The Australian National University
Faculty of Engineering and Information Technology
Department of Engineering

Theory and Design of Higher Order Sound Field Recording

Michael C. T. Chan

June 2003

Student Number: **3221999**
Supervisor: **Dr. Thushara Abhayapala**

A thesis submitted in part fulfilment of the degree of Bachelor of Engineering at the Australian National University

Acknowledgements

The work presented in this thesis would not have been possible without the support I've received during this project. I would like to express my gratitude to:

- Dr. Thushara Abhayapala for his excellent supervision. His guidance, encouragement and assistance made this project an invaluable experience and created a positive impact towards my education in general.
- The Department of Telecommunications Engineering at the Research School of Information Sciences and Engineering (RSISE) for providing facilities to assist in the completion of this thesis. Thanks to Ms. Lesley Cox for her assistance.
- Dr. Shankar Kalyanasundaram for his support and advice, especially for his recommendation to start early, as this had made a huge difference.
- David for proof reading this thesis and providing me with constructive feedback and suggestions.
- My sister Fiona, for her statistical inspiration.
- My parents for their continual support, guidance, encouragement and patience. Special thanks to my father for proof reading this thesis.
- My friends for their encouragement. Special thanks to Howard for sharing his honours year experience and giving me some helpful hints. Also, thanks to Ada for her support and encouragement.

Abstract

This thesis investigates the theory and design of recording the spherical harmonic components of an arbitrary three-dimensional sound field. The approach taken is based on capturing the spherical harmonic components using a spherical microphone. This choice led to the investigation of various design issues including the inherent limitations of the spherical microphone, discretization of the theoretical continuous spherical microphone into a microphone array and associated spatial aliasing problems, calibration errors of these microphones and signal processing issues. A fourth order microphone design was presented and analysed, which allowed the verification, integration and evaluation of the design issues mentioned earlier, in the context of this design. Overall, the design was capable of recording a frequency range of $[340, 3400]$Hz, thus finding applications in future teleconferencing systems. The design can also be directly applied to beamforming applications.

Glossary of Definitions

$\mathbb{R}$	real numbers
i	the imaginary unit, denoted $\sqrt{-1}$
(x, y, z)	point in right-handed Cartesian coordinate system
(r, θ, ϕ)	point in spherical coordinate system, with distance r from the origin, elevation angle θ and azimuth angle ϕ
$\mathbf{x}$	vector notation of a point in $\mathbb{R}^3$ space
$\hat{\mathbf{x}}$	unit vector in the direction of $\mathbf{x}$
$(\cdot)^*$	complex conjugation
$(\cdot)'$	first derivative
$\lvert \cdot \rvert$	absolute value
$\lVert \cdot \rVert$	Euclidean distance
δ_{nm}	Kronecker delta function: $\delta_{nm} = 0$ for $n \neq m$, $\delta_{nm} = 1$ for $n = m$
$\mathrm{sgn}(\cdot)$	signum function: $\mathrm{sgn}(x) = 0$ for $x = 0$, $\mathrm{sgn}(x) = x/\lvert x \rvert$ for $x \neq 0$
$Z \sim \mathcal{N}(\mu, \sigma)$	random variable Z is Gaussian distributed with mean μ and standard deviation σ
PDF	probability density function
DFT	discrete Fourier transform

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In real life, sound is perceived to originate from all directions and distances. For example, in a quiet park setting, our auditory system is able to accurately determine the location of a sound source in three-dimensional space, such as the whistling of a bird (this is referred to as *localisation*). As another example, consider a live musical performance. Being able to perceive the positions of individual instruments present at a live performance contributes to the enjoyment of the performance. Other aspects that contribute to enjoyment comes from sound that does not directly originate from sound sources. These include sound that is reflected from objects and surfaces, as well as reverberation (sound not perceived to originate from any particular direction), both of which depend on the characteristics of the venue [1]. These contribute a sense of space and envelopment to the listener. Therefore, both direct and indirect sounds are important acoustic aspects that capture the event.

In order to reproduce the auditory sensation of being at an event, for example, one would require surround sound technology, which exhibits certain characteristics. Most importantly is the ability to record and reproduce sound such that sources are perceived to be localised. Another is the ability to preserve the indirect and reverberant effects of the event. One way is to capture the sound pressure at each point in a given three-dimensional space, known as a *sound field*. If one can record and reproduce the sound field, one or more listeners will be able to experience the recorded event as if they were actually present. This thesis shall explore the recording problem and will not address the reproduction of sound fields. Such a recording system is referred to as a three-dimensional audio recording system [2].

There are many applications of three-dimensional audio recording systems. We have already described one above, that is, to capture an acoustic event. To some extent, they could also be applied to the recording of movie sound tracks, broadcast drama and virtual reality situations [1]. Perhaps the most promising future application, however, is its potential use in *telepresence*. This is the idea that interaction between parties at remote locations can be made more effective and efficient if it makes use of as much perceptual information as possible [3]. That is, to emulate real interaction that occurs between people in the physical world. An artist's impression of its application to teleconferencing in the future is given in Figure 1.1 (reproduced from [4]). Three-dimensional audio is one such tool which could be used to record and reproduce voices of remote parties so that their voices appear to originate from particular locations in a room [3]. Alternatively, it is possible to "steer" the three-dimensional audio recording system towards a desired sound source, such as a person speaking, whilst attenuating

undesired sounds and ambient noise. This technology is known as *beamforming*.

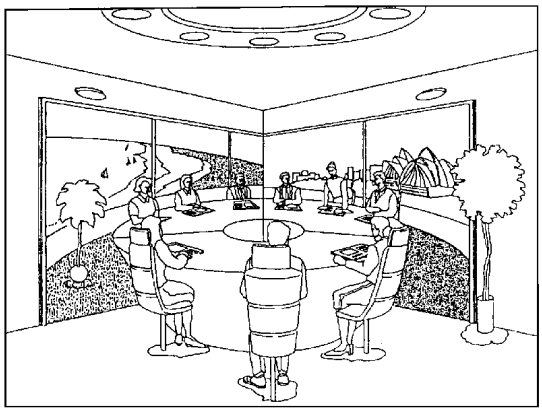


Figure 1.1: An artist's impression of a futuristic teleconference.

1.1 Background and Motivation

Existing surround sound recording and reproduction technology has room for improvement in terms of providing a sense of realism, in the way we described above. The well known two-channel stereo sound technology in popular use today, makes it possible to record information about sound sources and reproduce this information so that these sources can be localised within a range of angles about a horizontal plane from the listener's perspective [5]. This was first introduced by Alan Dower Blumlein, who was issued a patent for his work in 1931 (Refer to [6]). However, stereo has the limitation that it cannot be used as a surround sound system simply because it cannot envelope the listener with two speakers.

Dolby attempted to fix this problem with Dolby Surround technology[1]. Although this is a sound reproduction technology, we discuss it here because its deficiencies are partly the motivations for technologies to be described in this thesis. In Dolby Surround systems, the front left, front right and front centre speakers are used to create a highly localised sound image, whilst two rear surround speakers are used to create an ambient sound field [8,9]. This turns out to be especially effective in a cinema setting where the idea is to direct dialogue towards the front in order to coordinate well with the picture. Reverberant sound is directed to the rear surround channels to envelope the audience in sound, as if one was actually at the event, such as on the sidelines of a football stadium, for example [1,9]. However, because Dolby Surround systems are targeted for use primarily in the cinema, it was intentionally designed so that localisation of sound sources is imprecise. This means that it is not suitable for accurate reproduction of an acoustic event.

The technologies that attempt to solve this problem is based on the recording and reproduction of the original sound field. One is called wave field synthesis, where work has been published in [10]. We will not pursue this technology further in this thesis. Instead, we will discuss another called ambisonics, which is based on the work by a group of British researchers at the Mathematical Institute of Oxford [11], most notably, the work in [12] by Michael Ger-

[1]For a brief coverage of the development of Dolby Surround technology, the reader is referred to [7].

zon. Work by Bamford in [8] suggests that the first order ambisonic reproduction system is superior to both Dolby Surround and stereo in terms of localisation. The theory uses a mathematical representation of a sound field based on spherical harmonic components. Once the spherical harmonic components have been recorded, these serve as an interface to the reproduction problem. The *more* spherical harmonic components (or said another way, the higher the order) of the sound field that are recorded, the *greater* the directional effects and more *precise* localisation of sound sources [1]. However, the recording would require a greater number of channels [8]. A commercial microphone exists called the SoundField microphone, which is capable of recording the first order spherical harmonic components [13].

However, a drawback of ambisonics is that it is inherently limited by the way the spherical harmonic components of the sound field are recorded. This is because the theory requires that the sound field is captured by microphones coincident at a single point in space. In practice, placing discrete microphones coincidently at a single point is not possible, so they are placed in a regular geometric arrangement and the spherical harmonic components are obtained by taking linear combinations of the signals from these microphones. The requirement of having to use regular geometric arrangements makes design of higher order microphones difficult, especially because regular polyhedra are few. For example, microphones are placed at the vertices of a regular dodecahedron arrangement (the regular polyhedron with the greatest number of vertices [14]) in the design of a second order sound field microphone, as presented in [1].

An alternative approach, also based on recording the spherical harmonic components of a sound field, was taken by Abhayapala and Ward in [2] and Elko and Meyer in [15]. At a given radius (as opposed to recording at a single point in ambisonics), they used a theoretical continuous spherical microphone (that is, a microphone capable of picking up the sound pressure at any point on a spherical surface) to capture the spherical harmonic components via signal processing. The advantage of this approach is that it is not limited by the need to use fixed geometrical arrangements. Instead, it requires a discrete approximation to the continuous spherical microphone in a practical design [2]. Both papers demonstrated a microphone design capable of capturing up to and including third order spherical harmonic components of a sound field. Abhayapala and Ward were able to demonstrate that their design was able to capture a frequency range of [340,3400]Hz. Elko and Meyer have commercialised their third order microphone called the EigenMike [16]. More recently, there has been work done by Laborie, Bruno and Montoya [17], which also uses this alternative approach. A third order system was also considered and in addition, their design did not restrict microphones to be placed in a spherical arrangement. The main drawback of these systems is that they have been limited to third order designs.

1.2 Problem Statement

In this thesis, we extend the work already done using this alternative approach by investigating design issues of recording higher order spherical harmonic components of a sound field and by designing a higher order microphone that preserves the frequency range of [340,3400]Hz achieved in [2]. We will present an analysis of this design to provide better insight in relation to its performance (more details provided in the next section). Beamforming has been investigated in a number of sources, most relevant being [18] and [15]. Thus, this topic will only be lightly

mentioned in this thesis. Sound field reproduction will not be covered and for information regarding this, work has been published in [19] and further investigated in [20].

1.3 Contributions

The work presented in this thesis has made the following main contributions to sound field recording:

- Analysis of the role of rigid spheres in improving microphone design.
- Analysis of various microphone arrangements, which are applicable to microphone array design.
- A model for analysing the error due to inexact positioning of microphones in an array was presented.
- A set of error measures was proposed for error analysis of spherical microphones.
- Design and analysis of a fourth order spherical microphone.

1.4 Outline of Thesis

This thesis shall take on a structure which is roughly divided into three phases, namely, theory, design and simulation. The content of a chapter tends to rely on the previous chapters and so to get the most out of this thesis, we recommend that it be read in a chapter-by-chapter basis. The content shall be described below.

- **Chapter 2:** provides an overview of the theory. It gives a brief presentation of the solution to the wave equation in spherical coordinates and the resulting modal analysis tools that shall be used to represent arbitrary sound fields. Spherical scattering theory is introduced. It consolidates the theory presented in this chapter and sets out the recording problem that is to be investigated in the rest of this thesis.
- **Chapter 3:** identifies the various issues relating to the design of a practical microphone array. It covers the inherent limitations, modal aliasing, discretization of the continuous spherical microphone into a microphone array, calibration errors and signal processing issues.
- **Chapter 4:** applies the material addressed in Chapter 3 to produce a practical design and presents error measures and simulation results to assess the design's performance, both quantitatively and qualitatively.

All items in this thesis that were produced via numerical calculations are documented in Appendix H. This gives an index to the software item, which is located on the attached CD-ROM.

Chapter 2

Theory of Spherical Harmonic Representation of Sound Field

2.1 Introduction

An arbitrary sound field can be modelled as a three-dimensional scalar field. This field quantifies the sound pressure at every point in an arbitrary coordinate system. All homogeneous sound fields (ie. no sound sources inside the region of interest) satisfy the well-known classical wave equation[1] given by

$$\nabla^2 s = \frac{1}{c^2}\frac{\partial^2 s}{\partial t^2}, \tag{2.1}$$

where ∇^2 is the Laplacian operator, which can be expanded according to the desired coordinate system and c is the speed of wave propagation in the medium of interest. The sound field is denoted as $s = s(\mathbf{x}, t)$ at point $\mathbf{x} \in \mathbb{R}^3$ and at time t [18].

We chose to use the spherical coordinate system to exploit spherical symmetry. We shall use the convention where the elevation angle (from the vertical axis) is represented by θ and the azimuth angle is represented by ϕ. The radial coordinate shall be represented by r. This is shown in Figure 2.1.

[1] The derivation of this is accessible in many sources including [18, 21–23].

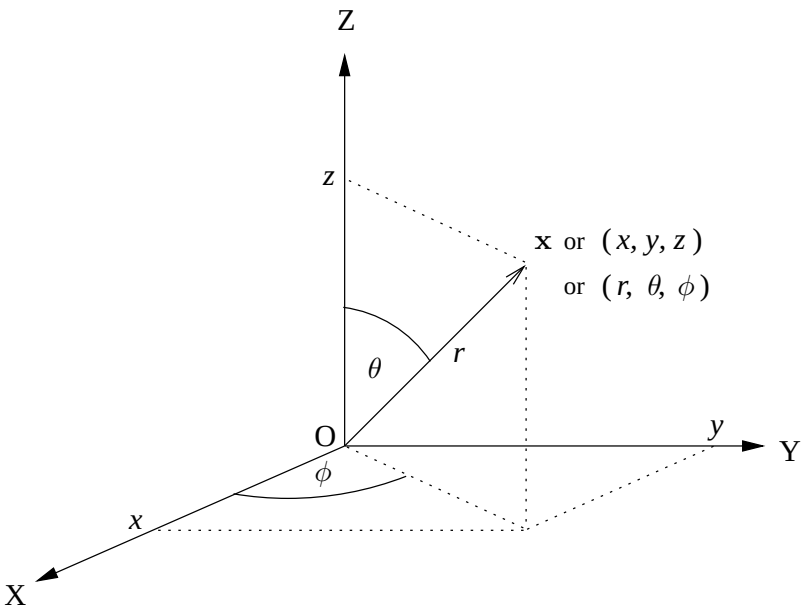


Figure 2.1: The relationship between the spherical coordinate system and the right-handed Cartesian coordinate system.

The spherical coordinate (r, θ, ϕ) has an equivalent right-handed Cartesian coordinate (x, y, z) and they are related by:

$$\begin{aligned} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta. \end{aligned}$$

The direction implied by (θ, ϕ) can be represented by $\hat{\mathbf{x}}$, where $\hat{\mathbf{x}} = \mathbf{x}/\|\mathbf{x}\|$.

2.2 General Solution to the Wave Equation in Spherical Coordinates

In this section, we give a concrete background to the mathematics used to represent an arbitrary sound field in a region of space. As will be discussed below, the form of this solution is too general for our purposes but it will give insight to the representation of a homogeneous sound field given in the next section, which we will use in the remaining part of this thesis.

Referring back to the wave equation, (2.1), we chose to expand the Laplacian operator into spherical coordinates. This was shown in [24] as

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\Big(r^2\frac{\partial}{\partial r}\Big) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\Big(\sin\theta\frac{\partial}{\partial\theta}\Big) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}.$$

By substituting this back into (2.1), we obtain the wave equation in spherical coordinates as

$$\frac{1}{r^2}\frac{\partial}{\partial r}\Big(r^2\frac{\partial s}{\partial r}\Big) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\Big(\sin\theta\frac{\partial s}{\partial\theta}\Big) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 s}{\partial\phi^2} = \frac{1}{c^2}\frac{\partial^2 s}{\partial t^2}. \tag{2.2}$$

This equation is a partial differential equation in four independent variables (namely the spatial coordinates, r, θ, ϕ and temporal coordinate, t) and can be solved using the method of

separation of variables [25]. This is done by assuming the solution of (2.2) to be of the form

$$s(r,\theta,\phi,t) = R(r)\Theta(\theta)\Phi(\phi)T(t). \tag{2.3}$$

Since literature is widely available on its solution in spherical coordinates, rather than showing all the working, the separated differential equations and their solutions shall be explained below. Firstly, the separated differential equation for the temporal variable, t, is given by

$$\frac{1}{T(t)}\frac{1}{c^2}\frac{d^2T(t)}{dt^2} = -k^2. \tag{2.4}$$

The assumption has been made that the separation of variables constant, $-k^2$, is negative because this leads to the desired time solutions corresponding to propagating and standing waves, whereas decaying exponentials would be modelled if a positive constant was assumed [23]. Equation (2.4) has the general solution

$$T(t) = T_1 e^{i\omega t} + T_2 e^{-i\omega t}, \tag{2.5}$$

where

$$\omega = kc. \tag{2.6}$$

Note that $\omega = 2\pi f$ is the temporal frequency. In this thesis, we assume that the medium of propagation is homogeneous and therefore c is constant. Since k is proportional to ω, we shall also refer to k as the temporal frequency.[2] We chose to set T_2 to zero since $e^{-i\omega t}$ represents a wave propagating backwards in time and so has no physical meaning [26]. Thus

$$T(t) = T_1 e^{i\omega t}. \tag{2.7}$$

By substituting the time solution of (2.7) into (2.3), and then substituting the expression into (2.1), we obtain

$$\nabla^2 u + k^2 u = 0, \tag{2.8}$$

where $u = u(r,\theta,\phi) = R(r)\Theta(\theta)\Phi(\phi)$. Equation (2.8) is known as the Helmholtz equation. This equation describes the spatial solution of the wave equation. A similar method of separation of variables is used to solve it, where we use $-m^2$ as the separation of variables constant. The differential equation relating to the azimuth angle ϕ is

$$\frac{1}{\Phi(\phi)}\frac{d^2\Phi(\phi)}{d\phi^2} = -m^2, \tag{2.9}$$

which has the general solution,

$$\Phi(\phi) = \Phi_1 e^{im\phi} + \Phi_2 e^{-im\phi}. \tag{2.10}$$

Both solutions are kept. In both cases, the constant, m, must be an integer to ensure continuity and periodicity of $\Phi(\phi)$ [27, p.185]. The differential equation relating to the radial coordinate

[2] In this thesis, we will use the variables ω, f and k interchangeably, where appropriate.

r is given by

$$\frac{d}{dr}\Big(r^2\frac{dR(r)}{dr}\Big) + \big[k^2r^2 - n(n+1)\big]R(r) = 0. \tag{2.11}$$

Here, the separation of variables constant has been written as $n(n+1)$ since (2.11) has continuous solutions only if n is an integer [23, p.380]. This equation is known as the spherical Bessel differential equation, which has two types of solutions, namely, $j_n(x)$ (called the spherical Bessel function of the first kind) and $y_n(x)$ (called the spherical Bessel function of the second kind). The general solution to the spherical Bessel differential equation is

$$R(r) = R_1\, j_n(kr) + R_2\, y_n(kr). \tag{2.12}$$

The differential equation relating to the elevation angle θ is

$$\sin\theta\frac{d}{d\theta}\Big(\sin\theta\frac{d\Theta(\theta)}{d\theta}\Big) + \big[n(n+1)\sin^2\theta - m^2\big]\Theta(\theta) = 0. \tag{2.13}$$

This equation is known as the associated Legendre differential equation and has two types of solutions, namely, $P_n^m(x)$ (called the associated Legendre function of the first kind) and $Q_n^m(x)$ (called the associated Legendre function of the second kind).[3] The general solution to the associated Legendre differential equation is

$$\Theta(\theta) = \Theta_1\, P_n^m(\cos\theta) + \Theta_2\, Q_n^m(\cos\theta), \tag{2.14}$$

where n is a non-negative integer and $-n \leq m \leq n$. However, we are interested in solutions which model sound phenomena, and so we require solutions which are unique, continuous and finite for $-1 \leq \cos\theta \leq 1$. Since $Q_n^m(x)$ are not finite at $\cos\theta = \pm 1$, we chose to set Θ_2 to zero [23, p.384].

The angle functions of (2.10) and (2.14) are combined into single functions called *spherical harmonics* and are defined as

$$Y_{nm}(\hat{\mathbf{x}}) = \sqrt{\frac{2n+1}{4\pi}\frac{(n-|m|)!}{(n+|m|)!}}P_{n|m|}(\cos\theta)e^{im\phi}, \tag{2.15}$$

where n is a non-negative integer and $-n \leq m \leq n$. Here, n is referred to as the order and m is referred to as the mode of the spherical harmonic [2]. These functions are orthonormal and satisfies

$$\int Y_{nm}(\hat{\mathbf{x}})Y^*_{n'm'}(\hat{\mathbf{x}})d\hat{\mathbf{x}} = \delta_{nn'}\delta_{mm'}, \tag{2.16}$$

where n and n' index the order, and m and m' index the mode of the corresponding spherical harmonic functions. Note that the integration is performed over the unit sphere.[4] This is referred to as the *orthonormality property* [2].

Any solution to the wave equation of (2.1) can therefore be written as

$$S(\mathbf{x};k) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n}\big[A_{nm}(k)j_n(k\|\mathbf{x}\|) + B_{nm}(k)y_n(k\|\mathbf{x}\|)\big]Y_{nm}(\hat{\mathbf{x}}), \tag{2.17}$$

[3] For more information about Legendre polynomials, refer to [24, 27, 28].

[4] That is, $d\hat{\mathbf{x}} = d(\cos\theta)d\phi = \sin\theta d\theta d\phi$ and the limits are for θ, from 0 to π and for ϕ, from 0 to 2π.

where the coefficients, $A_{nm}(k)$ and $B_{nm}(k)$, are generally complex valued. Note that we have taken the frequency domain representation of $s(\mathbf{x}, t)$, where the temporal dependence is implied by the frequency dependence of these coefficients [27, p.186].

2.3 A Homogeneous Sound Field

The form of (2.17) is too general for our purposes. Consider a sound field in a region $\Omega \in \mathbb{R}^3$, where all sources are outside this region, as shown in Figure 2.2.

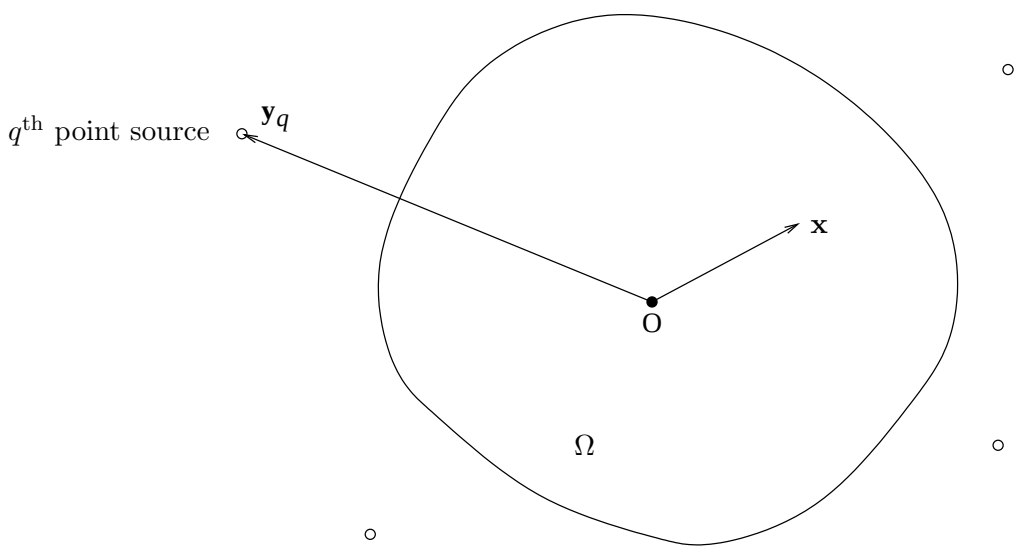


Figure 2.2: The sound field of interest Ω, with respect to the point sources.

This is known as a homogeneous sound field. Suppose there are Q point sources located at $\mathbf{y}_q \notin \Omega$, $q = 1, \ldots, Q$, and let the q^{th} source radiate a signal $v_q(t)$. Then, the sound field at a point $\mathbf{x} \in \Omega$ is given by

$$s(\mathbf{x}; t) = \sum_{q=1}^{Q} \frac{1}{\|\mathbf{y}_q - \mathbf{x}\|} \, v_q\Big(t - \frac{\|\mathbf{y}_q - \mathbf{x}\|}{c}\Big). \tag{2.18}$$

Here, we assume free space wave propagation. Taking the temporal Fourier transform of (2.18) and using the time-shift property yields

$$S(\mathbf{x}; k) = \sum_{q=1}^{Q} \frac{e^{-ik(\|\mathbf{y}_q - \mathbf{x}\|)}}{\|\mathbf{y}_q - \mathbf{x}\|} \, V_q(k), \tag{2.19}$$

where $V_q(k)$ is the Fourier transform of $v_q(t)$, $S(\mathbf{x}; k)$ is the Fourier transform of $s(\mathbf{x}; t)$ and $k = 2\pi f/c$. We use the identity

$$\frac{e^{-ik(\|\mathbf{y}_q - \mathbf{x}\|)}}{\|\mathbf{y}_q - \mathbf{x}\|} = 4\pi(-i)k \sum_{n=0}^{\infty} \sum_{m=-n}^{n} h_n^{(2)}(k\|\mathbf{y}_q\|) Y_{nm}^*(\hat{\mathbf{y}}_q) j_n(k\|\mathbf{x}\|) Y_{nm}(\hat{\mathbf{x}}), \tag{2.20}$$

for $\|\mathbf{y}_q\| > \|\mathbf{x}\|$, where $h_n^{(2)}(x) = j_n(x) - in_n(x)$ (called the spherical Hankel function of the second kind) [29, p.30], to write

$$\begin{aligned} S(\mathbf{x};k) &= \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \underbrace{\Big(\sum_{q=1}^{Q} 4\pi(-i)kh_n^{(2)}(k\|\mathbf{y}_q\|)Y_{nm}^*(\hat{\mathbf{y}}_q)V_q(k)\Big)}\, j_n(k\|\mathbf{x}\|)Y_{nm}(\hat{\mathbf{x}}) \\ S(\mathbf{x};k) &= \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \gamma_{nm}(k) j_n(k\|\mathbf{x}\|)Y_{nm}(\hat{\mathbf{x}}), \end{aligned} \tag{2.21}$$

where the under-braced term $\gamma_{nm}(k)$ is constant with respect to the position $\mathbf{x}$. We represent these frequency dependent expressions as

$$\gamma_{nm}(k) = \sum_{q=1}^{Q} 4\pi(-i)kh_n^{(2)}(k\|\mathbf{y}_q\|)Y_{nm}^*(\hat{\mathbf{y}}_q)V_q(k), \tag{2.22}$$

which are the corresponding coefficients of the terms, $j_n(k\|\mathbf{x}\|)Y_{nm}(\hat{\mathbf{x}})$, called *modes* or *modal components*.[5] These terms are indexed by each order n which has corresponding modes $-n \leq m \leq n$.[6] Therefore, we call these coefficients, the *modal coefficients*. Equation (2.21) is referred to as the *synthesis equation* and this representation is only valid for homogeneous sound fields. From here onwards, all sound fields that we consider will be assumed to be homogeneous. Note that (2.21) implies that a homogeneous sound field is represented using an infinite number of modal components and corresponding modal coefficients.

If we now multiply both sides of (2.21) by $Y_{n'm'}^*(\hat{\mathbf{x}})$ and integrate both sides over the unit sphere, we obtain

$$\int S(\mathbf{x};k)Y_{n'm'}^*(\hat{\mathbf{x}})d\hat{\mathbf{x}} = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \gamma_{nm}(k)j_n(k\|\mathbf{x}\|)\int Y_{nm}(\hat{\mathbf{x}})Y_{n'm'}^*(\hat{\mathbf{x}})d\hat{\mathbf{x}}.$$

Then, using the orthonormality property of (2.16), the modal coefficients can be written as

$$\gamma_{nm}(k) = \frac{1}{j_n(k\|\mathbf{x}\|)}\int S(\mathbf{x};k)Y_{nm}^*(\hat{\mathbf{x}})d\hat{\mathbf{x}}, \tag{2.23}$$

provided that $j_n(k\|\mathbf{x}\|) \neq 0$. Equation (2.23) is referred to as the *analysis equation*. The material outlined in this section is referred to as *modal analysis* [18].

2.4 Scattering from a Rigid Sphere

In spherical microphone design, there are advantages of using a rigid spherical scatterer (which serves as an obstruction), as will be discussed in Section 3.3. When a rigid spherical scatterer is placed inside the sound field, the sound field observed will differ with the situation where there is no scatterer. Hence, we need to model this phenomenon. This is a scattering problem

[5] Note that the term *mode* has been overloaded and refers also to the index, m. Therefore, in this thesis, we will use the term modal components for this context.

[6] From here onwards, when we refer to modal components of order n, we implicitly refer to all the modes for $-n \leq m \leq n$.

and involves splitting the total pressure field S into a sum of two parts. That is,

$$S(r,\theta,\phi;k) = S_{\mathrm{i}}(r,\theta,\phi;k) + S_{\mathrm{scat}}(r,\theta,\phi;k) \tag{2.24}$$

where S_{i} is the incident field (the field that would be measured had the scatterer not been present) and S_{scat} represent the scattered field, which is a new quantity that measures the change in the incident field as a result of the presence of the scatterer.

To obtain a boundary condition for this situation, we note that the total radial velocity must vanish on the surface of a rigid sphere. Therefore, the boundary condition for a rigid sphere scatterer of radius R, can be expressed as

$$\frac{\partial}{\partial r}\Big(S_{\mathrm{i}}(r,\theta,\phi;k) + S_{\mathrm{scat}}(r,\theta,\phi;k)\Big)\bigg|_{r=R} = 0, \tag{2.25}$$

as shown in [27, p.228]. Now, from Section 2.3, the incident sound field can be represented by (2.21) and explicitly in terms of spherical coordinates,

$$S_{\mathrm{i}}(r,\theta,\phi;k) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \gamma_{nm}(k) j_n(kr) Y_{nm}(\theta,\phi). \tag{2.26}$$

The scattered field, however, is different because it consists of outgoing waves. This can be generally represented as

$$S_{\mathrm{scat}}(r,\theta,\phi;k) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \zeta_{nm}(k) h_n^{(1)}(kr) Y_{nm}(\theta,\phi), \tag{2.27}$$

where $h_n^{(1)} = j_n(x) + i n_n(x)$ (called the spherical Hankel function of the first kind) [27]. $\zeta_{nm}(k)$ represents the modal coefficients for the scattered field. Now, by taking the derivative of (2.26) and (2.27) with respect to r, evaluating the derivatives at $r = R$ and then substituting them into the boundary condition of (2.25), we obtain the relationship,

$$\zeta_{nm}(k) \quad = \quad -\gamma_{nm}(k)\frac{j_n^{'}(kR)}{h_n^{(1)'}(kR)}, \tag{2.28}$$

where $(\cdot)'$ indicates the first derivative. By substituting (2.26) and (2.27) into (2.24), and using (2.28), the sound field, when obstructed by a rigid spherical scatterer is given by

$$S(r,\theta,\phi;k) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \gamma_{nm}(k)\Big(j_n(kr) - \frac{j_n^{'}(kR)}{h_n^{(1)'}(kR)} h_n^{(1)}(kr)\Big) Y_{nm}(\theta,\phi). \tag{2.29}$$

In this thesis, we shall denote the radial term in (2.29) as

$$b_n(k\|\mathbf{x}\|) \triangleq j_n(k\|\mathbf{x}\|) - \frac{j_n^{'}(kR)}{h_n^{(1)'}(kR)} h_n^{(1)}(k\|\mathbf{x}\|), \tag{2.30}$$

where the expression have been rewritten in terms of $\|\mathbf{x}\|$, which is the convention we have adopted. Note that (2.30) is implicitly a function of the radius of the rigid spherical scatterer, R.

2.5 Analysis and Synthesis of a Sound Field

Analogous to the Fourier series coefficients for a periodic temporal signal, the modal coefficients $\gamma_{nm}(k)$ in (2.21) contain all the information about an arbitrary sound field. Therefore, if we can record them, it is theoretically possible to reconstruct the acoustic environment.

We turn our attention to recording these coefficients by considering the form of the analysis equation (2.23). It is possible to capture the modal coefficients of a sound field by evaluating this integral over an arbitrary surface, for each n and $-n \leq m \leq n$ desired. This means that it is only necessary to evaluate the sound field over this surface rather than at each and every point in the sound field.

Alternatively, if we decide to evaluate the modal coefficients over a spherical surface (by constraining $\|\mathbf{x}\| = r$, where $r \neq 0$ is an arbitrary constant radius), which has within it a rigid spherical scatterer of radius $0 \leq R \leq r$, the resultant sound field is now given by (2.29). Therefore, the analysis equation becomes

$$\gamma_{nm}(k) = \frac{1}{b_n(k\|\mathbf{x}\|)} \int S(\mathbf{x}; k) Y_{nm}^*(\hat{\mathbf{x}}) d\hat{\mathbf{x}}. \tag{2.31}$$

Thus from this equation, we are similarly able to extract the modal coefficients $\gamma_{nm}(k)$.

2.6 Problem Formulation

In this thesis, we specifically chose to examine the recording of the modal coefficients over a spherical surface, using (2.23) and (2.31). Spherical geometry was chosen because it is relatively easier to analyse and we can consider issues dealing with the radial and angular representations[7] of the sound field separately. These issues are the focus of the next chapter.

[7] Refer to Section 2.2 for information about these.

Chapter 3

Theory and Design of Spherical Microphone

3.1 Introduction

In the previous chapter, we have decided to capture the modal coefficients using a spherical surface. Therefore, we turn our attention to the design issues related to this decision in this chapter. Firstly, before we embark on analysing intimate details of the microphone itself, we present the implications of our inability to capture an infinite set of modal coefficients. Then, we move on to analyse the inherent limitations of recording from a continuous spherical microphone. Because such a microphone cannot be practically implemented, we look at issues of discretizing it by using a microphone array, which introduces problems associated with insufficient spatial sampling. In practice, these discrete microphones cannot be placed exactly in the location on the sphere as desired, and the calibration errors associated with inexact positioning are examined. Finally, we look at the signal processing that is required in a practical system.

3.2 Finite Order Design

The synthesis equation of (2.21) is of the form of an infinite sum, which means an infinite number of orders of modal components. With a continuous spherical microphone, it is possible to extract all these components exactly (Refer to Section 3.4) but in terms of practicality, the fact that such a microphone cannot be implemented and that the signal processing power we have at our disposal is finite, means that we can only consider recording a finite number of orders of modal coefficients.

Fortunately, this fact does not halt us from continuing further. One interesting point to note about the synthesis equation is that it was found in [18, pp.46-7] that the lower orders are most significant while the higher orders give the finer details to a sound field. This is analogous to frequency domain analysis of temporal signals since the lower frequencies tend to be the most significant while the higher frequencies add detail to the signal. This means that *order truncation* (limiting to a finite number of orders) of an arbitrary sound field can be performed to approximate the infinite sum representation of (2.21). To demonstrate the increased accuracy of the approximation when more orders are considered in the representation

of a plane wave sound field, consider Figure 3.1.

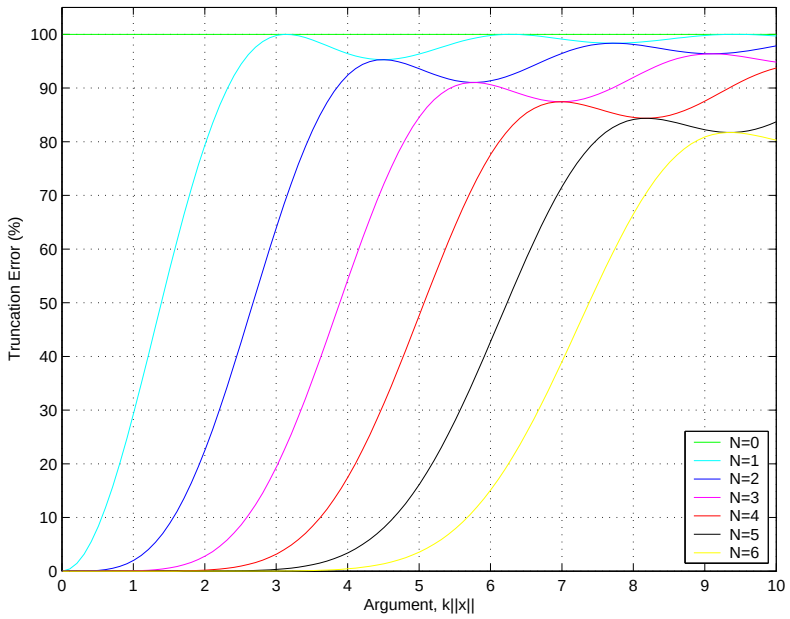


Figure 3.1: Truncation error of a plane wave sound field when only the modal components of orders $0 \leq n < N$ are considered, for various values of N.

This figure shows the decreasing error due to order truncation (we refer to this error as *truncation error*) as more orders of modal components are considered. At lower values of the argument, $k\|\mathbf{x}\|$ (that is, for the plane wave sound field of frequency k, at a given radius from the origin, $\|\mathbf{x}\|$) the decrease in truncation error is apparent. Nevertheless, the error increases rapidly to 100% as $k\|\mathbf{x}\|$ is increased and this limits the size of the region where high spatial accuracy can be achieved. In other words, the spatial accuracy decreases rapidly as $k\|\mathbf{x}\|$ increases. Fortunately, the temporal content of the sound field is independent of truncation error and this means that the temporal content can still be intelligible even when this error is large. This is analogous to listening to a stereo recording far from the speakers placed in a standard stereo set up, such that one perceives little spatial content (like a mono recording). For details of the process of producing Figure 3.1, the reader is referred to Sections 4.2 and 4.3.

3.3 Nature of Spherical Microphone

As introduced earlier in Section 1.1, we idealise the microphone as a continuous spherical sensor, which we call a continuous spherical microphone. This idealisation allows us to evaluate the integral of (2.23) exactly for modal coefficients of any order. We retain this idealisation within this section. Consider the situation where we use a continuous spherical microphone of radius $\|\mathbf{x}\| = r$ to record the sound field. We assume that the continuous spherical microphone is transparent with respect to the sound field, such that it does not disturb the sound field in any way. We shall call this an open spherical configuration.

Now, observe the radial solution $j_n(k\|\mathbf{x}\|)$ in the expression of (2.21). This is plotted in Figure 3.2. Notice that for small values of frequency k, the values of $j_n(kr)$ for $n \geq 2$ are small. In addition, $j_n(kr)$ has zeros for all $n \geq 0$. At the corresponding frequencies k, the modal components of the corresponding orders n, will be "perceived" by the continuous spherical microphone to be small or equal to zero, respectively. Therefore, they will be difficult to record at those frequencies. We would like to record over a band of frequencies say for modal components of orders $0 \leq n < N$. However, the presence of these zeros in the radial solution means that we are greatly restricted in the range of frequencies for which an accurate recording can be made.

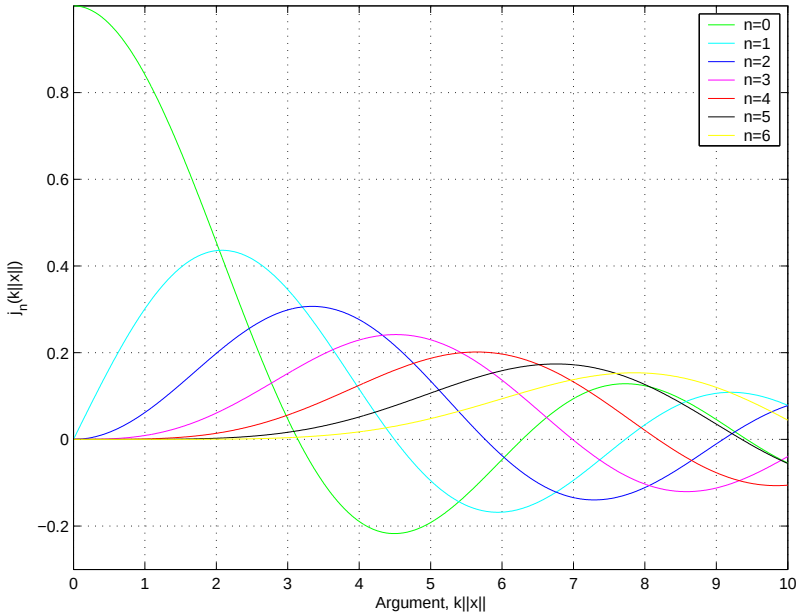


Figure 3.2: Plots of spherical Bessel functions, $j_n(k\|\mathbf{x}\|)$, for various orders n.

Fortunately, the restriction by the zeros of the spherical Bessel functions can be alleviated if we decide to use a rigid spherical configuration, which means that the surface of the continuous spherical microphone of radius $\|\mathbf{x}\| = r$ coincides with a rigid spherical scatterer (ie. a rigid sphere). In this case, the rigid spherical scatterer interacts and alters the sound field. As shown in Section 2.4, the radial solutions are represented by (2.30). These are plotted in Figure 3.3.

Notice that there are no zeros in the functions, $b_n(k\|\mathbf{x}\|)$. The other noticeable advantage of these functions is that at lower values of k and for $n > 0$, $b_n(kr)$ is approximately 3dB greater than $j_n(kr)$, as demonstrated in Figure 3.4. This is due to diffraction over the rigid sphere [15]. At higher values of frequency k, the scattering effects will become more prominent compared to diffractional effects.

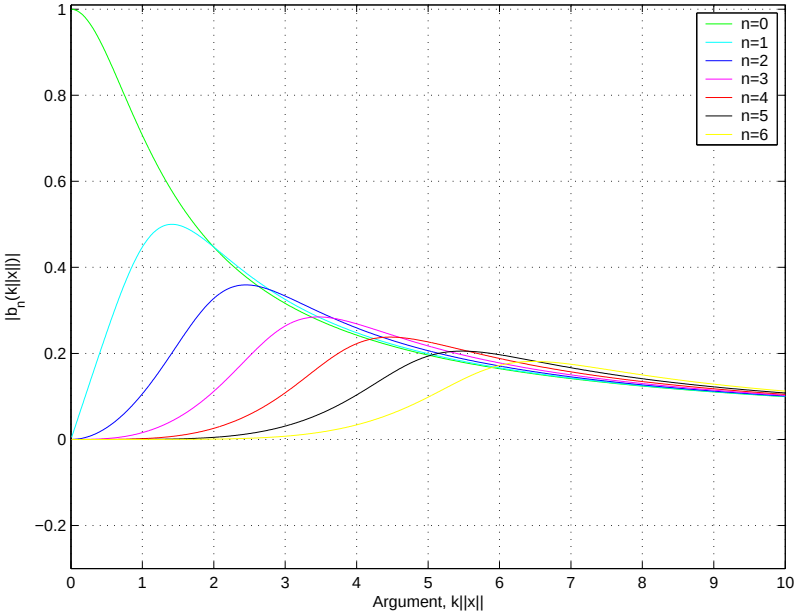


Figure 3.3: Plots of $b_n(k\|\mathbf{x}\|)$ for various orders n.

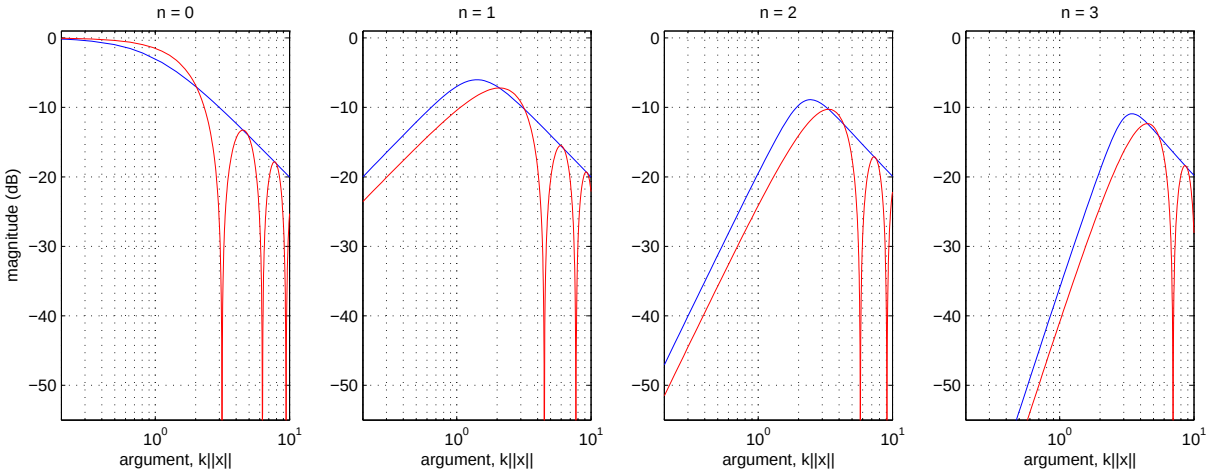


Figure 3.4: Comparison between $j_n(k\|\mathbf{x}\|)$ (in red) and $b_n(k\|\mathbf{x}\|)$ (in blue).

There is a relationship between the open and rigid spherical configurations, which is expected since a rigid spherical scatterer of zero radius is effectively having an open spherical configuration. Firstly, in order to determine this relationship, we need to consider the function $b_n(k\|\mathbf{x}\|)$ in (2.30). In addition, we also need to consider the recurrence relationship for the spherical Bessel functions,

$$j_n^{'}(kR) = j_{n-1}(kR) - \frac{n+1}{kR} j_n(kR), \tag{3.1}$$

and the small argument approximation,

$$h_n^{(1)'}(kR) \approx \frac{i(n+1)(2n-1)!!}{(kR)^{n+2}}, \tag{3.2}$$

both of which are given in [27, p.197]. Using these, we obtain

$$\lim_{R\to 0} \frac{j_n'(kR)}{h_n^{(1)'}(kR)} = 0. \tag{3.3}$$

From this result and (2.30), we have

$$\lim_{R\to 0} b_n(k\|\mathbf{x}\|) = j_n(k\|\mathbf{x}\|). \tag{3.4}$$

To demonstrate this relationship, the radius of the continuous spherical microphone r was fixed while the radius of the rigid sphere R, was gradually reduced. The response of $b_n(kr)$ is plotted in Figure 3.5 for different values of R.

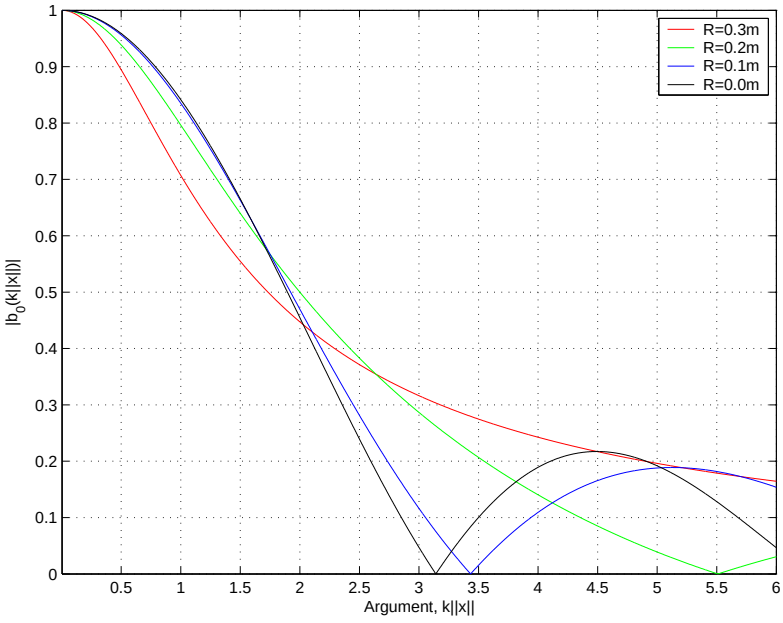


Figure 3.5: Radial response $b_0(k|\mathbf{x}\|)$ for a fixed continuous spherical microphone at radius $\|\mathbf{x}\| = r = 0.3$m, for various radii of the rigid sphere R.

Notice that in the presence of a rigid spherical scatterer, whether the continuous spherical microphone coincides with it or not, the first zero of the radial response $b_0(k|\mathbf{x}\|)$, is shifted to a larger value of $k\|\mathbf{x}\|$ than in the case of an open spherical configuration. In terms of design, this means that a larger range of frequencies can be accurately recorded when either a rigid or an intermediate spherical configuration is used, compared to an open spherical configuration.

3.4 Modal Aliasing

Modal aliasing occurs when the higher order modal components of the sound field are recognised as lower order modal components. This is analogous to temporal aliasing of signals, where higher frequency components are recognised as lower frequency components. Modal aliasing occurs due to insufficient spatial sampling when the theoretical continuous spherical microphone is approximated with a discrete microphone array [2, 15]. The analysis we present in this section is independent of whether we use the open or rigid spherical configurations, described in Section 3.3.

To see how the effects of aliasing are introduced, consider (2.23). The integral can be evaluated with a continuous spherical microphone but since it is not possible to implement one, this can be approximated by a discrete microphone array, say with Q ideal omni-directional point sensor microphones with real weights w_q, distributed over the surface of a sphere of radius R, using the trapezoidal integration method [30]. In Section 3.5, we will show examples that demonstrate that a finite value of Q can only resolve a finite number of modal components of a sound field. For now, assume that Q is sufficiently large and that the microphones are arranged in a way, such that this approximation is *exact* for resolving the modal coefficients of a sound field for orders $0 \leq n < N$. We denote this order truncated sound field that contains only these modal components as $S_{0:(N-1)}(\mathbf{x}; k)$. Then, the modal coefficients for orders $0 \leq n < N$ are

$$\gamma_{nm}(k) = \frac{1}{j_n(kR)} \sum_{q=1}^{Q} S_{0:(N-1)}(R\hat{\mathbf{x}}_q; k) Y_{nm}^*(\hat{\mathbf{x}}_q) w_q = \frac{1}{j_n(kR)} \int S_{0:(N-1)}(R\hat{\mathbf{x}}; k) Y_{nm}^*(\hat{\mathbf{x}}) d\hat{\mathbf{x}}. \tag{3.5}$$

where $j_n(kR)$ is replaced with $b_n(kR)$ if a rigid spherical configuration is considered. However, when the microphone array is exposed to the total sound field $S(\hat{\mathbf{x}}; k)$, the modal coefficients we are concerned with are corrupted by the sound field due to higher order modal components, which we denote as $S_{N:\infty}(\mathbf{x}; k)$. Now, the recorded modal coefficients will be

$$\begin{aligned} \hat{\gamma}_{nm}(k) &= \frac{1}{j_n(kR)} \sum_{q=1}^{Q} \big(S_{0:(N-1)}(R\hat{\mathbf{x}}_q; k) + S_{N:\infty}(R\hat{\mathbf{x}}_q; k) \big) Y_{nm}^*(\hat{\mathbf{x}}_q) w_q & (3.6) \\ &= \gamma_{nm}(k) + \epsilon_{nm}(k), & (3.7) \end{aligned}$$

where, $\epsilon_{nm}(k)$ represents the error due to modal aliasing [2]. This is given by

$$\epsilon_{nm}(k) = \frac{1}{j_n(kR)} \sum_{q=1}^{Q} S_{N:\infty}(R\hat{\mathbf{x}}_q; k) Y_{nm}^*(\hat{\mathbf{x}}_q) w_q. \tag{3.8}$$

By substituting the expanded form of $S_{N:\infty}(R\hat{\mathbf{x}}_q; k)$ of (2.21) into (3.8), we have

$$\epsilon_{nm}(k) = \frac{1}{j_n(kR)} \sum_{n'=N}^{\infty} \sum_{m'=-n'}^{n'} \gamma_{n'm'}(k) j_{n'}(kR) \times \sum_{q=1}^{Q} Y_{n'm'}(\hat{\mathbf{x}}_q) Y_{nm}^*(\hat{\mathbf{x}}_q) w_q. \tag{3.9}$$

Thus far, we can see that modal aliasing is introduced when we use an approximation to the continuous spherical microphone. To observe the case of the continuous spherical microphone,

this is equivalent to taking

$$\lim_{Q\to\infty} \epsilon_{nm}(k) = \frac{1}{j_n(kR)} \sum_{n'=N}^{\infty} \sum_{m'=-n'}^{n'} \gamma_{n'm'}(k) j_{n'}(kR) \times \int Y_{n'm'}(\hat{\mathbf{x}}) Y_{nm}^*(\hat{\mathbf{x}}) d\hat{\mathbf{x}} \;=\; 0, \qquad (3.10)$$

where the orthonormality property of (2.16) was used. Therefore, there will be no error due to modal aliasing for all orders n. This is to say that the continuous spherical microphone is able to resolve all orders exactly. Since practical situations require a finite number microphones, we can conclude that aliasing is inevitable.

By considering the form of (3.9), it can be observed that it is difficult to quantify the effect of modal aliasing. We can, however, restrict the amount of modal aliasing that occurs. Firstly, Figures 3.2 and 3.3 show that for each incremental increase in n, the magnitudes in the baseband regions of $j_n(k\|\mathbf{x}\|)$ and $b_n(k\|\mathbf{x}\|)$, respectively, are significantly smaller. Secondly, (3.9) shows that the error due to modal aliasing comes from higher order modes, specifically the modal components of order $n \geq N$. Since the modal components are of the form $j_n(k\|\mathbf{x}\|)Y_{nm}(\hat{\mathbf{x}})$ or $b_n(k\|\mathbf{x}\|)Y_{nm}(\hat{\mathbf{x}})$ (depending on the spherical configuration used, refer to Section 3.3), we can control the magnitude of these for orders $n \geq N$ by restricting the magnitudes of $j_n(k\|\mathbf{x}\|)$ and $b_n(k\|\mathbf{x}\|)$ for $n \geq N$, respectively.

We can specify an upper bound $(k\|\mathbf{x}\|)_u$, such that in the interval $k\|\mathbf{x}\| = [0, (k\|\mathbf{x}\|)_u]$, the modal components for $n \geq N$ are at least DdB smaller in magnitude compared to the $(N-1)^{\text{th}}$ modal component. Table 3.1 gives argument values for $D = 10$dB for a range of N. Note that for the case of $b_n(kr)$, we assume that $R = r$ (the case of a spherical microphone coincident with a spherical scatterer). Any intermediate configuration have a value that lies between the values of the two configurations.[1] In terms of design, increasing D generally decreases modal aliasing but reduces the value of $(k\|\mathbf{x}\|)_u$.

N	Value of $(k\|\mathbf{x}\|)_u$ for $j_n(k\|\mathbf{x}\|)$	Value of $(k\|\mathbf{x}\|)_u$ for $b_n(k\|\mathbf{x}\|)$
1	0.89	0.55
2	1.47	1.22
3	2.05	1.91
4	2.63	2.59
5	3.21	3.25
6	3.78	3.87
7	4.36	4.46
8	4.94	5.03
9	5.51	5.59
10	6.09	6.16

Table 3.1: Value of $(k\|\mathbf{x}\|)_u$, such that the magnitude of $j_n(k\|\mathbf{x}\|)$ and $b_n(k\|\mathbf{x}\|)$ for $n \geq N$ are at least 10dB smaller than $j_{N-1}(k\|\mathbf{x}\|)$ and $b_{N-1}(k\|\mathbf{x}\|)$, respectively.

Now, let us review our assumption where Q was sufficiently large enough to support (3.5) for $0 \leq n < N$. This can now be specified as a constraint on the positions of the omni-directional point sensor microphones and their weights. This is called the *orthonormality constraint* and

[1] This follows from the relationship between the two configurations, (3.4), as shown in Section 3.3.

is given by [15] as

$$\sum_{q=1}^{Q} Y_{n'm'}(\hat{\mathbf{x}}_q) Y_{nm}^*(\hat{\mathbf{x}}_q) w_q = \delta_{nn'}\delta_{mm'}, \tag{3.11}$$

for $0 \leq n < N$. This constraint ensures that the Q microphones are able to resolve all $\gamma_{nm}(k)$ exactly for this range of n, hence the use of the term, orthonormality. The analysis of (3.11) is not straightforward, and has not been presented in the literature. The approach that shall be taken is to analyse the degree that a given arrangement of microphones satisfies (3.11). This will be deferred to the next section.

3.5 Microphone Arrangements and Weights

As was stated earlier, the orthonormality constraint of (3.11) is difficult to analyse and to directly derive a set of suitable points and weights on a sphere would not be straightforward. However, the form of (2.23) suggests that intuitively, the goal is to find an integration approximation method for a function over a unit sphere, so that we can resolve the modal coefficients of orders $0 \leq n < N$ as accurate as possible using Q microphones. Fortunately, there are numerous ways to arrange a finite number of microphones in a spherical array to provide approximations to the continuous spherical microphone for this purpose. The extent to which these arrangements satisfy (3.11) gives an indication as to the level of accuracy in which the modal coefficients for these orders can be resolved.

We will look at specific configurations in detail, namely, the Gaussian and trigonometric quadrature arrangements as well as the truncated icosahedron arrangement, then lightly introduce cubature arrangements. The merits of each differ and in terms of design, we would chose the one which best fits our needs. Some of these merits include, for example, scalability, physical realisability and efficiency, as will be covered below. In analysing these arrangements, we assume that the microphones are ideal omni-directional point sensors. The positions of the microphones are intrinsically characteristic of the arrangement under consideration, but since we are approximating an integration over the unit sphere, the weights of any arrangement must sum to the surface area of the unit sphere, ie.

$$\sum_{q=1}^{Q} w_q = 4\pi. \tag{3.12}$$

In order to evaluate the performance of an arrangement of microphones, (3.11) shall be numerically analysed with respect to the arrangement. That is, we shall evaluate

$$C_{nm,n'm'} = \sum_{q=1}^{Q} Y_{n'm'}(\hat{\mathbf{x}}_q) Y_{nm}^*(\hat{\mathbf{x}}_q) w_q, \tag{3.13}$$

where we consider $Y_{n'm'}(\hat{\mathbf{x}})$ as *input* spherical harmonic components to be tested against $Y_{nm}^*(\hat{\mathbf{x}})$, the spherical harmonic components which we would like to *resolve*. The ideal is to achieve

$$C_{nm,n'm'} = \delta_{nn'}\delta_{mm'} \tag{3.14}$$

for $0 \leq n, n' < N$ since this corresponds to the right-hand side of orthonormality constraint of

(3.11). As an error measure, we evaluate the upper bound of the deviation from this ideal as

$$\Delta C_{\max} \triangleq \max\left|C_{nm,n'm'} - \delta_{nn'}\delta_{mm'}\right|, \tag{3.15}$$

for $0 \leq n, n' < N$. Essentially, this gives an indication of the extent of integration approximation error.

Equation (3.13) will be presented as a plot of $C_{nm,n'm'}$ as a function of n', m' and n, m. The n', m' and n, m axes will be labelled according to the ascending order of modes as shown in Table 3.2. Colour will be arbitrarily used in these types of plots for ease of reading only.

mode	1	2	3	4	5	6	...	9	...	16	...	25	...
n, m	0,0	1,-1	1,0	1,1	2,-2	2,-1	...	2,2	...	3,3	...	4,4	...

Table 3.2: The ascending order of modes.

3.5.1 Gaussian and Trigonometric Quadrature Arrangement

The Gaussian and trigonometric quadrature arrangements are based on optimal methods of evaluating definite integrals. Firstly, consider a sound field $S_{0:(N-1)}(\mathbf{x}; k)$, that has been order truncated to include modal components of orders $0 \leq n < N$. Then, by the definition of spherical harmonics in (2.15), the highest order of all the modal components in $S_{0:(N-1)}(\mathbf{x}; k)$ is $N-1$. This means that the highest degree of the associated Legendre function of the first kind that we would be concerned with is $N-1$, in terms of the variable $x = \cos\theta$. To evaluate the integral of (2.23), it is necessary to be able to evaluate a definite integral of the form

$$I_{nm} = \int_{-1}^{1} P_{nm}(x) P_{n'm'}(x) dx$$

for all $0 \leq n, n' < N$. The integrand is a polynomial of degree at most $2N-2$ in x. For such an integral, it is possible to evaluate this exactly with N sample points located in $[-1, 1]$, using the Gaussian-Legendre quadrature [31, 32].[2] The points and corresponding weights for some values of N can be found in Table A.1. Translating back to the θ variable, we obtain corresponding sample points in $[0, \pi]$.

As for the ϕ variable, it is necessary to be able to approximate an integral of the form

$$I_m = \int_{0}^{2\pi} h(\phi) e^{im\phi} d\phi$$

for all $-N < m < N$, given that $h(\phi)$ is of the form

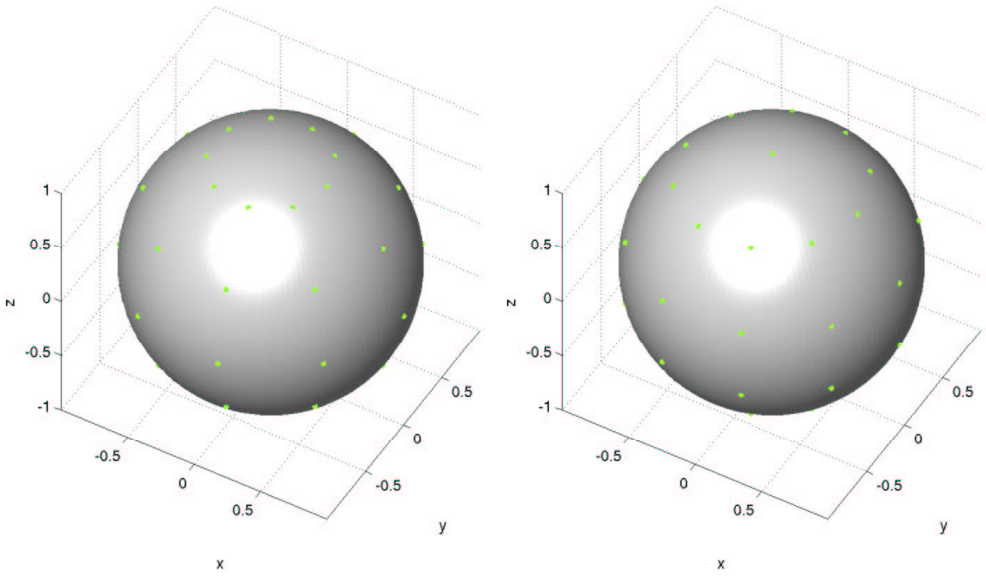
$$h(\phi) = \sum_{|m|<N} a_m e^{im\phi},$$

where a_m are generally complex coefficients. This can be done by using trigonometric quadratures[3], where $K = 2(N-1)+1$ points are equally spaced in $[0, 2\pi]$ with equal weights of $\frac{1}{K}$.

[2] Refer to Appendix A.1 for a more detailed treatment.
[3] Refer to Appendix A.2 for a more detailed treatment.

Therefore, the total number of points required for this arrangement is $2N^2 - N$, to satisfy (3.11) for the spherical harmonic components of orders $0 \leq n < N$. The main advantage of Gaussian and trigonometric quadratures is its scalability since an arrangement can be derived for any value of $N \geq 2$.[4] One drawback to this arrangement is that for higher order designs, the increased packing near the poles may become a practical issue. Fortunately, Mohlenkamp has suggested an approximation termed, "thinning the grid", where it is possible to neglect some points which are very close to the pole [32].

As a way of demonstrating the design process, we chose to obtain a specific arrangement such that the orthonormality constraint of (3.11) holds for all $0 \leq n < 5$. This implies that $N = 5$ and so the number of points required on the θ coordinate is 5, while the number of points required on the ϕ coordinate is 9, a total of 45 points. Figure 3.6(a) shows this arrangement of microphones on the unit sphere. To approximate the integral over the unit sphere with this arrangement, the weights, w_q, were assigned to the products of the weights of the θ and the ϕ coordinates and then scaled such that (3.12) holds.



(a) $K = 9$ with no poles.

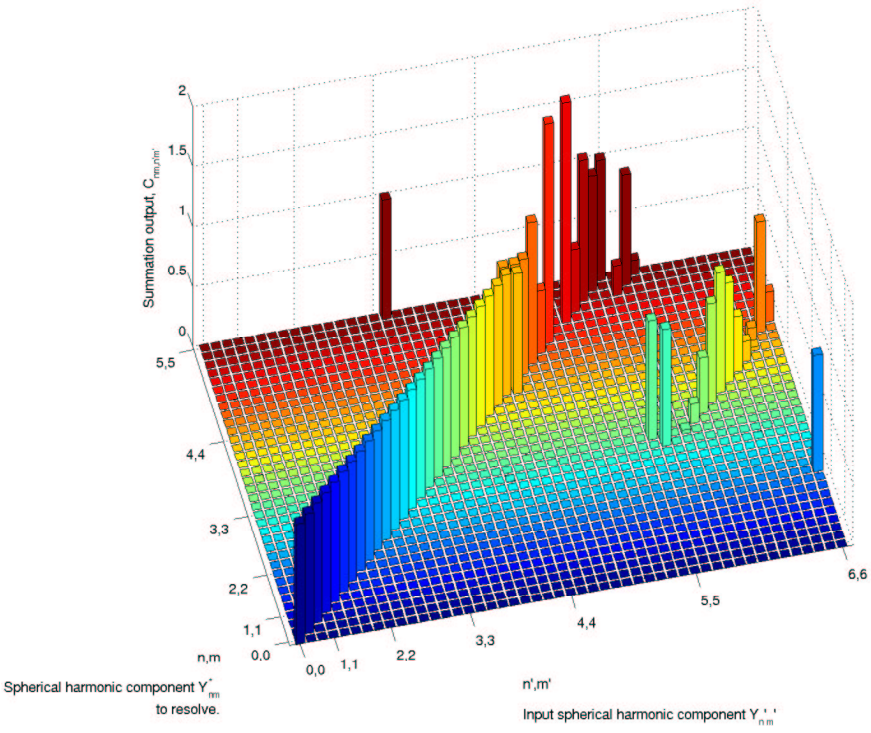
(b) $K = 10$ with two poles.

Figure 3.6: Positions of microphones on the unit sphere (shown in green) for the Gaussian and trigonometric quadratures with $N = 5$.

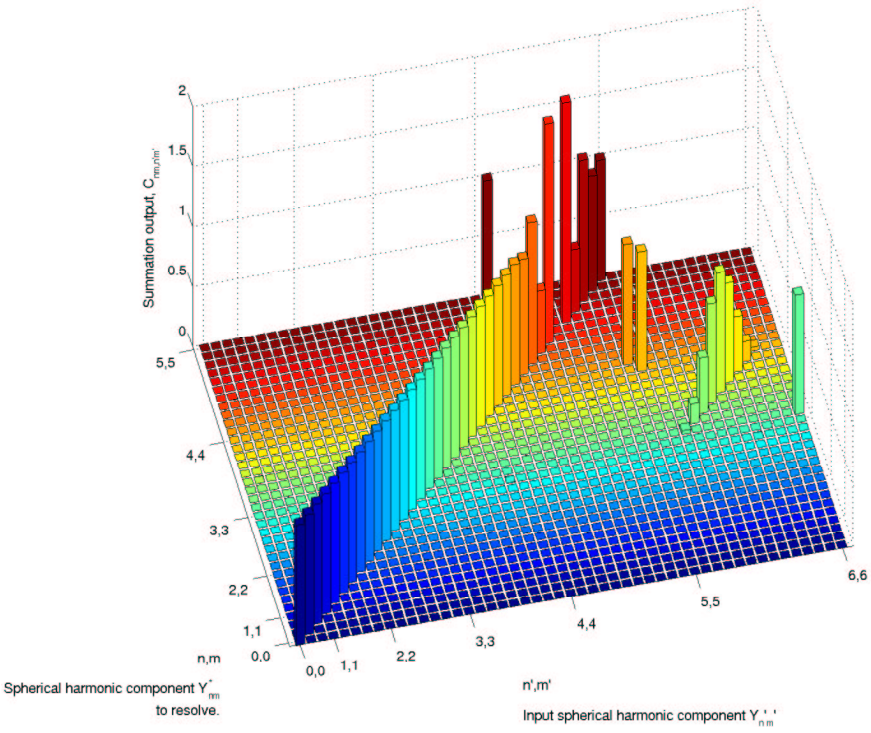
Equation (3.13) was evaluated and plotted in Figure 3.7(a) as a function of n', m' and n, m. The characteristic diagonal row at unity height for spherical harmonic components of orders $0 \leq n < 5$ suggests that these can be resolved at a high level of accuracy. This means that the corresponding modal components can also be resolved to a high level of accuracy.[5] In fact, for this case, $\Delta C_{\max}$ was found to be less than 0.001. One interesting observation is that very few 5^{th} order spherical harmonic components (ie. $Y_{n'm'}(\hat{\mathbf{x}})$ for $n' = 5$) alias *into* components of orders $0 \leq n < 5$ that we are interested in. In fact, only the spherical harmonic component inputs $Y_{n'm'}(\hat{\mathbf{x}})$ for $n', m' = 5, 5$ and $n', m' = 5, -5$ cause such modal aliasing.

[4]Note that the nature of Gaussian quadratures requires that $N \geq 2$ (Refer to Appendix A.1).

[5]Recall that each modal component is a spherical harmonic component multiplied by a radially dependent factor. Here, we assume that the radially dependent factor is sufficiently large (refer to Section 3.3).



(a) $K = 2(N-1)+1$ microphones along the ϕ coordinate.



(b) $K = 2N$ microphones along the ϕ coordinate.

Figure 3.7: Plot of $C_{nm,n'm'}$ given by (3.13) for $n' < 7$ and $n < 6$, for the Gaussian and trigonometric quadrature arrangement with $N = 5$ and no poles.

The above observation motivates the investigation of improving the spatial sampling of the microphone arrangement so that we can obtain no modal aliasing due to spherical harmonic components of order $n = 5$. The advantage of this improvement is that the spherical harmonic components that cause modal aliasing are those of order $n \geq 6$. This allows us to chose a larger value of $(k\|\mathbf{x}\|)_u$ when controlling the magnitudes of the corresponding modal components that cause modal aliasing (refer to Section 3.4). The θ coordinate shall be left unchanged since the quadrature is exact, but along the ϕ coordinate, instead of 9 microphones, the effectiveness of 10 microphones shall be investigated. The positions of this arrangement of microphones is characteristically the same as Figure 3.6(a) and therefore, will not be plotted. Figure 3.7(b) is a plot of $C_{nm,n'm'}$ as a function of n', m' and n, m, for this new arrangement. Notice that the modal aliasing caused by the spherical harmonic components $n', m' = 5, 5$ and $n', m' = 5, -5$ are now removed. $\Delta C_{\max}$ was also found to be less than 0.001. In general, this improvement can be achieved for any order design N, by adding an additional microphone to the ϕ coordinate so that $K = 2N$.

The next challenge is to maintain the performance achieved thus far and to improve efficiency by using less microphones. To do this, we propose that on the θ coordinate, we choose a Gaussian quadrature that must include as two of its sample points, the two end points, $x = 1$ and $x = -1$. These points are conveniently located at the poles of the sphere (ie. $x = 1$ corresponds to $\theta = 0$, and $x = -1$ corresponds to $\theta = \pi$).[6] It can be numerically shown that polynomials of degree up to $2N - 1$ require $N + 1$ points if the two end points are used. Therefore, following from our previous design of $N = 5$, we now require 6 sample points on the θ coordinate. In general, this arrangement requires $2N^2 - 3N + 3$ points in order to be able to resolve all the spherical harmonic components of orders $0 \leq n < N$. The reduction in the number of sample points required is due to spherical geometry, where only one sample point is required at each pole. Each pole is weighted K times more than if it were not an end point, which is as if there were K separate sensors at the same point. This arrangement requires 38 sample points. Again, to obtain no modal aliasing due to spherical harmonic components of order $n = 5$, an additional microphone shall be added to the ϕ coordinate, which brings the total number of sample points for this arrangement to 42. This arrangement is shown in Figure 3.6(b).

Visually, the plot of $C_{nm,n'm'}$ for this arrangement is characteristically identical to Figure 3.7(b) and will not be reproduced. It was found that $\Delta C_{\max} = 0.018$, which is more than a magnitude larger than for the arrangements considered previously. Therefore, we can observe a tradeoff between achieving better performance (lower $\Delta C_{\max}$) and improving efficiency (smaller number of microphones). This can be explained as the result of changing the level of spatial sampling used. A summary of the number of points required under different conditions is given in Table 3.3. No advantages was found when only one pole was used and shall not be further investigated. Overall, the total number of points required in Gaussian and trigonometric quadrature arrangements is approximately one half of the number of points required for the equiangular arrangements suggested in [2].

[6] Refer to Appendix A.1.

Condition	Points required
ϕ coordinate	$K = 2(N-1)+1$
ϕ coordinate with no aliasing due to N^{th} order spherical harmonic components	$K = 2N$
θ coordinate with no end points	N
θ coordinate with two end points	$N+1$
Arrangement with no poles	$2N^2 - N$
Arrangement with no poles, with no aliasing due to N^{th} order spherical harmonic components	$2N^2$
Arrangement with two poles	$2N^2 - 3N + 3$
Arrangement with two poles, with no aliasing due to N^{th} order spherical harmonic components	$2N^2 - 2N + 2$

Table 3.3: Number of points required for the Gaussian and trigonometric quadrature under various conditions. In each case, the aim is to be able to resolve all the spherical harmonic components of orders $0 \leq n < N$.

3.5.2 Truncated Icosahedron Arrangement

Originally suggested by Elko and Meyer in [15], the position of microphones are placed in the centre of the faces of a truncated icosahedron, or better known informally as the soccer ball. This is shown in Figure 3.8.

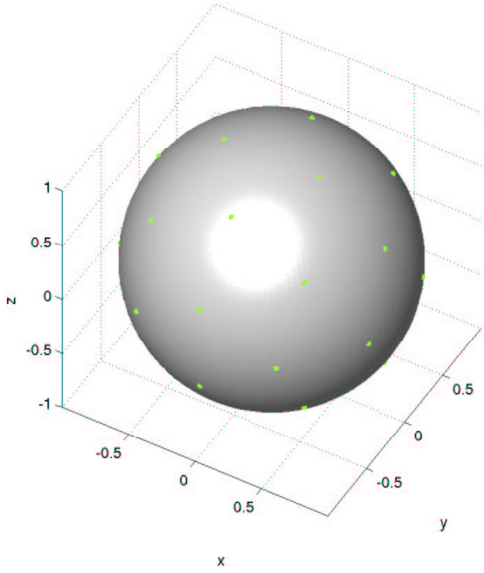


Figure 3.8: Positions of microphones on the unit sphere (shown in green) for the truncated icosahedron arrangement.

Notice that the microphones are more equally spaced on the surface of the sphere than those proposed in Section 3.5.1. In fact, as shall be shown below, this arrangement is more efficient for an $N = 5$ design. However, unlike the Gaussian and trigonometric quadrature, this arrangement lacks scalability and thus cannot be adapted for higher order designs. Having obtained the points corresponding to the truncated icosahedron arrangement, a set of weights

would need to be determined to approximate an integral over the unit sphere. Elko and Meyer did not address this issue in [15] so one could only assume that equal weights that satisfied (3.12) were used in their work. Evaluating (3.15) for this case, gave $\Delta C_{\max} = 0.028$. This turns out to be about two magnitudes higher than for the Gaussian and trigonometric arrangements. This is largely a result of using inappropriate weights, as shall be shown below.

Instead of equal weights, a set of weights shall be derived based on the scheme that the weight for a point is proportional to the average of the distances to all the neighbouring points. In our calculation, we shall deal with distances along the surface of a truncated icosahedron rather than the actual Euclidean distances for ease of calculation. The symmetry of this arrangement allows for this simplification. Firstly consider a pentagonal and a hexagonal face laid out in a plane, as shown in Figure 3.9. This shows two sample points located at the centre of each face.

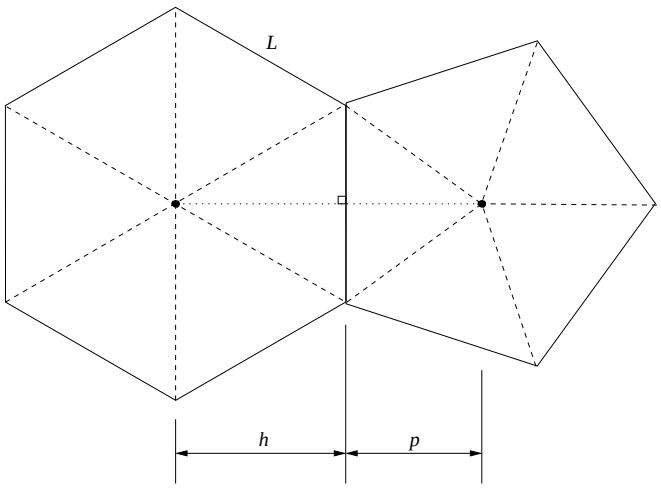


Figure 3.9: Layout of a hexagonal and pentagonal face connected by a common edge.

Let h denote the length from the centre of the hexagonal face to the perpendicular intercept at the edge. Similarly, let p be the length from the centre of the pentagonal face to the perpendicular intercept at the edge. Therefore, the length between any two points is either:

- $(h + p)$, when one point lies in the centre of a hexagonal face and the other lies in the centre of a pentagonal face, or

- $2h$, when both points lie in the centre of a hexagonal faces.

Now note that each of the five neighbouring points, of a point located at the centre of a pentagonal face, are located at the centres of hexagonal faces. Also, three of the six neighbouring points, of a point located at the centre of a hexagonal face, are located at the centres of hexagonal faces, while the rest are located at the centres of pentagonal faces. With this information, if we denote $\bar{d}_h$ as the average distance to neighbouring points from a point located at the centre of a hexagonal face, and $\bar{d}_p$ as the average distance to neighbouring points from a point

located at the centre of a pentagonal face, the ratio

$$\frac{\bar{d}_p}{\bar{d}_h} = \frac{\frac{1}{5}[5(h+p)]}{\frac{1}{6}[3(h+p)+3(2h)]} = 0.9459. \tag{3.16}$$

The expressions, $h = \frac{1/2}{\tan 30^\circ}L$ and $p = \frac{1/2}{\tan 36^\circ}L$, where L is the edge length are derived from basic trigonometry, and were substituted into (3.16). The final step is to normalise the weights such that (3.12) holds, to approximate integrals over the unit sphere. Based on the weighting scheme adopted, the weights w_h and w_p for points located at the centre of a hexagonal and pentagonal faces, respectively, are given by

$$w_h = \frac{\bar{d}_h}{20\bar{d}_h + 12\bar{d}_p} 4\pi = 0.4008 \tag{3.17}$$

and

$$w_p = \frac{\bar{d}_p}{20\bar{d}_h + 12\bar{d}_p} 4\pi = 0.3791, \tag{3.18}$$

where (3.16) was used and given that there are 20 hexagonal faces and 12 pentagonal faces on a truncated icosahedron.

The plot of $C_{nm,n'm'}$ for the truncated icosahedron arrangement using the weights given in (3.17) and (3.18), is shown in Figure 3.10. For this arrangement $\Delta C_{\max} = 0.008$, which is around one-third smaller than the value obtained with equal weights. By comparing Figure 3.10 with Figure 3.7(b), we find that the performance in terms of orthogonality and modal aliasing are similar. However, the truncated icosahedron arrangement requires only 32 sample points, compared with 42 and 50 for the Gaussian and trigonometric arrangements ($N = 5$, $K = 10$) with two poles and no poles, respectively.

Related to this arrangement are uniform discretizations, which place the points on the vertices of a platonic solid. However, there are only a few of these solids and none can support more than 30 points [32]. Therefore, they are only suitable for lower order designs.

3.5.3 Cubature Arrangement

Sample points and corresponding weights were produced by Fliege in [33] for integration over a sphere based on cubature formulae. However, in the evaluation of $C_{nm,n'm'}$, the efficiency was found to be significantly poorer that the arrangements discussed above. Therefore, further analysis of these arrangements shall not be performed. However, when designing for high orders, the Gaussian and trigonometric quadrature method poses practical problems of microphone packing near the poles, while the truncated icosahedron arrangement is not extensible. Cubature arrangements can be used in this situation and would be a topic for further research.

3.5.4 Sphere Packing

Sets of points based on maximising the minimal distance between the points on a sphere were proposed in [34], but no attempt was made to provide weights for integration over a sphere. These will not be further investigated in this thesis, but could be the topic of future work.

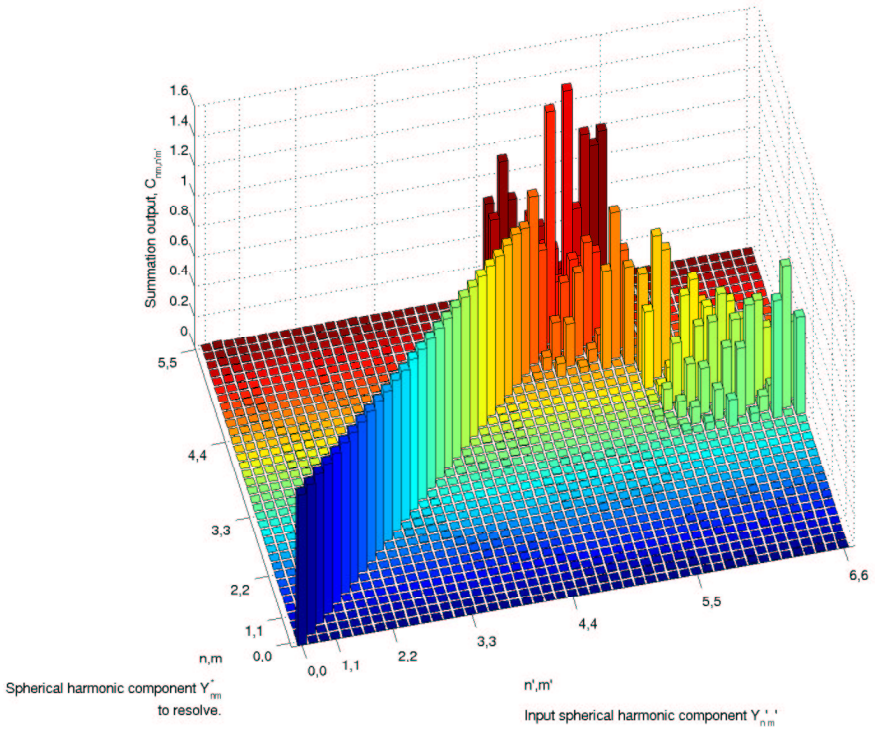


Figure 3.10: Plot of $C_{nm,n'm'}$ given by (3.13) for $n' < 7$ and $n < 6$ for the icosahedron arrangement.

3.6 Calibration Errors

The microphone arrangements presented in Section 3.5 achieved their specified performance when the microphones are positioned exactly. However, in practice this is not possible. In this section, we analyse the effect of calibration errors due to inexact microphone positioning.

Although we do not expect the Q microphones to lie in their actual positions $\mathbf{x}_q$ for $q = 1, \ldots, Q$, we would like to specify them in practice, to be "very likely" to lie somewhere within a specified distance $L_{\max}$ units *along the surface* of the unit sphere from their actual positions. Consider the q^{th} microphone position, $\mathbf{x}_q$. We can parameterise a relative offset from $\mathbf{x}_q$ by using a distance variable L_q units along the surface of the unit sphere at an angle A_q, relative to $\mathbf{x}_q$.[7]

For this single microphone position, we would like to model calibration error by assigning L_q and A_q to random variables. For the case of L_q, we firstly consider the Gaussian distributed random variable, $Z \sim \mathcal{N}(0, \sigma)$, where the mean is zero and standard deviation is σ. Let $f_Z(z)$ be the corresponding probability density function (PDF) of Z. Then, we assign

$$L_q = g(Z) = |Z|, \tag{3.19}$$

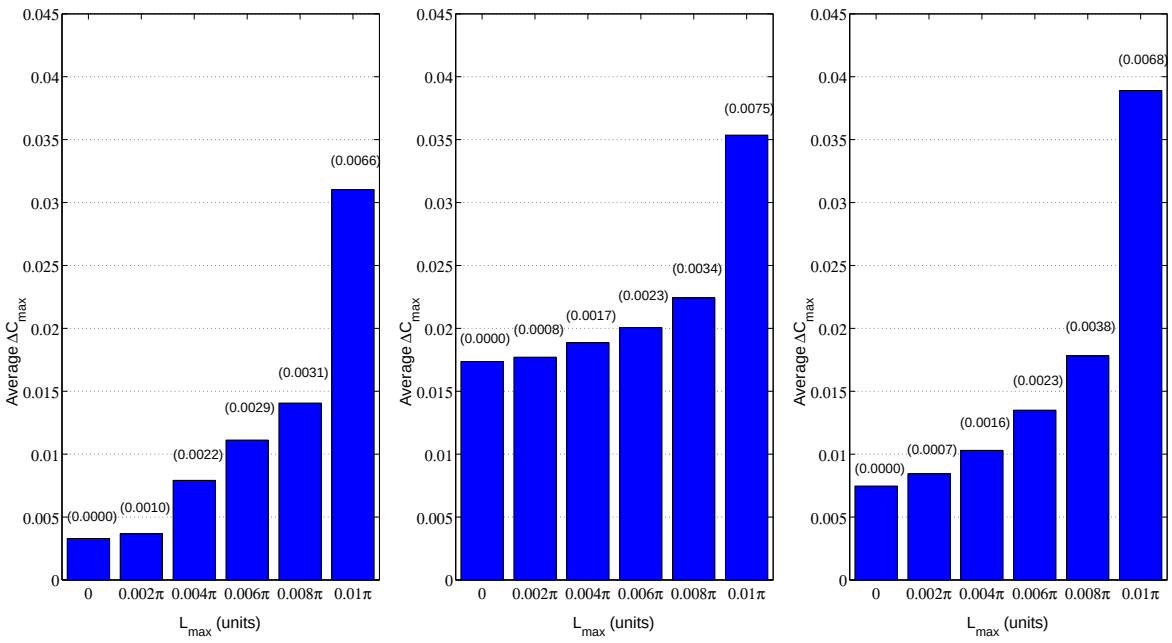
[7]Figure C.1 in Appendix C indicates the reference for A_q.

since we are only concerned with non-negative values for L_q. Therefore, the PDF of L_q is

$$f_{L_q}(l_q) = \begin{cases} 2f_Z(l_q), & \text{if } l_q \geq 0 \\ 0, & \text{otherwise} \end{cases} . \tag{3.20}$$

The proof is shown in Appendix B. Now, it is possible to be specific about the meaning of the phrase, "very likely", as used above. As an example of a specification, we would like to have approximately 99.7% of the area under the PDF of L_q to lie in $[0, L_{\max}]$, and therefore, we assign $\sigma = L_{\max}/3$.[8] As for A_q, we simply want to assign it to a uniformly distributed random variable in $[0, 2\pi)$. In order to "randomise" our microphone position in preparation for calibration error calculations, $\mathbf{x}_q$ was offset by a distance L_q units along the surface of the unit sphere at angles A_q. Appendix C shows how this was performed. This was repeated for $q = 1, \ldots, Q$.

To assess the effect of calibration errors, for the microphone arrangements described in Section 3.5, we considered various values of $L_{\max}$ between 0 and 0.01π. For each $L_{\max}$, we evaluated 40 samples of $\Delta C_{\max}$ in (3.15) and obtained from these, the mean and standard deviation of $\Delta C_{\max}$. The means (and standard deviations in parentheses) are plotted in Figures 3.11(a) to 3.11(c) for each arrangement. The sets of $\Delta C_{\max}$ can be modelled by the gamma distribution, as shown Appendix D.



(a) Gaussian and trigonometric, with no poles.

(b) Gaussian and trigonometric, with two poles.

(c) Truncated Icosahedron.

Figure 3.11: Plots of means of $\Delta C_{\max}$ for various values of $L_{\max}$ (and standard deviation in parentheses), for three microphone arrangements. ($N = 5$, $K = 10$ for the two Gaussian and trigonometric arrangements).

We can observe that both the mean and standard deviation of $\Delta C_{\max}$ increases steadily

[8] 99.7% of the area under a Gaussian distribution corresponds to three standard deviations [35]. Conveniently, this also holds for (3.20).

with $L_{\max}$ for all arrangements until a large increase occurs at $L_{\max} = 0.01\pi$. This observation suggests that to keep calibration error under a narrow allowable range, it is important to specify small values for $L_{\max}$. These results can be generalised to be applied to positioning of microphones on a sphere of any radius r. That is, if we decide on a value of $L_{\max}$ that corresponds to our desired performance of $\Delta C_{\max}$, then we need to specify

$$L_{\max}^{'} = L_{\max} r, \tag{3.21}$$

for our spherical microphone array of radius r. The relationship, $s = r\theta$, was used where θ is the angle subtended by an arc of length s in a sector of side length r.

3.7 Signal Processing

Thus far, we have considered the various issues related to recording the modal coefficients of a sound field. Specifically, it was shown in Sections 3.4 and 3.5 that the continuous spherical microphone could be approximated by a discrete microphone array so that the modal coefficients could be approximately recorded according to the equation,

$$\hat{\gamma}_{nm}(k) \approx \frac{1}{b_n(kR)} \sum_{q=1}^{Q} S(R\hat{\mathbf{x}}_q; k) Y_{nm}^*(\hat{\mathbf{x}}_q) w_q. \tag{3.22}$$

Now, we shall present this equation in a block diagram form, as shown in Figure 3.12, to allow the reader to appreciate the signal processing involved in recording these modal coefficients. In the process of translating (3.22) into Figure 3.12, the equation should be viewed as a frequency domain representation. For example, frequency dependent terms are viewed as transfer functions of temporal filters. The terms, w_q and $Y_{nm}^*(\hat{\mathbf{x}}_q)$ for $q = 1, \ldots, Q$, are frequency independent and so appear as constant gains, but the terms $1/b_n(kR)$, are frequency dependent and so they are transfer functions of temporal filters.

The modal coefficient outputs, $\hat{\gamma}_{nm}(k)$, are functions of frequency so they can also be viewed as generally complex signals in the time domain. Therefore, we can apply temporal low pass or band pass filters to these signals so to selectively filter out unwanted frequency components. This is useful because, as shown in Section 3.3, the recorded coefficients become inaccurate at the lower frequencies. These temporal filters can be used to select different frequency bands from microphone arrays of different radii, to obtain an accurate recording over a larger frequency band.

3.8 Directional Beamformer

Having obtained the modal coefficients, $\hat{\gamma}_{nm}(k)$, we can use these directly in beamforming applications. We specify a *beampattern*, which is much like a spatial filter that can be defined to amplify or attenuate the signal as a function of θ and ϕ. The beampattern can be fully encapsulated by a set of weights, which are applied to the modal coefficients. In essence, beamforming involves taking a weighted sum of the modal coefficients to steer the microphone array towards a particular direction. Further details of beamforming are abundant in the literature and a few have been mentioned earlier in Section 1.2.

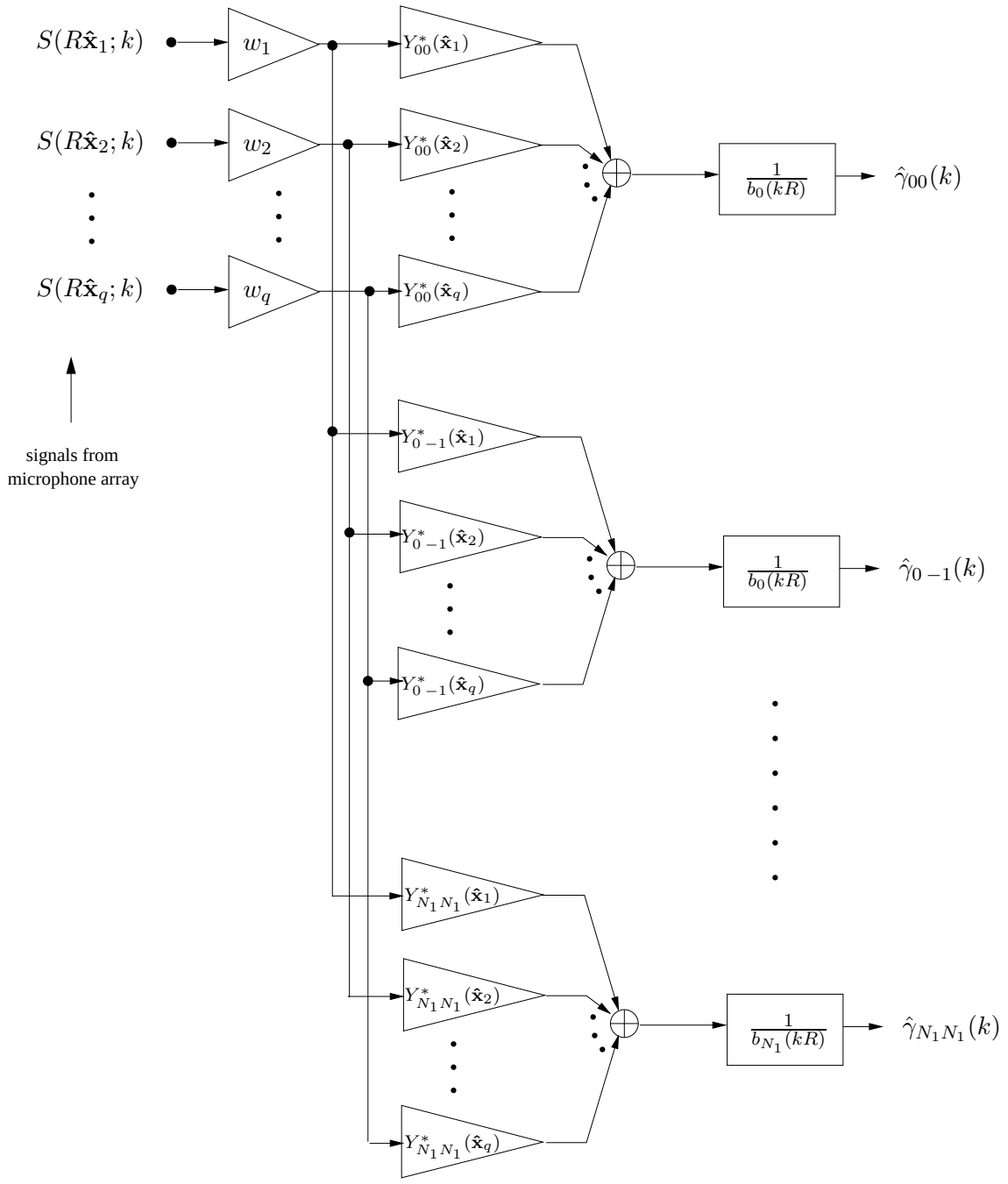


Figure 3.12: Block diagram showing the signal processing required to resolve the modal coefficients. If the coefficients for orders $0 \leq n < N$ are desired, then we set $N_1 = (N-1)$.

3.9 Summary

We have presented the main design issues that would need to be considered in spherical microphone design. However, we are yet to use them in a meaningful context, such as applying them together in a specific design. This will be the focus of the next chapter.

Chapter 4

Simulation of a Fourth Order Microphone Design

4.1 Introduction

Chapter 3 dealt with design issues based around spherical microphone design. In order to verify, integrate and evaluate them as a whole, we produce a specific design and simulate it in Mathworks MATLAB in this chapter. Before we can move forward in assessing a design's performance, we begin by proposing some quantitative error measures as well as a standard in which to apply them. In terms of the design itself, we will take on an iterative process, which means that the details of its performance will be presented along the way. Finally, further quantitative and qualitative analysis will be presented.

4.2 Error Measures

There are a few sources of error in which we would be interested in assessing a design. These are listed below and are explained in the indicated sections:

1. Truncation error (Section 4.2.1)
2. Error of recording (Section 4.2.2)
3. Aliasing error (Section 4.2.3)

All these errors characterise the spatial quality of the sound field as opposed to temporal quality and therefore, do not affect the intelligibility of the sound. The first error measure, the truncation error, differs from the other two errors because it is a general assessment of order truncation of sound fields. The last two measures assess characteristics which are intrinsic to the microphone design itself. Both ignore the effects of order truncation in order to assess other sources of error. Notice that all these error measures vary with the frequency component k of the sound field and the distance from the origin, $\|\mathbf{x}\|$. All plots of these errors will be displayed in percentages.

Specifically, the last two measures employ the concept of a recorded sound field, $\tilde{S}(\mathbf{x}; k)$. The tilde in this expression denotes that the sound field was *perfectly reconstructed* from recorded modal coefficients. This means that no limitations relating to sound field reproduction are

considered and so this allows us to concentrate on the issues relating to recording only.[1] The subscript in an expression of the form $S_{0:(N-1)}(\mathbf{x};k)$ denotes a sound field that is order truncated to include modal components of orders $0 \leq n < N$ of the original sound field only. This notation is also applied to recorded sound fields to indicate the orders that were recorded.

4.2.1 Truncation Error

Truncation error is an inherent limitation in any practical design and was described earlier in Section 3.2. The fact that a design is specified to record modal coefficients of orders $0 \leq n < N$ means that the loss of information regarding the orders $n \geq N$ will introduce errors. This error was used in [19] and is intuitively defined as

$$\epsilon_{\text{trunc}}(\|\mathbf{x}\|;k) = \frac{\int \left| S(\mathbf{x};k) - S_{0:(N-1)}(\mathbf{x};k) \right|^2 d\hat{\mathbf{x}}}{\int \left| S(\mathbf{x};k) \right|^2 d\hat{\mathbf{x}}}, \tag{4.1}$$

where $S(\mathbf{x};k)$ is the original sound field and $S_{0:(N-1)}(\mathbf{x};k)$ is the original sound field order truncated to include modal components of orders $0 \leq n < N$. Equation (4.1) can be simplified to

$$\epsilon_{\text{trunc}}(\|\mathbf{x}\|;k) = 1 - \frac{\sum_{n=0}^{N-1} \sum_{m=-n}^{n} \left| \gamma_{nm}(k) j_n(k|\mathbf{x}\|) \right|^2}{\sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left| \gamma_{nm}(k) j_n(k|\mathbf{x}\|) \right|^2}, \tag{4.2}$$

where $\gamma_{nm}(k)$ are the modal coefficients of the original sound field, $S(\mathbf{x};k)$. The proof of (4.2) is shown in Appendix E.3.

4.2.2 Error of Recording

The error of recording includes two sources of error. These are the error due to the inherent limitations of the use of a single spherical microphone array[2] and the error due to the integration approximation of the continuous spherical microphone. This is defined as

$$\epsilon_{\text{rec}}(\|\mathbf{x}\|;k) = \frac{\int \left| S_{0:(N-1)}(\mathbf{x};k) - \tilde{S}_{0:(N-1)}(\mathbf{x};k) \right|^2 d\hat{\mathbf{x}}}{\int \left| S_{0:(N-1)}(\mathbf{x};k) \right|^2 d\hat{\mathbf{x}}}, \tag{4.3}$$

where $S_{0:(N-1)}(\mathbf{x};k)$ is the original sound field order truncated to include modal components of orders $0 \leq n < N$, and $\tilde{S}_{0:(N-1)}(\mathbf{x};k)$ is the perfectly reconstructed sound field from the modal coefficients recorded from $S_{0:(N-1)}(\mathbf{x};k)$. Equation (4.3) can be simplified to

$$\epsilon_{\text{rec}}(\|\mathbf{x}\|;k) = \frac{\sum_{n=0}^{N-1} \sum_{m=-n}^{n} \left| \big(\gamma_{nm}(k) - \tilde{\gamma}_{nm}(k)\big) j_n(k\|\mathbf{x}\|) \right|^2}{\sum_{n=0}^{N-1} \sum_{m=-n}^{n} \left| \gamma_{nm}(k) j_n(k\|\mathbf{x}\|) \right|^2}, \tag{4.4}$$

[1] As stated in Chapter 1, sound field reproduction shall not be investigated in this thesis.
[2] Refer to Section 3.3.

where $\gamma_{nm}(k)$ are the modal coefficients corresponding to $S_{0:(N-1)}(\mathbf{x};k)$ and $\tilde{\gamma}_{nm}(k)$ are the recorded modal coefficients corresponding to $\tilde{S}_{0:(N-1)}(\mathbf{x};k)$. The proof of (4.4) is shown in Appendix E.1. The order truncation of the original sound field is necessary so to avoid including truncation errors.

4.2.3 Aliasing Error

Aliasing error is a measure of the extent of modal aliasing (Refer to Section 3.4). This error has been described theoretically in (3.8) but is not a practical measure. Hence we define this new error measure as

$$\epsilon_{\text{alias}}(\|\mathbf{x}\|;k) = \frac{\int \left|\tilde{S}_{0:(N-1)}(\mathbf{x};k) - \hat{S}_{0:(N-1)}(\mathbf{x};k)\right|^2 d\hat{\mathbf{x}}}{\int \left|S_{0:(N-1)}(\mathbf{x};k)\right|^2 d\hat{\mathbf{x}}}, \tag{4.5}$$

where $S_{0:(N-1)}(\mathbf{x};k)$ is the original sound field order truncated to include modal components of orders $0 \leq n < N$, $\tilde{S}_{0:(N-1)}(\mathbf{x};k)$ is the perfectly reconstructed sound field from the modal coefficients recorded from $S_{0:(N-1)}(\mathbf{x};k)$ and $\hat{S}_{0:(N-1)}(\mathbf{x};k)$ is the perfectly reconstructed sound field from the modal coefficients recorded from the original sound field, $S(\mathbf{x};k)$. Equation (4.5) can be simplified to

$$\epsilon_{\text{alias}}(\|\mathbf{x}\|;k) = \frac{\sum_{n=0}^{N-1}\sum_{m=-n}^{n}\left|\big(\tilde{\gamma}_{nm}(k) - \hat{\gamma}_{nm}(k)\big)j_n(k\|\mathbf{x}\|)\right|^2}{\sum_{n=0}^{N-1}\sum_{m=-n}^{n}\left|\gamma_{nm}(k)j_n(k\|\mathbf{x}\|)\right|^2}, \tag{4.6}$$

where $\gamma_{nm}(k)$ are the modal coefficients corresponding to $S_{0:(N-1)}(\mathbf{x};k)$, $\tilde{\gamma}_{nm}(k)$ are the recorded modal coefficients corresponding to $\tilde{S}_{0:(N-1)}(\mathbf{x};k)$ and $\hat{\gamma}_{nm}(k)$ are the recorded modal coefficients corresponding to $\hat{S}_{0:(N-1)}(\mathbf{x};k)$. The proof of (4.6) is shown in Appendix E.2.

4.3 Test Sound Field

All the error measures in Section 4.2 are dependent on the input sound field $S(\mathbf{x};k)$. In order to compare designs, these will be calculated with respect to a predefined standard input sound field. We assign $S(\mathbf{x};k)$ to be an arbitrary plane wave. A reason for this choice is so to ensure that the temporal frequency content of the sound field is known. Since the microphone arrangements are designed to approximate (2.23) for $0 \leq n < N$, which is an integral over a sphere, we assume that a plane wave travelling in a particular direction would be representative of a plane wave from any direction. We expect that this is true for the error of recording but not for the aliasing error because higher order modal components are introduced in the calculation of aliasing error. Therefore, the characteristics observable from the aliasing error results are representative but the numerical values should be taken as indicative only.

The plane wave can be approximated using a point source, $\mathbf{y}_q$, at a large distance from the origin, whilst the angular location of the point source can be arbitrarily chosen. Specifically, we arbitrarily assign $\mathbf{y}_q$ to $(r,\theta,\phi) = (1000, \pi/2, \pi/3)$. The sound field is mathematically given

by (2.20).

The modal coefficients of the original plane wave sound field, $\gamma_{nm}(k)$, can be obtained from (2.22), and the recording of the modal coefficients $\tilde{\gamma}_{nm}(k)$ from this plane wave can be obtained by performing a simulation on Mathworks MATLAB using the design under consideration. The frequency k of the plane wave is a parameter which can be varied to assess the design's response to frequency.

4.4 Initial Design

We begin by outlining a specification for our design. Firstly, we would like to extend beyond the capabilities of existing designs (refer to Section 1.1) by choosing to record the modal components of the sound field for orders $0 \leq n < 5$. That is, we would like to design a fourth order microphone. The portion of the sound field which we would like to restrict our attention to is a spherical region of radius 0.5m. A rigid spherical configuration shall be used where the continuous spherical microphone coincides with the rigid spherical scatterer (Refer to Section 3.3). We would like to record a $10:1$ frequency range of $f = [340, 3400]$Hz.

In order to approximate the continuous spherical microphone, we chose to use the truncated icosahedron arrangement since this arrangement satisfies the orthonormality constraint of (3.11) for $N = 5$ and possesses the ability to prevent aliasing from 5^{th} order modal components. Furthermore, it is the most efficient of the options described previously, requiring only 32 microphones (refer to Section 3.5.2). We assume that there are no calibration errors (Section 3.6) in our design. This will be analysed separately in Section 4.6.4.

To specify the level of accuracy desired, we will set two constraints. For a given microphone radius r, these constraints limit the range of frequencies k which are recorded. The subscript u shall denote the upper bound and the subscript l shall denote the lower bound of a quantity. The constraints are

1. $(kr)_u$ is assigned the value given in Table 3.1 for $N = 6$ to minimise modal aliasing ($N = 6$ was chosen because no 5^{th} order modal components cause modal aliasing into modal components of orders $0 \leq n < 5$).[3]

2. $(kr)_l$ is assigned the minimum value of kr such that $\epsilon_{\text{rec}}(\|\mathbf{x}\|; k) \leq \epsilon_{\text{rec,max}}$ holds, to limit the error of recording.[4] For argument sake, let $\epsilon_{\text{rec,max}}$ =5%.

These constraints imply a level of accuracy of recording the modal coefficients for the frequency range, $[k_l, k_u]$ or equivalently, $[f_l, f_u]$.[5]

The analysis of the first constraint directly affects the design of the radius of the microphone. To minimise modal aliasing from orders $n \geq 6$, we chose $(kr)_u = 3.87$ (which corresponds to

[3] For further explanation for the requirement of this constraint, refer to Section 3.4.
[4] For further explanation for the requirement of this constraint, refer to Section 3.3.
[5] k and f are related by (2.6).

$N = 6$). Therefore,

$$k_u r = 3.87 \tag{4.7}$$

$$r = \frac{3.87}{k_u}$$

$$r = \frac{3.87c}{2\pi f_u}, \tag{4.8}$$

where (2.6) was used. Substituting $f_u = 3400$Hz into (4.8) gives a radius of $r = 6.2$cm.[6] The application of this constraint means that an upper bound has been placed on the aliasing error within the frequencies in $[0, f_u]$. The actual level of modal aliasing that occurs is quantified in Section 4.6.3.

We now address the second constraint. We no longer have the freedom to vary r since we have already chosen a value above. Therefore, this constraint will specify a value of f_l. The error of recording given in (4.4), was calculated for this design as a function of f, and is plotted in Figure 4.1 at various radii from the origin.[7]

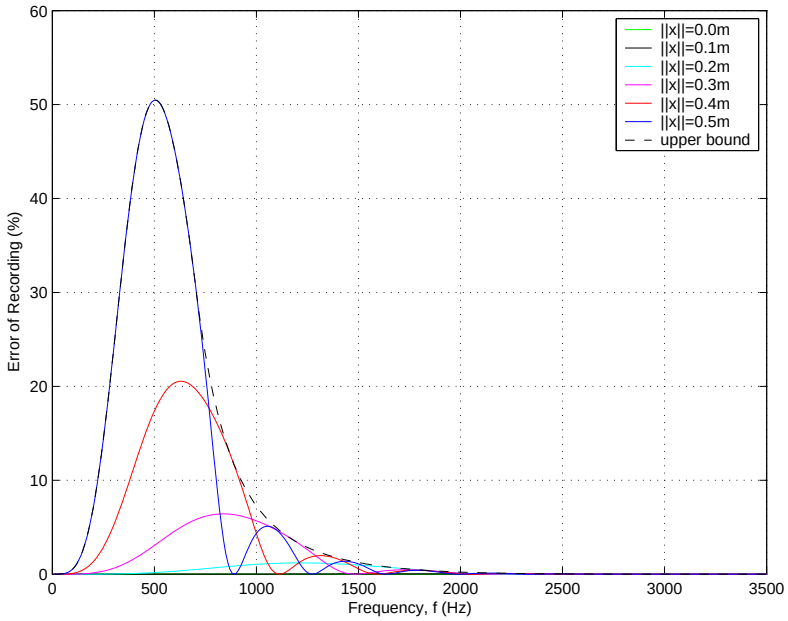


Figure 4.1: Error of recording at various radii, for microphone array at $r = 6.2$cm with coincident rigid spherical scatterer. The dotted black line indicates the upper bound to error for $0 \leq \|\mathbf{x}\| \leq 0.5$m.

The peaking behaviour of the curves can be explained by the fact that as we move towards lower frequencies, the magnitude of $b_n(k\|\mathbf{x}\|)$ for $n \geq 2$ becomes small. This means that the error increases due to the absence of modal components at these orders. However, the lowering of frequency means that the variation in space within a given area is slower and at zero frequency, there would be no variation in space. This causes a reduction in error and the peaking behaviour can be explained as the interaction between the two. However, this peaking

[6] We assume the speed of propagation through air is $c = 342\text{ms}^{-1}$.

[7] The frequency variable f was used instead of k because it is more intuitive and meaningful.

behaviour is characteristic of the truncated icosahedron arrangement *and only* when there are no calibration errors, as will be shown in Section 4.6.4.

Note that in Figure 4.1, only $0 \leq \|\mathbf{x}\| \leq 0.5$m were considered. However, the obvious trend is that when we look towards the lower frequencies and move further away from the origin, the error of recording rises rapidly. By considering this trend and the fact that only a spherical region of radius 0.5m is considered, this allows us to assign a frequency f_l in which to ensure that the second constraint is satisfied. By observation, we chose $f_l = 1100$Hz.

Therefore, the microphone design satisfies the specifications and constraints outlined above for the frequencies in the range $f = [f_l, f_u] = [1100, 3400]$Hz.

4.5 Finalisation of Design

Thus far, the design does not satisfy the intended frequency range. The approach to solving this problem is to employ a second microphone array at a larger radius in order to capture the lower frequencies in $f = [340, 1100]$Hz. This time, the only option is to use an open spherical microphone, which encases the inner microphone array. This is an intermediate configuration as described in Section 3.3. Again, the truncated icosahedron arrangement is chosen for its efficiency in recording modal coefficient of orders $0 \leq n < 5$.

To determine a suitable radius, we look at the first constraint to determine a suitable value of $(k'r')_u$. The value suggested in Table 3.1 corresponding to $b_n(k\|\mathbf{x}\|)$ is no longer suitable for this configuration. Knowing the relationship between $j_n(k\|\mathbf{x}\|)$ and $b_n(k\|\mathbf{x}\|)$ (refer to Figure 3.5), we chose to be conservative by assigning the value of $(k'r')_u$ that corresponds to $j_n(k\|\mathbf{x}\|)$. However, there is a zero in the function $j_0(k\|\mathbf{x}\|)$ at $k\|\mathbf{x}\| = 3.14$ (refer to Figure 3.2), so instead, we must assign $(k'r')_u$ to a smaller value of 3.14. Thus, we require

$$k'_u r' = 3.14 \tag{4.9}$$

$$r' = \frac{3.14}{k'_u}$$

$$r' = \frac{3.14c}{2\pi f'_u}. \tag{4.10}$$

Substituting $f'_u = 1100$Hz into (4.10) gives a radius of $r' = 15.5$cm.

The error of recording for this microphone array is shown in Figure 4.2. In a similar way to the procedure employed in Section 4.4, a value of $(k'r')_l$ can be found. Based on the second constraint and the small magnitudes displayed in Figure 4.2, it is possible to assign $(k'r')_l$ to zero. Therefore, the microphone design satisfies the specifications and the two constraints described above, for the frequencies in the range $f = [340, 1100]$Hz.[8] Although it is possible to include the baseband frequencies to give a total frequency range of $f = [f'_l, f'_u] = [0, 1100]$Hz, Section 4.6.4 will show that this is not possible when there are calibration errors.

[8]Note that this holds only when a spherical region of radius 0.5m is considered.

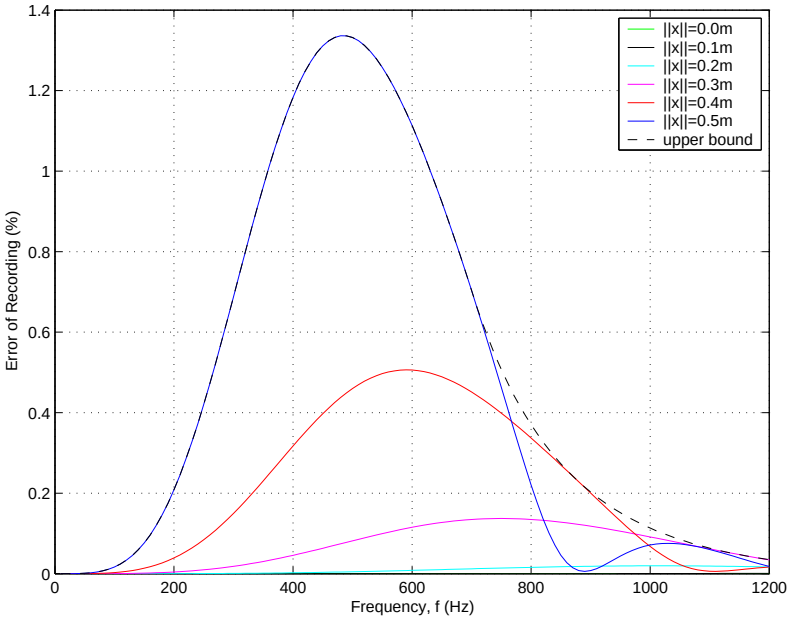


Figure 4.2: Error of recording at various radii, for microphone array at $r' = 15.5$cm with a rigid spherical scatterer at $R = 6.2$cm. The dotted black line indicates the upper bound to error for $0 \leq \|\mathbf{x}\| \leq 0.5$m.

A design of a double array microphone has been given, which satisfies the specifications and constraints in Section 4.4. A total of 64 microphones were placed in two spherical locations. The frequency response of the system satisfies the specified range of $[340, 3400]$Hz. Details of how the two frequency bands can be combined are given in Section 3.7. In the next section, other errors affecting the design are examined. From here onwards, we differentiate the two microphone arrays by identifying them as either the inner (Section 4.4) or outer (this section) microphone array.

4.6 Quantitative Analysis

4.6.1 Truncation Error

As described in Sections 3.2 and 4.2.1, truncation error is inherent in any practical design. Up until now, the design process and error analysis described in Sections 4.4 and 4.5 has been implicitly concerned with recording up to and including 4^{th} order modal components of the sound field. We now turn our attention to the truncation error, which seeks to determine the extent of error due to the absence of the higher orders. A plot of the truncation error is shown in Figure 4.3. From this plot it is evident that the magnitude of the truncation error is quite significant, and there is a limited frequency band and a limited size of the spherical region in which a certain level of spatial accuracy can be attained. Again, we stress that this affects the spatial quality of the sound field and not the temporal quality.

The only strategy to reduce the truncation error is to design for recording of higher orders. However, designing higher order microphone arrays is difficult, which was the underlying theme that came through in Chapter 3. For example, the number of microphones and also the signal

processing requirement increases in the order of N^2. Issues related to the calculation of this error are described in Appendix G. Refer to Section 3.2 for more details.

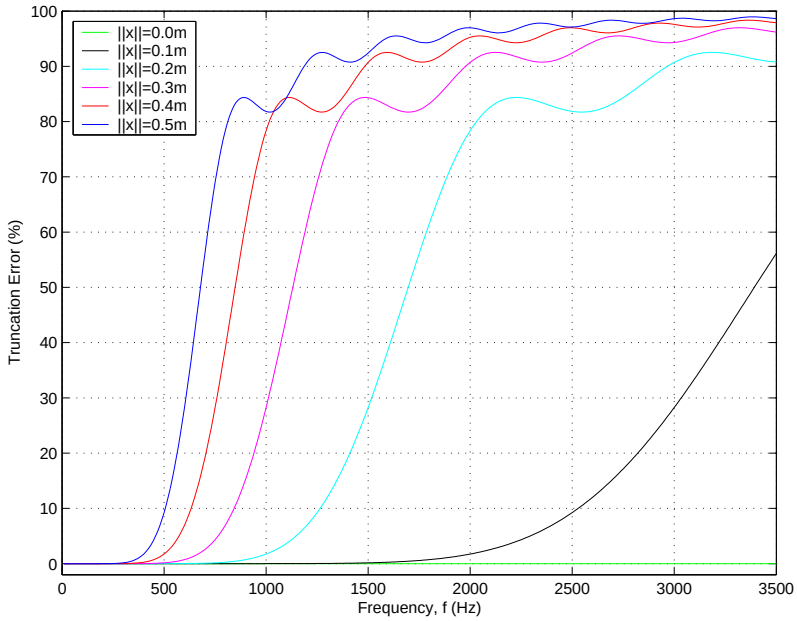


Figure 4.3: Truncation error at various radii for any design that records up to and including 4^{th} order modal components.

4.6.2 Error due to Integration Approximation

The various microphone arrangements give integration approximations designed to be accurate up to a particular order, as described in Section 3.5. If higher order modal components are present, the higher spatial frequency content causes error in resolving the modal coefficients of interest, which we distinctively identify as aliasing error (Refer to Section 3.4). However, when no higher orders are present, as is the case when measuring the error of recording, the error due to integration approximation can be observed.

Revisiting the error of recording shown in Figure 4.1, the relatively small error in the frequency range [3000,3500]Hz would lead one to conclude that the error due to integration approximation is negligible.[9] This agrees with the high level of orthogonality of the microphone arrangements as analysed in Section 3.5. Since, this error assesses the truncated icosahedron arrangement by itself, it applies to both microphone array designs.

4.6.3 Robustness to Aliasing

We will now assess the effectiveness of the second constraint, described in Section 4.4. That is, we would like to look at the aliasing error due to the higher order modal components of the sound field. The aliasing error of (4.6) was evaluated and plotted in Figure 4.4, for the inner microphone array discussed in Section 4.4.

[9] We can make this conclusion because the error of recording is made up of two errors, one of which is the error due to integration approximation. Since the sound field is order truncated, this error must be constant with respect to frequency k (because spatial frequencies are capped at a maximum). Therefore, the smallest observable error in Figure 4.1 gives an upper bound to the error due to integration approximation.

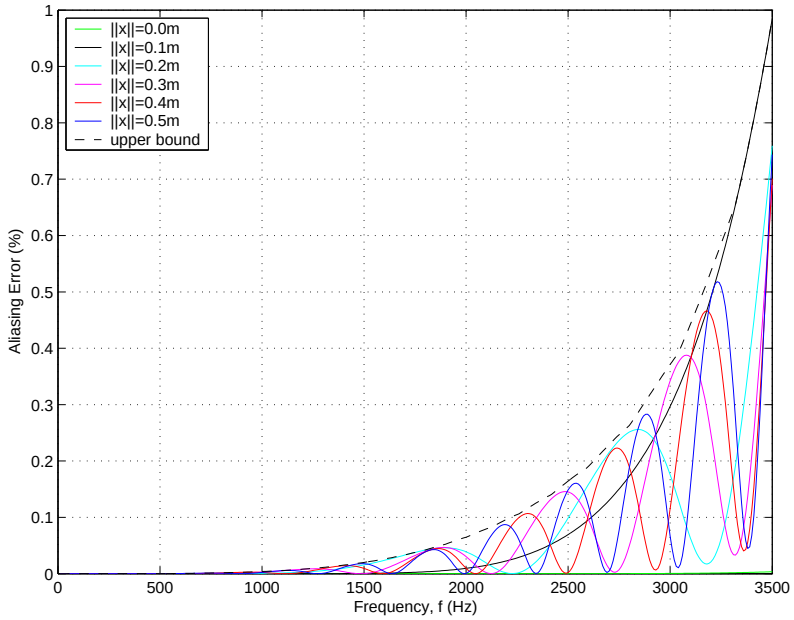


Figure 4.4: Total aliasing error at various radii, for the inner microphone array. The dotted black line indicates the upper bound to error for $0 \leq \|\mathbf{x}\| \leq 0.5$m.

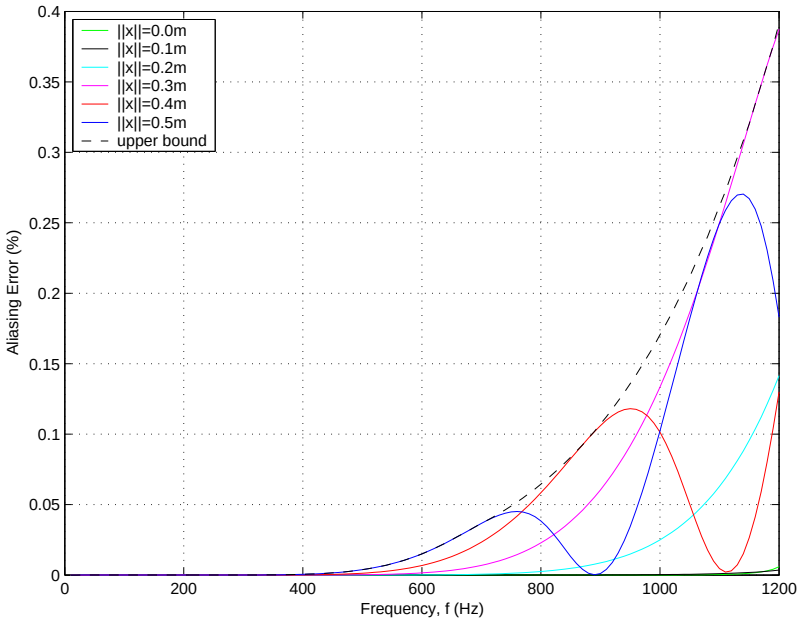


Figure 4.5: Total aliasing error at various radii, for the outer microphone array. The dotted black line indicates the upper bound to error for $0 \leq \|\mathbf{x}\| \leq 0.5$m.

It can be inferred from the figure that the aliasing error is bounded by 1.0% for our frequency range of interest. Also, it can be observed that the upper bound of the aliasing error rises rapidly with frequency, which justifies the need to limit the value of $(kr)_u$. The calculation details of the aliasing error is explained in more detail in Appendix F.

Similarly, for the outer microphone array discussed in Section 4.5, the aliasing error was

evaluated and plotted in Figure 4.5. This shows that the aliasing error is bounded by 0.4% for our frequency range of interest. Again, it can be observed that the aliasing error rises rapidly with frequency. The process of calculating this error is described in Appendix F. However, in both cases, the numerical values should be interpreted with some caution since they correspond specifically to the test sound field that was used, as highlighted in Section 4.3.

4.6.4 Effect of Calibration Error

In our design in Sections 4.4 and 4.5, we specified no calibration error. In this section, we use the model presented in Section 3.6 to assess the implications of this simplification. Specifically, we re-analyse both microphone arrays by using the truncated icosahedron arrangement with calibration error specified at $L_{\max} = 0.002\pi$.[10] The truncation error is by default, independent of the calibration error whilst the aliasing error was observed to be independent of the calibration error for this particular situation. The error of recording, however, changed dramatically, as shown for the inner and outer microphone arrays in Figures 4.6 and 4.7, respectively.

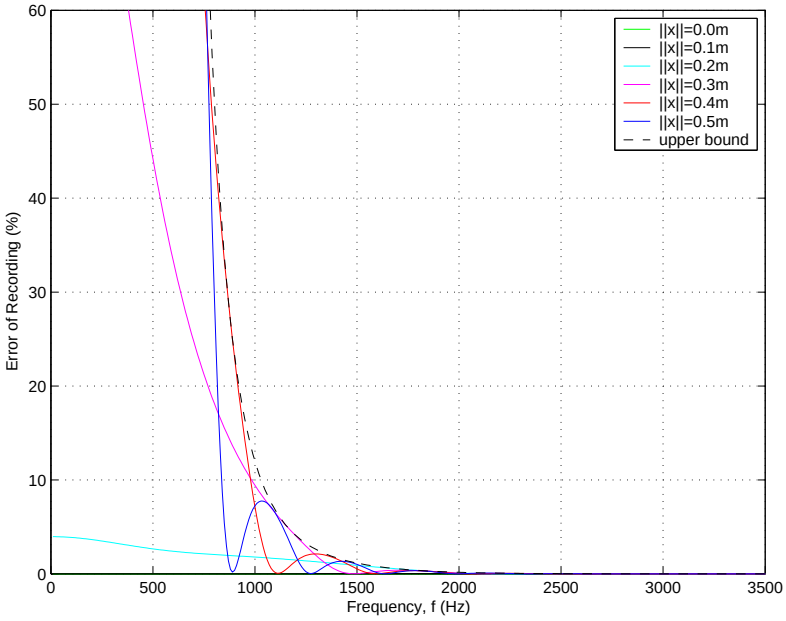


Figure 4.6: Error of recording at various radii $\|\mathbf{x}\|$, for the inner microphone array with calibration error specified at $L_{\max} = 0.002\pi$. The dotted black line indicates the upper bound to error for $0 \leq \|\mathbf{x}\| \leq 0.5$m.

[10] Refer to Appendix H for a copy of the "randomised" points used. These same points were used for both microphone array simulations in this section.

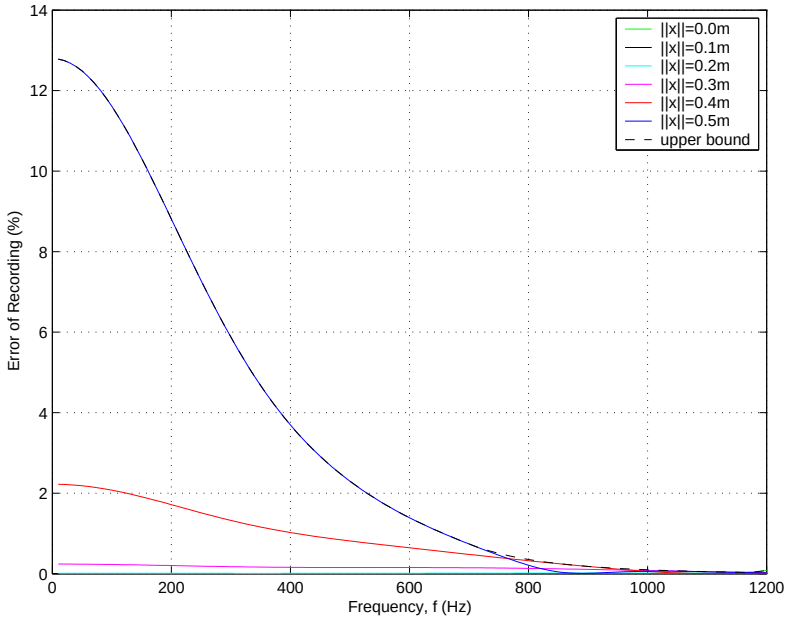


Figure 4.7: Error of recording at various radii $\|\mathbf{x}\|$, for the outer microphone array with calibration error specified at $L_{\max} = 0.002\pi$. The dotted black line indicates the upper bound to error for $0 \leq \|\mathbf{x}\| \leq 0.5$m.

From these, we can see that the peaking behaviour has disappeared and instead, there is a rapid rise in error at the lower frequencies. Generally, the error has increased throughout the frequencies displayed for both cases. Therefore, during the design process, the parameter f_l of the second constraint will generally be larger if calibration error is modelled. Thus, we will be left with a smaller frequency range for accurate recording. As a comparison, f_l would be changed from 1100Hz to 1170Hz for the inner microphone array, whilst f_l would be changed from 0Hz to 340Hz. Although these values are specific to the "randomised" points that were used, what can be gained from Figures 4.6 and 4.7 is the change in the characteristics of the error of recording as a result of calibration error.

4.7 Qualitative Analysis

In this section, we provide visualisations to demonstrate the recordings obtained from the design specified in Sections 4.4 and 4.5. It is not possible to present a three-dimensional sound field on paper, so instead, we produce two-dimensional plots of a sound field on the plane, $z = 0$m, for the region, $|x| \leq 0.5$m and $|y| \leq 0.5$m. Nor can we effectively show the sound field over time so each plot we present can be thought of as a snapshot taken at an instant in time. Furthermore, complex valued sound fields were used throughout this thesis for mathematical convenience, but in reality they are real valued. Therefore, it is sufficient to observe the real part of the sound field (so long as sensible snapshots are taken so that they are representative of its characteristics). To assess the performance of a design, snapshots are given at various frequencies. In each of the snapshots, the following sound fields are shown:

(a). Plane wave test sound field, as described in Section 4.3.

(b). Sound field in (a), which has been order truncated to include modal components of orders $0 \leq n < 5$.

(c). Sound field in (b), which accounts for the presence of the rigid spherical scatterer.

(d). Perfectly reconstructed[11] sound field using the modal coefficients recorded from (c).

A colour bar is included to indicate the value of the sound field. The values are meaningful if one keeps in mind that the test sound field described in Section 4.3 is used in all the snapshots. Note that the values inside the rigid spherical scatterer are invalid.

4.7.1 Recordings From the Inner Microphone Array

The snapshots in this section were obtained via simulation of the design specified in Section 4.4. In each of the snapshots, the spherical microphone array is coincident with the rigid spherical scatterer and is outlined in cyan.

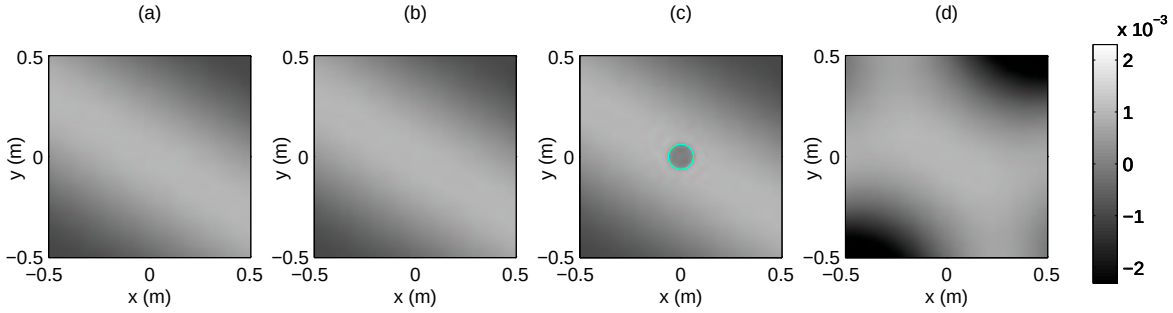


Figure 4.8: Inner microphone array recording of a plane wave of 0.25kHz.

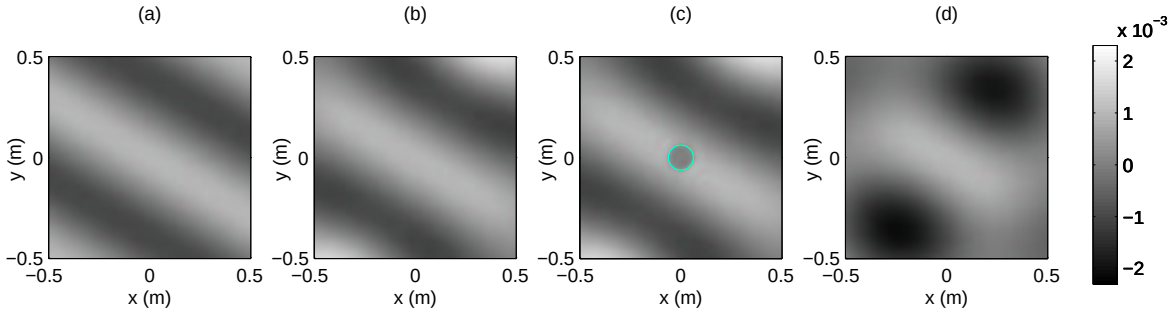


Figure 4.9: Inner microphone array recording of a plane wave of 0.5kHz.

[11] Refer to Section 4.2 for details of this concept.

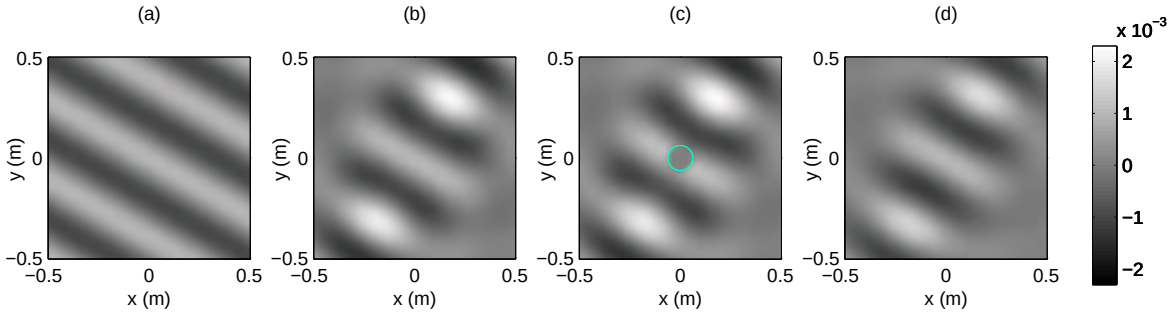


Figure 4.10: Inner microphone array recording of a plane wave of 1kHz.

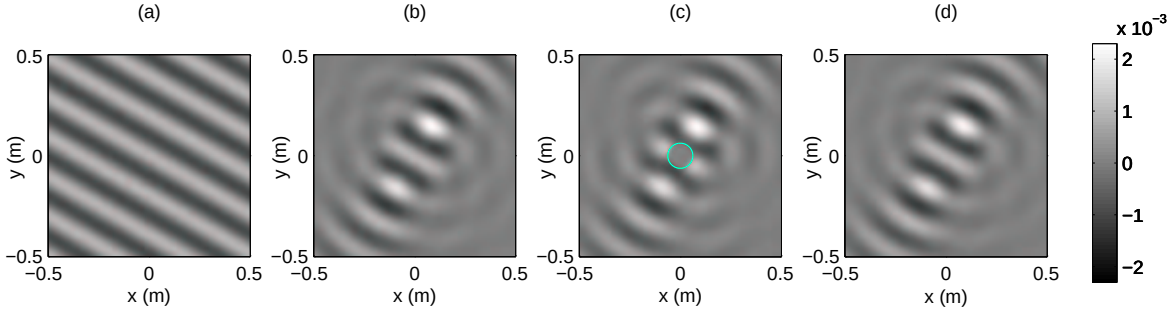


Figure 4.11: Inner microphone array recording of a plane wave of 2kHz.

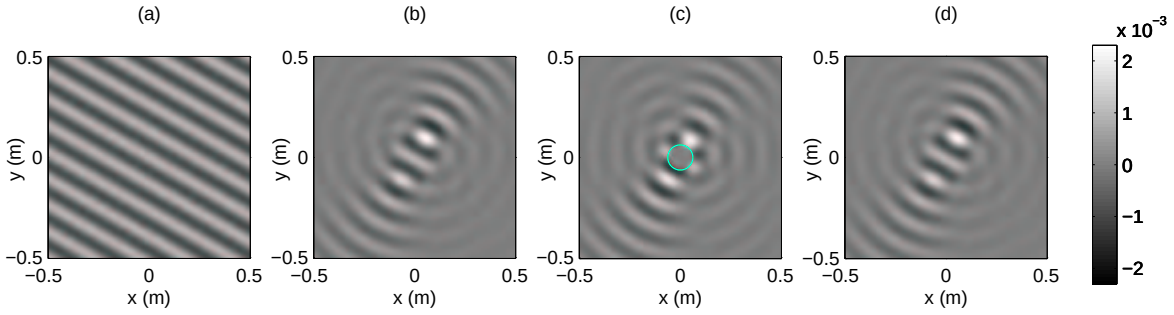


Figure 4.12: Inner microphone array recording of a plane wave of 3kHz.

Firstly, consider parts (a) and (b) of the Figures 4.8 to 4.12. Qualitatively, we can see from these snapshots that as frequency is increased, the resemblance of (b) to (a) reduces rapidly. Furthermore, the further away from the origin we consider, the lower the accuracy. These demonstrate the effects of order truncation error as described in Section 4.6.1.

Now, consider the sound field with a rigid spherical scatterer from which recordings are taken, shown in (c). Note that we can only just begin to make out the scattering effects on the sound field at 3kHz. With respect to the radius of the rigid spherical scatterer, the diffractional effects are dominant for frequencies below 3kHz. Refer to Section 3.3 for more details.

To assess the accuracy of recording using the inner microphone array, qualitatively, we

compare the resemblance of (d) to (b). We can see that the level of resemblance possesses the characteristic peaking behaviour, where qualitatively, we have the largest discrepancy at 0.5kHz. This corresponds to the error of recording described in Section 4.4 and shown in Figure 4.1. If calibration errors were factored into the design, the discrepancy would be much greater at 0.25kHz and 0.5kHz. Overall, we can qualitatively observe that there is a high level of accuracy at least for frequencies in the range 1kHz to 3kHz, as we had expected when designing this inner microphone array in Section 4.4.

4.7.2 Recordings From the Outer Microphone Array

The results in this section were obtained via simulation of the design specified in Section 4.5. In each of the snapshots, the location of the microphone array is indicated by the blue dot-dashed circle. The rigid spherical scatterer is outlined in green.

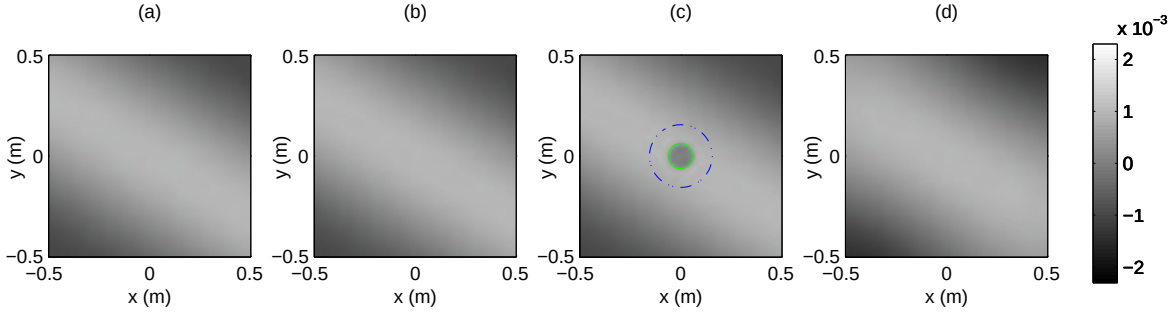


Figure 4.13: Outer microphone array recording of a plane wave of 0.25kHz.

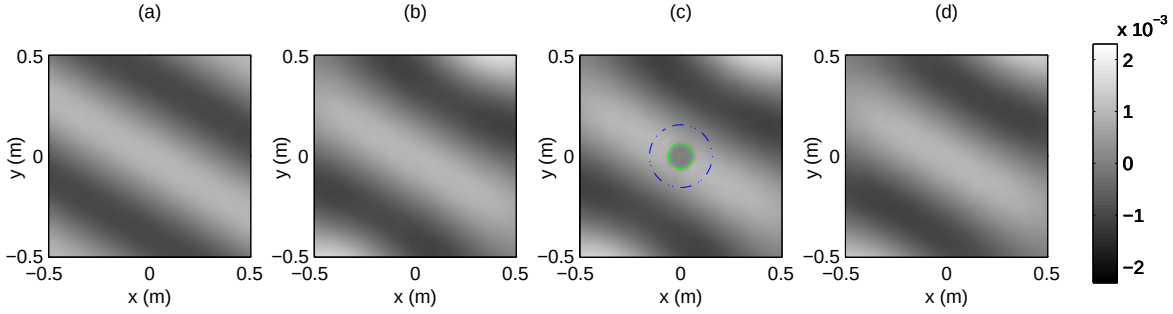


Figure 4.14: Outer microphone array recording of a plane wave of 0.5kHz.

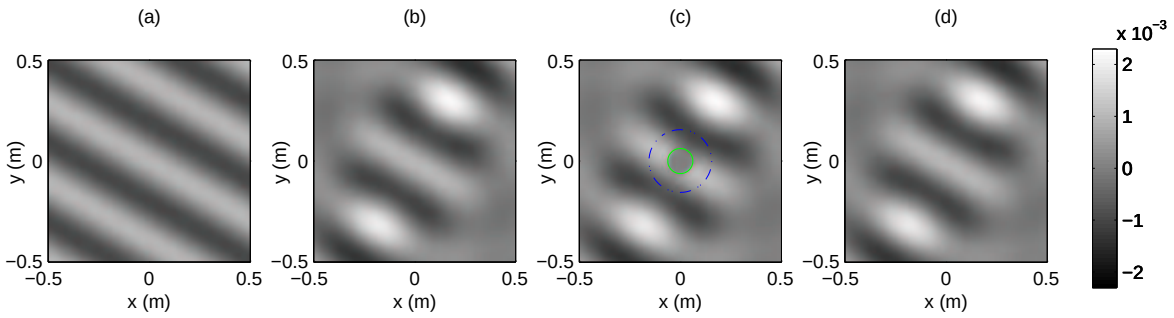


Figure 4.15: Outer microphone array recording of a plane wave of 1kHz.

As with the inner microphone array in Section 4.7.1, we can observe the effects of truncation error.

Qualitatively, the most striking observation is the relatively higher accuracy of recording with the outer microphone array compared to the inner microphone array at 0.25kHz (Figure 4.13 compared to Figure 4.8) and 0.5kHz (Figure 4.14 compared to Figure 4.9). Observe the higher level of resemblance of (d) to (b) for the outer microphone recordings. This corresponds to the small magnitudes shown in the error of recording in Figure 4.2. Again, were calibration errors factored into the design, the increased error of recording would affect the lower frequencies, such as at 0.25kHz.

Therefore, from what has been presented in both Section 4.7.1 and this section, we can chose to obtain the higher frequency content of the sound field from the inner microphone array and the lower frequency content of the sound field from the outer microphone array via temporal filtering, as described in Section 3.7. From this qualitative treatment, we can see that we are able to record the modal coefficients of orders $0 \leq n < 5$ of a sound field in the frequency band [340,3400]Hz, as desired.

4.8 Summary

We have integrated and applied the design issues presented in Chapter 3 to produce a fourth order spherical microphone design. The design has achieved less than 5% error of recording and aliasing error in the frequency range $[340, 3400]$Hz. Furthermore, during the analysis of calibration error, we found an increase in the error of recording for all frequencies observed. Generally, calibration error was found to reduce the allowable frequency range.

Chapter 5

Conclusions and Future Research

In summary, the work presented in this thesis has made the following important contributions to sound field recording:

- Consolidated theory between the theoretical continuous spherical microphone with and without the use of a rigid sphere, which allowed the use of intermediate configurations between the two. This theory paves the way towards the design of multi microphone array systems.
- Application and evaluation of the Gaussian and trigonometric quadrature integration methods to approximate the continuous spherical microphone using a microphone array. These arrangements can be used in a practical spherical microphone design. They was found to be relatively efficient and is scalable.
- A model for analysing the error due to inexact positioning of microphones in an array (or calibration error) was presented. Analysing calibration error is essential when a physical microphone is implemented.
- A set of error measures was proposed for error analysis of spherical microphones. Specifically, the most important ones were error of recording and aliasing error. These would serve as a basis for future work in microphone design.
- Design and analysis of a fourth order spherical microphone. This is a novel design, which extends those presented in the literature by its ability to record one order higher. This gives increased directional accuracy and more precise localisation of sound sources, compared to existing third order designs, whilst maintaining the same frequency band.

However, as with existing designs, there are some limitations in our design. Firstly, it is only accurate at recording the band of frequencies required for speech, which is around 0.2kHz to 3.4kHz [18]. This is promising for applications such as teleconferencing, but there is still a fair way to go before we can capture high fidelity music recordings, where we are looking at a range of frequencies up to, say 15kHz. The second concern is the level of spatial accuracy achieved in our design. Although improved over existing designs, our fourth order microphone provides limited spatial resolution, where the lack of accuracy of localisation is very much noticeable as we move further away from the origin. Overall, we expect that our design is especially useful at capturing reverberant and ambient sound, which is instantly applicable to

existing surround sound systems like Dolby Surround. Beamforming is another area in which our design is directly applicable.

From the treatment presented in this thesis, the challenge for future design improvements includes increasing the spatial accuracy and extent while at the same time, enlarging the frequency range in which the sound field can be captured. The following items have been identified as areas in which future work could be done:

- The design could incorporate theory suggested in [17] to remove the restriction of placing microphones on a sphere. This is expected to increase the range of frequencies which we can record.

- The rigid spherical scatterer assumes perfectly rigid material. This is not achievable in practice since all materials have some degree of elasticity. Modelling this could be the focus of future work.

- The microphones were idealised as omni-directional point sensors. This is not the case in practice and the implications could be investigated. We anticipate that large area microphones could be designed so as to reduce the effects of modal aliasing.

- Analysis of cubature and sphere packing arrangements for use in higher order microphone design.

- The error analysis presented in this thesis used a single plane wave as the test sound field. Further analysis by evaluating the error measures with different test sound fields could be performed. Possible sound fields could include one with a linear combination of plane waves, or a realistic sound field, which has the characteristics expected in practical acoustic events.

- A prototype of the system could be built so to test out the theory and design issues we have outlined in this thesis. The other important benefit of a prototype is to assess qualitatively from a listener's perspective.

Appendix A

Gaussian and Trigonometric Quadratures

A.1 Gaussian Quadratures

Gaussian quadratures are methods which allow the exact evaluation of definite integrals, where the integrand is a polynomial. The Gaussian-Legendre quadrature is a specific Gaussian quadrature method and it is the most appropriate if there is a need to minimise the total number of points. For further details, the reader is referred to [31] and [32]. Specifically, this quadrature is concerned with evaluating definite integrals on $[-1, 1]$, where the polynomial integrand is restricted to a maximum degree of, say, $2N - 1$. Then, the Gaussian-Legendre quadrature specifies the location of the N sample points required, $\{x_k\}_{k=1}^{N}$, and their corresponding weights, $\{w_k\}_{k=1}^{N}$, such that

$$\int_{-1}^{1} f(x)dx = \sum_{k=1}^{N} w_k f(x_k). \tag{A.1}$$

The function, $f(x)$ can be written in the form of what is known as a Lagrange interpolating polynomial [30]. For example, the equation of a line can be written as

$$f(x) = \frac{x - x_2}{x_1 - x_2} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2). \tag{A.2}$$

If we take the integral of $f(x)$ in the form of (A.2), we obtain

$$\begin{aligned} \int_{-1}^{1} f(x)dx &= \int_{-1}^{1} \Big(\frac{x - x_2}{x_1 - x_2}\Big)dx \times f(x_1) + \int_{-1}^{1} \Big(\frac{x - x_1}{x_2 - x_1}\Big)dx \times f(x_2) \\ &= w_1 f(x_1) + w_2 f(x_2), \end{aligned} \tag{A.3}$$

where (A.3) is in the form of (A.1). This procedure can be extended to polynomials of any degree.

Now, consider a polynomial $f(x)$ of degree at most $2N - 1$ in the variable x. We can show that we require N points for the exact quadrature of the definite integral of $f(x)$ between $x = -1$ and $x = 1$. Firstly, we obtain a polynomial by taking $f(x)$ and dividing it by the Legendre polynomial, $P_N(x)$. Then, the resulting polynomial must be of degree $N - 1$, so it

can be written in quotient-remainder form as

$$\begin{aligned} \frac{f(x)}{P_N(x)} &= P_{N-1}(x) + \frac{r_{N-1}(x)}{P_N(x)} \\ \therefore f(x) &= P_N(x)P_{N-1}(x) + r_{N-1}(x), \end{aligned} \tag{A.4}$$

where $r_{N-1}(x)$ represents a polynomial of degree at most $N-1$. The integral of $f(x)$ is

$$\int_{-1}^{1} f(x)dx = \int_{-1}^{1} P_N(x)P_{N-1}(x)dx + \int_{-1}^{1} r_{N-1}(x)dx \tag{A.5}$$

$$= \int_{-1}^{1} r_{N-1}(x)dx \tag{A.6}$$

where the orthogonality property of Legendre polynomials[1] has been used. Since $r_{N-1}(x)$ is of degree at most $N-1$, the quadrature can be evaluated exactly with N sample points by using a Lagrange interpolating polynomial. However, in terms of the quadrature of A.1, we cannot blindly assume that we can use this property. Instead, we generally need $2N$ sample points to evaluate the quadrature as

$$\int_{-1}^{1} f(x)dx = \sum_{k=1}^{2N} w_k P_N(x_k)P_{N-1}(x_k) + \sum_{k=1}^{2N} w_k r_{N-1}(x_k). \tag{A.7}$$

We need $2N$ sample points as a result of the degree of the integrand of the first term on the right-hand side of (A.5). We can, however, assume the orthogonality property of Legendre polynomials if we choose $\{x_k\}_{k=1}^{N}$ to be the N zeros of $P_N(x)$ in $[-1, 1]$ [32]. This is because the first term on the right-hand side of (A.7) vanishes as a result. Thus, we are left with a total of N sample points.

Software exists which calculates the points and weights for a given degree of the integrand.[2] Table A.1 lists the sample points and corresponding weights for some values of N. Table A.2 lists the sample points when the two end points, $x = -1$ and $x = 1$ are specifically to be used.

[1] $\int_{-1}^{1} P_n(x)P_m(x)dx = 0$, where $n \neq m$. Refer to [**?**] for more details.

[2] `sgausq` is a program which calculates sample points and corresponding weights for a given value of N. It also allows the user to specify whether one or both end points should be used. This can be found at http://www.netlib.org/go/. Refer to Appendix H.

N	Sample Points, x_k	Weights, w_k
2	-0.5773502692	1.0000000000
	0.5773502692	1.0000000000
3	-0.7745966692	0.5555555556
	0.7745966692	0.5555555556
	0.0000000000	0.8888888888
4	±0.8611363116	0.3478548451
	±0.3399810436	0.6521451549
5	±0.9061798459	0.2369268851
	±0.5384693101	0.4786286705
	0.0000000000	0.5688888888
6	±0.9324695142	0.1713244924
	±0.6612093865	0.3607615730
	±0.2386191861	0.4679139346

Table A.1: Gaussian-Legendre quadrature sample points and corresponding weights for some values of N.

A.2 Trigonometric Quadratures

Trigonometric quadratures are covered in the literature under the subject of the discrete Fourier transform (DFT). This is used to approximate definite integrals with trigonometric expressions in the integrand. The approximation is done via a discrete sum using the trapezoidal rule. Specifically, K equispaced samples are evaluated between 0 and 2π [30]. The quadrature is given by

$$I_m = \int_0^{2\pi} h(\phi) e^{im\phi} d\phi \approx (\frac{2\pi}{K}) \sum_{n=0}^{K-1} h(\phi_n) e^{im\phi_n} \tag{A.8}$$

for all $-N < m < N$ and where $\phi_n = \frac{2\pi n}{K-1}$. Note that we must use at least $K = 2(N-1)+1$ samples because this is the number of independent values of I_m which we wish to resolve. To ensure that aliasing does not occur, we must ensure that $h(\phi)$ is of the form

$$h(\phi) = \sum_{|m|<N} a_m e^{im\phi},$$

where a_m are real coefficients. This is due to the fact that only K samples of $h(\phi)$ are taken in the approximation. Note that in (A.8), the weights are all equal to $2\pi/K$ since the sample points are equally spaced.

N	Sample Points, x_k	Weights, w_k
2	-0.9999999400	0.3333333430
	0.0000000035	1.3333333700
	0.9999998810	0.3333334920
3	-0.9999999400	0.1666658070
	-0.4472136200	0.8333334330
	0.4472136500	0.8333354590
	1.0000000000	0.1666662100
4	-1.0000000000	0.0999998376
	-0.6546537280	0.5444444420
	0.0000000007	0.7111114260
	0.6546540860	0.5450564030
	0.9999995830	0.0993885547
5	-1.0000000000	0.0666665956
	-0.7650553580	0.3784751590
	-0.2852315900	0.5548588630
	0.2852315900	0.5548909900
	0.7650979160	0.3826575580
	0.9999575020	0.0624518767
6	-1.0000000000	0.0476189330
	-0.8302239180	0.2768260540
	-0.4688487650	0.4317453210
	0.0000000800	0.4876189830
	0.4688486750	0.4318937660
	0.8303149940	0.2818730180
	0.9999089240	0.0424235985

Table A.2: Gaussian-Legendre quadrature sample points and corresponding weights for some values of N, when the two end points are specifically to be used. Note that $N + 1$ points are required.

Appendix B

Proof of PDF for L_q

In Section 3.6, we were interested in the PDF of the distance random variable L_q from the q^{th} microphone position, $\mathbf{x}_q$. We began with the random variable, $Z \sim \mathcal{N}(0, \sigma)$, with mean zero and standard deviation σ, which has a PDF of

$$f_Z(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-z^2/2\sigma^2}. \tag{B.1}$$

Then, we defined

$$L_q = g(Z) = |Z|, \tag{B.2}$$

which is a function of Z. This means that we can determine the PDF of L_q if for all l_q, the equation $g(z) = l_q$ has a countable number of solutions z_i and for all these solutions, $g^{'}(x_i)$ exists and is non-zero. Then we can use the following relationship, as given in [36],

$$f_{L_q}(l_q) = \sum_i \frac{f_Z(z_i)}{|g'(z_i)|}. \tag{B.3}$$

For $l_q \geq 0$, we have two solutions to the equation, $g(z) = l_q$, ie. $z_1, z_2 = \pm l_q$. Hence,

$$\begin{aligned} f_{L_q}(l_q) &= \frac{f_Z(z_1)}{|\text{sgn}(z_1)|} + \frac{f_Z(z_2)}{|\text{sgn}(z_2)|} \\ f_{L_q}(l_q) &= f_Z(l_q) + f_Z(-l_q) \\ f_{L_q}(l_q) &= 2f_Z(l_q), \end{aligned} \tag{B.4}$$

due to symmetry and the fact that $\text{sgn}(\cdot)$ is the signum function, which has the property [37]

$$\text{sgn}(x) = \frac{d|x|}{dx}. \tag{B.5}$$

Now, for $l_q < 0$, there are no solutions to $g(z) = l_q$ and therefore, $f_{L_q}(l_q) = 0$. By writing in a more compact fashion, we obtain (3.20), as required.

Appendix C

Random Positioning of Microphones

In order to calculate calibration errors as set out in Section 3.6, we consider a point $\mathbf{x}_q = (\theta_q, \phi_q)$. We first solve the problem of shifting this point a distance l_q along the surface of the unit sphere, from its origin position. The problem can be generalised as finding the points on a circle of radius l_q along the surface of the unit sphere, centred at $\mathbf{x}_q$. We define a *local* coordinate system, shown in Figure C.1, where the S_2 axis is along the surface of the unit sphere in the direction of decreasing elevation angle, and the S_1 axis points in the direction of increasing azimuth angle (Refer to Figure 2.1). The radial direction is out of the page at the point $\mathbf{x}_q$. We will discuss the effect at the poles below.

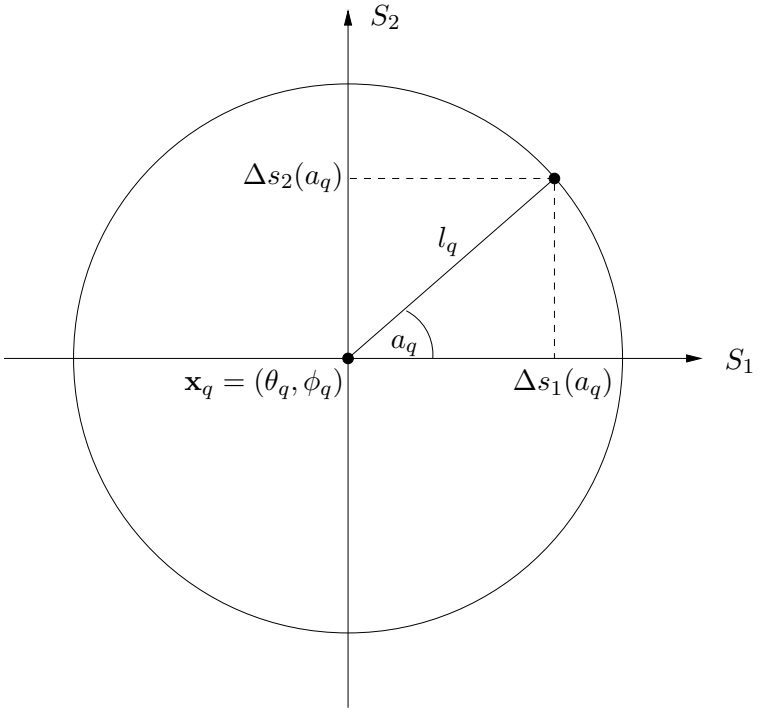


Figure C.1: Surface coordinate system, where the origin corresponds to a point $\mathbf{x}_q$ on the unit sphere. A circle of radius l_q is drawn on the surface centred at this point. The angular parameter is represented by a_q. The radial direction is out of the page at $\mathbf{x}_q$.

We used the term, *local* earlier, meaning that it is only valid in a small region, say $l_q \leq 0.1\pi$, since the curvature of the unit sphere means that this coordinate system is only an approximation to describing a circle of radius l_q along the surface of the sphere, centred about

$\mathbf{x}_q$. Fortunately, we are only concerned with such points and so this approximation is sufficient for our purposes. We will refer to S_1 and S_2 as the *surface coordinates*.

Before we can continue, we need to find the a relationship between the *change in* the spherical coordinates and the *change in* the surface coordinates.

$$\Delta S_1 \leftrightarrow \Delta\theta \tag{C.1}$$
$$\Delta S_2 \leftrightarrow \Delta\phi. \tag{C.2}$$

We are interested in the change in these coordinates because, as will be shown below, we will deal with relative rather than absolute quantities.

C.1 Mapping Between Spherical Coordinates and Surface Coordinates

Let Δs_2 be a length along the S_2 coordinate and Δs_1 be a length along the S_1 coordinate. Assume that the angle $\Delta\theta$ subtends an arc of length Δs_2 on the unit sphere in the S_2 coordinate. By the relationship, $s = r\theta$, which relates the angle θ with the arc length s, we have

$$\Delta s_2 = \Delta\theta. \tag{C.3}$$

However, when we consider the S_1 coordinate, the relationship between Δs_1 and $\Delta\phi$ is not the same. This is because an angle $\Delta\phi$ subtends an arc of length Δs_1 on the unit sphere in the S_1 coordinate, which is dependent on elevation θ - an angle $\Delta\phi$ subtends an arc of length Δs_1, which is smaller at any other elevation compared to at the equator, $\theta = \pi/2$. To quantify this, we take a slice parallel to the x-y plane at elevation $\theta = [0, \pi)$, $\theta \neq \pi/2$, and we observe that the radius of the resulting circle on this plane has been multiplied by a factor of $\sin\theta$, compared to the case where $\theta = \pi/2$ (from basic geometry). Therefore, the circumference of the circle on this plane is reduced by the same factor. Thus, we have

$$\Delta s_1 = \Delta\phi \sin\theta, \tag{C.4}$$

which is dependent on elevation θ.

C.2 Equation of a Circle on the Surface of Unit Sphere

Having determined the relationship between the surface and spherical coordinates, it is possible to determine the equation of a circle on the surface of the sphere. We will use the parameter a_q to define the circle as parametric equations, $\Delta s_1(a_q)$ and $\Delta s_2(a_q)$, corresponding to coordinates, S_1 and S_2, respectively. Therefore, we have

$$\Delta s_1(a_q) = l_q \cos a_q \tag{C.5}$$
$$\Delta s_2(a_q) = l_q \sin a_q. \tag{C.6}$$

By using the relationships of (C.3) and (C.4), we obtain the corresponding parametric equations in terms of angular values,

$$\Delta\theta(a_q) = l_q \sin a_q \tag{C.7}$$

$$\Delta\phi(a_q) = \frac{l_q \cos a_q}{\sin(\theta_0 + \Delta\theta(a_q))}. \tag{C.8}$$

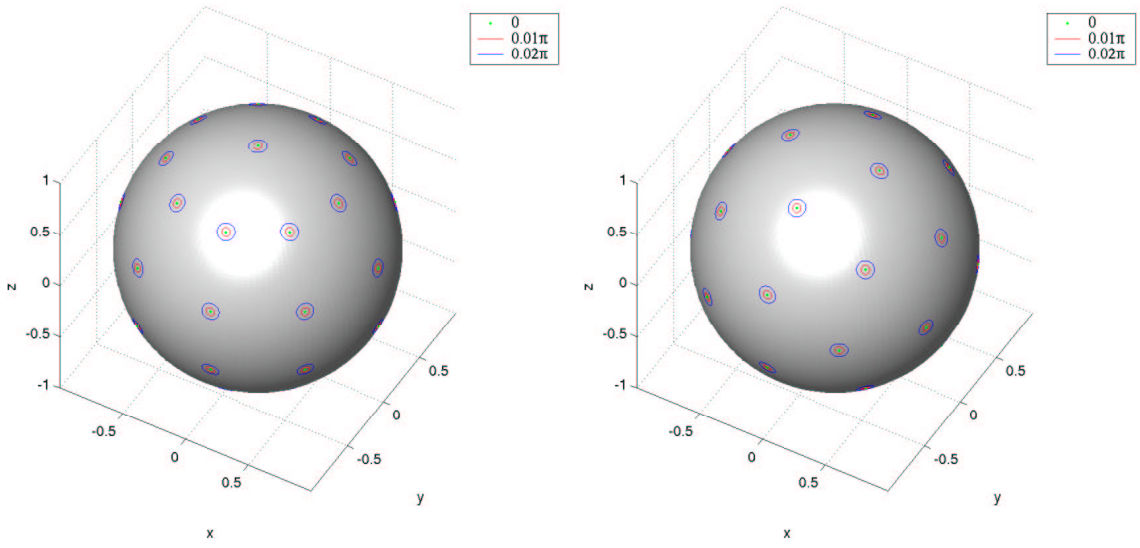
Therefore, the parametric equations of the circle in spherical coordinates are

$$\theta(a_q) = \theta_q + \Delta\theta(a_q) \tag{C.9}$$

$$\phi(a_q) = \phi_q + \Delta\phi(a_q). \tag{C.10}$$

These become invalid whenever the point $\mathbf{x}_q$ is too close to, or at the poles, because θ takes on values outside its domain, $[0, \pi]$. (C.9) and (C.10) will draw figure-of-eight like contours. Perhaps a more sophisticated coordinate system would alleviate this problem, but this will be satisfactory for our purposes since it it largely affects only the Gaussian arrangement with the two poles, described in Section 3.5.1. Instead, we solve this problem by assigning $\Delta\theta(a_q) = l_q$ and $\Delta\phi(a_q) = a_q$ for the two points at the poles.

Figure C.2 shows the circles centred about each microphone position at various values of $l_q = l$, where l is constant, for two microphone arrangements.



(a) Gaussian and trigonometric quadratures with $N = 5$ and $K = 10$, two poles.

(b) Truncated icosahedron arrangement.

Figure C.2: Two arrangements with plots of circles at various values of l centred about points.

C.3 Randomising the Positions

To implement the random variables L_q and A_q, $q = 1, \ldots, Q$ in Mathworks MATLAB according to Section 3.6, we can use the functions, `randn` and `rand`, respectively. The former is a random variable Z_1, with a standard Gaussian distribution, while the latter is a random variable Z_2, which is uniformly distributed in $[0, 1]$. Therefore, we set

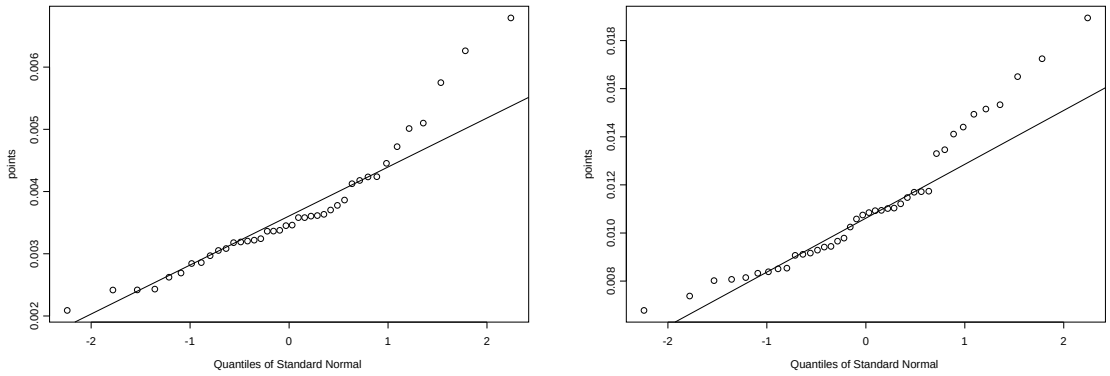
```
Z1 = randn (1);   % Standard Gaussian RV, 1 by 1 matrix output.
Z2 = rand (1);    % Uniform RV, 1 by 1 matrix output.
sigma = Lmax / 3; % s.d. for 99.7% corresponding to Lmax.
Lq = abs (Z1) * sigma;
Aq = Z2 * 2*pi;
```

where the relationship $L_q = \mu + Z_1\sigma$ was used, and μ is the mean (zero in our case) and σ is the standard deviation [38]. By assigning $l_q = L_q$ and $a_q = A_q$ and substituting into (C.9) and (C.10), we obtain the randomised microphone positions in spherical coordinates.

Appendix D

Distribution of Calibration Error

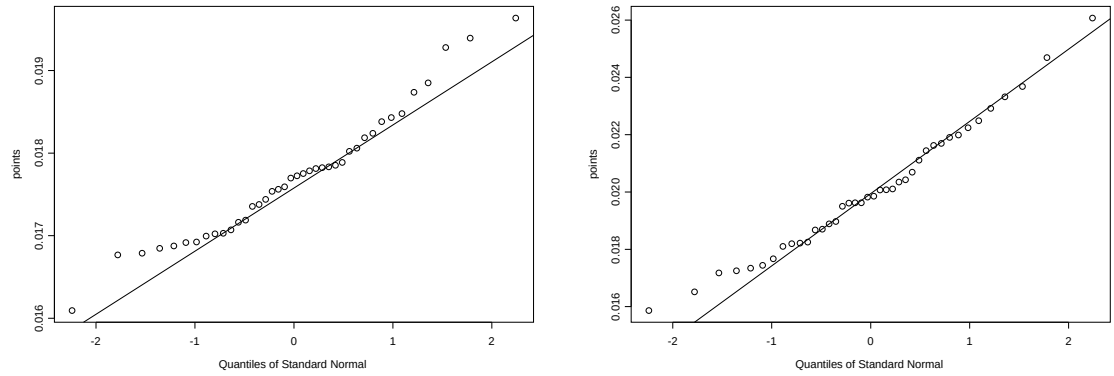
Calibration error was shown in Section 3.6 using $\Delta C_{\max}$ of (3.15) as the measure. The means and standard deviations were obtained from sets of 40 samples and the normal probability plots[1] were plotted using the S-Plus 6.1 statistical software package[2], for each $L_{\max}$ and each arrangement analysed. The plots with $L_{\max}$ of 0.002π and 0.006π are displayed in Figures D.1 and D.2 for the Gaussian and trigonometric quadrature arrangements (both with $N = 5$ and $K = 10$) without poles and with two poles, respectively. Figure D.3 corresponds to the truncated icosahedron arrangement.



(a) $L_{\max} = 0.002\pi$ (b) $L_{\max} = 0.006\pi$

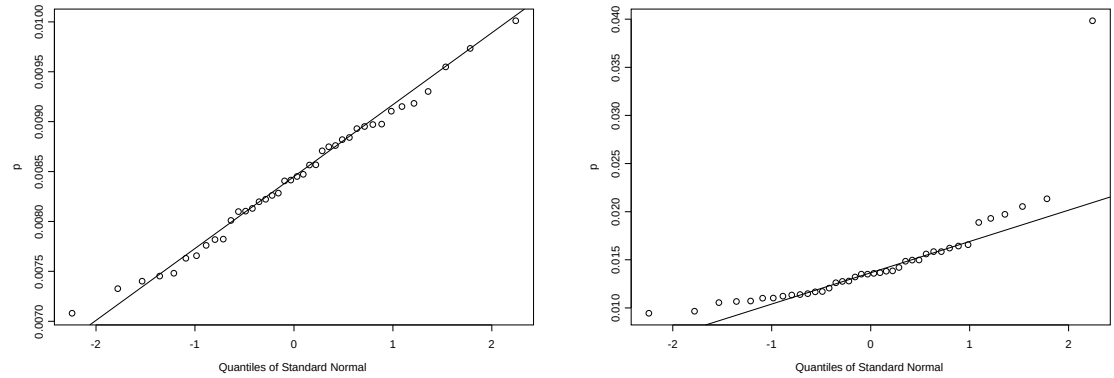
Figure D.1: Normal probability plot of 40 samples of $\Delta C_{\max}$ for the Gaussian and trigonometric arrangement with $N = 5$ and $K = 10$ with no poles.

[1]Refer to [39] for further details.
[2]Refer to [40] for further details.



(a) $L_{\max} = 0.002\pi$ (b) $L_{\max} = 0.006\pi$

Figure D.2: Normal probability plot of 40 samples of $\Delta C_{\max}$ for the Gaussian and trigonometric arrangement with $N = 5$ and $K = 10$ with two poles.



(a) $L_{\max} = 0.002\pi$ (b) $L_{\max} = 0.006\pi$

Figure D.3: Normal probability plot of 40 samples of $\Delta C_{\max}$ for the truncated icosahedron arrangement.

From all of these, we can see that the distributions deviate slightly away from linear behaviour in a concave-up fashion. This indicates a slight skewed-right deviation from the Gaussian distribution [39]. This skewed-right behaviour is typical of data that are always non-negative, as is in this case, which suggests that the gamma distribution would be a suitable model [38]. The gamma distribution parameters α and β are related to the mean μ and standard deviation σ by $\mu = \alpha\beta$ and $\sigma^2 = \alpha\beta^2$ [38].

Appendix E

Proof of Error Equations

The proof of each of the alternative forms of error measures in Section 4.2 shall be shown in this appendix. Each makes extensive use of the property of complex numbers, $|z|^2 = zz^*$.

E.1 Proof of Error of Recording

The error of recording is given by (4.3). To begin with, the original sound field order truncated to $0 \leq n < N$, can be expanded as

$$S_{0:(N-1)}(\mathbf{x}; k) = \sum_{n=0}^{N-1} \sum_{m=-n}^{n} \gamma_{nm}(k) j_n(k\|\mathbf{x}\|) Y_{nm}(\hat{\mathbf{x}}), \tag{E.1}$$

based on (2.21). Similarly, a sound field, which has been perfectly reconstructed from modal coefficients recorded from $S_{0:(N-1)}(\mathbf{x}; k)$, can be expanded as

$$\tilde{S}_{0:(N-1)}(\mathbf{x}; k) = \sum_{n=0}^{N-1} \sum_{m=-n}^{n} \tilde{\gamma}_{nm}(k) j_n(k\|\mathbf{x}\|) Y_{nm}(\hat{\mathbf{x}}), \tag{E.2}$$

where $\tilde{\gamma}_{nm}(k)$ are the recorded modal coefficients from the sound field, $\tilde{S}_{0:(N-1)}(\mathbf{x}; k)$. The integrand in the denominator of the error of recording, (4.3), can be simplified to

$$\begin{aligned}
\left|S_{0:(N-1)}(\mathbf{x}; k)\right|^2 &= S_{0:(N-1)}(\mathbf{x}; k) S^*_{0:(N-1)}(\mathbf{x}; k) \\
\left|S_{0:(N-1)}(\mathbf{x}; k)\right|^2 &= \sum_{n=0}^{N-1} \sum_{m=-n}^{n} \gamma_{nm}(k) j_n(k\|\mathbf{x}\|) Y_{nm}(\hat{\mathbf{x}}) \\
&\qquad \times \sum_{n'=0}^{N-1} \sum_{m'=-n'}^{n'} \gamma^*_{n'm'}(k) j_{n'}(k\|\mathbf{x}\|) Y^*_{n'm'}(\hat{\mathbf{x}})
\end{aligned}$$

$$\begin{aligned}
\left|S_{0:(N-1)}(\mathbf{x}; k)\right|^2 = \sum_{n=0}^{N-1} \sum_{m=-n}^{n} \sum_{n'=0}^{N-1} \sum_{m'=-n'}^{n'} & \gamma_{nm}(k) \gamma^*_{n'm'}(k) j_n(k\|\mathbf{x}\|) j_{n'}(k\|\mathbf{x}\|) \\
& \times Y_{nm}(\hat{\mathbf{x}}) Y^*_{n'm'}(\hat{\mathbf{x}}),
\end{aligned} \tag{E.3}$$

when (E.1) is substituted and noting that $j_n(k\|\mathbf{x}\|)$ is a real valued function. Now, integrating (E.3) over the unit sphere gives

$$\begin{aligned}\int \left|S_{0:(N-1)}(\mathbf{x};k)\right|^2 d\hat{\mathbf{x}} &= \sum_{n=0}^{N-1}\sum_{m=-n}^{n}\sum_{n'=0}^{N-1}\sum_{m'=-n'}^{n'} \gamma_{nm}(k)\gamma^*_{n'm'}(k) j_n(k\|\mathbf{x}\|) j_{n'}(k\|\mathbf{x}\|) \\ &\qquad \times \int Y_{nm}(\hat{\mathbf{x}}) Y^*_{n'm'}(\hat{\mathbf{x}}) d\hat{\mathbf{x}} \\ \int \left|S_{0:(N-1)}(\mathbf{x};k)\right|^2 d\hat{\mathbf{x}} &= \sum_{n=0}^{N-1}\sum_{m=-n}^{n} \left|\gamma_{nm}(k) j_n(k\|\mathbf{x}\|)\right|^2, \end{aligned} \tag{E.4}$$

where the orthonormality property of (2.16) was used. We now consider the numerator of (4.3). In much the same way, we can write

$$\int \left|S_{0:(N-1)}(\mathbf{x};k) - \tilde{S}_{0:(N-1)}(\mathbf{x};k)\right|^2 d\hat{\mathbf{x}} = \sum_{n=0}^{N-1}\sum_{m=-n}^{n} \left|\Big(\gamma_{nm}(k) - \tilde{\gamma}_{nm}(k)\Big) j_n(k\|\mathbf{x}\|)\right|^2, \tag{E.5}$$

where the orthonormality property of (2.16) was used. Substituting both (E.4) and (E.5) into (4.3) yields (4.4) as required.

E.2 Proof of Aliasing Error

Aliasing error is given by (4.5). This proof largely follows from the concepts used in Appendix E.1. Based on (2.21), we can write

$$\hat{S}_{0:(N-1)}(\mathbf{x};k) = \sum_{n=0}^{N-1}\sum_{m=-n}^{n} \hat{\gamma}_{nm}(k) j_n(k\|\mathbf{x}\|) Y_{nm}(\hat{\mathbf{x}}), \tag{E.6}$$

where $\tilde{\gamma}_{nm}(k)$ are the recorded modal coefficients from the original sound field $S(\mathbf{x};k)$. Making use of (E.2) and (E.6) we obtain

$$\int \left|\tilde{S}_{0:(N-1)}(\mathbf{x};k) - \hat{S}_{0:(N-1)}(\mathbf{x};k)\right|^2 d\hat{\mathbf{x}} = \sum_{n=0}^{N-1}\sum_{m=-n}^{n} \left|\Big(\tilde{\gamma}_{nm}(k) - \hat{\gamma}_{nm}(k)\Big) j_n(k\|\mathbf{x}\|)\right|^2, \tag{E.7}$$

where the orthonormality property of (2.16) was used. Substituting (E.4) and (E.7) into (4.5) yields (4.6) as required.

E.3 Proof of Truncation Error

Truncation error, given by (4.1), was originally proposed by Abhayapala and Ward in [19]. This proof largely follows from the concepts used in Appendix E.1. By making use of (2.21) and (E.1), we can write

$$\begin{aligned}\int \left|S(\mathbf{x};k) - S_{0:(N-1)}(\mathbf{x};k)\right|^2 d\hat{\mathbf{x}} &= \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \left|\gamma_{nm}(k) j_n(k\|\mathbf{x}\|)\right|^2 \\ &\qquad - \sum_{n=0}^{N-1}\sum_{m=-n}^{n} \left|\gamma_{nm}(k) j_n(k\|\mathbf{x}\|)\right|^2 \end{aligned} \tag{E.8}$$

and

$$\int \left| S(\mathbf{x};k) \right|^2 d\hat{\mathbf{x}} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left| \gamma_{nm}(k) j_n(k\|\mathbf{x}\|) \right|^2. \tag{E.9}$$

where the orthonormality property of (2.16) was used. Substituting (E.8) and (E.9) into (4.1) yields (4.2).

Appendix F

Calculation of Aliasing Error

This section describes the calculation of aliasing error specifically for the inner microphone design described in Section 4.4. The same concepts are used in the calculation of aliasing error for the outer microphone array specified in Section 4.5, and so they will not be repeated here.

In relation to calculating the aliasing error, we are interested in recording a range of baseband frequencies. Therefore, due to the characteristics of $j_n(k\|\mathbf{x}\|)$ and $b_n(k\|\mathbf{x}\|)$, the most significant modal components responsible for modal aliasing are the lower orders in $n = [N, \infty]$. Figures F.1 though to F.4 show that the aliasing caused by each increasing order decreases approximately by an order of magnitude, with the exception of the 5^{th} order. This exception is due to the fact that the microphone arrangement used satisfies the orthonormality constraint of (3.11). Therefore, it is possible to approximate the total aliasing error by considering only the aliasing from modal component of orders 6 and 7, for example. The total aliasing error for the inner microphone array was given in Figure 4.4. The total aliasing error for the outer microphone array was given in Figure 4.5.

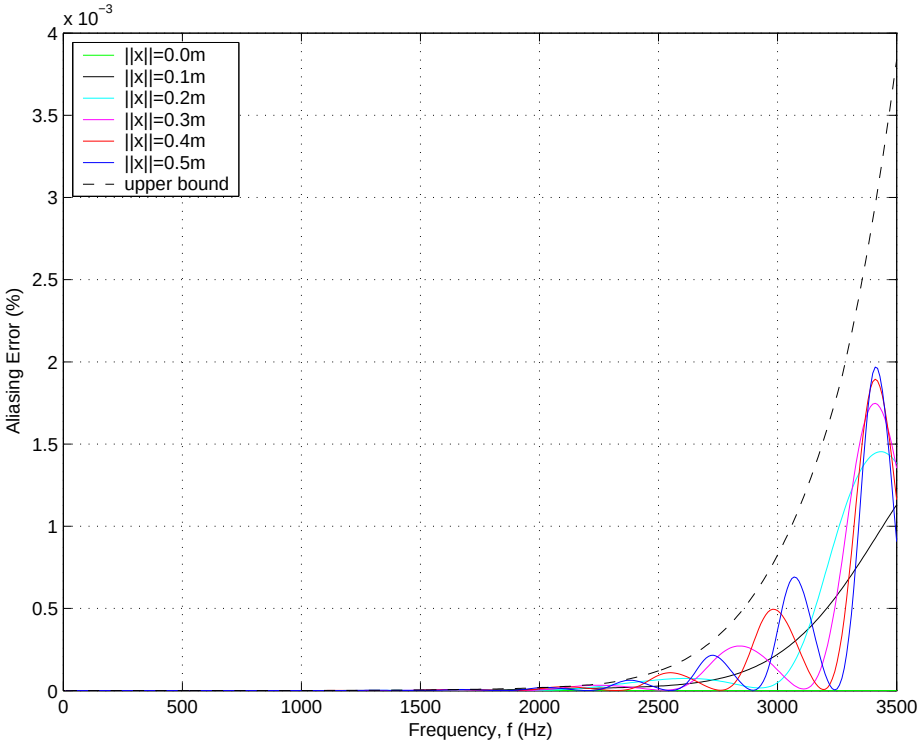


Figure F.1: Alias error from 5^{th} order modal components.

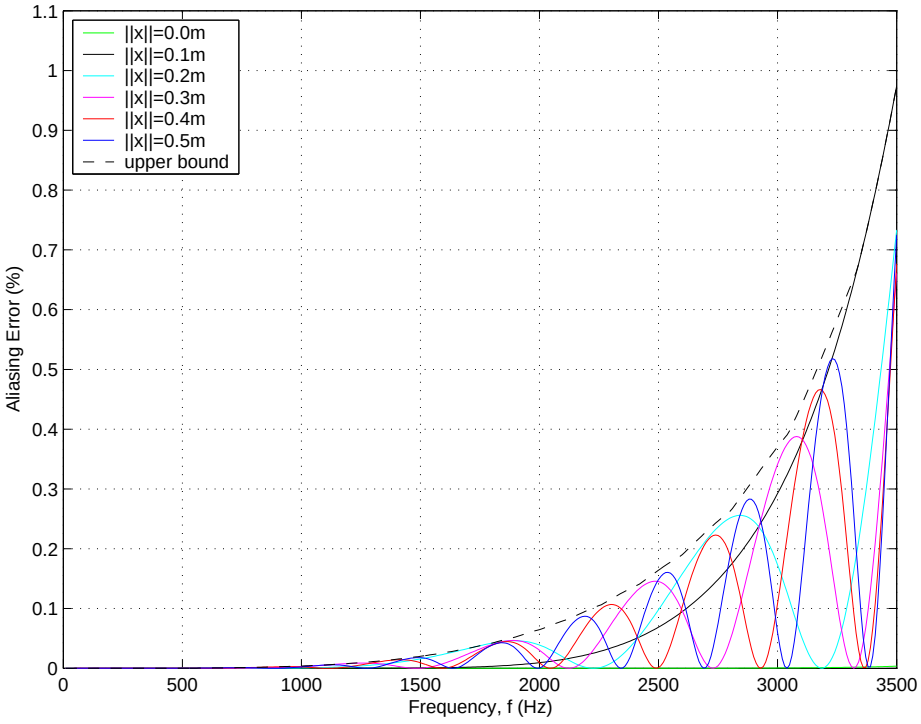


Figure F.2: Alias error from 6^{th} order modal components.

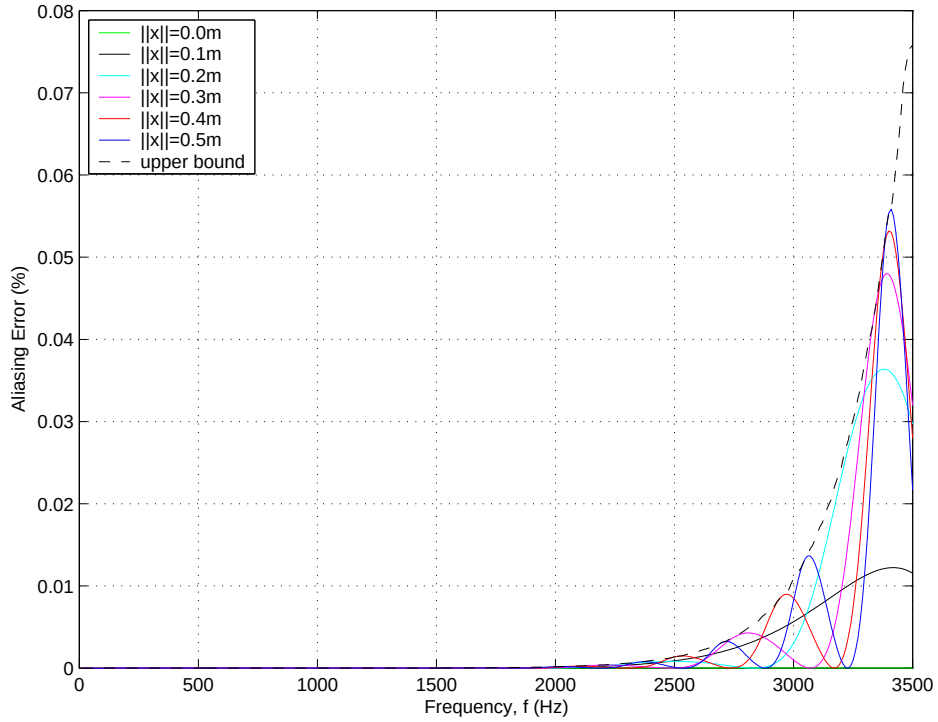


Figure F.3: Alias error from 7^{th} order modal components.

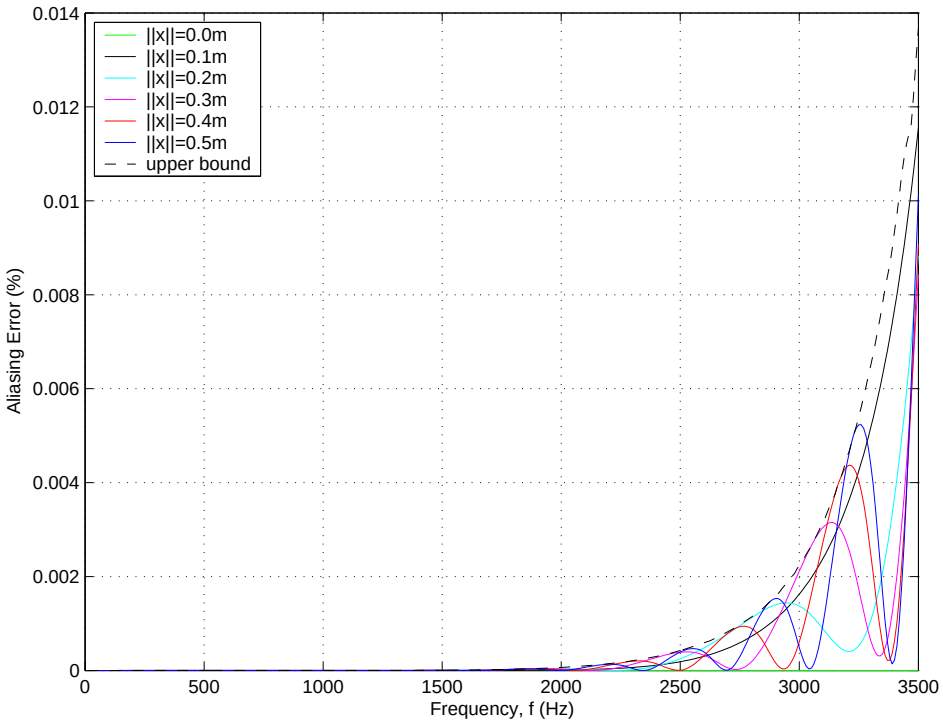


Figure F.4: Alias error from 8^{th} order modal components.

Appendix G

Calculation of Truncation Error

We need to make an approximation to (4.2) since the infinite sum in the denominator makes it unsuitable for direct numerical calculations. Firstly, we write it as

$$\epsilon_{\text{trunc}}(\|\mathbf{x}\|;k) = 1 - \frac{\sum_{n=0}^{N-1}\sum_{m=-n}^{n}\left|\gamma_{nm}(k)j_n(k|\mathbf{x}\|)\right|^2}{\int\left|S(\mathbf{x};k)\right|^2 d\hat{\mathbf{x}}}, \tag{G.1}$$

where $\gamma_{nm}(k)$ are the modal coefficients of the original sound field, $S(\mathbf{x};k)$. The test sound field used in calculating errors, as described in Section 4.3, can be written as

$$S(\mathbf{x};k) = \frac{e^{ik\|\mathbf{y}_s-\mathbf{x}\|}}{\|\mathbf{y}_s-\mathbf{x}\|}, \tag{G.2}$$

where $\mathbf{y}_s$ is the location of the point source (Refer to Section 2.3 for details). Now,

$$\begin{aligned}\left|S(\mathbf{x};k)\right|^2 &= S(\mathbf{x};k)S^*(\mathbf{x};k)\\ \therefore \left|S(\mathbf{x};k)\right|^2 &= \frac{1}{\|\mathbf{y}_s-\mathbf{x}\|^2}.\end{aligned} \tag{G.3}$$

If for all $\mathbf{x}$ under consideration, $\mathbf{y}_s \gg \mathbf{x}$ (that is, the point source is far away from the region we would like to record), then we can make the approximation of $\|\mathbf{y}_s-\mathbf{x}\| \approx \|\mathbf{y}_s\|$. Using this approximation we obtain

$$\left|S(\mathbf{x};k)\right|^2 \approx \frac{1}{\|\mathbf{y}_s\|^2}, \tag{G.4}$$

which is a constant with respect to $\hat{\mathbf{x}}$. It is now possible to approximate the integral in the denominator of (G.1) as

$$\begin{aligned}\int\left|S(\mathbf{x};k)\right|^2 d\hat{\mathbf{x}} &\approx \frac{1}{\|\mathbf{y}_s\|^2}\int d\hat{\mathbf{x}}\\ \int\left|S(\mathbf{x};k)\right|^2 d\hat{\mathbf{x}} &\approx \frac{4\pi}{\|\mathbf{y}_s\|^2}.\end{aligned} \tag{G.5}$$

Appendix H

Index to Files on CD-ROM for Items in Thesis

Table H.1 references each item presented in this thesis with the corresponding files that produced the item. These can be located via the file location indicated on the CD-ROM. Note that the MATLAB scripts that have long file names may pose a problem when attempting to execute them on a Windows version of MATLAB.

Table H.1: Index to files on CD-ROM.

	Item(s)	Description	File location on CD-ROM
1	Figure 3.1	Truncation error as a function of N	simulation/misc/ order_based_truncation_error.m
2	Figure 3.2	Plot of $j_n(k\|\mathbf{x}\|)$	simulation/misc/ plot_spherical_bessel.m
3	Figure 3.3	Plot of $b_n(k\|\mathbf{x}\|)$	simulation/misc/plot_scatter.m
4	Figure 3.4	Comparison between $j_n(k\|\mathbf{x}\|)$ and $b_n(k\|\mathbf{x}\|)$	simulation/misc/ compare_sphere.m
5	Figure 3.5	$b_0(k\|\mathbf{x}\|)$ as R is varied	simulation/misc/compare_radii.m
6	Table 3.1	Maximum argument values for reducing modal aliasing	simulation/misc/ scatter_order_argument.m
7	Figure 3.6(a)	Gaussian quadrature points on the unit sphere	simulation/array/ plot_gaussian_array.m
8	Figure 3.6(b)	Gaussian quadrature points on the unit sphere with two poles	simulation/array/ plot_gaussian_ends_array.m
9	Figure 3.7(a)	Plot of $C_{nm,n'm'}$ for the Gaussian quadrature arrangement	simulation/array/ check_gaussian_orthogonality.m
10	Figure 3.7(b)	Plot of $C_{nm,n'm'}$ for the improved Gaussian quadrature arrangement	Refer to Item 9
11	Figure 3.8	Icosahedron points on the unit sphere	simulation/array/ plot_icosahedron_array.m

	Item	Description	File location on CD-ROM
12	Figure 3.10	Plot of $C_{nm,n'm'}$ for the icosahedron arrangement	simulation/array/check_icosahedron_orthogonality.m
13	Figure 3.11	Calibration error (actual samples used, in simulation/array/misplacement.xls)	simulation/array/plot_misplace.m and Items 9 and 12
14	Figure 4.1	Error of recording for inner array	simulation/record/icosahedron_recording_aliasing_error.m
15	Figure 4.2	Error of recording for outer array	Refer to Item 14
16	Figure 4.3	Truncation error for $N = 5$ design	simulation/record/truncation_error.m
17	Figure 4.4	Total aliasing error for the inner array	Refer to Item 14
18	Figure 4.5	Total aliasing error for the outer array	Refer to Item 14
19	Random points on truncated icosahedron	Generated by Item 12 with the parameter `Lmax` set to 0.002π	simulation/record/random_icosahedron_Lmax0.002pi.mat
20	Figures 4.6 and 4.7	Error of recording with calibration error	Refer to Items 14 and 19
21	Figures 4.9 to 4.12	Inner array sound field recordings	simulation/record/icosahedron_recording.m
22	Figures 4.13 to 4.15	Outer array sound field recordings	Refer to Item 21
23	Table A.1 and A.2	Calculation of Gaussian quadrature points (wrapper program to `sgausq.f`)	simulation/array/nodes/nodes.f
24	Figure C.2	Circles on surface of sphere centred about truncated icosahedron points	simulation/array/plot_icosahedron_circles.m
25	Figures D.1 to D.3	Normal probability plots	`qqnorm` and `qqline` functions in S-Plus 6.1
26	Figures F.1 to F.4	Aliasing error	Refer to Item 14

References

[1] P. Cotterell, "On the theory of the second-order soundfield microphone," Ph.D. dissertation, Instrumentation and Signal Processing Research Group, Department of Cybernetics, University of Reading, Reading, UK, 2002. [Online]. Available: http://www.personal.rdg.ac.uk/∼shr97psc/Thesis.html

[2] T. Abhayapala and D. Ward, "Theory and design of higher order sound field microphones using spherical microphone array," in *Acoustics, Speech and Signal Processing, IEEE International Conference*, vol. 2, Orlando, FL, USA, May 2002, pp. 1949–1952.

[3] M. Evans, A. Tew, and J. Angus, "Spatial audio teleconferencing - which way is better?" in *International Community for Auditory Display, International Conference*, CA, USA, 1997.

[4] M. Miyoshi and N. Koizumi, "Research on acoustics for future telecommunications services," *Applied Acoustics*, vol. 36, pp. 307–326, 1992.

[5] C. Kyriakakis, "Fundamental and technological limitations of immersive audio systems," in *Proc. IEEE*, vol. 86, no. 5, May 1998, pp. 307–326.

[6] A. D. Blumlein, "Improvements in and relating to sound-transmission, sound-recording and sound-reproducing systems," UK Patent 394 325, 1931.

[7] J. Hull. (1999) Surround sound past, present, and future: A history of multichannel audio from mag stripe to dolby digital. White Paper. Dolby Laboratories Inc. [Online]. Available: http://www.dolby.com/ht/430.l.br.9904.surhist.pdf

[8] J. Bamford, "An analysis of ambisonic sound systems of first and second order," Ph.D. dissertation, University of Waterloo, Waterloo, Canada, 1995.

[9] F. E. Toole, "Direction and space, the final frontiers: How many channels do we need to be able to believe that we are "there"?" White Paper, Harman International Industries, Inc., 2000.

[10] D. de Vries and M. M. Boone, "Wave field synthesis and analysis using array technology," in *Proc. 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New York, USA, 1999.

[11] R. Elen, "Ambisonics: The surround alternative," in *Surround 2001 Conference*, Beverly Hills, CA, USA, 2001.

[12] M. A. Gerzon, "Periphony: With-height sound reproduction," *J. Audio Eng. Soc.*, vol. 21, no. 1, pp. 2–10, 1973.

[13] SoundField Research. (2003) SPS422 Microphone System. [Online]. Available: http://www.soundfield.com/prod01.htm

[14] A. Bogomolny. (2003) Regular polyhedra. [Online]. Available: http://www.cut-the-knot.com/do_you_know/polyhedra.shtml

[15] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in *Acoustics, Speech and Signal Processing, IEEE International Conference*, vol. 2, Orlando, FL, USA, May 2002, pp. 1781–1784.

[16] ——, "Electroacoustic systems for 3-d audio - a report from the pittsburgh meeting," *Newsletter Acoustic Soc. America*, vol. 12, no. 3, pp. 2–3, 2002.

[17] A. Laborie, R. Bruno, and S. Montoya. (2003) Immersivesound research. [Online]. Available: http://www.immersivesound.com/research.php

[18] T. Abhayapala, "Modal analysis and synthesis of broadband nearfield beamforming arrays," Ph.D. dissertation, Australian National University, Canberra, Australia, 1999. [Online]. Available: http://www.rsise.anu.edu.au/∼thush/publications/isspa_03_b.pdf

[19] T. Abhayapala and D. Ward, "Reproduction of a plane-wave sound field using an array of loudspeakers," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 6, pp. 697–707, 2001.

[20] D. Excell, "Reproduction of a 3d sound field using an array of loudspeakers," Honours Thesis, Department of Engineering, FEIT, Australian National University, Canberra, Australia, 2003.

[21] D. Hall, *Basic Acoustics*. New York: John Wiley & Sons, 1987.

[22] L. J. Ziomek, *Fundamentals of Acoustic Field Theory and Space-Time Signal Processing*. Boca Raton: CRC Press, 1995.

[23] E. Skudrzyk, *The Foundations of Acoustics: Basic Mathematics and Basic Acoustics*. New York: Springer-Verlag, 1971.

[24] E. W. Weisstein. (1999) Spherical coordinates. [Online]. Available: http://mathworld.wolfram.com/SphericalCoordinates.html

[25] ——. (1999) Separation of variables. [Online]. Available: http://mathworld.wolfram.com/SeparationofVariables.html

[26] L. Kinsler, A. Frey, A. Coppens, and S. James, *Fundamentals of Acoustics*. New York: John Wiley & Sons, 1999.

[27] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. London: Academic Press, 1999.

[28] E. W. Weisstein. (1999) Legendre polynomial. [Online]. Available: http://mathworld.wolfram.com/LegendrePolynomial.html

[29] D. Colton and R. Kress, "Inverse acoustic and electromagnetic scattering theory, 2nd edition," ser. Applied mathematical sciences. New York: Springer, 1998.

[30] P. Davis and P. Rabinowitz, *Methods of Numerical Integration.* New York: Academic Press, 1975.

[31] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing.* New York: Cambridge University Press, 1988.

[32] M. Mohlenkamp. (2002) A user's guide to spherical harmonics. [Online]. Available: http://www.math.ohiou.edu/∼mjm/research/uguide.pdf

[33] J. Fliege. (1999) Integration nodes for sphere. [Online]. Available: http://www.mathematik.uni-dortmund.de/lsx/fliege/nodes.html

[34] N. J. A. Sloane, R. H. Hardin, and W. D. Smith. (2000) Spherical codes. [Online]. Available: http://www.research.att.com/∼njas/packings/

[35] S. Holmes. (2002) The normal distribution. [Online]. Available: http://www-stat.stanford.edu/∼susan/courses/s141/node13.html

[36] J. Proakis and M. Salehi, *Communication Systems Engineering, 2nd Edition.* New Jersey: Prentice Hall, 2002.

[37] N. Andersson. (2001) The signum function. [Online]. Available: http://www.maths.soton.ac.uk/staff/Andersson/MA361/node14.html

[38] D. Wackerly, W. Mendenhall, and R. Scheaffer, *Mathematical Statistics with Applications, 6th Edition.* CA, USA: Thomson Learning, 2002.

[39] NIST/SEMATECH. (2003) Normal probability plot in nist/sematech e-handbook of statistical methods. [Online]. Available: http://www.itl.nist.gov/div898/handbook/eda/section3/normprpl.htm

[40] Insightful Corporation, *S-Plus 6 for Windows guide to statistics.* Seattle, Wash.: Insightful Corporation, 2001.