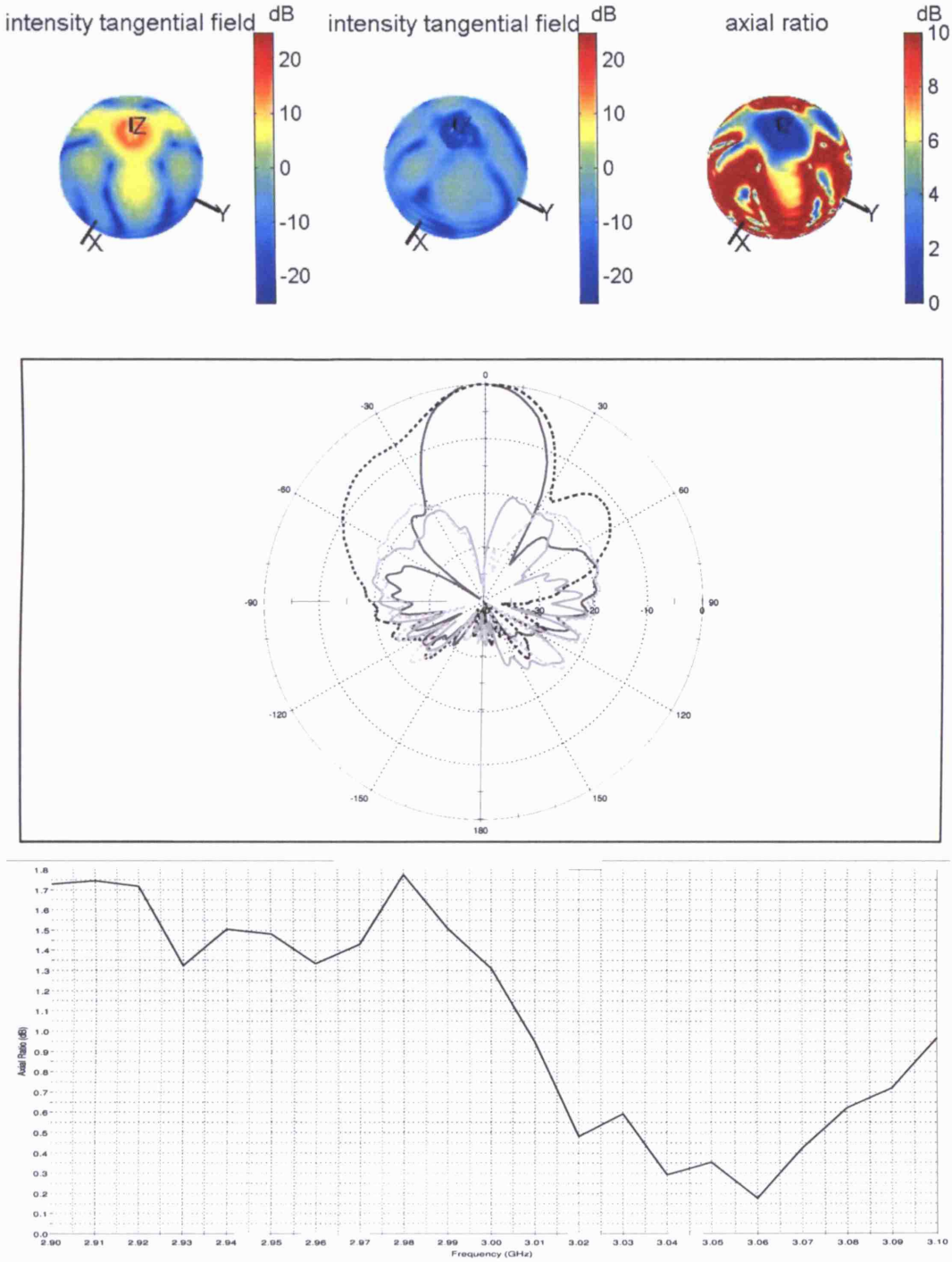
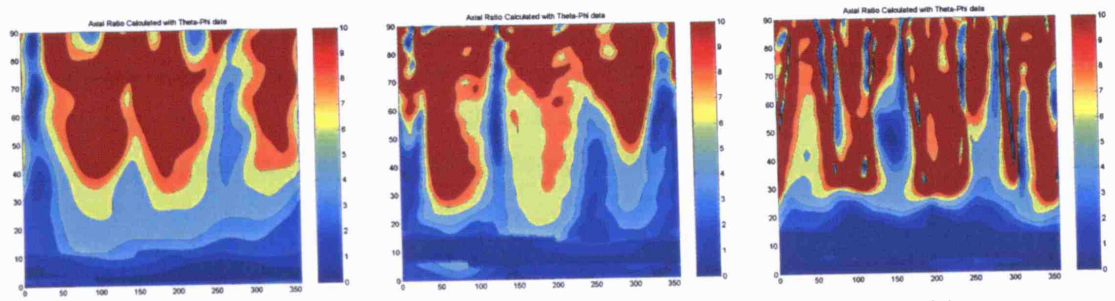


Figure 9.22: Measured data from an active subarray of four tiles. The data plots are equivalent to those in figure 9.20.



(a) (b) (c)

Figure 9.23: Polar maps of the axial ratio for the single tile (a), the tile with neighbours on the sphere (b) and the entire subarray (c). The horizontal axis corresponds with azimuthal angle and the vertical with elevation away from NP, so that the bottom line is the NP point on the sphere and the top line is the equatorial circle for which $z = 0$. The triangular geometry of the subarrays is clearly visible from the three North-South directed "branches" in the axial ratio maps.

sphere of the antenna. The vertical axis has the elevation angle from 0 to 90 degrees, and the horizontal angle represents the azimuthal angle (both with respect to the antenna, not the measurement range).

9.5.3 Results Obtained From Post-Processing

In section 9.5.2 we have obtained measurement data for a single tile as well as a subarray of four tiles. We will use the measured patterns as RGPs to predict the performance of the entire array antenna upon completion. Strictly speaking, the data obtained from measurements on the entire subarray represents no RGP as there were no neighbouring tiles present.

The simple formula 6.34 will be used here to synthesise an array excitation. Note that this will result in patterns that achieve maximum directivity in some desired direction, regardless of spatial bandwidth limitations. We can therefore expect significant sidelobes. Also, the sidelobes that exist as a result of aliasing (the effect of which can be reduced by aiming at a broader beam and less directivity) will generally vary with scan angle, as was explained in section 8.1.5.

To obtain the array RGP, the procedure described below is followed. It is by no means presented as an optimal or elegant procedure, but it is a good quick and dirty approach to predict the array performance.

- Numerically calculate the CPST of the measured element RGP. In the current implementation a simple numerical projection was used to this effect, that should be replaced with a smarter algorithm in the future. The result of this projection is shown in figure 9.24 for the single tile embedded in a subarray and mounted on the metal sphere. The expansion for the entire subarray is also shown in the same figure. The higher degree modes decay around a degree of $l = 25$ as could be expected from the size of the array antenna. For higher degrees however, more smaller peaks are found in the spectral domain, which were not included in the calculations. Calculations were performed up to a degree $L_{max} = 28$. In plot (e) of figure 9.24 it is verified that this is acceptable.

- The expansions obtained in the previous step are rotated to the positions of the various

tiles or subarrays, using the equations in section 4.2. If an array of separately controlled tiles is considered, it is impossible to make the tiles coincident with a rotation alone. Nevertheless, an approximate rotation can be used, as is derived in appendix B. In the same appendix, the positions of all tiles and subarray phase centres are tabulated.

- The rotated expansion is reconstructed in the spatial domain, in the direction of maximal directivity alone. From this reconstruction, the excitation coefficient of the tile or subarray is obtained according to equation 6.34.
- From the linearity of the CPST it follows that the rotated expansions can be summed, after multiplication with the excitation coefficients of the associated tile or subarray.
- Finally the hence obtained expansion is reconstructed in the spatial domain to give the array RGP.

The array geometry was designed to be separable in 60 subarrays as well as in 240 tiles, as was explained in section 8.2. The procedure above was carried out for both the configuration with 60 subarrays and the configuration with 240 tiles. For each configuration, the beam was steered to 4 different angles of elevation: 0, $\frac{\pi}{12}$, $\frac{\pi}{6}$ and $\frac{\pi}{2}$. For the RGPs of the subarrays and tiles the measured data were used from the previous section, together with the position information given in appendix B. The results are shown in figures 9.25 and 9.26 for the subarray and tile configurations respectively. Vertical cuts through the patterns for $\phi = 0$ are shown in figure 9.27.

In section 7.3 two ways were explained to reduce or even remove aliasing errors, at the expense of a reduced directivity. The easier way is to filter out the higher degree spherical modes from the radiation pattern. With reference to the procedure presented above to obtain the array RGP from the element RGP, this is simply achieved by reconstructing only the desired lower modes in the very last step of the procedure. From the introduction to this chapter in section 9.1 we know that for the case of the subarray configuration, the aliasing folds back all the way to the lowest degree and cannot be removed entirely. For the case of the 240 separately controlled tiles, we have seen that the aliasing can be removed by filtering out the degrees above $L_{max} = 4$. Both configurations have been tested this way and the results are shown in figures 9.28 and 9.29 for the subarray and tile configurations respectively. Vertical cuts through the patterns for $\phi = 0$ are shown in figure 9.30. For both configurations the reconstruction was performed up to a maximum degree of $L_{max} = 4$.

Another way to remove aliasing errors that was explained in section 7.3 is by filtering the field prior to sampling it. The mechanisms of spherical convolution established in chapter 7 can be used to work out an ideal pattern for the subarrays. An in-depth treatment of this topic is necessary, as will be repeated in the section on future work in chapter 10. It could for instance be found that the subarrays are too directive to work well together on the curved sphere. A quick simulation can be performed that shows the effect of altering the spatial bandwidth of the element RGPs. Hereto the calculations in the first steps of the procedure on page 262 were simply limited to a lower degree. The results shown in figure 9.31 were obtained by limiting

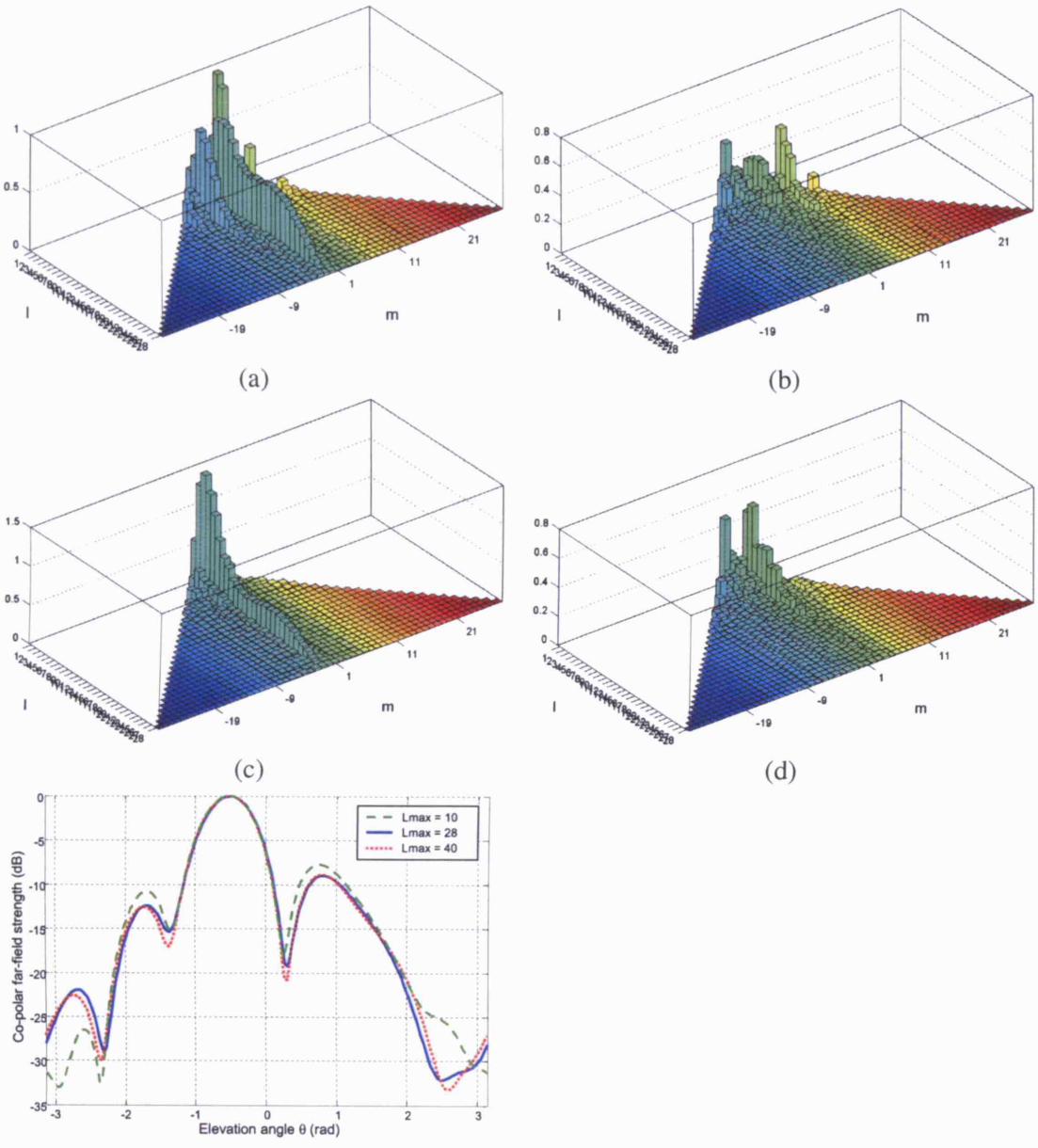


Figure 9.24: In (a), the RHCP and in (b) the LHCP CPSTs are illustrated for the RGP of one subarray. The coefficients were obtained by numerical projection of measured data of the subarray on the CPST basis functions. The measured data for the subarray was illustrated in figure 9.22. Figures (c) and (d) show the RGP for a single tile. The data used to obtain these expansions was illustrated in figure 9.21. As the subarrays are more directive than a single tile, the spatial bandwidth of their radiation pattern is also higher. The spatial bandwidth of the subarrays may be too high to be used effectively in an array configuration. The procedure explained on page 262 was followed to calculate the array pattern with the beam scanned to an elevation of $\frac{\pi}{6}$ off the z-axis with the RGPs of the subarrays calculated to a maximum degree of 10, 28 and 40. It is seen from the results in plot (e) that the extensive extra calculations to go up to degree 40 do not carry much extra information.

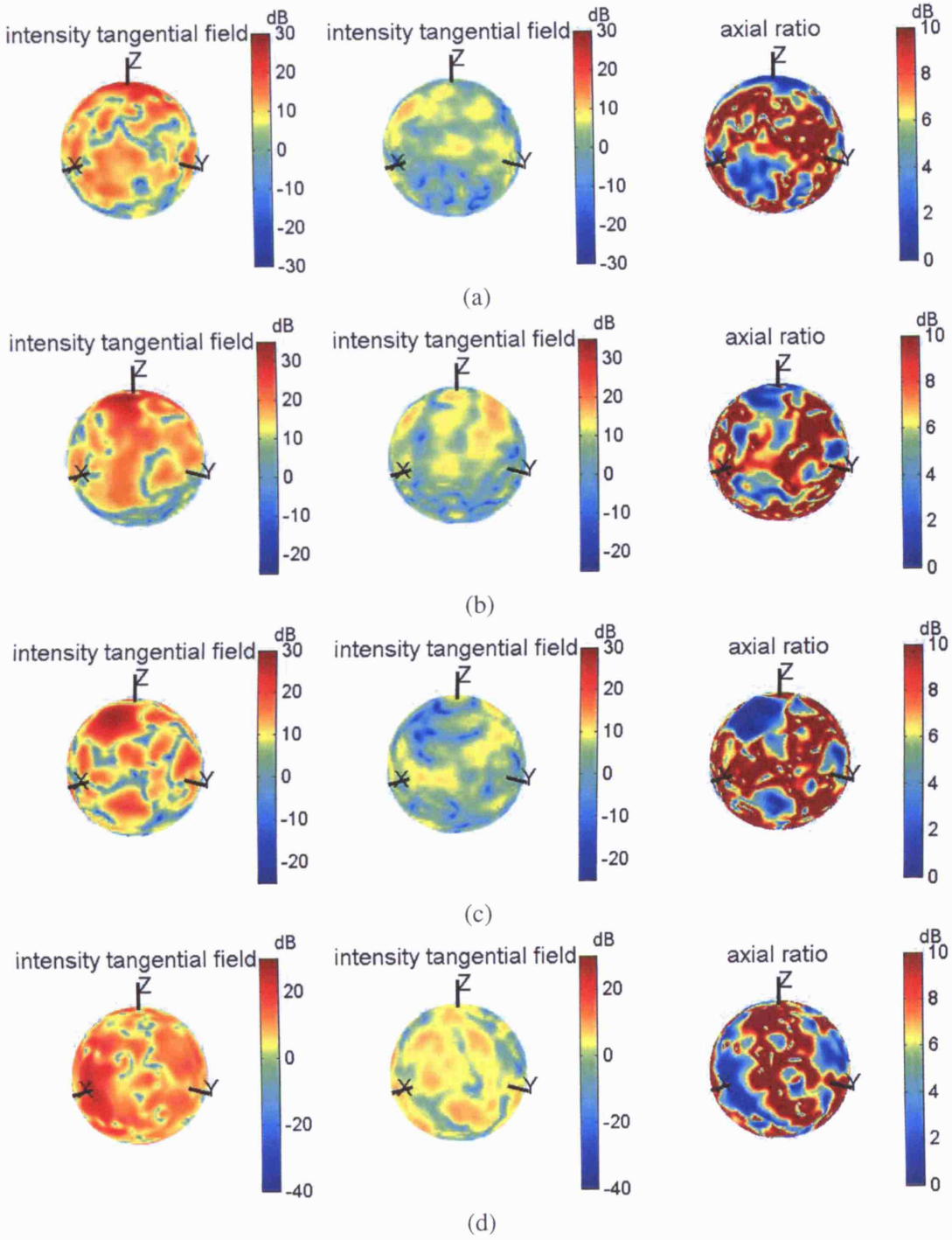


Figure 9.25: Four different array RGP are calculated based on the measurement data of subarray shown in figure 9.22. The subarray positions are obtained and tabulated in appendix B. The array pattern is steered to four different elevations: $\theta = 0$ in (a), $\theta = \frac{\pi}{12}$ in (b), $\theta = \frac{\pi}{6}$ in (c) and $\theta = \frac{\pi}{2}$ in (d). The pattern was optimised entirely to maximise directivity in the given directions, which partially explains the irregular patterns. The few but highly directive subarrays cause the principle of uniform resolution not to hold in this case and the pattern looks better in (c) than in (b). Vertical cuts through the patterns are shown in figure 9.27.

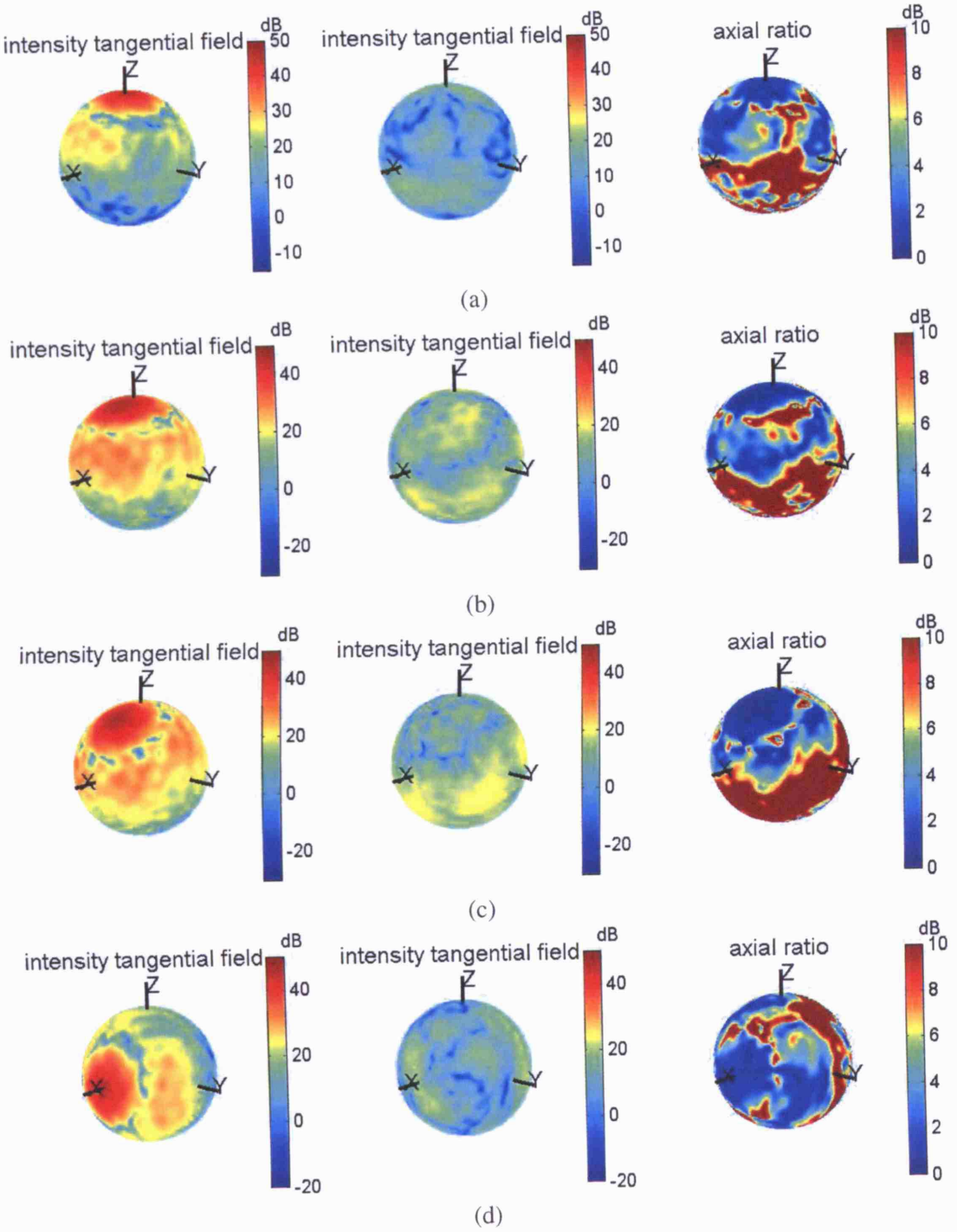
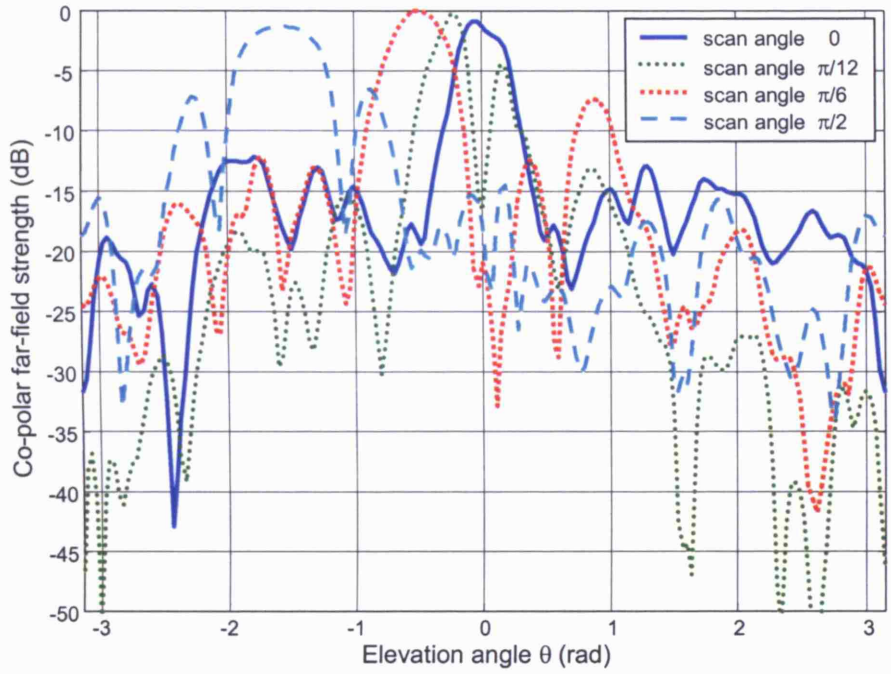
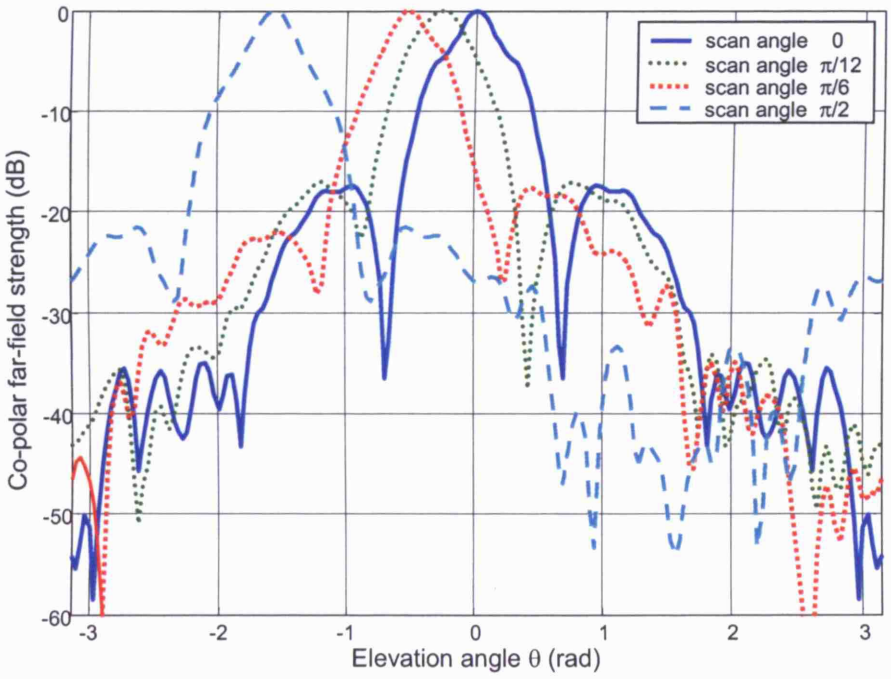


Figure 9.26: Four different array RGP are calculated based on the measurement data of an embedded tile shown in figure 9.21. The tile positions are obtained and tabulated in appendix B. The array pattern is steered to four different elevations: $\theta = 0$ in (a), $\theta = \frac{\pi}{12}$ in (b), $\theta = \frac{\pi}{6}$ in (c) and $\theta = \frac{\pi}{2}$ in (d). The pattern was optimised entirely to maximise directivity in the given directions, and it may be possible to reduce the sidelobe levels at an expense of a broader beamwidth. As there are 4 times more sampling positions that for the subarray configuration shown in figure 9.25, the radiation pattern suffers less from aliasing and looks smoother and more stable with scan angle. Vertical cuts through the patterns are shown in figure 9.27.



(a)



(b)

Figure 9.27: The figure shows two-dimensional cuts through the patterns in figures 9.25 (a) and 9.26 (b). For the 60-subarray configuration in (a), the array RGP is clearly irregular and unstable with scan angle, due to severe aliasing effects. Also the patterns in (b) suffer from aliasing effects, but to a lesser extent. Here at least the main beam remains intact, even for a very large scan angle of $\frac{\pi}{2}$, which is the key advantage spherical array antennas have over any other type of antenna.

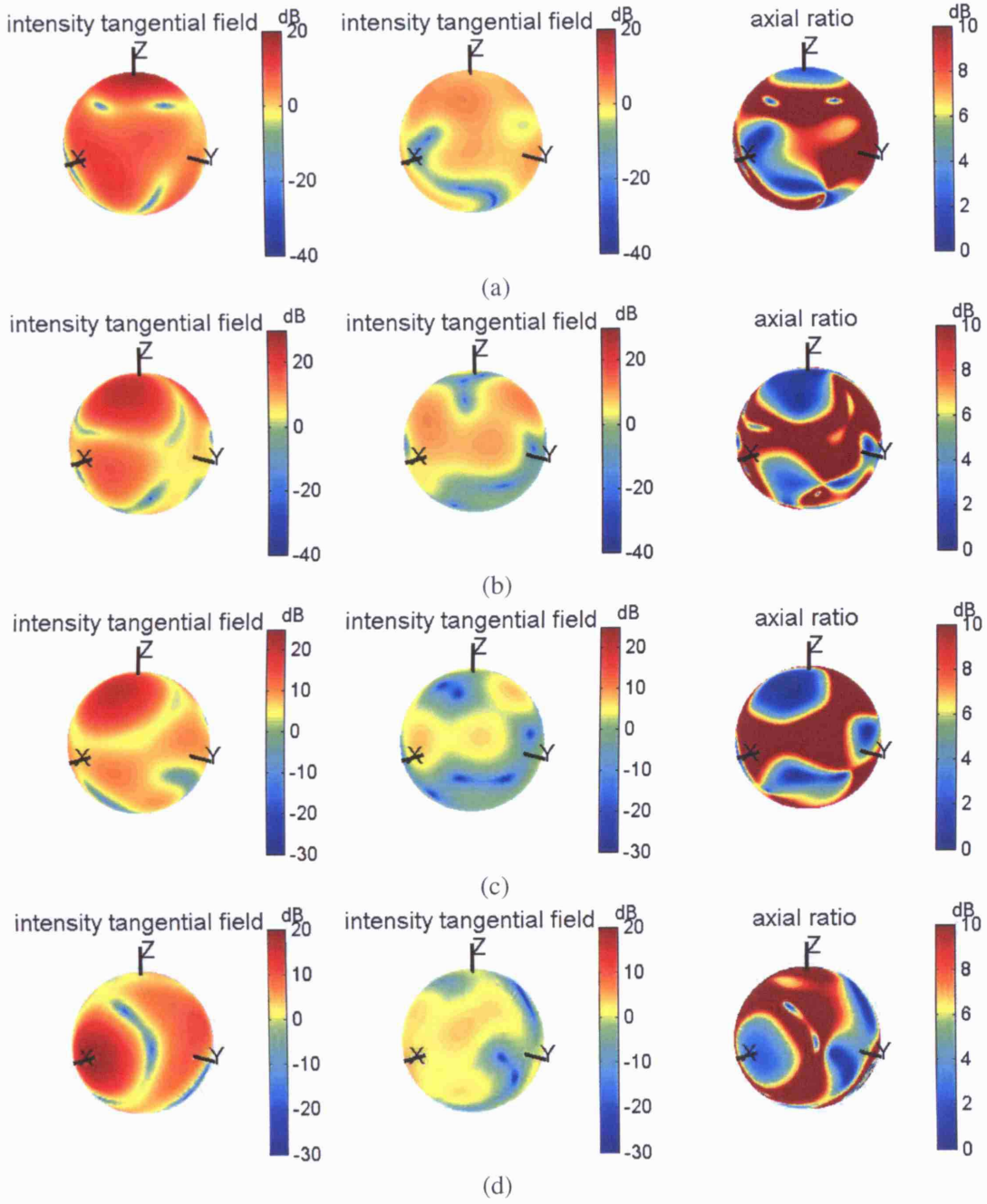


Figure 9.28: The procedure with which the array RGPs shown in figure 9.25 were derived is repeated, but the CPST modes for a degree above $L_{max} = 4$ were filtered out. In comparison with figure 9.25, the smoothing effect of leaving out higher degree modes is clearly visible. The figure shows that despite the fact that aliasing effects can not entirely be removed for the 60-subarray configuration, the trade-off between directivity and aliasing errors can still be made. Vertical cuts through the patterns are shown in figure 9.30.

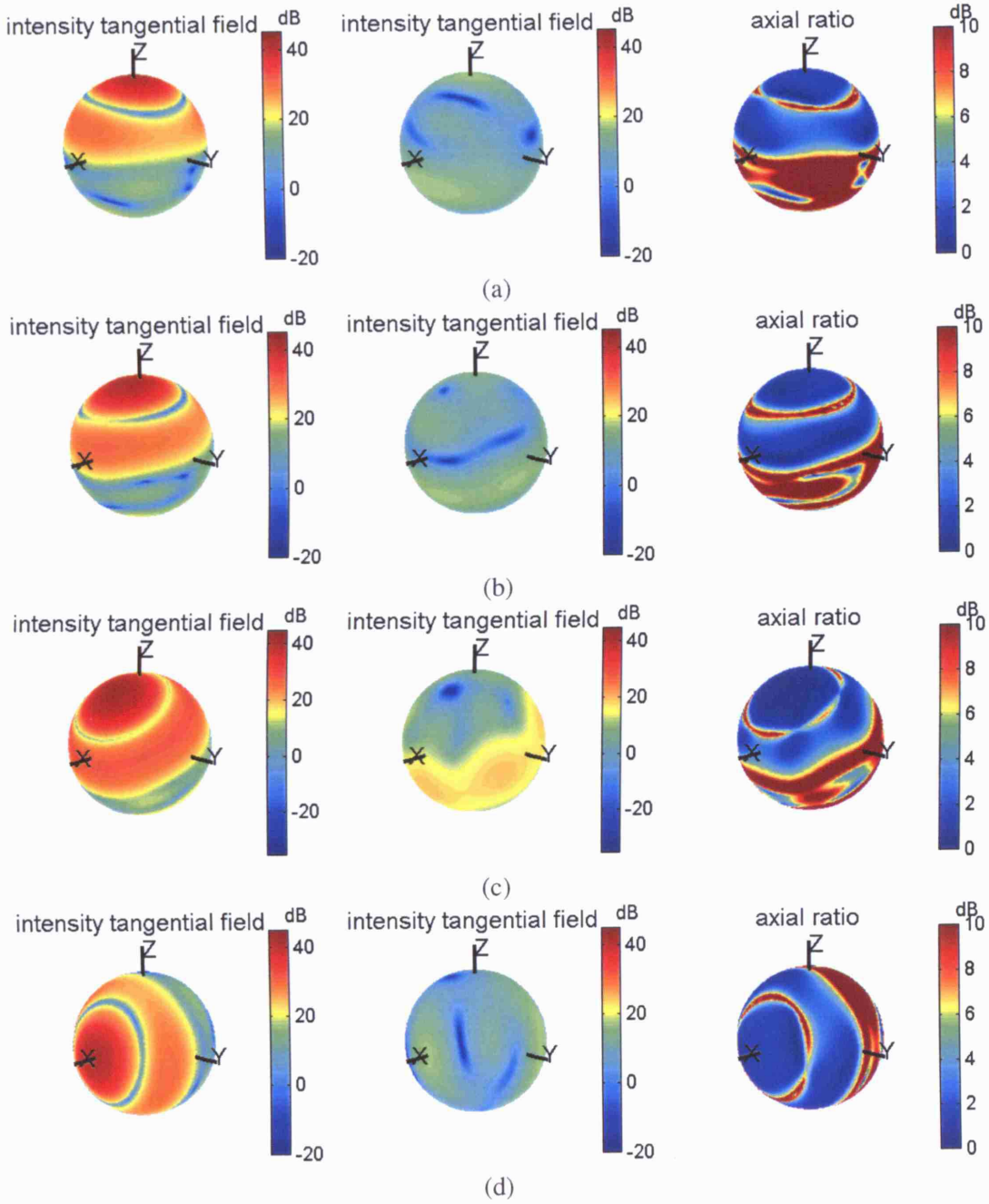
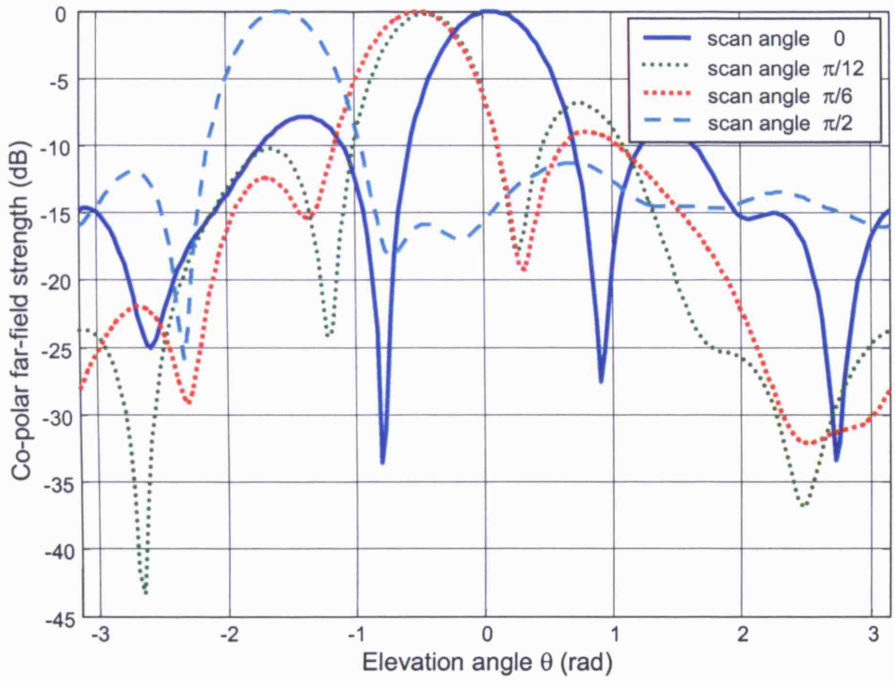
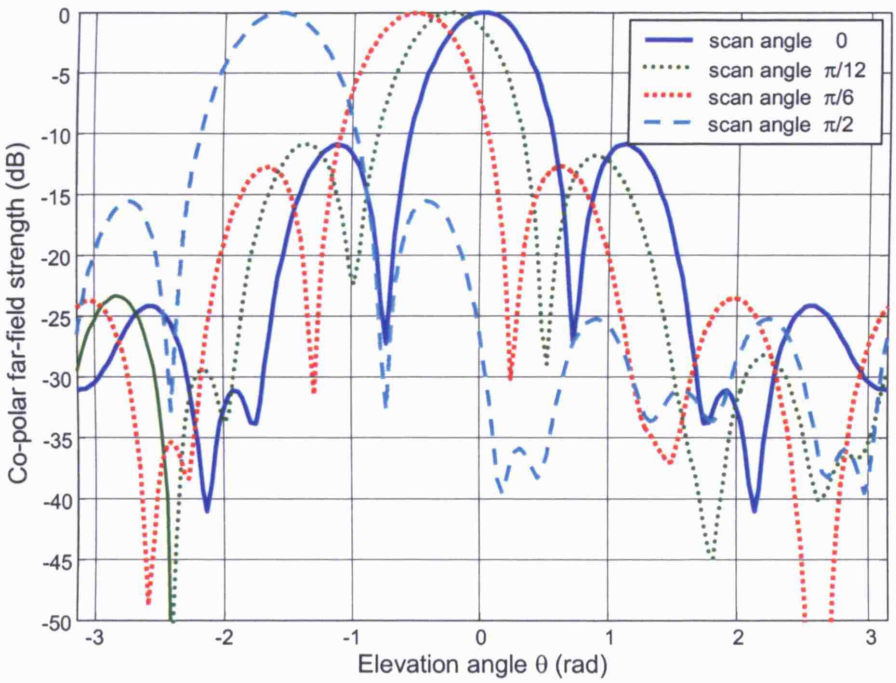


Figure 9.29: The procedure with which the array RGPs shown in figure 9.26 were derived is repeated, but the CPST modes for a degree above $L_{max} = 4$ were filtered out. As could be expected for the 240-tile configuration, the performance becomes rather good after filtering out the aliasing effects. Also note the excellent polarisation purity in almost an entire hemisphere. Vertical cuts through the patterns are shown in figure 9.30.



(a)



(b)

Figure 9.30: The figure shows two-dimensional cuts through the patterns in figures 9.28 (a) and 9.29 (b). In comparison to figure 9.27 it is clear that directivity has been reduced, but the patterns are much better behaved. The patterns for the 240-tile configuration in (b) has become almost stable with scan angle, which indicates that aliasing effects have successfully been removed. For the patterns in (a) this can not be achieved entirely. There are significant sidelobes left in the pattern, which can perhaps be reduced by employing a smoother filter, rather than the step-filter that was used to discard the higher degree modes. Note that, just as everywhere in this text, also these plots are obtained from a single polarisation.

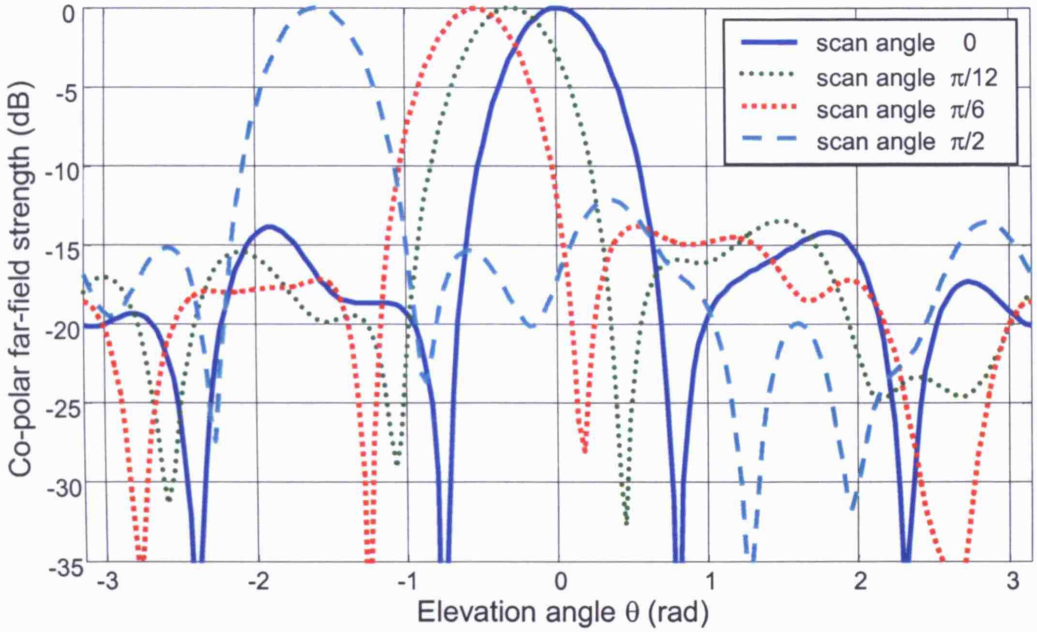


Figure 9.31: A vast improvement over the patterns shown in figure 9.25, also for the configuration with 60 subarrays and for the same scan angles, is obtained here by cutting of the element RGPs shown in figure 9.24 (a) and (b) to a maximum degree of $L_{max} = 5$. Also the reconstruction of the array RGP was limited to the same degree. This shows that the combination of a clever signal processing with a a proper design for the subarray radiation patterns can improve the performance of the spherical array with less elements than would otherwise be required. For the case of LISA it will further be investigated how far the performance can be pushed with only 60 channels, a reduction by 4 to the number of tiles and by 12 to the total number of radiating elements. This is important, because the key to produce low-cost spherical array antennas will lie in the correct combination of the proposed subarraying technique with a good overall array performance.

both the element RGPs as well as the reconstruction of the array RGP to a maximum degree of $L_m = 5$. Although the results shown cannot be obtained in reality as it is impossible to completely remove the higher degree modes from the field, the well-behaved plots in figure 9.31 suggest that it is indeed the higher-degree modes that caused the distorted patterns in the previous figures. The plots in figure 9.24 suggest that the problems arise from the higher degree modes in the patterns of the tiles and even more so from the subarrays. The subarraying as it is currently implemented does not have the desired anti-aliasing effect, but is aggravating the problem. With the aid of the convolution theorem in chapter 7 it may be possible to synthesise a better pattern for the subarray, but it is possible that having only 60 subarrays is surpassing the limits of what can be done with an array of this size and only a limited amount of controls. Further research is necessary to establish this. On the other hand, the scenario where the 240 tiles are controlled separately already demonstrates the unique properties of the spherical array antenna, and satisfactory results are obtained with filtering techniques that can effectively be implemented.

A final experimental study was made to check how well the array can produce low-gain patterns. Radiating equally in all directions with the same polarisation is impossible. By allowing some zeros in the pattern, the array should however be capable of producing a low-gain pattern of

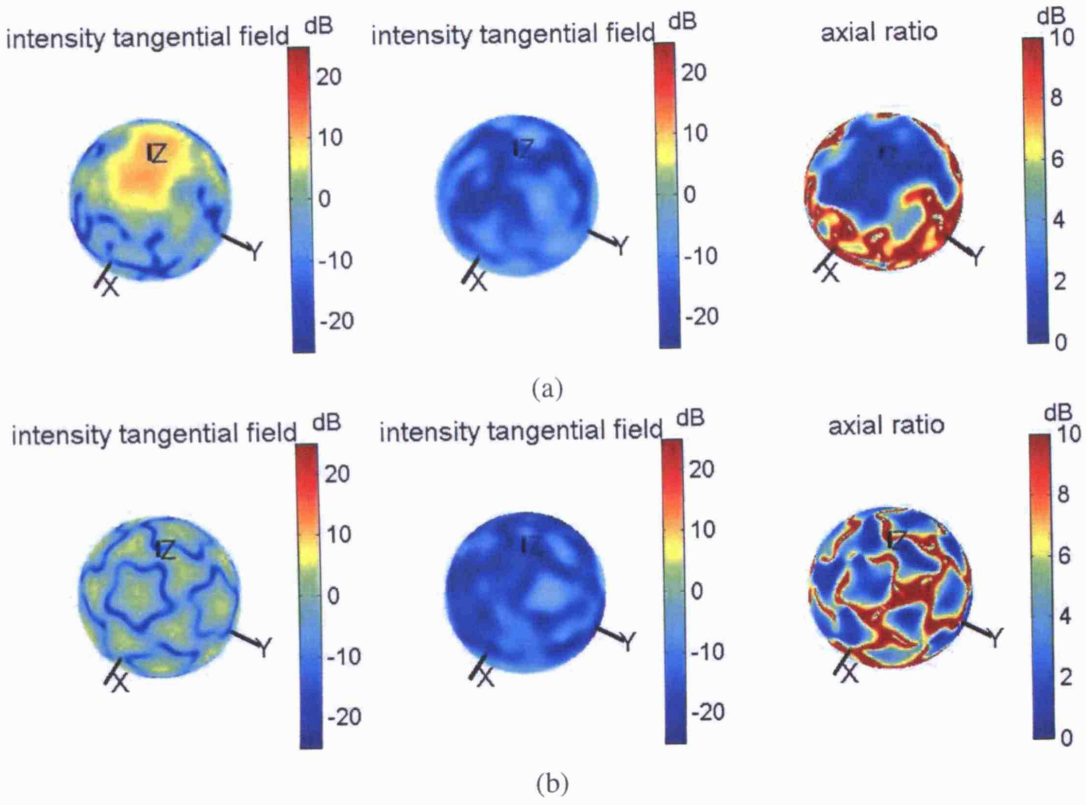


Figure 9.32: Two different low-gain patterns were synthesised. both patterns were created by feeding the subarrays with signals of equal amplitude. In (a) the phase were adjusted to align the elements of each subarray with the polarisation definition at the subarray phase centre. In (b) the subarrays were equally excited.

one single polarisation with very low cross-polarisation. Two different array excitations have been tried to achieve this. Both excitations excite all subarrays with uniform amplitude, but they differ in the phasing of the signals. As all elements are circularly polarised, changing the phase of the feeding signal changes the alignment of their phasors. This was used to align the elements of each subarray to the definition of co-polarisation at the phase centre of the subarray. As everywhere in this text, Ludwig's third definition for co- and cross-polarisation is used, that was defined in section 2.4.3. The reason to use that polarisation was explained in section 6.1. If this is effectuated, the phasors of the elements will align with respect to some favoured direction, taken to be the positive z-axis. Therefore a low-gain radiation pattern is expected in the direction of the positive z-axis, with very low cross-polarisation, that increase towards SP. The result is shown in plot (a) of figure 9.32.

In the second approach, all subarrays were equally excited in phase and amplitude. The positioning with icosahedral symmetry is left to set the alignment of the array element phasors. The symmetry will make that no direction is favoured, but in some directions the signals will cancel out and cause crests in the radiation pattern. The result is shown in plot (b) of figure 9.32.

9.6 Summary and Conclusions

In this chapter the proceedings were reported of the ongoing development of LISA, a prototype spherical array antenna. A number of experiments have been outlined, and a brief description of the measurement facility was given. A good understanding of the measurement facility is especially important as it is foreseen that a close interaction between the measurement system and the AUTwill be implemented in future experiments.

In the spirit of the general approach taken in this thesis, LISAs hardware will be controlled by a set of signals that correspond to a spherical mode excitation. Various signal processing schemes will be implemented in software. The system-level architecture to implement this approach has been shown and the subsystems were described with a varying level of detail.

Most attention was given to the RF front-end, consisting chiefly of the array panels with NOSCIPRA elements and an integrated RF circuit to feed the elements. Each of the panels has 3 radiating elements and a single connector, so that in all future configurations, the atomic radiation patterns that constitute the overall array RGP will be the RGPs of the 240 array panels. When this thesis drew to a closure, the RF front-end was fully developed and a test batch of array panels was produced and tested.

The results of a first set of experiments with the available panels has shown excellent performance of the panels and subarrays. The results were also used to get a more realistic prediction of the array performance. As expected, the performance of the completed array antenna will vary, depending on the several possible configurations it can be used in. As the minimal configuration with 60 subarrays offers a major cost and complexity reduction over its alternative configurations, future research will most probably concentrate on the advancement of the array performance in this minimalistic configuration. It was suggested how this could be achieved.

The work reported in this chapter indirectly contributes to all parts of the hypothesis in section 1.3.1, as it allows empirical research as a method to investigate the proposed SMT approach towards spherical array antennas. Moreover, the efforts put in the development of the various subsystems of LISA contribute in many ways to the field of antenna design in general. The expertise build up by the actual design and manufacturing of LISA will ultimately lead to a fair assessment of the last point of the hypothesis regarding the economical viability of array antennas for specific applications. As this will greatly depend on the trade-off between performance and the lowest number of active channels, it is too early to make any meaningful statements on that matter.

Chapter 10

Conclusions and Future Work

10.1 Conclusions

Firstly, the hypothesis that lay at the start of this work and was formulated in the introduction is repeated here:

- By using, and where necessary extending, Spherical Modal Theory (SMT), the electromagnetic analysis of spherical antennas can be simplified. The simpler description can be used in a computer model to simulate the radiation properties of a large spherical array antenna. Because the model is based on analytically obtained formula's, the simulations are faster than can be achieved with numerical EM solvers.
- Spherical modal theory allows an accurate integrated approach towards the analysis and synthesis of radiation patterns from a spherical array antenna, with inclusion of mutual coupling effects. With SMT the mutual coupling effects can be taken into account at once, in both the analysis and synthesis of a radiation pattern, and therefore disappear as unwanted, unknown, and compromising electromagnetic effects.
- Using SMT, the design and performance anticipation of spherical array antennas can be reduced to a few simple formula's, as a quicker and easier alternative for detailed analysis.
- Based on SMT, a framework can be constructed in analogy with phase mode theory for circular array antennas, that facilitates the development of advanced signal processing techniques for spherical array antennas.
- Spherical array antennas offer an economical way of achieving true omnidirectional beamsteering. With methods based on SMT the costs of spherical array antennas can be kept low by concentrating information of all array elements in a fewer number of channels. With the intelligence of the antenna residing in the signal processing stage, a cost reduction and reliability gain is achieved over switching or commutating solutions with complex hardware and/or expensive MicroElectroMechanical System (MEMS) technology.

In chapter 2 Spherical Modal Theory (SMT) has been thoroughly reviewed. Several modal expansions were reviewed, laying the theoretical foundations for the rest of this thesis. An

overview of different SMT expansions was given in table 2.1. An essential new development was the Circularly Polarised Spherical Transform (CPST). This expansion separates the Right Hand Circular Polarisation (RHCP) and Left Hand Circular Polarisation (LHCP) components of the far-field in two sets of expansion coefficients. This greatly simplifies the formulation of the theory to deal with spherical array antennas and the CPST was therefore used extensively throughout this thesis. A particular definition for the polarisation, defined in section 2.4.3 was adhered to.

Chapter 3 focused on the electromagnetic properties of one single antenna element on a large conducting sphere. The Spherical Wave Expansion (SWE) of such antenna was obtained for two distinct types of antennas: the aperture and the patch-type antenna. Also the design of a new kind of antenna, called Circularly Polarised, Non-Sequentially Coupled, Single-Feed, Annular-Ring Microstrip Antenna (NOSCIPRA), was presented as the element of choice for LISA, a prototype spherical array antenna that is currently still under development. For a circular aperture antenna, the comparison was made between the theoretically predicted antenna pattern and a simulated one and an impressive agreement was found between the two. The fast and accurate analysis of a single radiating element on a conducting sphere is a first example where SMT can be successfully applied to simplify and rapidly solve an electromagnetic problem regarding a spherical antenna, which proves the first point of the hypothesis.

The next level of complexity was to consider a pair of elements and calculating how much they electromagnetically couple into each other. In chapter 4 expressions were obtained analytically for the coupling coefficients, as a function of the positions of the two elements and the SWE coefficients of their radiated field. Hereto it was necessary to implement field rotations with the aid of SMT, an important concept that was later re-used in other chapters as well.

A practical spherical array antenna will ultimately consist of a large number of radiating elements, distributed over the sphere. In chapter 5 a model was presented, the scattering matrix description, in which the results obtained in the previous chapters were combined into one single description of an entire spherical array antenna. The model can quite simply be adapted to spherical array antennas with all kinds of feeding structures, radiating elements and so forth and permits an accurate calculation of the radiated field for a given array excitation, with inclusion of mutual coupling effects. The model has been implemented in software and is fully operational, albeit that some parts of the code could be implemented more efficiently and some modules to extend its capabilities yet have to be written. With the development, implementation and testing of the model, point two of the hypothesis is partially accepted, with the sidenote that more work is necessary to develop fast and accurate procedures to synthesise an array excitation. The ingredients necessary to do so have been developed and tested, chiefly in chapter 7, but they have not yet been integrated with the implementation of the scattering matrix description.

In chapter 6 a number of parameters that are commonly used to describe the performance of an antenna were obtained for the spherical array antenna by placing them in the context of SMT. The focus lay on the establishment of simple design formula's, in the spirit of the third point of the hypothesis in section 1.3.1. A detailed treatment of the polarisation definition already used in previous chapters was given and the choice for this particular polarisation was justified. The

adoption of a suitable polarisation and the definition of the associated CPST in chapter 2 were of key importance to render the derivations in this chapter in their simplest form. Much attention was given to the maximum attainable directivity as well as to the beamwidth, sidelobe levels and aperture efficiency of the spherical array antenna. Also, a preliminary sampling theorem was intuitively formulated, that was given more attention in chapters 7 and 8. There is scope for future work to complete the few presented formula's with more performance characteristics, such as for instance the quantification of various error terms.

In chapter 7 a framework for advanced signal processing techniques was established. Formula's were presented to backproject the CPST to any finite radius. The abstract machinery of distributions was used to formally obtain a sampling theorem for both scalar functions and vector fields tangential to the sphere that relates the spatial bandwidth of the radiation pattern to the sampling degree of the sampling distribution, with abstraction from the actual distributions used. Again using the CPST, the convolution theorem was generalised to vectorfields tangential to the sphere, which is the most important tool for an efficient implementation of various spatial filters such as those used in direction finding algorithms or anti-aliasing filters. From a system-level perspective, a processing chain was laid out for a practical antenna system. This chapter focused on the fourth point of the hypothesis. Also this point is close to being flagged as true, although there is plenty of scope to develop more advanced signal processing schemes with the tools provided in chapter 7.

With a spherical geometry, the relation between the spatial bandwidth of a function to be sampled and the required number of samples to reconstruct the function without errors is much more complicated than for circular or periodic functions. Therefore the sampling distribution was introduced as an abstract tool that could be used in the previous chapters, without consideration for the complex relationships between its spatial bandwidth and the collection of samples. It was found that not only the number of samples were important, but also their position and the weights, if any, used to multiply the sampled values with. All this matter was the subject of chapter 8. Apart from an in-depth study of this topic, three main sampling strategies were suggested to be used on a spherical array antenna, depending on whether the total number of elements is very large (typically over a hundred), medium range (from 20 to about a hundred) or small (20 and less). For 20 or less elements, the vertices of one of the platonic solids would be used as sampling points, whereas for a medium range distribution, a t-design would be used. For a very large distribution it may be advisable to use subarraying, a technique that limits the number of active channels to less than the number of elements. Also the technique of subarraying is a more complicated matter on the sphere, and received a lot of attention in chapter 8.

In chapter 9 the development of LISA, a prototype spherical array antenna was described. A system-level description was given, and the design of the subsystems was reported with a varying level of detail. Most attention was given to the RF front-end, containing the array panels with three radiating NOSCIPRA elements. A first batch of array panels was designed, manufactured and tested and the results from measurements were used to forecast the performance of the fully populated array antenna. The total number of radiating elements on LISA is 720. It

will be a challenge to make the array work well with as few as 60 channels, but some hopeful results were obtained that point in that direction. An alternative configuration is to use the 240 tiles as subarrays, in which case the antenna can be made to perform well without much difficulty. Not before the LISA array is completed and operational, can solid conclusions be drawn regarding the real performance limitations, and therefore the last point of the hypothesis cannot be judged objectively at this point in time. The approach in this thesis was purposefully kept fairly generic and concentrated on methods rather than applications. A focused development towards a particular application may shed some light on the actual cost versus performance trade-off that spherical array antennas may bring to the market.

10.2 Future Work

The research undertaken in this PhD programme will be continued beyond the scope of this thesis. The aim is to investigate all aspects of the spherical array antenna and demonstrate the findings with a fully operational prototype spherical array antenna. Currently, another student has started a PhD programme that will continue this work. This chapter aims to ensure the continuity of the overall project with an outline of suggestions for future work that would complement the work reported in this thesis quite nicely. This outline solely reflects the opinion of the author at the time of writing. It may serve as a perhaps useful guideline for the ongoing research, but is also likely to be adjusted to new priorities as the research continues and new insights come to light.

10.2.1 Signal Processing

Several tasks are yet on the to-do list for future research, but the most extensive list of tasks that can be readily taken on in the near future pertains to the processing of the array signals. Perhaps this is also the most interesting part, as there is a lot to be gained in terms of performance and costs of the spherical array antenna. With the theory established in chapter 7 at hand, it is also thought that there is fundamentally little in the way to start developing more advanced signal processing schemes.

One of the more immediate, and still low-level, tasks would be to find the optimal subarray pattern to achieve the best array performance. This may involve building in a reconfiguration capability at subarray level, that can change the relative phase and amplitudes of the 4 tiles within a subarray, relative to a single feeding signal. There are of course lots of other possibilities of things that can be done, but from the findings in chapter 9 it is clear that a clever solution will need to be found if the array is to operate well with 60 controls instead of 240. Part of this work could be to quantify the impact of subarray directivity on overall array directivity and obtain appropriate rules of thumb or formula's for the parameter η_f, that was introduced in section 6.3.3. Also the plane-wave analysis mentioned in chapter 7 should be rather straightforward to implement. An interesting variant to this is to focus the array to a finite radius, so that it is not a plane-wave analysis but a curved wavefront analysis that is performed. In the environment of the anechoic chamber this may be used to locate the transmitting probe.

At medium level, some well established techniques could be drawn from circular array phase

mode processing and generalised to be applied to spherical array antennas. For example, Schelkunoff-type null steering could be implemented , where the array pattern is synthesised or analysed based on its nulls. To this effect, it may be worth exploring whether the Legendre addition theorem can be used to write the expanded sum of the radiation pattern in the form of a multiplication of spherical distances to null-coordinates. Another example is to investigate whether patterns without sidelobes can be formed with an excitation that is the equivalent of the binomial excitation on a circular array antenna.

At a higher level, the array antenna could be turned into a truly smart antenna, capable of performing high level tasks with tracking, beamforming, multi-beam and capabilities the like. The concept of smart antennas was not touched upon in this thesis, but interesting and powerful applications may be found where high-level signal processing meets the instantaneous omnidirectional sensing capability of the spherical array antenna.

Finally, another idea that could be given some more thought is the possibility of active Radar Cross Section (RCS) alteration and/or cancellation of the metal sphere to masquerade or hide the antenna in a hostile environment.

10.2.2 Analysis and Modelling

The implementation of the scattering matrix description can be improved by adding functionality for the patch-type elements. This is straightforward for elements where the cavity model is valid, as the required formula's were already obtained in chapter 3. However an even more powerful approach could be to step away a little bit from the pure analytical approach and devise hybrid simulation tools that can model more complex elements as the NOSCIPRA element, with increasing level of accuracy. The analysis can then be compared with results from measurements. A good candidate for a hybrid simulation tool could for instance combine a Finite Element Method (FEM) applied in a thin spherical region surrounding the complex structures with array panels, with Galerkin's Method of Moments (MoM) outside the spherical surface. With the spherical modes as basis functions, the latter boils down to application of the SMT based analysis applied in this thesis.

An interesting continuation of the work done on spherical array synthesis with the scattering matrix description would be to break down the applied correction for the mutual coupling effects in the spherical harmonic basis, and investigate whether the correction coefficients show an invariance towards rotation of the radiation pattern. If that were true, than the mutual coupling correction can be applied at the level of spherical modes, similar to what has been done with circular arrays[2]. As the rotational invariance of these coefficients would depend on each array element seeing an identical radiation environment, this would require a uniform array element distribution. For a large number of elements this cannot be achieved. However, with the subarraying techniques adopted in chapter 8, the radiation environment of each element may have enough similarity for the rotational invariance of the modal mutual coupling correction coefficients.

10.2.3 Development

The development of LISA is ongoing and the first task at hand is obviously to finish the planned developments as described in chapter 9. However, in future work even more ambitious developments could be investigated and tried out. An interesting possibility is to develop some kind of element or subarray that can be placed in an almost ad hoc way on a spherical surface. The array would then autocalibrate itself so that it works out how the subarrays need to be operated. A difficulty is that the subarrays would probably need to carry a full-blown receiver on board, so that the demodulated signal can be transmitted with a wireless link for instance. Furthermore, the subarrays could be made out of a flexible substrate to make them absolutely conformal. It is not expected that the ideas suggested above will be implemented anytime soon, but it is nice to know that the scope for future work is only limited by our imagination.

A more immediate, but not less demanding, step forwards could be to move to a wideband system, or at least a system with some bandwidth. In this thesis only spot-frequency operation has been considered, although it was mentioned already in the introduction that only arrays with a metal sphere in the middle would be considered because they offer more signal bandwidth in the near-field to far-field transform.

10.2.4 Fundamental Theory

Depending on the likings of future researchers, there are also more fundamental ideas to be explored and tasks to be done that require to delve a little deeper into theory. Perhaps one of the more urgent tasks in this category is the quantification of errors of all kinds to give answers to questions as "How would an average phase error of 8° on the radiating elements impact the radiation pattern?".

A completely different but intriguing idea is to use the volume of the sphere for spatial filtering. Instead of placing the radiating elements immediately above a conducting sphere, they could be placed on a thick dielectric sphere with a metal ball of smaller radius in the centre. The metal sphere in the centre is expected to act as a spatial low-pass filter on the modes that can propagate to the far-field to reduce aliasing for instance. A variant to this theme is to have spherical layers of different dielectric constant, with the spatial filter acting differently at different EM frequencies. This idea may be of interest for multi-band operation.

Of practical importance is to consider deviations of the groundplane from a perfect full sphere to other related shapes such as a hemisphere on a conducting plane for instance. This particular example would be of interest for an antenna application on the roof of a vehicle.

Finally, some attention is given in appendix D to an analog implementation of the Spherical Fourier Transform (SFT) transform. Although the author has some doubts about the feasibility and the value of exploring this idea much further, it may in some cases be a good idea to transfer some of the processing to the IF or RF stage, for instance to apply a spatial filter and thereby limiting the number of channels.

10.2.5 Applications

The prototype developed in this PhD project could be taken further towards a real operational array antenna that is capable of establishing and maintaining communication with a LEO satellite. Alternatively, a specific application could be targeted in cooperation with an industrial partner. An interesting application that comes to mind is an antenna for a micro-satellite that can establish communication with a ground station, regardless of its attitude control operations. Several possible applications were already mentioned in the introduction to this thesis.

Appendices

Appendix A

Convex Polyhedra

This appendix reviews the 3-dimensional solids or polyhedra that can be used to approximate a sphere with a simpler construction. It is likely that for most spherical array designs a faceted array based on one of these solids described here will be chosen over a perfectly spherical shape, because metal spheres are costly and difficult to manufacture.

More specifically, this appendix illustrates and classifies the convex uniform polyhedra and their dual solids. The uniform polyhedra are polyhedra with identical polyhedron vertices. They have regular polygons as faces and show a high degree of symmetry. More importantly, they have their vertices lying on a spherical surface, which makes them suitable to serve as an approximation of the sphere. In fact, to implement a spherical array antenna, we want the phase centres of all subarrays or facets to lie on a sphere, with the faces of the approximating solid tangent to a sphere in these phase centres. Moreover, the dual solids are face-uniform. That is why we will look at the dual solids of the convex uniform polyhedra as the candidate antenna shapes, while the vertices of the convex uniform polyhedra serve as phase centre distributions for the subarrays.

There are two kinds of uniform convex polyhedra: the regular or platonic solids and the semi-regular or archimedian solids (with the semi-regular including the quasi-regular solids). The dual solids of the platonic solids are again platonic solids. The dual solids of the archimedian solids are the catalan solids. Hence, the antenna shape will take the form of a platonic or a catalan solid. There are five platonic solids, and 13 archimedian solids.

All illustrations in this appendix were made using a shareware programme called Poly v.1.11. For more information on the geometry of solids, the following references can be useful: [138][139][140][141][142][143]. In addition, an extensive amount of non peer reviewed information is available on the following websites: www.wikipedia.org and mathworld.wolfram.com.

For each of the solids, the symmetry group will be given, as this is important for the construction of antennas with the subarraying technique. The finite subgroups of the rotation group can be classified as follows [144]:

- C_k: the cyclic group (e.g. pyramid)
- D_k: the dihedral group (e.g. prism)

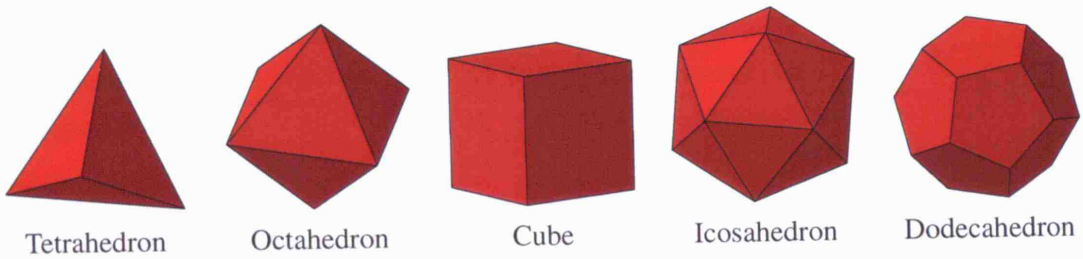


Figure A.1: The five platonic solids

- T: the tetrahedral group (e.g. tetrahedron)
- O: the octahedral group (e.g. cube)
- I: the icosahedral group (e.g. icosahedron)

The first two groups are related with rotations in 2 dimensions and not very interesting in the context of this thesis. When a group symbol gets a subscript h, as in I_h, this denotes the full achiral group. Without the subscript, the symbol refers to the chiral subgroup.

Some of the solids below are mentioned with a * after their name. These solids are chiral solids - they exist in two enantiomorph forms that are each others mirror image but cannot be mapped onto one another by translations and rotations alone.

A Platonic Solids

The Platonic solids, or regular solids, are regular convex polyhedra, with regular convex polygon faces that are all congruent. There are five such solids, illustrated in figure A.1. The dual solid of a platonic solid is again a platonic solid.

Of the platonic solids, the tetrahedron has tetrahedral symmetry. The octahedron and the cube have octahedral symmetry. The platonic solids with icosahedral symmetry are the icosahedron and the dodecahedron. The dual solid of the tetrahedron is the tetrahedron - it is self-dual. The cube and the octahedron are each others dual solid. The same is true for the icosahedron and the dodecahedron.

B Archimedian Solids

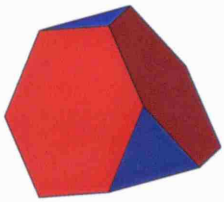
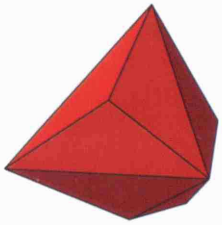
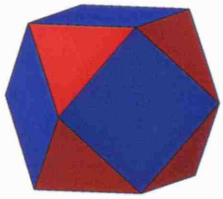
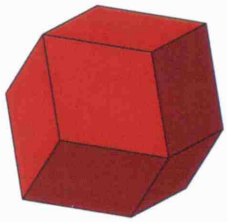
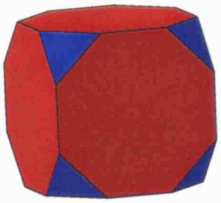
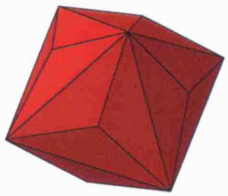
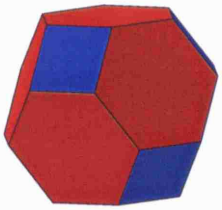
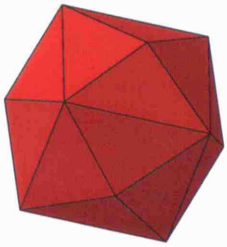
The archimedian solids, or semi-regular solids, have two or more sets of regular congruent faces. They are illustrated in table A.1.

Table A.1: The archimedian and catalan solids are listed. The column headed with V contains the number of vertices of the archimedian solid, which is also the number of faces of the dual catalan solid. The column headed with F contains the number of faces of the archimedian solid and the number of vertices of the dual catalan solid.

Archimedian Solid	**V**	**F**	**S**	**Catalan Solid**

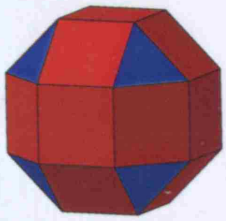
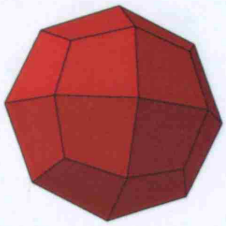
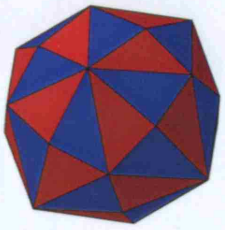
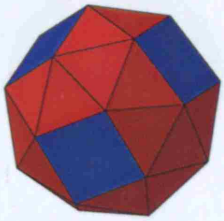
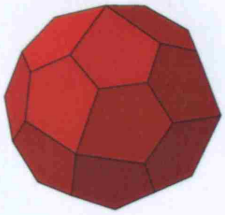
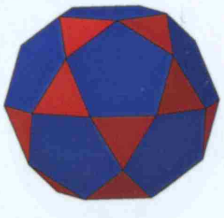
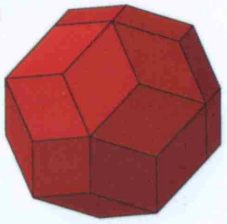
Continued on next page

Table A.1 – continued from previous page

Archimedian Solid	V	F	S	Catalan Solid
truncated tetrahedron 	12	8	T_d	triakis tetrahedron 
cuboctahedron 	12	14	O_h	rhombic dodecahedron 
truncated cube 	24	14	O_h	triakis octahedron 
truncated octahedron 	24	14	O_h	tetrakis hexahedron 

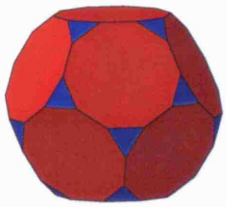
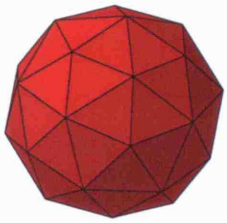
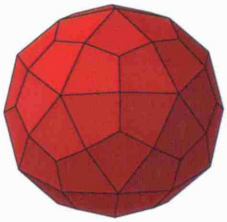
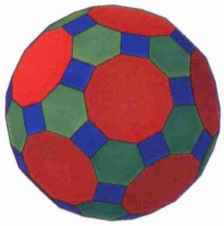
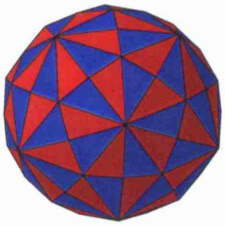
Continued on next page

Table A.1 – continued from previous page

Archimedian Solid	V	F	S	Catalan Solid
rhombicuboctahedron 	24	26	O_h	deltoidal icositetrahedron 
rhombitruncated cuboctahedron 	48	26	O_h	disdyakis dodecahedron 
snub cuboctahedron* 	24	38	O	pentagonal icositetrahedron* 
icosidodecahedron 	30	32	I_h	rhombic triacontahedron 

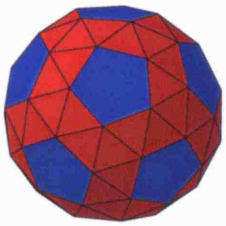
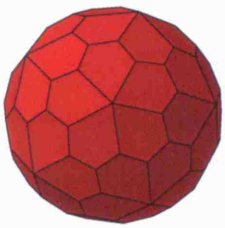
Continued on next page

Table A.1 – continued from previous page

Archimedian Solid	V	F	S	Catalan Solid
truncated dodecahedron 	60	32	I_h	triakis icosahedron 
truncated icosahedron 	60	32	I_h	pentakis dodecahedron 
rhombicosidodecahedron 	60	62	I_h	deltoidal hexecontahedron 
rhombitruncated icosidodecahedron 	120	62	I_h	disdyakis triacontahedron 

Continued on next page

Table A.1 – continued from previous page

Archimedian Solid	V	F	S	Catalan Solid
snub icosidodecahedron* 	60	92	I	pentagonal hexecontahedron* 

Appendix B

Array Element Positions

Table B.1 lists the Euler angles for the 240 phase centres of the array panels. Each of the panels features three radiating elements. Four tiles constitute one subarray. Every four entries the phase centre of a subarray is given between horizontal lines. These centres coincide with the phase centres of the centre tile within every one subarray.

Apart from the Euler angles, the table lists the radii of the phase centres of the tiles. Indeed, these points do not lie on the surface of a sphere. The phase centres of the subarrays however do lie on a spherical surface of radius $R = 420mm$.

The positions are given with three Euler angles, rather than two, so that the orientation and position of the tiles are specified. Care should be taken, because not all tiles are tangential to the same sphere. The first two Euler angles, together with the radius, specify the phase centre of each tile exactly. When applied to a reference vector in $(0, 0, R)$ and aligned with the x-axis, 240 vectors are obtained that are almost aligned with the tiles, but not exactly. These vectors are however all tangential to the sphere and can therefore be used in an algorithm to superimpose the far-fields from each tile, as was done in section 9.5.3. The vectors are shown in figure B.1. To obtain all vertices of the tiles exactly, the 60 icosahedral rotations, indicated in the table between horizontal lines, can be applied to the vertices in table B.2.

Table B.1: Listing of 240 sets of Euler angles and radii that specify the positions of the 240 tiles of the array antenna.

Tile No	ξ	θ	ϕ	R
1	2.73	-0.88	-0.28	420.00
2	2.89	0.91	2.60	420.13
3	2.65	0.71	2.97	420.40
4	2.65	1.04	3.01	420.13
5	0.41	0.88	-2.86	420.00
6	-2.65	1.04	-3.01	420.13
7	-2.65	0.71	-2.97	420.40
8	-2.89	0.91	-2.60	420.13
9	-1.88	-0.55	0.72	420.00

Continued on next page

Table B.1 – continued from previous page

Tile No	ξ	θ	ϕ	R
10	-1.92	0.76	-2.37	420.13
11	1.58	-0.52	0.36	420.40
12	0.96	-0.43	1.07	420.13
13	3.14	0.18	-3.14	420.00
14	2.02	-0.20	1.14	420.13
15	0.00	0.37	-3.14	420.40
16	-2.02	-0.20	-1.14	420.13
17	0.84	-0.74	1.25	420.00
18	-2.39	-0.54	1.35	420.13
19	-2.42	-0.87	1.43	420.40
20	-2.11	-0.83	0.99	420.13
21	-2.17	1.35	-0.86	420.00
22	0.96	1.15	-0.83	420.13
23	0.95	1.46	-0.71	420.40
24	1.01	1.46	-1.04	420.13
25	-1.57	1.57	0.18	420.00
26	1.57	1.39	0.09	420.13
27	1.57	1.57	0.37	420.40
28	1.57	1.75	0.09	420.13
29	2.07	-1.17	-1.96	420.00
30	-1.00	-1.06	-2.15	420.13
31	1.99	1.08	1.37	420.40
32	-1.07	-1.37	-1.95	420.13
33	1.26	0.55	0.72	420.00
34	1.22	-0.76	-2.37	420.13
35	-1.57	0.52	0.36	420.40
36	0.96	-0.43	-2.07	420.13
37	2.17	1.35	0.86	420.00
38	-1.01	1.46	1.04	420.13
39	-0.95	1.46	0.71	420.40
40	-0.96	1.15	0.83	420.13
41	0.00	-1.75	-1.57	420.00
42	-3.12	-1.66	-1.39	420.13
43	-3.14	-1.94	-1.57	420.40
44	3.12	-1.66	-1.75	420.13
45	0.98	-1.35	-0.86	420.00
46	0.96	1.15	2.31	420.13
47	-2.19	-1.46	-0.71	420.40

Continued on next page

Table B.1 – continued from previous page

Tile No	ξ	θ	ϕ	R
48	-2.13	-1.46	-1.04	420.13
49	-3.14	-1.39	-1.57	420.00
50	0.02	-1.49	-1.75	420.13
51	0.00	-1.20	-1.57	420.40
52	-0.02	-1.49	-1.39	420.13
53	-2.30	-2.41	-1.25	420.00
54	-2.39	2.60	1.79	420.13
55	0.72	-2.27	-1.43	420.40
56	1.03	-2.31	-0.99	420.13
57	-0.41	-2.26	0.28	420.00
58	2.89	-2.23	0.54	420.13
59	-0.49	2.43	-2.97	420.40
60	-0.49	2.10	-3.01	420.13
61	0.34	-1.22	0.43	420.00
62	-2.88	-1.07	0.59	420.13
63	-2.82	-1.39	0.49	420.40
64	0.42	1.20	-2.93	420.13
65	-0.98	1.79	2.28	420.00
66	-1.01	-1.69	-1.04	420.13
67	-0.95	-1.68	-0.71	420.40
68	-0.96	-1.99	-0.83	420.13
69	-0.00	2.96	3.14	420.00
70	-1.13	-2.94	-1.14	420.13
71	-3.14	2.78	3.14	420.40
72	1.13	-2.94	1.14	420.13
73	-1.08	-1.97	1.96	420.00
74	2.14	-2.08	2.15	420.13
75	1.99	-2.06	1.77	420.40
76	2.07	-1.77	1.95	420.13
77	-0.84	-0.74	1.90	420.00
78	-1.03	0.83	-0.99	420.13
79	2.43	-0.87	1.71	420.40
80	2.39	-0.54	1.79	420.13
81	-1.57	1.57	-2.96	420.00
82	1.57	1.39	-3.06	420.13
83	-1.57	-1.57	0.37	420.40
84	1.57	1.75	-3.06	420.13
85	2.30	-2.41	1.25	420.00

Continued on next page

Table B.1 – continued from previous page

Tile No	ξ	θ	ϕ	R
86	2.11	2.31	-2.16	420.13
87	-0.71	-2.27	1.43	420.40
88	-0.75	-2.60	1.35	420.13
89	0.34	1.93	-0.43	420.00
90	-2.88	2.07	-0.59	420.13
91	-2.82	1.75	-0.49	420.40
92	-2.72	1.95	-0.21	420.13
93	0.41	0.88	0.28	420.00
94	-2.65	1.04	0.13	420.13
95	-2.65	0.71	0.17	420.40
96	-2.89	0.91	0.54	420.13
97	3.14	0.18	0.00	420.00
98	2.02	-0.20	-2.01	420.13
99	0.00	0.37	0.00	420.40
100	1.13	0.20	-1.14	420.13
101	0.84	-0.74	-1.90	420.00
102	-2.39	-0.54	-1.79	420.13
103	0.72	0.87	1.43	420.40
104	1.03	0.83	0.99	420.13
105	-2.07	-1.17	-1.18	420.00
106	-2.07	1.37	1.95	420.13
107	1.16	-1.08	-1.37	420.40
108	1.00	-1.06	-0.99	420.13
109	2.81	1.22	2.71	420.00
110	-0.42	1.20	2.93	420.13
111	2.83	-1.39	-0.49	420.40
112	2.88	-1.07	-0.59	420.13
113	-2.81	-1.93	-0.43	420.00
114	0.26	-2.07	-0.59	420.13
115	-2.82	1.75	2.65	420.40
116	-2.72	1.95	2.93	420.13
117	-0.98	-1.35	0.86	420.00
118	2.13	-1.46	1.04	420.13
119	2.19	-1.46	0.71	420.40
120	2.18	-1.15	0.83	420.13
121	2.81	-1.93	0.43	420.00
122	2.72	1.95	-2.93	420.13
123	2.83	1.75	-2.65	420.40

Continued on next page

Table B.1 – continued from previous page

Tile No	ξ	θ	ϕ	R
124	-0.26	-2.07	0.59	420.13
125	-2.07	-1.17	1.96	420.00
126	1.07	-1.37	1.95	420.13
127	-1.98	1.08	-1.37	420.40
128	-2.14	1.06	-0.99	420.13
129	2.07	-1.17	1.18	420.00
130	-1.00	-1.06	0.99	420.13
131	-1.15	-1.08	1.37	420.40
132	2.07	1.37	-1.95	420.13
133	-0.41	0.88	-0.28	420.00
134	2.89	0.91	-0.54	420.13
135	2.65	0.71	-0.17	420.40
136	2.65	1.04	-0.13	420.13
137	2.81	1.22	-0.43	420.00
138	-0.42	1.20	-0.21	420.13
139	-0.32	1.39	-0.49	420.40
140	-0.26	1.07	-0.59	420.13
141	-2.07	1.97	1.18	420.00
142	1.07	1.77	1.19	420.13
143	-1.98	-2.06	-1.77	420.40
144	-2.14	-2.08	-2.15	420.13
145	0.98	1.79	0.86	420.00
146	-2.18	1.99	0.83	420.13
147	-2.19	1.68	0.71	420.40
148	-2.13	1.69	1.04	420.13
149	-2.73	2.26	2.86	420.00
150	0.49	2.10	3.01	420.13
151	0.50	2.43	2.97	420.40
152	-2.89	-2.23	-0.54	420.13
153	-0.84	2.41	1.25	420.00
154	-1.03	-2.31	-2.16	420.13
155	-0.71	-2.27	-1.71	420.40
156	-0.75	-2.60	-1.79	420.13
157	0.00	-1.75	1.57	420.00
158	-3.12	-1.66	1.75	420.13
159	-3.14	-1.94	1.57	420.40
160	3.12	-1.66	1.39	420.13
161	-1.88	2.59	2.42	420.00

Continued on next page

Table B.1 – continued from previous page

Tile No	ξ	θ	ϕ	R
162	-1.92	-2.39	-0.77	420.13
163	1.58	2.62	2.78	420.40
164	-2.18	-2.71	-1.07	420.13
165	1.57	1.57	-0.18	420.00
166	-1.57	1.75	-0.09	420.13
167	-1.57	1.57	-0.37	420.40
168	-1.57	1.39	-0.09	420.13
169	-2.07	1.97	-1.96	420.00
170	-2.07	-1.77	1.19	420.13
171	1.16	2.06	-1.77	420.40
172	-2.14	-2.08	0.99	420.13
173	-0.34	1.93	0.43	420.00
174	2.72	1.95	0.21	420.13
175	2.83	1.75	0.49	420.40
176	2.88	2.07	0.59	420.13
177	-1.26	0.55	-0.72	420.00
178	2.18	0.43	-1.07	420.13
179	1.57	0.52	-0.36	420.40
180	1.92	0.76	-0.77	420.13
181	-0.84	-0.74	-1.25	420.00
182	2.11	-0.83	-0.99	420.13
183	2.43	-0.87	-1.43	420.40
184	2.39	-0.54	-1.35	420.13
185	1.57	1.57	2.96	420.00
186	-1.57	1.75	3.06	420.13
187	-1.57	1.57	2.78	420.40
188	-1.57	1.39	3.06	420.13
189	-2.17	-1.79	0.86	420.00
190	0.96	-1.99	0.83	420.13
191	0.95	-1.68	0.71	420.40
192	1.01	-1.69	1.04	420.13
193	-3.14	-1.39	1.57	420.00
194	0.02	-1.49	1.39	420.13
195	0.00	-1.20	1.57	420.40
196	-0.02	-1.49	1.75	420.13
197	-0.98	1.79	-0.86	420.00
198	-1.01	-1.69	2.10	420.13
199	2.19	1.68	-0.71	420.40

Continued on next page

Table B.1 – continued from previous page

Tile No	ξ	θ	ϕ	R
200	2.18	1.99	-0.83	420.13
201	-2.81	1.22	0.43	420.00
202	0.26	1.07	0.59	420.13
203	0.32	1.39	0.49	420.40
204	0.42	1.20	0.21	420.13
205	2.07	1.97	1.96	420.00
206	-1.00	2.08	2.15	420.13
207	-1.15	2.06	1.77	420.40
208	2.07	-1.77	-1.19	420.13
209	1.88	2.59	-2.42	420.00
210	2.18	-2.71	1.07	420.13
211	1.57	-2.62	0.36	420.40
212	1.92	-2.39	0.77	420.13
213	0.84	2.41	-1.25	420.00
214	0.75	-2.60	1.79	420.13
215	0.72	-2.27	1.71	420.40
216	1.03	-2.31	2.16	420.13
217	-2.73	2.26	-0.28	420.00
218	0.49	2.10	-0.13	420.13
219	0.50	2.43	-0.17	420.40
220	0.26	2.23	-0.54	420.13
221	2.73	2.26	0.28	420.00
222	-0.26	2.23	0.54	420.13
223	-0.49	2.43	0.17	420.40
224	-0.49	2.10	0.13	420.13
225	1.88	2.59	0.72	420.00
226	-0.96	2.71	1.07	420.13
227	-1.57	2.62	0.36	420.40
228	-1.22	2.39	0.77	420.13
229	-0.00	2.96	-0.00	420.00
230	-1.13	-2.94	2.01	420.13
231	-3.14	2.78	-0.00	420.40
232	-2.02	2.94	1.14	420.13
233	-1.88	2.59	-0.72	420.00
234	-1.92	-2.39	2.37	420.13
235	1.58	2.62	-0.36	420.40
236	-2.18	-2.71	2.07	420.13
237	1.88	-0.55	-0.72	420.00

Continued on next page

Table B.1 – continued from previous page

Tile No	ξ	θ	ϕ	R
238	-0.96	-0.43	-1.07	420.13
239	1.57	0.52	2.78	420.40
240	1.92	0.76	2.37	420.13

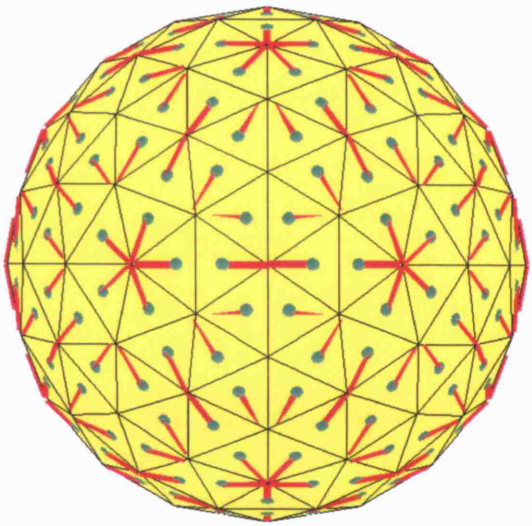


Figure B.1: The 240 tiles of the array antenna are shown, together with their 240 phase centres (in green). Also shown are 240 rotated versions of a reference vector (in red) with the rotations in table B.1. Without translations, these vectors can not possibly coincide with the tiles, but the vectors illustrate the pure rotations that can be used when superimposing the far-fields from the tiles.

Table B.2: Four times three vertices of the four tiles in one reference subarray of the array antenna. Applying the 60 icosahedral rotations specified by the highlighted Euler angles in table B.1 to the vertices in this list, will result in 240 triplets of vertices, one for each tile of the array antenna. It is not possible to construct the array with 240 rotations applied to a single tile.

Tile	X	Y	Z
centre	-44.41	-81.51	420.00
centre	-44.40	81.53	420.00
centre	76.14	0.01	420.00
top	71.52	-153.13	394.50
top	-44.41	-81.51	420.00
top	76.14	0.01	420.00
bottom right	-155.96	0.01	397.11
bottom right	-44.40	81.53	420.00
bottom right	-44.41	-81.51	420.00
bottom left	71.52	153.15	394.49
bottom left	-44.40	81.53	420.00
bottom left	76.14	0.01	420.00

Appendix C

Summary of MSc Thesis

This appendix briefly reviews some selected parts of my M.Sc. thesis[30] in which a preliminary design study of the spherical array was undertaken. The selected parts had a direct impact on the starting point of this PhD thesis and are therefore briefly reviewed. Each section covers a specific conclusion that was drawn from that earlier work.

C.1 A Spherical Harmonic Excitation Gives Rise to a Corresponding Harmonic in the Far-Field Pattern

In this section the Far-Field from a radiating sphere is obtained by calculating the Far-Field of each spherical harmonic that excites the sphere. The formulation used here is for scalar harmonics and inadequate for a detailed analysis of an electromagnetic spherical array antenna. It is included here because it is simple and elegant, and can be used for a detailed analysis of non-electromagnetic arrays (e.g. for SONAR applications). It also shows clearly the equivalence between the plane-wave expansion and the Far-Field transfrom.

Consider an excitation function, defined as a complex function over two variables: the angles of elevation and azimuth. This excitation function can be expanded in a basis of spherical harmonic terms. The question answered below, is how each of these harmonic terms will contribute to the far-field pattern of the excited sphere.

Let $E(\theta', \phi')$ be the excitation on the sphere. A prime denotes coordinates or entities affiliated with the excitation function.

$$E(\theta', \phi') = Y_{l'}^{m'}(\theta', \phi')$$

Where Y_l^m represents the (l, m) spherical harmonic. Using the far-field approximation we can calculate the far-field pattern as

$$D(\theta, \phi) = \frac{1}{4\pi} \iint_{s'} Y_{l'}^{m'}(\theta', \phi')\, e^{j\beta a \cos\psi} ds' \tag{C.1}$$

The plane waves in the far-field can be expanded into a sum of partial waves using *Bauers's formula.*

$$e^{j\beta a \cos\psi} = \sum_{l=0}^{\infty} (2l+1)\, j^l\, \hat{\mathrm{J}}_l(\beta a)\, \mathrm{P}_l(\cos\psi)$$

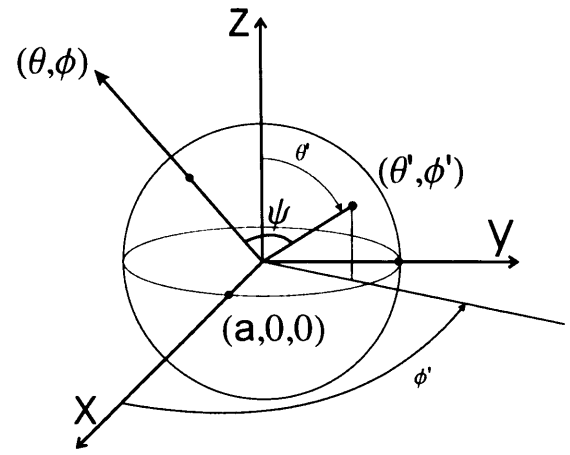


Figure C.1: Conventional spherical coordinates are used, as indicated in the figure.

Where $\hat{\mathrm{J}}_l$ is a spherical Bessel function of the first kind and P_l a Legendre polynomial, both of order l. We can also use the *addition theorem* to expand the Legendre zonal harmonics into spherical harmonics:

$$\mathrm{P}_l(\cos\psi) = \frac{4\pi}{2l+1} \sum_{m=-l}^{l} Y_l^{m^*}(\theta', \phi') Y_l^m(\theta, \phi)$$

The far-field becomes:

$$D(\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} j^l \, \hat{\mathrm{J}}_l(\beta a) \, Y_l^m(\theta,\phi) \underbrace{\iint_{s'} Y_{l'}^{m'}(\theta',\phi') \, Y_l^{m^*}(\theta',\phi') \, ds'}_{\delta_{ll'}\delta_{mm'}}$$

Where we used the orthogonality of spherical harmonics over the sphere. Finally, we get for the far-field:

$$D(\theta,\phi) = j^{l'} \, \hat{\mathrm{J}}_{l'}(\beta a) \, Y_{l'}^{m'}(\theta,\phi) \tag{C.2}$$

From the last equation we recognise that we find the applied spherical harmonic back in the far-field. The harmonic is multiplied with a scaling factor. The scaling factor is a spherical Bessel function of order l, which in turn is a function of the ratio $\frac{a}{\lambda}$, where a is the radius of the sphere and λ the signal wavelength.

A discrete formulation was also given, and a scalar sampling theorem was formulated for the equiangular distribution. As we no longer consider to use an equiangular sampling grid as a viable option for spherical array antennas, this is not repeated here.

C.2 Isotropic Radiators Limit the Bandwidth of the Spherical Array

In this thesis it was immediately assumed that a spherical array antenna would be implemented as a set of radiating elements on a PEC sphere. The reason for this is that it was concluded

already from the M.Sc. thesis that directive array elements would improve the operational bandwidth of the array antenna. This is explained in this section.

Equation (C.2) puts a severe limitation on the signal bandwidth that can be used. In order to achieve a far-field with the same variations in azimuth and elevation as the excitation function, we need to choose a ratio of $\frac{a}{\lambda}$ for which all the harmonic terms, present in the excitation function, are attenuated the same. This is hard to achieve over a broad band of wavelengths as can be seen from figure C.2. The problem arises from the oscillatory behaviour of the spherical Bessel function. For an harmonic of a certain order the $\frac{a}{\lambda}$-ratio has to be chosen very carefully to avoid the zeros of the spherical Bessel function. This is true for every desired harmonic component in the far-field, and they all have their zeros for different values of $\frac{a}{\lambda}$. Even without deviations from this operational wavelength, the harmonic components within the excitation function will be attenuated differently in the far-field, causing a disturbed pattern. This can be corrected for, but the disturbance will vary significantly with a deviation from the centre frequency. This is what causes the limitation on the signal bandwidth that can be used.

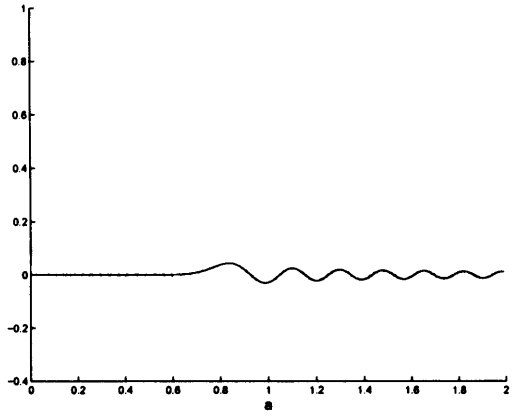


Figure C.2: In the figure the spherical Bessel functions order 0 to 3 are shown in grey, together with the Bessel function of order 32 in black. The functions are plotted versus the radius of the sphere with $\lambda = 0.15\ m$.

Physically, the problem arises from a strong coupling between oppositely placed radiating parts of the spherical array and has been observed with circular arrays as well[145]. The coupling between diametrically positioned elements is strongly dependent on the $\frac{a}{\lambda}$-ratio and the slightest variation from the design-wavelength will change the behaviour of the radiating surface completely. In the derivation of equation (C.2) isotropic radiators were assumed. This suggests that using directional element patterns can overcome this problem.

The effect of having directional elements, above a PEC sphere for instance, was simulated with a raised-cosine pattern. With reference to figure C.1, it was defined as:

$$\left(\frac{1}{2}\right)(1+\cos\psi) = \cos^2\left(\frac{\psi}{2}\right) \tag{C.3}$$

It was found that applying this pattern to the array elements neutralised the effect of different attenuation by the Bessel factors, so that the far-field becomes an unweighted sum of the spherical harmonics in the excitation function. This is shown in figure C.3. Experiments were carried out over wide bandwidths and statistics were gathered to show that a wideband operation could be achieved with, but not without, directive elements. Having a metal sphere in the middle of the array, automatically assures this is the case.

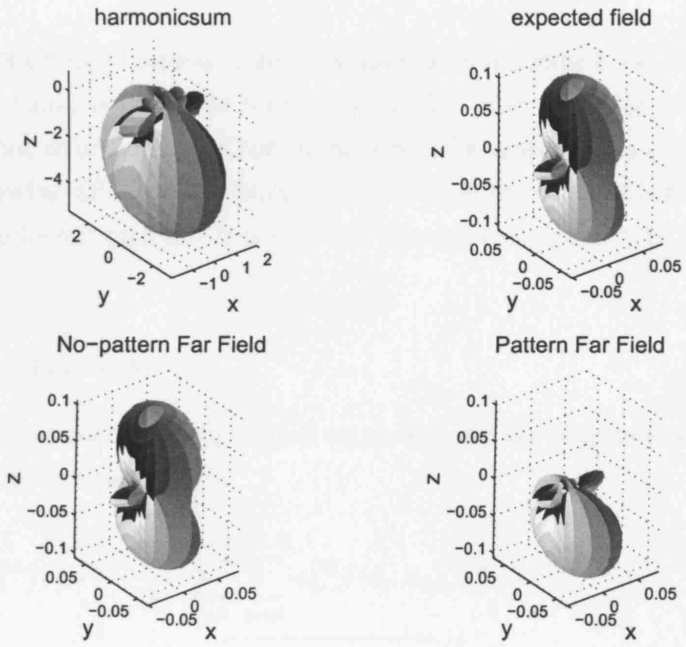


Figure C.3: The figure shows the effect of directional element patterns. The upper-left plot shows an excitation constituted of the sum of harmonics up to order 6. The upper-right plot shows the same sum with the Bessel attenuation factors applied. Bottom-left is a plot of the far-field, calculated with isotropic elements and a sampling grid of bandwidth $b = 32$, with a wavelength of 15 cm and a radius $a = 0.5$ m. The far-field calculation is repeated in the bottom-right plot, but with raised-cosine element patterns. We see the latter far-field is much closer to the original sum of harmonics, indicating that with directive array elements, we get rid of the differences between attenuation levels for different modes.

Appendix D

An Analog SFT Implementation

In this appendix, attention is turned to an early attempt to implement the SFT in an RF stage, just as the Butler matrix is known to implement an FFT. The appendix was copied from my M.Sc. thesis, and the efforts were not continued within the scope of this PhD programme. The development an analog SFT implementation did not fit in the development of LISA and was therefore given the lowest priority. It is included because the idea may be picked up again in future work.

D.1 System overview

We start with the spherical transform defined on an equiangular distribution. This formulation suggest a way for implementation:

$$\hat{f}(l,m) = \langle f, Y_l^m \rangle = \frac{\sqrt{2\pi}}{2b} \sum_{i=0}^{2b-1} \sum_{k=0}^{2b-1} a_i^{(b)} f(\theta_i, \phi_k)\, \bar{Y}_l^m(\theta_i, \phi_k) \tag{D.1}$$

$$\begin{aligned} \text{with} \quad \bar{Y}_l^m &= (-1)^m \sqrt{\frac{(2l+1)(l+m)!}{4\pi(l-m)!}}\, P_l^{-m}(\cos\theta_i)\, e^{-jm\phi_k} \\ &= (-1)^m \sqrt{\frac{(2l+1)(l+m)!}{4\pi(l-m)!}}\, (-1)^m \frac{(l-m)!}{(l+m)!}\, P_l^m(\cos\theta_i)\, e^{-jm\phi_k} \\ &= \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}\, P_l^m(\cos\theta_i)\, e^{-jm\phi_k} \end{aligned}$$

The design goal here is to simplify the processing that needs to be done in the RF-circuitry with minimum overall complexity. The transforming component is now implemented before the signals are down-converted and digitised. Therefore the design goal can be rephrased as: "minimise the RF-processing by doing as much as possible of the processing in the digital circuitry, without augmenting the number of down-convertors and A/D-convertors."

The weighting factors $a_i^{(b)}$ take into account the non-uniformity of the equiangular distribution. If we have a preprocessing stage that converts a more uniform distribution into an equiangular one, than these factors will more or less cancel out with the interpolation process. Anyway, we will assume the weighting factors are incorporated in a preprocessing stage. The multiplication with the normalising constant can evenly be done after digitisation. We can now rewrite

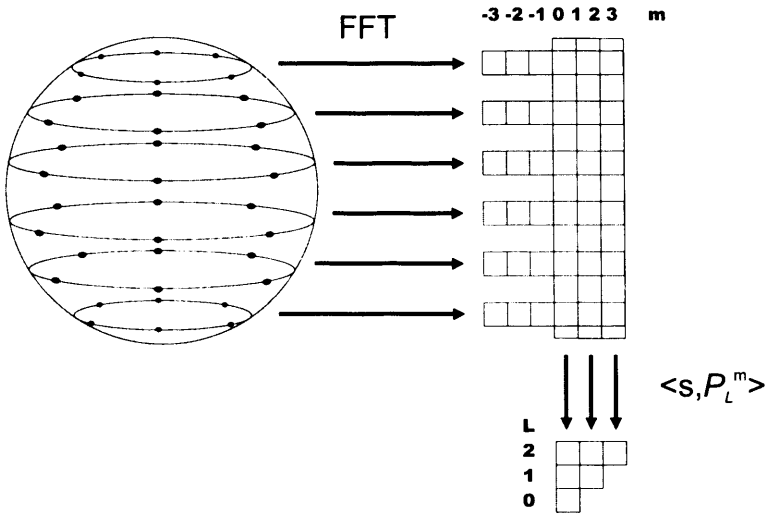


Figure D.1: The figure shows schematically the two stages in which a spherical Fourier transform is obtained. First, a bank of Butler matrices acts on the signals originating from azimuthal circles on the sphere. This stage is the projection of the incoming signal vectors onto the circular Fourier basis functions. The output vectors are recombined as indicated and fed into the second stage. The second stage projects its incoming signal vectors onto the associated Legendre functions. This Legendre transform can further be simplified as explained in the text.

equation D.1 and split it into two stages of processing:

$$\hat{f}(l,m) = \sum_{i=0}^{2b-1}\sum_{k=0}^{2b-1} f(\theta_i,\phi_k)\, P_l^m(\cos\theta_i)\, e^{-jm\phi_k} \qquad 0 \leqslant |m| \leqslant l < b$$

$$s[i] = \sum_{k=0}^{2b-1} f(\theta_i,\phi_k)\, e^{-jm\phi_k} \qquad m = -b+1,...,b-1 \tag{D.2}$$

$$\hat{f}(l,m) = \langle s, P_l^m \rangle = \sum_{i=0}^{2b-1} s[i]\, P_l^m(\cos\theta_i) \qquad l = m, m+1,...,b-1 \tag{D.3}$$

The first stage of processing, equation D.2, is the circular Fourier transform and can be processed with the FFT-algorithm. An implementation for RF-signals is well known and called a *Butler matrix.* We now have broken down the processing into a stage that consists of a bank of Butler matrices (one for each θ_i) and a second stage that has to perform a projection of its input vectors onto the associated Legendre polynomials. An overview of this system is given in figure D.1. To implement the second stage a fast algorithm for the Legendre transform has been developed in [54] and [55]. As we will see, the details of the algorithm became irrelevant in this text, but the idea is a divide and conquer strategy by making use of the recurrence property of the Legendre functions. With some minor modifications this algorithm has been implemented as a SIMULINK network. As the numerical algorithm is the best currently known in the sense of computational complexity to calculate a spherical Fourier transform, it was thought that it would also result in the least complex system when implemented as a circuit. Although the results obtained from the simulations with the resulting network were excellent, it was still too complicated to implement with RF-components. Another sub-optimal algorithm was therefore

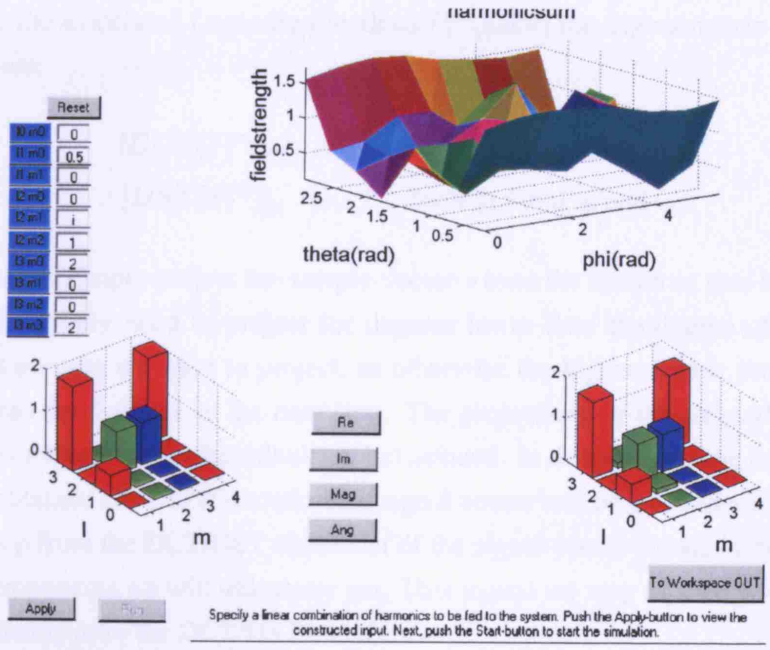


Figure D.2: The figure shows the interface from a SIMULINK simulation. A linear (complex) combination of harmonics can be applied as input to the network. The input is defined on a 8×8 equiangular grid, so no more than 16 spherical harmonics can be resolved without ambiguity. The 6 harmonics for which $m < 0$ are not included. The input constructed from the harmonic combination on the left is shown in the uppermost plot. The output is a set of signals, each of which is the response of the system to a spherical harmonic component within the applied input. The output of the network is shown on the right below.

implemented as well as a SIMULINK network, resulting already in a much simpler system. The reason that this can be true, is that the number of inputs and outputs of the algorithm to implement is not too large, while the first algorithm is shown to be optimal in the limit of a large number of harmonics to be calculated. There is another reason why the second implementation will be the first choice for an RF-application. This will be explained in the next paragraph. In figure D.2 the interface of the simulation using the second SIMULINK network is shown. The simulation showed that the algorithm is indeed capable of finding the correct spherical harmonic expansion in real-time. The reason why the second implementation lends itself very well for an implementation with RF-components, is that the design goal mentioned earlier is optimally met. Suppose we want to resolve the harmonic components up to some degree L. Than L is of course also the minimum number of controls and therefore the absolute minimum number of convertors[1] that any system will have. In the presented implementation, the minimum amount of processing will be done in the analog processing stage, transferring the computations mostly to the digital circuitry. This will be achieved with only the minimum number of L convertors. To achieve this, a clever trick is used, based on the orthogonality of the cosine and the sine base:

$$\begin{aligned} \langle s, P_l^m \rangle &= \langle DCT(s), DCT(P_l^m) \rangle \\ \langle s, P_l^m \rangle &= \langle DST(s), DST(P_l^m) \rangle \end{aligned}$$

[1] To be entirely exact, we would have $2L$ convertors for complex signals.

furthermore, as the associated Legendre functions $P_l^m(\cos\theta)$ are trigonometric polynomials of degree l, we have:

$$\begin{aligned} [DCT(P_l^m)]_k &= 0 \quad \text{for } k \geqslant l \text{ and m even} \\ [DST(P_l^m)]_k &= 0 \quad \text{for } k \geqslant l \text{ and m odd} \end{aligned}$$

This means, we can simply project the sample-vector s onto the cosine or sine base, depending on m. Hereby we only need to project for degrees lower than the degree of the associated Legendre function onto we want to project, as otherwise the higher degree projections would be projected onto the 0-vector in the next step. The projections of the Legendre polynomials onto the cosine or sine base can be calculated beforehand. In the next step the inner products of them with the obtained DCT/DST-transformed signal vector will be calculated. But the number of terms we keep from the DCT/DST transform of the signal vector equals already the number of harmonic components we will ultimately get. That means we may as well down-convert and digitise immediately after the DCT/DST transforms.

The whole processing is carried out in three steps: a bank of butler matrices, a bank of DCT/DST transforms, and digital processing. The latter can implement a more sophisticated algorithm or just an inner product with prestored vectors, in fact that is of no concern anymore. The complexity of the calculations we can implement digitally is no limiting factor. That is why the details of the algorithm to compute the spherical equivalent of an FFT have been omitted from this section.

D.2 Performing the DCT/DST transform

The principle of computing a discrete cosine transform over a set of incoming signals is very simple. We will start off with a definition:

The discrete cosine transform $s[k] = DCT\,(x[n])$ is defined as:

$$\begin{aligned} s[k] &= \omega(k) \sum_{n=0}^{N-1} x[n] \cos\left(\frac{\pi(2n+1)k}{2N}\right) \qquad k = 0, ..., N-1 \\ \omega(k) &= \begin{cases} \frac{1}{\sqrt{N}} & k = 0 \\ \sqrt{\frac{2}{N}} & 1 \leqslant k \leqslant N-1 \end{cases} \end{aligned}$$

Hence, for each output signal, we have to sum the incoming signals after we scaled them by some cosine factor $\cos\alpha$. The cosine factor can conveniently be rewritten as $\cos\alpha = \frac{1}{2}(e^{\alpha j} + e^{-\alpha j})$. This is nothing else than two counterbalanced phase shifts of the incoming signals. We can implement this with a variant of the Blass-matrix (which is some sort of beamforming network, [2][9]), but adapted to provide both a positive and a negative phase delay from the reference. This is illustrated in figure D.3. For the discrete sine transform, the principle is only slightly different and depicted in the figure as well. The next step is to translate the principle of operation into a real working design. The designing task was done with the Advanced Design System (ADS) software from Agilent. For our application a centre frequency of 3 GHz was

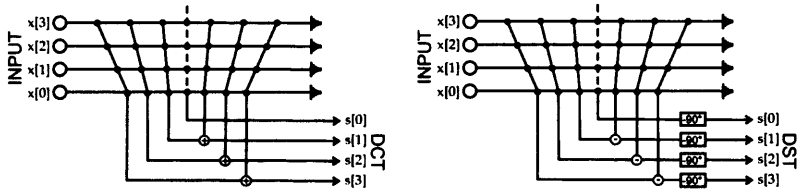


Figure D.3: Principle of a circuit that performs a DCT transform (left) and a DST transform (right). The connecting black dots will be implemented with power dividers. One signal line is replaced with a dashed line to indicate the phase reference.

chosen. For a design at this frequency, it makes sense to opt for a design based on microstrip technology. The design of the circuit depicted in the figure with microstrip technology is not as simple as it may seem and several difficulties presented themselves.

The design goal is to achieve a high degree of accuracy over a wide bandwidth. The black dots, call them nodes, in the schematic (figure D.3) represent power dividers. That means that, at every node, part of the signal power is taken from the feeding line. To achieve an accurate system, every power divider on a feeding line should take out an equal amount of power. As the signal gets weaker down the line, a greater portion has to be taken out. The last node is in fact the leftover, while the before-last node is a 3dB splitter, etc. Therefore the power dividers had to be carefully designed and each was separately optimised. For large coupling ratios, Wilkinson dividers are more appropriate, while for small coupling ratios coupled lines would be used. To simplify the design a small number of outputs was chosen ($N = 3$), so that all the dividers could be implemented with Wilkinson dividers.

From the schematic it seems that every divider is immediately connected to the node underneath. This would assume a four-port, with an ideal characteristic that can mathematically be shown to be impossible to achieve (from the S-matrix). One could imagine that the signal coming from a higher node combines its power with the signal taken out by the lower node and so on. Now the power on the line grows larger while moving towards the output. Furthermore, multiple reflections would occur between the nodes with imperfect impedance matches. It is clear that this would not be a viable approach. Instead, every signal originating from a node on a feeding line, comes out separately and all the appropriate signals are combined afterwards. This is shown in figure D.4. The problem with this circuit is the large number of cross-overs. The circuit cannot be implemented with an entirely planar structure and therefore substantial effort has been spent on designing a low-loss via-structure through a multi-layered board with signal layers on each side and a ground-plane between them, but the design never met acceptable standards.

Another stringent requirement in the design of the component is to keep the accuracy of the relative phase differences between the signals over a wide bandwidth. Therefore all the path-lengths from one input to the point where the signals get combined back together need to be approximately the same (within one or two wavelengths). This was easily achieved by adding some multiple of a wavelength to the path-lengths where needed.

To calculate the DCT different sums have to be calculated over the input signals, multiplied

with cosine factors. The cosine factors were substituted with sums of two phase-shifted versions of each signal. The mathematical computations are implemented as operations on voltage levels. In a system that would perform its operations in direct correspondence with the mathematical computations, no power dividers would be used for the nodes. Instead the nodes would be voltage probes with an output voltage that follows the voltage level on the feeding line. Such probes need to be very high input-impedance amplifiers with a negative feedback loop. In their simplest form, these so-called buffers are implemented as emitter-followers with one transistor. The signals are then combined back together with a device to add their voltages. The circuit would have become a complex active system, while this is unnecessary. That an active circuit is unnecessary is indicated by the orthonormality of the DCT-matrix. It should be possible to design the component with a power-conserving circuit. That is why power-dividers were used.

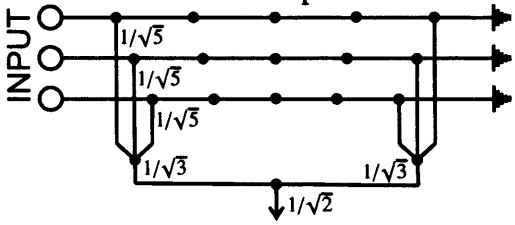


Figure D.4: The power dividers and combiners introduce an attenuation on the voltage amplitudes. Due to the symmetry of the system, the total attenuation is the same for every path and the attenuation can be corrected for by adjusting the normalisation factors at the output.

By implementing the nodes with power dividers, the total power may be preserved, but for N outputs, the amplitude of the voltage signal coming from each node will be a factor $(\sqrt{2N-1})^{-1}$ of the input voltage amplitude. The original voltage level could be restored by changing the impedance but, as this factor is the same for all signals, 50 Ω-lines can be used, as long as the constant attenuation is corrected for afterwards. The power combiners introduce the same kind of amplitude attenuation, as shown in figure D.4. For every possible path from input to output, the total attenuation (in voltage, not in power) is the same, except for the signals originating from the phase reference who lack the last power combiner. These signals form the DC-component at the output. As explained in section D.1, the DCT-component is the last one before the signals are digitised. The adjustment of the normalisation factors at the output can therefore be done after digitisation.

The power combiners are also Wilkinsons. They are introduced into the design as theoretical subcomponents. It is not possible to design an N-way Wilkinson with a planar structure if $N \geqslant 3$. For this reason the design drawing in figure D.6 does not include the power combiners. In the schematic (figure D.5) it is seen that the outputs are formed by summing the appropriate signals together. It might be a good idea to design the whole component as a 3-dimensional structure in the future.

A planar microstrip structure has been designed to test whether it would be feasible to implement the main part of the processing, the DCT transform, in an RF system. Three feeding lines have been designed, as shown in in the layout in figure D.6. Each of these lines has four Wilkinson power dividers, splitting the power in five equal parts. The Wilkinsons have an extra $\frac{\lambda}{4}$-section at each output to gain accuracy over a wider bandwidth. The design formula's of the Wilkinsons are given in table D.1. The actual design values deviate slightly from the theoretical values as every divider has been separately optimised with an electromagnetic moment optimisation program, included in ADS. The resistors from table D.1 are not shown in the layout

$\frac{P_3}{P_2} = k^2$	Z_{c2} $Z_0\sqrt{(k^2+1)k}$	Z_{c3} $Z_0\sqrt{\frac{k^2+1}{k^3}}$	Z_{c4} $Z_0\sqrt{k}$	Z_{c5} $Z_0\sqrt{\frac{1}{k}}$	R $Z_0\frac{k^2+1}{k}$	s_{12}/s_{13}
1/4	39.529	158.114	35.355	70.711	125	0.894/0.447
1/3	43.869	131.607	37.992	65.804	115.5	0.866/0.5
1/2	51.494	102.988	42.045	59.46	106	0.816/0.577
1	70.71	70.71	50	50	100	0.707/0.707

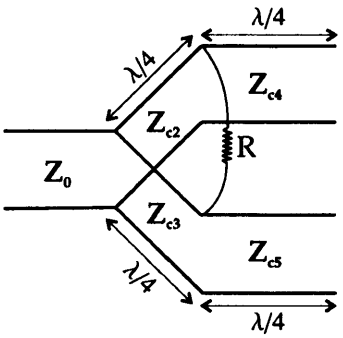


Table D.1: The table specifies the design-values for the impedances indicated in the schematic and the value of the resistor. This is done for each of the 4 power dividers needed to split the input power in five equal parts. In the rightmost column the design goal for each power divider is expressed with two s-parameters. The third was to minimise reflection.

figure. Their function is to dissipate uneven mode excitations flowing from right to left in the figure. As we are interested in the s-parameters from left to right and each of the three lines is designed for minimal reflections, the resistors are not of any importance. It is essential that the relative phase differences between the outputs of each line remain constant over the signal bandwidth. Therefore 50 Ω-line sections, an integer number of wavelengths long, have been added to make the path-lengths approximately equal.

The substrate is chosen to be a standard RT-Duroid product with a thickness of 1.5 mm and an electrical constant of $\epsilon_r = 2.213$. Metal losses were not included in the simulations. In the layout in figure D.6, the three feeding lines are shown. The relative phase differences between the outputs from a line are given in table D.2. The line lengths have to be corrected with a 90° phase shift to account for the phase shifts in the Wilkinsons. Figures D.7 and D.7 and table D.3 summarise the results obtained from simulations with the designed circuit. Graphs are shown from simulations on the *s-parameters*. The parameter s_{ij} gives the output voltage amplitude at port j, normalised to the input at port i, both normalised to an impedance of 50 Ω and with all other ports terminated with their characteristic impedance. Hence we can interpret these parameters as the calculated DCT of a 1 at input i and zero at all the others. Different input vectors could be (redundantly) analysed with linear combinations of the s-parameters. For this first approach the results are satisfactory and show the principle works. As mentioned before, the total circuit can not be implemented as an entirely planar microstrip structure. However, the results obtained here are stimulating to continue the search for the wide-band, RF, DCT transformer. Especially for a larger number of inputs/outputs, such a design will continue to be a challenge. When using this component within the framework of a spherical Fourier transform, one can profit from the advantage that fewer outputs are needed as not the whole DCT needs to be calculated, but only some lower order terms.

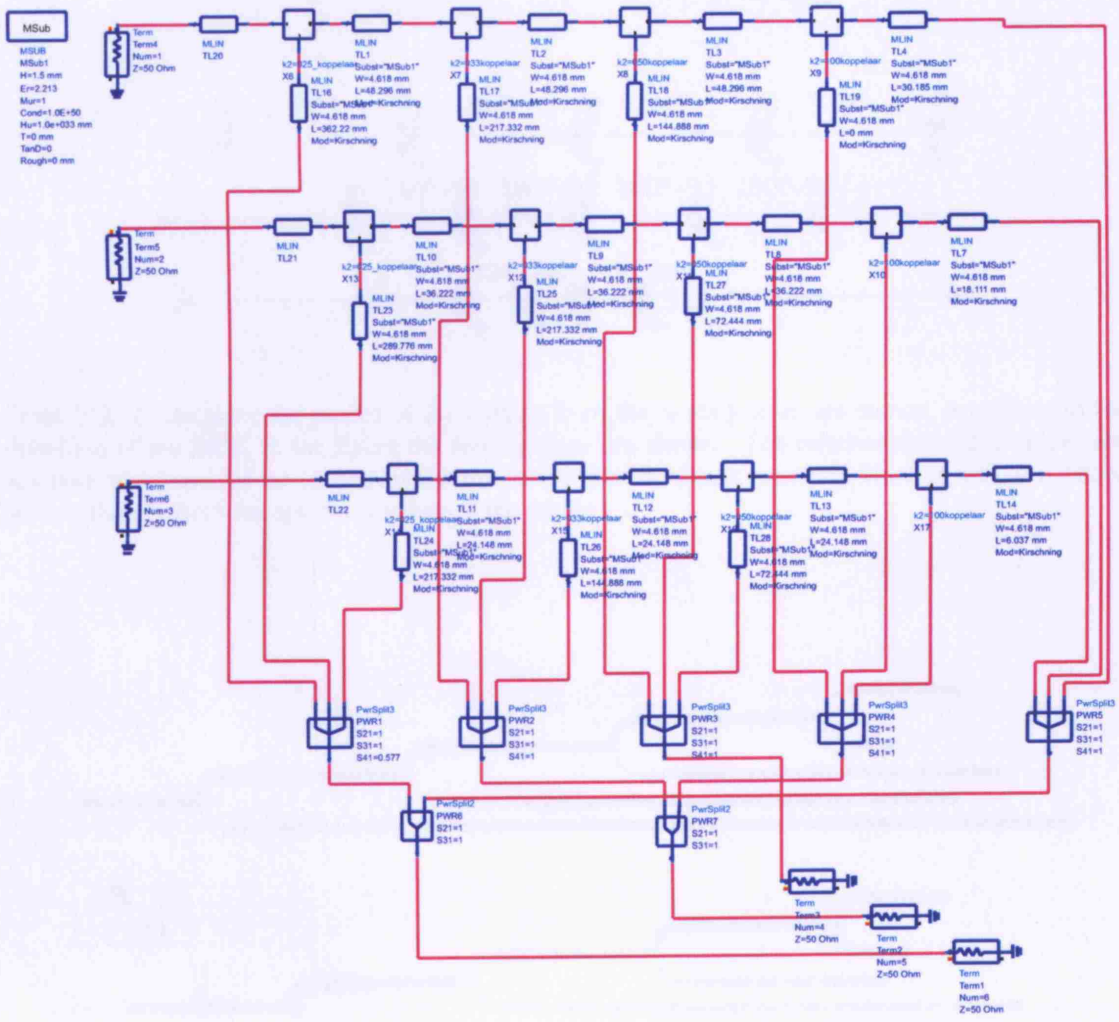


Figure D.5: This schematic was used for simulation. It shows 3 feeding lines with their dividers and the outputs at the bottom. The combiners at the bottom are theoretical summations. The feeding lines have been simulated separately, both with this model-based schematic and with EM moment simulations. The results from the latter were in perfect agreement with the former.

$\frac{(2n+1)k}{2N}\pi$	$\frac{k\pi}{6}$	$\frac{3k\pi}{6}$	$\frac{5k\pi}{6}$
$N=2$	$n=0$	$n=1$	$n=2$
$k=0$	0	0	0
$k=1$	$\frac{\pi}{6}$	$\frac{\pi}{2}$	$\frac{5\pi}{6}$
$k=2$	$\frac{\pi}{3}$	π	$\frac{5\pi}{3}$

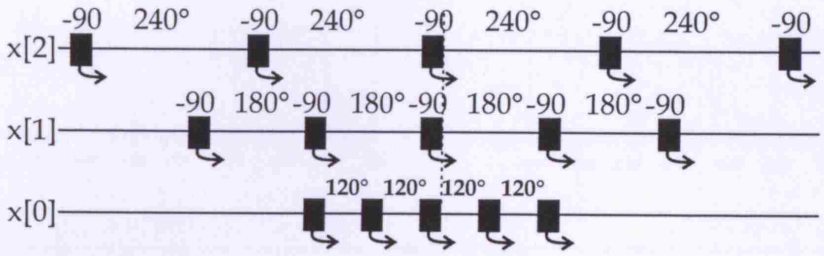


Table D.2: In the table the phases of the outputs from the feeding lines are shown, according to the definition of the DCT. In the figure the feeding lines are shown. The relative phase differences are adjusted to account for the phase shifts of the power dividers (black squares). Here as well, the values deviate slightly from the optimised values in the design.

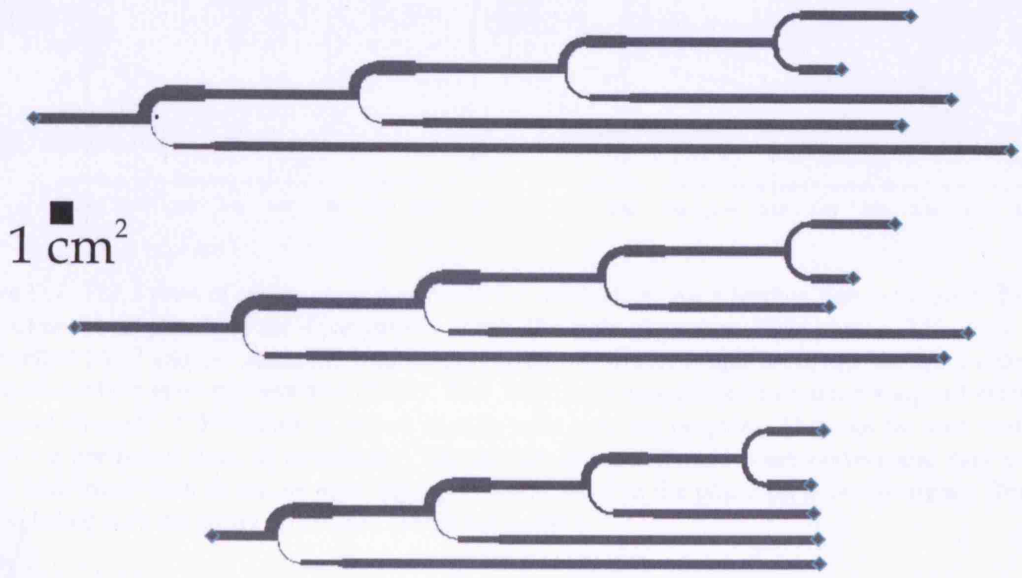


Figure D.6: This microstrip layout shows the 3 feeding lines from the schematic in figure D.5. A square cm. is shown as a reference scale. The substrate has a thickness of 1.5 mm. With this circuit already satisfying results were obtained, however some amelioration could still augment the performance. It is more likely that future designing hours can better be spend on a complete 3D-approach than on refining the performance of this microstrip layout.

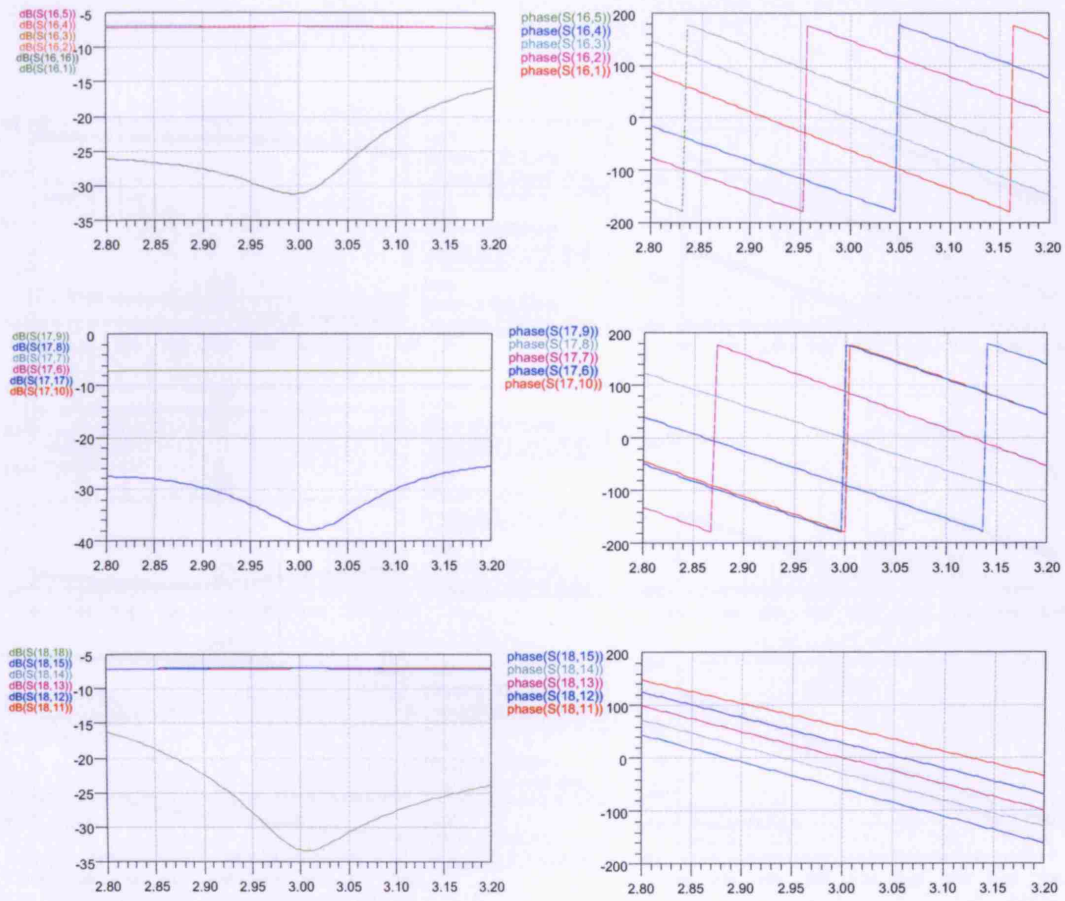


Figure D.7: The 3 rows of graphs show the simulation results from each feeding line separately. The left plots show the amplitude of the s-parameters in dB. The right plots show their phases. The input ports are labelled 16,17 and 18. From the s-parameters S(16,16), S(17,17) and S(18,18), we see a reflection on the lines which is mostly less then -20 dB. Also, over the whole plotted frequency range of 400MHz, the signal strength of the inputs is indeed equally split over the outputs. This can be seen from the coinciding horizontal lines of amplitude. The relative phase differences are correct and stay correct over a wide bandwidth as can be seen from the parallel lines in the phase plots on the right. This was accomplished with the added 50 Ω-lines before the outputs.

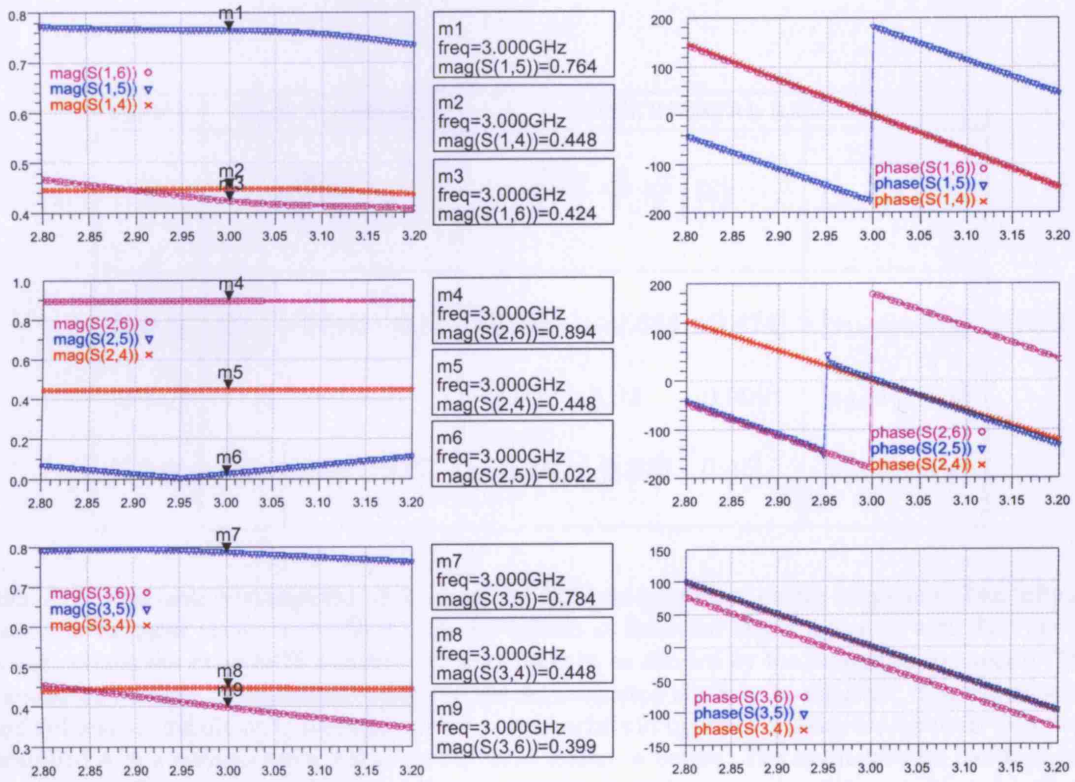


Figure D.8: The graphs show the s-parameters obtained from simulating the complete circuit shown in figure D.5. The input ports are now labelled 1,2 and 3 and the outputs 4,5 and 6. Ideally the lines in the amplitude plots on the left would be perfectly horizontal to keep high accuracy over a wide bandwidth. We see that some are indeed quite horizontal. The markers give the exact values at design frequency, which are evaluated in table D.3.

INPUT	EXACT VALUES	EXPERIMENTAL RESULTS	
x	$\frac{\mathrm{DCT}(x)}{\left[\frac{1}{\sqrt{3}}\ \frac{2}{\sqrt{3}}\ \frac{2}{\sqrt{3}}\right]}$	$\left[\sqrt{5}\ \frac{\sqrt{5}}{2}\ \frac{\sqrt{5}}{2}\right]$ [S]	S
[0 0 1]	[1 −0.866 0.5]	[1.002 −0.854 0.474]	$[s_{14}\ s_{15}\ s_{16}]$
[0 1 0]	[1 0 −1]	[1.002 0.02 −0.999]	$[s_{24}\ s_{25}\ s_{26}]$
[1 0 0]	[1 0.866 0.5]	[1.002 0.876 0.45]	$[s_{34}\ s_{35}\ s_{36}]$

Table D.3: This table evaluates the accuracy of the DCT calculation at centre frequency. The leftmost column is the input vector to the DCT with the indices as indicated in the figure of table D.2. In the second column the exact DCT transform is given, pointwise divided by the normalisation vector. This is necessary because, as explained in the text, the normalisation needs to be adjusted. The values in the third column are the obtained s-parameters from the markers in figure D.8. They are however pointwise multiplied with a normalisation vector for the same reason as before. The normalisation includes only the $\sqrt{5}$-factor from the power dividers and not the factors from the power combiners, indicated in figure D.4, because the theoretical components used in the schematic to recombine the signals do not attenuate the voltages. All values calculated by the circuit are within a few percent of the ones who should be obtained theoretically.

Bibliography

[1] L. Josefsson and P. Persson, *Conformal Array Antenna Theory and Design.* IEEE Press Series on Electromagnetic Wave Theory, Wiley-Interscience, 2006.

[2] A. Rudge, K. Milne, A. Olver, and P. Knight, eds., *The handbook of antenna design*, vol. 2 of *IEE Electromagnetic waves series.* Peter Peregrinus Ltd / IEE, 1983.

[3] H. Chireix, "Antennes à rayonnement zénithal réduit," *L'Onde Electrique*, vol. 15, pp. 440–456, 1936.

[4] D. Davies, "A transformation between the phasing techniques required for linear and circular aerial arrays," *Proc. IEE*, vol. 112, no. 11, pp. 2041–2045, 1965.

[5] IEE, *Broadband nulls from a circular array*, April 1985.

[6] IEE, *Broadband pattern synthesis from a circular array*, April 1989.

[7] H. Griffiths and R. Eiges, "Sectoral phase modes from circular antenna arrays," *Electronics Letters*, vol. 28, no. 17, pp. 1581–1582, 1992.

[8] R. Eiges and H. Griffiths, "Mode-space spatial spectral estimation for circular arrays," *IEE Proc. Radar, Sonar Navigation*, vol. 141, no. 6, pp. 300–306, 1994.

[9] J. Drabowitch, A. Papiernik, H. Griffiths, J. Encinas, and B. Smith, *Modern Antennas.* Springer, 2 ed., 1998.

[10] G. H. Knittel, "Choosing the number of faces of a phased-array antenna for hemisphere scan coverage," *IEEE Transactions on antennas and propagation*, vol. 13, pp. 878–882, November 1965.

[11] H. E. Schrank, "Basic theoretical aspects of spherical phased arrays," in *Phased Array Antennas* (A. A. Oliner and G. H. Knittel, eds.), pp. 323–327, Artech —House, Inc., June 1970.

[12] D. L. Sengupta, T. M. smith, and R. W. Larson, "Radiation characteristics of a spherical array of circularly polarized elements," *IEEE Transactions on antennas and propagation*, vol. 16, January 1968.

[13] J. Juntunen, "Radiation properties of spherical arrays," in *European Microwave Conference, 4th*, September 1974.

[14] S. Horiguchi, T. Ishizone, and Y. Mushiake, "Radiation characteristics of spherical triangular array antenna," *IEEE Transactions on antennas and propagation*, vol. 33, no. 4, pp. 472–476, 1985.

[15] A. K. Chan, A. Ishimaru, and R. A. Sigelman, "Equally spaced spherical arrays," *Radio Science*, vol. 3, no. 5, pp. 401–404, 1968.

[16] M. Hoffman, "Conventions for the analysis of spherical arrays," *IEEE Transactions on antennas and propagation*, pp. 390–393, 1963.

[17] B. P. Kumar and G. R. Branner, "The far-field of a spherical array of point dipoles," *IEEE Transactions on antennas and propagation*, vol. 42, April 1994.

[18] R. MacPhie, "The element density of a spherical antenna array," *IEEE Transactions on antennas and propagation*, January 1968.

[19] R. Stockton and R. Hockensmith, "Application of spherical arrays - a simple approach," in *Antennas and propagation Society International Symposium*, pp. 202–205, June 1977.

[20] J. Verhaevert, emmanuel Van lil, and A. V. de Capelle, "Three-dimensional monopole antenna for direction of arrival determination," *IEE Proceedings Microwave, Antennas and Propagation*, vol. 151, no. 2, pp. 121– 126, 2004.

[21] Z. Sipus and S. Skocic, "Application of vilenkin's addition theorem in the analysis of spherical arrays and periodic structures," in *Proceedings of The European Conference on Antennas and Propagation*, November 2006.

[22] N. Burum, S. Rupcic, and Z. Sipus, "Theoretical and experimental study of spherical arrays," *IEEE MELECON*, May 2004.

[23] Z. Sipus, N. Burum, and J. Bartolic, "Moment method analysis of rectangular microstrip antennas on spherical structures," in *IEEE Antennas and Propagation Society Internationsl Symposium*, vol. 3A, pp. 126–129, July 2005.

[24] P. Knott, "Design of triple patch antenna element for double curved conformal antenna arrays," in *Proceedings of The European Conference on Antennas and Propagation*, November 2006.

[25] M. Huang and S. Tan, "Spherical antenna arrays for phase mode processing," in *Proc. IEEE International Symposium on Micrwave, Antenna, Propagation and EMC Technologies for Wireless Communications*, 2005.

[26] M. Huang and S. Tan, "An improved spherical antenna array for wideband phase mode processing," *Progress In Electromagnetics Research*, vol. 66, pp. 27–40, 2006.

[27] M. Huang and S. Tan, "Spheroidal phase mode processing for antenna arrays," in *IEEE International workshop on Antenna Technology: Small Antennas and Novel Metamaterials*, March 2005.

[28] E. D. Witte, H. Griffiths, and P. Brennan, "Phase mode processing for spherical antenna arrays," *Electronics Letters*, vol. 39, October 2003.

[29] B. Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Transactions on Speech and Audio processing*, vol. 13, no. 1, 2005.

[30] E. D. Witte, "Spherical array antennas for leo satellite communications," thesis, University College London, Department of Electronic & Electrical Engineering, 2003.

[31] P. J. Oleski, "Space ground link subsystem," web IF-03-11, Air Force Research Laboratory, April 2004.

[32] S. Liu, B. Tomasic, and J. Turtle, "The geodesic sphere phased-array antenna for satellite operation and air/space surveillance," in *Proc. 2001 Antenna Applications Symposium*, September 2001.

[33] B. Tomasic, J. Turtle, S. Liu, R. Schmier, S. Bharj, and P. Oleski, "The geodesic dome phased array antenna for satellite control and communications - subarray design, development and demonstration," in *IEEE International Symposium on Phased Array Systems and Technology*, October 2003.

[34] B. Tomasic, J. Turtle, and S. Liu, "Spherical arrays - design considerations," in *18th International Conference on Applied Electromagnetics and Communications*, pp. 1–8, October 2005.

[35] J. Thompson, "Phased array antenna for satellite communications to small tactical platforms," in *IEE Colloquium on Military Satellite Communications*, 2000.

[36] D. Brandwood, "Phase steering for modified dodecahedral array for mobile satcoms," in *IEE Colloquium on Electronic Beam Steering*, 1998.

[37] I. Craddock, A. Preece, J. Leendertz, M. Klemm, R. Benjamin, and R. Nilavalan, "Development of a hemi-spherical wideband antenna array for breast cancer imaging," in *Proc. European Conference on Antennas and Propagation*, 2006.

[38] M. Scott, "Sampson mfr active phased array antenna," in *IEEE International Symposium on Phased Array Systems and Technology*, pp. 119–123, October 2003.

[39] V. Jamnejad, J. Huang, B. Levitt, T. Pham, and R. Cesarone, "Array antennas for jpl/nasa deep space network," in *IEEE Aerospace Conference Proceedings*, vol. 2, pp. 911–921, 2002.

[40] B. Rafaely and M. Park, "Super-eesolution spherical microphone arrays," in *Proceedings of Convention of Electrical and Electronics Engineers in Israel*, pp. 424–427, September 2004.

[41] B. Rafaely, "Phase-mode versus delay-and-sum spherical microphone array processing," *IEEE Signal Processing Letters*, vol. 12, October 2005.

[42] E. D. Witte, P. Brennan, and H. Griffiths, "Mutual coupling analysis of a spherical array antenna," in *Antennas, Radar, and Wave Propagation*, IASTED, pp. 324–329, July 2005.

[43] E. D. Witte, P. Brennan, and H. Griffiths, "Spherical wave expansion for a circular aperture in a large conducting sphere," in *Proceedings Loughborough Antennas & Propagation Conference*, pp. 344–347, April 2005.

[44] E. D. Witte, P. Brennan, and H. Griffiths, "Sampling and interpolation on the sphere," in *Proceedings Loughborough Antennas & Propagation Conference*, pp. 121–124, April 2006.

[45] E. D. Witte, L. Marantis, C. Rizzo, P. Brennan, and H. Griffiths, "A spherical near-field system for testing a spherical array of patch antennas," in *AMTA Europe Symposium*, May 2006.

[46] E. D. Witte, L. Marantis, K.-F. Tong, P. Brennan, and H. Griffiths, "Design and development of a spherical array antenna," in *Proceedings of The European Conference on Antennas and Propagation*, November 2006.

[47] L. Marantis, E. D. Witte, K. Tong, and P. Brennan, "Circularly polarized array of ring antennas using novel proximity-coupled feeds," in *Proceedings Loughborough Antennas & Propagation Conference*, April 2007.

[48] L. Marantis, E. D. Witte, and P. Brennan, "Spherical near-field measurements of a spherical array antenna," in *Antennas, Radar, and Wave Propagation*, IASTED, May 2007.

[49] P. Debye, "Der lichtdruck auf kugeln von beliebigem material," *Annalen der Physik*, vol. 30, pp. 57–136, 1909.

[50] G. Mie, "Beitrage zur optik truber medien, spexiell kolloidaler metallosungen," *Annalen der Physik*, vol. 25, pp. 377–445, 1908.

[51] W. Hansen, "A new type of expansion in radiation problems," *Physical review*, vol. 47, pp. 139–143, January 15 1935.

[52] W. Hansen, "Transformations useful in certain antenna calculations," *Journal of Applied Physics*, vol. 8, pp. 282–286, April 1937.

[53] C. Bouwkamp and H. Casimir, "On multipole expansions in the theory of electromagnetic radiation," *Physica*, vol. 20, no. 1-6, pp. 539–554, 1954.

[54] J. R. Driscoll and J. Dennis M. Healy, "Computing fourier transforms and convolutions on the 2-sphere," *Adv. Appl. Math.*, vol. 15, no. 2, pp. 202–250, 1994.

[55] D. M. H. Jr., D. Rockmore, P. Kostelec, and S. B. Moore, "Ffts for the 2-sphere - improvements and variations," tech. rep., Dartmouth College, Hanover, 1998.

[56] P. J. Kostelec and D. N. Rockmore, "Ffts on the rotation group," Tech. Rep. 03-11-060, Santa Fe Institute, 2003.

[57] G. A. Dilts, "Computation of spherical harmonic expansion coefficients via fft's," *Journal of Computational Physics*, vol. 57, pp. 439–453, 1985.

[58] N. Marcuvitz, *Waveguide Handbook*. McGraw-Hill Book Company, 1st ed., 1951.

[59] J. Stratton, *Electromagnetic Theory*. International Series in Physics, McGraw-Hill Book Company Inc., 1 ed., 1941.

[60] J. D. Jackson, *Classical Electrodynamics*. John Wiley & Sons, Inc., 3 ed., 1998.

[61] A. Rudge, K. Milne, A. Olver, and P. Knight, eds., *The handbook of antenna design*. No. 16 in IEE Electromagnetic waves series, Peter Peregrinus Ltd / IEE, 1983.

[62] J. Walter T. Grandy, *Scattering of Waves from Large Spheres*. Cambridge University Press, 2000.

[63] J. Hansen, *Spherical Near-Field Antenna Measurements*, vol. 26 of *IEE Electromagnetic Waves Series*. Peter Peregrinus Ltd. / IEE, 1988.

[64] R. Mailloux, *Phased array antenna handbook*. Artech House, 1994.

[65] H. Dachsel, "Fast and accurate determination of the wigner rotation matrices in the fast multipole method," *Journal of Chemical Physics*, vol. 124, no. 14, 2006.

[66] S. Trapani and J. Navaza, "Calculation of spherical harmonics and wigner d functions by fft. applications to fast rotational matching in molecular replacement and implementation into amore," *Acta Crystallographica*, vol. 62, pp. 262–269, 2006.

[67] M. A. Blanco, M. Florez, and M. Bermejo, "Evaluation of the rotation matrices in the basis of real spherical harmonics," *Journal of Molecular Structure*, vol. 419, pp. 19–27, 1997.

[68] C. H. Choi, J. Ivanic, M. S. Gordon, and K. Ruedenberg, "Rapid and stable determination of rotation matrices between spherical harmonics by direct recursion," *Journal of Chemical Physics*, vol. 111, pp. 8825–8831, November 1999.

[69] R. Alastair, *Quantum Mechanics*. IOP Publishing, 4th ed., 2002.

[70] P. Kostelec, D. K. Maslen, D. N. Rockmore, and J. Dennis Healy, "Computational harmonic analysis for tensor fields on the two-sphere." AFOSR contract DOD F4960-93-1-0567 NSF DMS Award 9404275.

[71] A. C. Ludwig, "The definition of cross polarisation," *IEEE Transactions on antennas and propagation*, January 1973.

[72] J. R. driscoll and J. Dennis M. Healy, "Asymptotically fast algorithms for spherical and related transforms," in *30th annual symposium on Foundations of Computer Science*, pp. 344–349, 1989.

[73] N. Sneeuw and R. Bun, "Global spherical harmonic computation by two-dimensional fourier methods," *Journal of Geodesy*, vol. 70, no. 4, pp. 224–242, 1996.

[74] L. Bailin and S. Silver, "Exterior electromagnetic boundary value problems for spheres and cones," *IRE Transactions on Antennas and Propagation*, pp. 9–16, January 1956.

[75] C. Chen, "Some remarks on exterior electromagnetic boundary value problems for spheres," *IEEE Transactions on antennas and propagation*, vol. 18, pp. 705–707, September 1970.

[76] K.-L. Wong, *Design of Nonplanar Microstrip Antennas and Transmission Lines*. John Wiley & Sons, Inc., 1999.

[77] R. Mittra, ed., *Computer Techniques For Electromagnetics*. Hemisphere Publishing Corporation, 1987.

[78] A. A. Kishk, "Analysis of spherical annular microstrip antennas," *IEEE Transactions on antennas and propagation*, vol. 41, no. 3, pp. 338–343, 1993.

[79] W. Tam, A. Lai, and K. Luk, "Annular ring microstrip antenna on a spherical body," *Microwave and Optical Technology Letters*, vol. 7, pp. 371–375, June 1994.

[80] K. Luk and W. Tam, "Patch antennas on a spherical body," *IEE Proceedings-H*, vol. 138, pp. 103–108, February 1991.

[81] W. Tam, "Resonance in spherical-circular microstrip structures," *IEEE Trans. Microwave Theory Tech.*, vol. 39, pp. 700–704, April 1991.

[82] C. A. Balanis, *Antenna Theory Analysis and Design*. John Wiley & Sons, Inc., 2 ed., 2005.

[83] D. Jackson, J. Williams, A. Bhattacharyya, R. Smith, S. Buchheit, and S. Long, "Microstrip patch designs that do not excite surface waves," *IEE Transactions on Antennas and Propagation*, 1993.

[84] V. Gonzalez-Posadas, D. Segovia-Vargas, E. Rajo-Iglesias, J. Vazquez-Roy, and C. Martin-Pascual, "Approximate analysis of short circuited ring patch antenna working at tm/sub 01/ mode," *IEEE Transactions on antennas and propagation*, vol. 54, pp. 1875–1879, June 2006.

[85] J. Davies, "A least-squares boundary residual method for the numerical solution of scattering problems," *IEEE Trans. Microwave Theory Tech.*, 1973.

[86] B. Rahman and J. Davies, "Analysis of optical waveguide discontinuities," *J. Lightwave Technol.*, 1988.

[87] K. Morishita, S. Inagaki, and N. Kumagai, "Analysis of discontinuities in dielectric waveguides by means of the least squares boundary residual method," *IEEE Trans. Microwave Theory Tech.*, 1979.

[88] R. Li, N. Bushyager, J. Laskar, and M. Tentzeris, "Determination of reactance loading for circularly polarized circular loop antennas with a uniform travelling-wave current distribution," *IEEE Transactions on antennas and propagation*, vol. 53, December 2005.

[89] J. Venkatesan and J. W.R. Scott, "Design of the double-y balun for use in gpr applications," in *Detection and Remediation Technologies for Mines and Minelike Targets IX, Proceedings of SPIE*, vol. 5415, pp. 383–398, September 2004.

[90] D. Delaune, J. T. S. Sumantyo, M. Takahashi, and K. Ito, "Circularly polarized rounded-off triangular microstrip line array antenna," *IEICE Transactions on Communications*, vol. E89-B, April 2006.

[91] Y.-L. Lee, T.-R. Chen, and J.-S. Row, "Cicularly polarized proximity-coupled microstrip antennas," *Microwave and Optical Technology Letters*, vol. 46, pp. 429–430, September 2005.

[92] J.-S. Row, "Design of aperture-coupled annular-ring microstrip antennas for circular polarization," *IEEE Transactions on antennas and propagation*, vol. 53, pp. 1779–1784, May 2005.

[93] K.-F. Tong, "A new single-feed proximity coupled circularly polarized square ring antenna," in *APMC Conference Proceedings*, 2006.

[94] H.-M. Chen and K.-L. Wong, "On the circular polarization operation of annular-ring microstrip antennas," *IEEE Transactions on antennas and propagation*, vol. 47, pp. 1289–1292, August 1999.

[95] B. K. Kanaujia and B. R. Vishvakarma, "Design considerations for the development of the annular ring microstrip antenna," *International Journal of Electronics*, vol. 89, no. 8, pp. 665–677, 2002.

[96] H. Liu and X.-F. Hu, "Input impedance analysis of a microstrip annular-ring antenna with a thick substrate," *Progress In Electromagnetics Research*, vol. 12, pp. 177–204, 1996.

[97] W. C. Chew, "A broad-band annular-ring microstrip antenna," *IEEE Transactions on antennas and propagation*, vol. 30, pp. 918–922, September 1982.

[98] D. M. Kokotoff, R. B. Waterhouse, C. R. Birtcher, and J. T. Aberle, "An annular ring coupled circular patch with enhanced performance," *Electronics Letters*, vol. 33, pp. 2000–2001, November 1997.

[99] G. Moernaut and G. Vandenbosch, "Simple pen and paper design of shorted annular ring antenna," *ELectronics Letters*, 2003.

[100] Y.-L. Liu, *Mutual Coupling in Spherical Arrays of Circular Waveguide Elements*. Dissertation, Polytechnic Institute of New York, June 1976.

[101] A. Hessel, Y.-L. Liu, and J. Shmoys, "Mutual admittance between circular apertures on a large conducting sphere," *Radio Science*, vol. 14, 1979 January.

[102] D. Winch and P. Roberts, "Derivatives of addition theorems for legendre functions," *Journal of Australian Mathematical Society*, vol. 37, pp. 212–234, 1995.

[103] N. Vilenkin, *Special Functions and the Theory of group Representations*. Amer Mathematical Society, 1968.

[104] A. Edmonds, *Angular momentum in quantum mechanics*. Princeton University Press, 2 ed., 1974.

[105] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in Fortran 77: The art of scientific computing*. Cambridge University Press, 1986.

[106] M. Abramowitz and I. A. Stegun, eds., *Handbook of mathematical functions*. National Bureau of Standards Applied Mathematics Series, United States Department of Commerce, 10 ed., December 1972.

[107] G. V. Borgiotti, "Mutual coupling analysis of a conformal array of elements on a cylindrical surface," *IEEE Transactions on antennas and propagation*, vol. 18, pp. 55–63, January 1970.

[108] D. M. Pozar, *Microwave Engineering*. John Wiley & Sons, Inc., 2 ed., 1998.

[109] C.-T. Tai and C. S. Pereira, "An approximate formula for calculating the directivity of an antenna," *IEEE Transactions on antennas and propagation*, 1976.

[110] I. Gelfand and G. Shilov, *Generalized Functions: Properties and Operations*, vol. 1. Academic Press, 1964.

[111] I. Gelfand and N. Y. Vilenkin, *Generalized Functions (Applications of Harmonic Analysis)*, vol. 4. Academic Press, 1964.

[112] B. Rafaely, B. Weiss, and E. Bachmat, "Spatial aliasing in spherical microphone arrays," *IEEE Transactions on Signal Processing*, 2007. accepted for publication.

[113] B. Rafaely, "Plane-wave decomposition of the sound field on a sphere by spherical convolution," *Journal of Acoustical Society of America*, vol. 116, pp. 2149–2157, October 2004.

[114] M. Park and B. Rafaely, "Sound-field analysis by plane-wave decomposition using spherical microphone array," *Journal of Acoustical Society of America*, vol. 118, November 2005.

[115] E. Saff and A. Kuijlaars, "Distributing many points on a sphere," *Mathematical Intelligencer*, vol. 19, no. 1, 1997.

[116] I. H. Sloan, "Polynomial interpolation and hyperinterpolation over general regions," *Journal of Approximation Theory*, vol. 83, pp. 238–254, 1995.

[117] R. Hardin and N. Sloane, "Mclaren's improved snub cube and other new spherical designs in three dimensions," *Discrete and Computational Geometry*, vol. 15, pp. 429–441, 1996.

[118] I. H. Sloan and R. S. Womersley, "The uniform error of hyperinterpolation on the sphere," *Advances in Multivariate Approximation*, vol. 107, pp. 289–386, 1999.

[119] N. Sneeuw, "Global spherical harmonic analysis by least squares and numerical quadrature methods in historical perspective," *Geophys. J. Int.*, vol. 118, pp. 707–716, 1994.

[120] J. Hansen, ed., *Spherical Near-Field Antenna Measurements*, distribution 4. No. 26 in IEE Electromagnetic Waves Series, Peter Peregrinus, 1988.

[121] B. Thye and T. Yeo, "Computing spherical transform and convolution on the 2-sphere." http://people.csail.mit.edu/ythomas/6869.pdf, May 2005.

[122] R. S. Womersley and I. H. Sloan, "How good can polynomial interpolation on the sphere be?," *Advances in Computational Mathematics*, vol. 14, no. 3, pp. 195–226, 2001.

[123] I. H. Sloan and R. S. Womersley, "Extremal systems of points and numerical integration on the sphere," *Adv. Comput. Math.*, vol. 21, no. 1-2, pp. 107–125, 2004.

[124] J. Cui and W. Freeden, "Equidistribution on the sphere," *Siam J. Sci. Comput.*, vol. 18, no. 2, pp. 595–609, 1997.

[125] W. Freeden, "On integral formulas of the (unit) sphere and their application to numerical computation of integrals," *Computing*, vol. 25, 1980.

[126] J. Goethals and J. Seidel, "The football," *Niew Archief voor Wiskunde*, vol. 3, no. 29, pp. 50–58, 1981.

[127] W. Neutsch, "Optimal spherical designs and numerical integration on the sphere," *Journal of Computational Physics*, vol. 51, pp. 313–325, 1983.

[128] L. Baggen, S. Holzwarth, W. Simon, and O. Litschke, "Phased array using the sequential rotation principle: analysis of coupling effects," in *IEEE International Symposium on Phased Array Systems and Technology*, pp. 571–576, October 2003.

[129] J. Huang, "A technique for an array to generate circular polarization with linearly polarized elements," *IEEE Transactions on antennas and propagation*, vol. 34, no. 9, 1986.

[130] E. Soliman, S. Brebels, E. Beyne, and G. Vandenbosch, "Sequential-rotation arrays of circularly polarized aperture antennas in the mcm-d technology," *Microwave and Optical Technology Letters*, vol. 44, March 2005.

[131] R. Ramirez, F. D. Flaviis, and N. Alexopoulos, "Single-feed circularly polarized microstrip ring antenna and arrays," *IEEE Transactions on antennas and propagation*, vol. 48, July 2000.

[132] IEEE, *A New 3-Way Power Divider with Various Output Power Ratios*, 1996.

[133] C. Chao, "A new odd number n-way power combiner/divider," in *Microwave Symposium Digest, IEEE MTT-S*, 1977.

[134] IEEE, *Miniaturization of 3- and 5-way Bagley Polygon power dividers*, 2005.

[135] F. Alhargan and S. Judah, "Circular and annular sector planar components of arbitrary angle for n-way power dividers/combiners," *IEEE Trans. Microwave Theory Tech.*, vol. 42, no. 9, 1994.

[136] M. Abouzahra and K. Gupta, "Multiport power divider-combiner circuits using circular-sector-shaped planar components," *IEEE Trans. Microwave Theory Tech.*, vol. 36, no. 12, 1988.

[137] M. Abouzahra and K. Gupta, "Multiple-port power dividers/combiners circuits using circular microstrip disc configuration," in *Microwave Symposium Digest, IEEE MTT-S*, 1987.

[138] P. R. Cromwell, *Polyhedra*. Cambridge University Press, 1997.

[139] M. J. Wenninger, *Polyhedron Models*. Cambridge University Press, 1971.

[140] H. Coxeter, *Regular Polytopes*. Macmillan, 1963.

[141] H. Cundy and A. Rollett, *Mathematical Models*. Oxford Clarendon Press, 1961.

[142] M. J. Wenninger, *Spherical Models*. Cambridge University Press, 1979.

[143] M. Wenninger, *Dual Models*. Cambridge University Press, 1983.

[144] R. Muradyan, "On discrete subgroups of the three-dimensional rotation group," *Theoretical and Mathematical Physics*, vol. 46, March 1981.

[145] R. C. Hansen, "Significant phased array papers," 1973.