

Reproduction of a 3D Sound Field Using an Array of Loudspeakers

David Excell – 3222940

Supervisor: Dr. Thushara Abhayapala

June 2003

A thesis submitted in part fulfilment of the degree of
Bachelor of Engineering (Honours) at
The Department of Engineering
Australian National University



THE AUSTRALIAN
NATIONAL UNIVERSITY

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university. To the best of the author's knowledge and belief, it contains no material previously published or written by another person, except where due reference is made in the text.

David Excell

18 June 2003

Abstract

In this thesis the fundamental problem in signal processing of controlling a sound field within a region of space is studied. Using spherical and cylindrical harmonics, reproduction systems in both two and three dimensions have been designed. From the design of the systems, equations are derived describing the performance of the system for arbitrary sound fields. The equations are defined in terms of the number of loudspeakers, the size of the listening area and the frequency content of the sound fields.

Results from numerical verification have been included within this thesis. The implementation of the 2D system is also included in this thesis demonstrating the ability of the designed system to reproduce 'real' sound fields. In the implementation process, a synthesis system was developed allowing conventionally recorded sound samples to be placed in 2D space forming a spatial sound field. In the process it is demonstrated that the reproduction system does not solely use amplitude panning techniques but also considers the phase and delay of the sound field components to recreate the sound field.

Acknowledgements

Without support and encouragement from a number of people and organisations the research presented in this thesis would not have been possible. I would like to especially thank the following:

- Dr. Thushara Abhayapala for his excellent supervision, guiding my efforts into producing a piece of research that I am proud of;
- The Research School of Information Sciences and Engineering (RSISE) for providing me with a Summer Research Scholarship;
- Both the Department of Telecommunications (RSISE) and the Department of Engineering (Faculty of Engineering and Information Technology) for providing the facilities enabling the completion of this research work;
- Tom and Michael for their proof reading of this thesis along with providing positive feedback;
- Glenn Dickins for his suggested solutions to problems encountered while working on this thesis; and
- My family for their continual support and guidance with everything I set out to achieve.

Finally, I would like to thank my girlfriend, Bronnie, whose understanding and support has been most appreciated while I have spent countless hours working towards this thesis.

Trademarks

AC-3, Dolby, Dolby Digital and Dolby Surround are trademarks of Dolby Laboratories Licensing Corporation. DTS is a registered trademark for Digital Theater Systems Incorporated. NXT and SurfaceSound are registered trademarks of NXT Public Limited Company. THX is a registered trademark of Lucasfilm Limited.

Contents

Glossary	vii
List of Figures	ix
1 Introduction	1
1.1 Thesis Motivation	1
1.2 Literature Survey	1
1.3 Commercial Opportunities	3
1.4 Thesis Outline	4
1.5 Contributions	5
2 Background Theory	6
2.1 Theory of Acoustics	6
2.2 Coordinate System	7
2.3 Modal Wave Equation Solution	7
2.4 Sound Field Control	9
2.5 Sound Sources	10
2.6 Problem Formulation	10
3 Sound Field Reconstruction	12
3.1 Three Dimensional Reconstruction	12
3.1.1 Exact Reproduction	12
3.1.2 Approximate Reproduction	14
3.1.3 Calculating the Loudspeaker Weights	17
3.1.4 Loudspeaker Placement	20
3.1.5 Alternative Loudspeaker Weight Calculation	20
3.2 Two Dimensional Reconstruction	22
3.2.1 Exact Reproduction	22
3.2.2 Approximate Reproduction	23
3.2.3 Calculating the Loudspeaker Weights	26
3.2.4 Loudspeaker Placement	26
4 Numerical Simulation	27
4.1 Three Dimensional Sound Field Reconstruction	27
4.1.1 Error Measurement	29
4.2 Error Equation Verification	29
4.3 Effect of Loudspeaker Weight Estimation	31

4.4	Comparison of Solutions	33
4.5	Summary of Results	34
5	Physical Implementation	35
5.1	Design	35
5.2	Sound Field Synthesis	36
5.3	Loudspeaker Signals	39
5.4	Computational Cost	40
5.5	Reproduction Accuracy	42
5.5.1	Modal Limitation	42
5.5.2	Loudspeakers	42
5.5.3	Miscellaneous	43
5.6	Listening Tests	44
5.7	Summary of Results	44
6	Sound Field Control	45
6.1	Non-uniform Loudspeaker Placement	45
6.2	Listening Area Repositioning and Duplication	46
7	Conclusions and Further Work	49
7.1	Conclusions	49
7.2	Further Research	50
A	Mathematical Proofs	51
B	Detailed Implementation Design and Setup	64
B.1	Detailed Implementation Design and Setup	64
B.2	Windowing Function	73
C	Listening Test Summary	75
C.1	Single Stationary Narrowband Sound Source	75
C.2	Single Rotating Narrowband Sound Source	75
C.3	Single Rotating Broadband Sound Source (Circle)	77
C.4	Single Rotating Broadband Sound Source (Square)	77
C.5	Multiple Stationary Broadband Sound Sources	78
D	MatLab Code Listing	79
D.1	Three Dimensional Simulation	80
D.2	Two Dimensional Implementation	86
E	Summary of Content on Attached Media	91
	Bibliography	92

Glossary

Symbol	Definition
2D	Two Dimensional
3D	Three Dimensional
\mathbb{R}	Real Numbers
\mathbb{Z}	Integers
a_l	Loudspeaker Weight
c	Speed of Sound
c	Radian
DFT	Discrete Fourier Transform
DSP	Digital Signal Processor
E	Normalised Expansion Error
f	Frequency (Hertz)
\mathfrak{S}	Fourier Transform
\mathfrak{S}^{-1}	Inverse Fourier Transform
$h_n^{(1)}(\cdot)$	Spherical Hankel Function of the First Kind
$h_n^{(2)}(\cdot)$	Spherical Hankel Function of the Second Kind
$H_n^{(1)}(\cdot)$	Ordinary Hankel Function of the First Kind
i	Imaginary Part of a Complex Number ($i = \sqrt{-1}$)
$j_n(\cdot)$	Spherical Bessel Function of the First Kind
$J_n(\cdot)$	Ordinary Bessel Function of the First Kind
m	Metre
$n_n(\cdot)$	Spherical Bessel Function of the Second Kind
$N_n(\cdot)$	Ordinary Bessel Function of the Second Kind
L	Number of Loudspeakers
$P_{nm}(\cdot)$	Associated Legendre Function
r	Radial Distance from the Origin
$S(x; k)$	Original Sound Field to be Reproduced
t	Time in Seconds.
$T(x; k)$	Reproduced Sound Field
\mathbf{x}	Matrix or Vector
$\hat{\mathbf{x}}$	Unit Vector in the direction of \mathbf{x}
x	Vector Length (i.e. $\mathbf{x} = x\hat{\mathbf{x}}$)
x, y, z	Components of a Vector in Cartesian Coordinates
$Y_{nm}(\cdot)$	Spherical Harmonics
\mathbf{y}	Sound Source Location

Symbol	Definition
y_l	Loudspeaker position
α	Sound Field Coefficient
δ	Kronecker-Delta Function
ϵ	Normalised Truncation Error
γ	Sound Surface Coefficient
θ	Elevation
ϑ	Plane-wave Elevation
ϕ	Azimuth
φ	Plane-wave Azimuth
ω	Radian Frequency
λ	Wavelength
Υ	Normalised Reproduction Error
*	Complex Conjugate
°	Degrees
$ \cdot $	Absolute Value or Complex Magnitude
$\ \cdot\ $	Euclidian Distance
$\lceil \cdot \rceil$	Integer Ceiling Operator

List of Figures

1.1	Thesis Outline	4
2.1	Three Dimensional Coordinate System	7
3.1	Spherical Bessel Function.	13
3.2	Normalised expansion error as plane-wave direction is varied.	16
3.3	Plots of formulated 3D error equations for the approximate reproduction solution. (a), (c) and (e) contain the mean normalised error observed from a thousand random fifty plane-wave sound fields, (b), (d) and (f) shows the standard deviation of the normalised error from these sound fields. The plane-wave frequency for all sound fields was 1 kHz.	18
3.4	The two figures show the layout of twenty-five and thirty-six loudspeakers placed in an uniform packing around a unit sphere.	20
3.5	Normalised Truncation Error for two plane-wave sound field.	23
3.6	Plots of formulated 2D error equations for the approximate reproduction solution. (a), (c) and (e) contain the mean error observed from a thousand random fifty plane-wave sound fields, (b), (d) and (f) contain the standard deviation of the normalised error from these sound fields. The plane-wave frequency for all sound fields was 1 kHz.	25
4.1	Sound field reproduction at height $z = 0.1\text{m}$	28
4.2	Sound field reproduction at height $z = 0\text{m}$	28
4.3	Illustration of the three errors defined in Section 3.1.2. The height of the sound field illustrated is $z = 0\text{m}$, the twenty-five loudspeakers were placed on a radius of $r = 1\text{m}$	30
4.4	Comparison of the reproduction errors measured via numeric simulation and using the error formulas presented in Section 3.1.2. The sound field consisted of fifty random plane-waves at 1 kHz.	30
4.5	Effect of varying the number of loudspeakers for a fixed number of sound field coefficients on the estimation of the loudspeaker weights. The reproduction errors at the four spherical radii were averaged from ten random fifty plane-wave sound fields at 1 kHz. The loudspeaker radius was 1m.	31
4.6	Reproduction error observed at different radii when the reproduction order (N) is varied for a fixed number of loudspeakers. Figure (a) shows the result for $L=25$ and Figure (b) shows the result for $L=36$. The results shown have been averaged from ten randomly selected sound fields containing fifty plane-waves at 1 kHz. The loudspeaker radius was 1m.	32

4.7	Comparison of the ‘matrix’ and ‘continuous’ methods for obtaining the loudspeaker weights. The results are averaged over twenty-five sound fields containing fifty random plane-waves at 1 kHz. The radius of the loudspeakers was 1m.	33
5.1	Basic System Implementation	35
5.2	Illustration of the conversion of sampled audio to data frames. Note: In (b) and (c) the blue and red lines have been shifted in amplitude by +0.1 and -0.1 respectively to ease readability.	38
5.3	Loudspeaker waveform properties	41
5.4	Expected modal error for the frequency range 0.1-10 kHz at the listening radius of $x_0 = 0.2\text{m}$	43
5.5	Expected reproduction error for the frequency range 0.1-10 kHz at the listening radius of $x = 0.2\text{m}$. The values shown have been averaged over twenty different sound fields consisting of twenty-five random plane-waves.	44
6.1	Effect of Non-uniform Loudspeaker Placement	46
6.2	Loudspeaker Coordinate Reference	46
6.3	Sound field reproduction at height $z = 0\text{m}$	47
6.4	Reproduction error as displacement of the listening area is increased.	48
A.1	Arbitrary Sound Field	51
B.1	Creamware TDAT Analog 16	65
B.2	Yamaha Natural Sound Stereo Amplifiers (AX-490)	65
B.3	Tannoy System 600 Loudspeaker	66
B.4	Listening Environment Layout	68
B.5	Cabling Diagram for the sound field reproduction system.	69
B.6	Photograph of the completed system.	70
B.7	Photograph of the equipment used to drive the speakers.	70
B.8	Screen Capture of the Scope software package	71
B.9	Screen Capture of the Cool Edit Pro Software.	72
B.10	AC-3 Window Function	74
C.1	Sound Field for Demonstration 1.	76
C.2	Sound Field for Demonstrations 2 and 3.	76
C.3	Sound Field for Demonstration 4.	77
C.4	Sound Field for Demonstration 5.	78
E.1	Contents of the attached compact disks.	91

Introduction

1.1 Thesis Motivation

The study of acoustics is almost as old as civilisation. In 3000 B.C. a Chinese philosopher tried to link the pitch of sound to the five elements; earth, water, air, fire and wind [1, p. 1]. Hearing is arguably the most important sense for many land-based animals providing awareness of danger in their surroundings.

The ability to control a sound field¹ within a given region of space is a fundamental problem in signal processing. The multi-frequency nature of sound adds complexity that is not commonly observed in other signal processing problems [2, p. 56]. The aim of this thesis is to present an auditory system that is capable of reproducing an arbitrary sound field within a region of space using an array of loudspeakers. Such a system aims to provide a solution to this fundamental signal processing problem. In the process, a method of defining the performance bounds of reproducing a given sound field is derived.

A system is presented in this thesis that overcomes the limitations of existing reproduction systems operating in either two or three dimensions. The performance of the designed system has been verified through numerical simulation. The effectiveness of the design in a real world situation has been demonstrated by a physical implementation of the two dimensional system. The flexibility of the designed system, to provide enhanced performance in specific applications, is also shown.

The general approach to solving the reproduction problem follows that proposed in [3] for a plane-wave sound field. In this work spherical harmonics were used to provide a mathematical description of a spatial sound field. The major extension to this work, presented in this thesis, is the ability to reproduce arbitrary sound fields and the further investigation of performance measures.

1.2 Literature Survey

Artificial sound reproduction systems, as opposed to musical instruments, began their existence when Bell and Watson invented the telephone in 1877 [4]. Over the last 125 years, sound reproduction systems have been improved in efficiency and quality. They

¹A sound field is defined as the variations in air pressure within a region of space.

are used in a variety of applications (distance education and entertainment) and are included in many electronic devices. Today a variety of methods to reproduce sound exist. An overview of the historically successful and current advances in sound reproduction are presented below.

The first reproduction system that placed sounds in a plane was developed by Blumlein in 1931 [2]. This stereo reproduction system used a combination of delays and level differences between the loudspeakers to create the effect. Stereo reproduction systems became popular in the home in 1958 with the Long Playing (LP) phonograph and Frequency Modulation (FM) broadcasting. Today the majority of commercial surround sound reproduction systems are based on Dolby Surround technology. This technology is still a minor extension of the original work by Blumlein as reproduced sounds are contained within two dimensional space [5].

The next major advancement in sound reproduction was the step towards sound field reproduction. Acoustic events generate sound fields that are described by pressure variations around the mean air pressure. Sound field reproduction systems attempt to reproduce these pressure variations. There are two basic approaches to solving this problem. The first is to recreate the sound field over a region of space (holographic reproduction) and the second is to recreate the sound field at the location of the ears (binaural or transaural reproduction)[6, p. 9]. In this thesis the first method of reproducing a sound field over a region of space is considered.

One of the first holographic systems was proposed by Grenzo in 1972 [7]. This was the first reproduction system to encode height information. His work was later developed as the 'ambisonics' system [8]. The ambisonics system uses first order spherical harmonics to represent and reproduce sound fields. A second order spherical harmonic representation has also been considered in [9]. The major disadvantages of the ambisonics systems are the assumptions that sources radiate plane-waves and that sound fields produced by loudspeakers are plane-waves at the listening position. For these assumptions to be met, the sources and loudspeakers must be placed far from the listener. This can often be impractical in physical applications [10].

Another solution to holographic reproduction is known as wave field synthesis. A sound field within a hall can be acquired by recording impulse responses along an array of microphones. The sound field is reproduced by feeding the impulse responses into an array of loudspeakers [11]. This technique has been applied to evaluate the acoustic properties of halls [12].

A solution to the holographic reproduction of a plane-wave sound field using spherical harmonic techniques is presented in [3]. The proposed solution removes the assumptions of the ambisonics system, modelling each loudspeaker as a point source and has no theoretical limit on the order of spherical harmonic reproduction. Performance bounds are given for an array of loudspeakers to reproduce a plane-wave sound field within a spherical region of space. This forms the bases to the solution given in this thesis for the posed problem.

The use of spherical harmonics for sound field analysis (also referred to as modal anal-

ysis) has been used in the development of beamforming microphone arrays [13]. Spherical harmonics are also used in the study of physical chemistry and computer graphics [14]. A detailed review of current sound field representation and reconstruction techniques along with a discussion of significant psycho-acoustical sound reproduction experiments is presented in [6]. The discussion focuses on a balance between continual development of accurately reproduced sound fields to the error and uncertainty of human sensory systems.

1.3 Commercial Opportunities

In this section, the system presented is compared and positioned to *commercial off-the-shelf* sound reproduction technology. The majority of devices available on the market today support technology produced or licensed to Dolby Laboratories.

Dolby Digital (also known as AC-3) was introduced in 1992 to provide six channels of sound for motion pictures. The six channel system is often referred to as a '5.1' system consisting of five discrete full frequency range channels: left, centre, right, left surround and right surround, plus a sixth '.1' channel for low-frequency effects. An additional two channels have been added to the system to provide rear-left and rear-right surround channels. The fundamental technology that Dolby Digital provides is a lossy compression algorithm for the conversion of analogue and digital data [5, 15].

Other acronyms associated with surround sound technology are DTS and THX. DTS is another multi-channel lossy encoding system created by DTS Technology that competes with Dolby Digital. THX is a set of criteria that applies to the design of equipment used in sound reproduction systems. A movie, audio decoder and amplifier, loudspeaker, or even a cinema must comply to these standards to be THX certified [16].

Current commercial reproduction systems provide the technology to store up to eight channels of digital data to be played through eight spatially located loudspeakers. These systems do not provide the audio producer with tools to record a spatial sound field or specify what audio should be heard through each loudspeaker. Current systems also assume that the loudspeakers are placed in pre-determined positions. This is generally not the case for 95% of home installations [17].

The research presented in this thesis enhances existing sound field reproduction technology by providing a specification of a spatial sound field, the synthesis of sound samples in space and the generation of loudspeaker audio signals. The proposed reproduction system requires that a description of the sound field be stored rather than the audio signals. The sound field description is then used in conjunction with the details of a listener's individual listening environment (the geometry and number of loudspeakers) to produce the loudspeaker audio signals. This allows the optimal listening experience to be individualised for each listener.

The description of the sound field required by the system is stored in a number of discrete channels. It is probable that existing Dolby Digital and DTS technology could be used to encode and store this data. The key modification required to existing decoders

to facilitate the proposed system is enhanced processing power. The processing power is required to calculate the loudspeaker audio signals. The decoder would also be required to allow the user to specify the details of their loudspeaker configuration.

The key improvements that the reproduction system proposed in this thesis exhibits over existing commercial reproduction systems is the provision of a method to represent and synthesise a spatial sound field. In addition to this a specification of a process to transform the sound field into a set of loudspeaker signals required for reproduction is presented.

Over the last two years new technology has begun replacing traditional ‘cube’ loudspeakers with flat panels. The SurfaceSound technology developed by NXT implements a distributed mode loudspeaker. Distributed mode loudspeakers have several driving systems located on a single diaphragm. This has the effect of replacing a single loudspeaker with many smaller loudspeakers. The benefit of this technology is that it provides uniform intensity and frequency balance over a much wider listening window [18]. If loudspeakers using the SurfaceSound technology were used in the system presented in this thesis the reproduced sound field would be improved as these loudspeakers are a closer match to the ideal model that was utilised.

1.4 Thesis Outline

The aim of this thesis has already been discussed in this introductory chapter. The structure of this thesis and the research directions taken are shown in Figure 1.1 below.

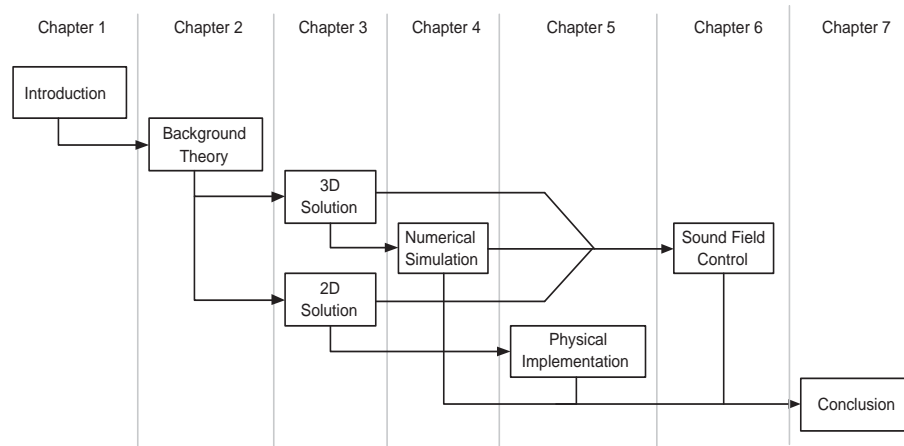


Figure 1.1: Thesis Outline

Chapter 2 introduces the theory used in the remainder of the thesis to generate a solution to the posed problem. This chapter also includes a mathematical description of the problem using the introduced terminology.

Chapter 3 provides a theoretical solution to the sound field reproduction problem. Equations are also introduced to describe the accuracy of the reproduced sound field.

Chapter 3 is divided into two sections: Section 3.1 describes the three dimensional reproduction solution and Section 3.2 describes the two dimensional reproduction solution.

Chapter 4 presents the results of numerical simulations used to verify the three dimensional solution described in Section 3.1.

Chapter 5 describes the design and construction of a physical implementation of the two dimensional system presented in Section 3.2. A solution is provided to the problem to synthesising a sound field and solving the theoretical equations for multiple frequencies.

Chapter 6 uses the theory presented in Chapter 3 and the simulation techniques in Chapter 4 to demonstrate the flexibility of the designed system within the context of two specific applications.

Chapter 7 summaries the results presented in this thesis. Also included are ideas for future research to extend the proposed solution.

1.5 Contributions

This project has made major contributions to sound field reproduction using spherical harmonic techniques. The following list identifies these major contributions:

- Design, simulation and implementation of a system capable of reproducing arbitrary sound fields. The reproduction system has a strong mathematical bases describing the spatial nature of a sound field. The demonstration of the implemented reproduction system has successfully verified the initial design.
- Detailed error analysis is performed for the designed reproduction system. In this analysis the existence of the ‘expansion error’ was discovered. This component of the reproduced sound field error has not been identified previously. Using the error analysis a method of selecting the number of loudspeakers for a particular listening area size is introduced.
- A method of synthesising a spatial sound field using spherical harmonics has been successfully designed and implemented. This method of sound field synthesis could be used in the production of motion picture sound tracks.
- The use of spherical harmonics to describe a spatial field is generally considered in narrowband applications. In this thesis a method of extending this technique for broadband applications has been demonstrated.

Background Theory

This chapter presents the underlying theory of acoustics used in the remainder of this thesis. The problem to be solved is formally described at the end of the chapter.

2.1 Theory of Acoustics

The study of acoustics is formulated from the science of sound or vibrations through solids, liquids and gasses. The sensation of hearing is a direct result of variations in air pressure vibrating the eardrum inside our ears. The classical wave equation describes the spatial and temporal characteristics of a wave function, therefore it can be used to describe an arbitrary sound field. The classical wave equation is given as [19, p. 14]

$$\nabla^2 p(\mathbf{x}, t) = \frac{1}{c^2} \frac{\partial^2 p(\mathbf{x}, t)}{\partial t^2} \quad (2.1)$$

where

$$\text{div grad} = \nabla^2 = \sum_{j=1}^D \frac{\partial}{\partial x_j^2},$$

$p(\mathbf{x}, t)$ is the pressure, \mathbf{x} is the position in \mathbb{R}^D , t is the time and c is the speed of sound. The wave equation essentially states that the amount of matter that flows into an elementary volume produces a corresponding change in density or pressure [1, p. 278].

There are an infinite number of solutions to the wave equation. In this thesis the wave equation is expressed in terms of modes. Modes are orthogonal functions of spatial coordinates obtaining mathematical properties facilitating the analysis of arbitrary sound fields. The modal representation expresses a sound field as a mathematical series. Truncation of the series provides a definite and necessary bound on the clarity of sound field representation and reproduction. Loss of clarity occurs as the distance from the centre of the sound field is extended.

2.2 Coordinate System

Throughout this thesis both two dimensional (2D) and three dimensional (3D) coordinate systems are used. The spherical coordinate system defines a point in 3D space by the distance from the origin (r), the elevation from the vertical axis (θ) and the azimuth (ϕ). This coordinate system is illustrated in Figure 2.1.

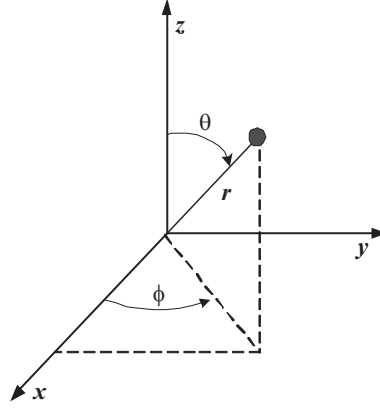


Figure 2.1: Three Dimensional Coordinate System

Spherical coordinates can be converted to cartesian coordinates using the relationships

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

The 2D coordinate system is obtained directly from the 3D system by restricting points to the plane $z = 0$ (or $\theta = \pi/2^\circ$). The 2D or cylindrical coordinate system represents points using the radial distance (r) and the azimuth (ϕ).

2.3 Modal Wave Equation Solution¹

The previously defined pressure field $p(\mathbf{x}; t)$ is dependant on space and time. The time-harmonic dependance of the sound field can be calculated to obtain its space and frequency based properties.

$$p(\mathbf{x}; t) = g(\mathbf{x}; \omega) e^{i\omega t} \quad (2.2)$$

where $g(\mathbf{x}; \omega)$ is the space dependant part of the wave field and ω is the radian frequency.

¹The majority of the content in this section has been adapted from [20, p. 35–39]

Substituting (2.2) into the time dependent wave equation (2.1) results in the *reduced wave equation* also known as the *Helmholtz Equation* [19, p. 15]

$$\nabla^2 g + k^2 g = 0 \quad (2.3)$$

where the wave number

$$k = \frac{2\pi f}{c} = \frac{\omega}{c} = \frac{2\pi}{\lambda}$$

and λ is the wavelength. The wave number (k) is a constant multiple of the radian frequency ω (as the speed of sound is assumed to be independent of its frequency) and is often referred to as the *frequency* within this thesis.

The Helmholtz equation (2.3) can be expressed in terms of spherical coordinates [21, p169]

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial g}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial g}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 g}{\partial \phi^2} + k^2 g = 0. \quad (2.4)$$

The solution to (2.4) for $g(r, \theta, \phi; k)$ can be written as linear combinations of the complex modes in the form [20, p.37] [22, p.194]

$$\sqrt{\frac{2}{\pi}} \begin{matrix} h_n^{(1)}(kr) \\ h_n^{(2)}(kr) \end{matrix} P_{n|m|}(\cos \theta) e^{im\phi}, \quad n \in \mathbb{Z}, m \in \mathbb{Z} \text{ and } |m| < n \quad (2.5)$$

where

$$\begin{aligned} h_n^{(1)}(x) &= j_n(x) + in_n(x) = \sqrt{\frac{\pi}{2x}} \left[J_{n+\frac{1}{2}}(\cdot) + iN_{n+\frac{1}{2}}(\cdot) \right] \\ h_n^{(2)}(x) &= j_n(x) - in_n(x) = \sqrt{\frac{\pi}{2x}} \left[J_{n+\frac{1}{2}}(\cdot) - iN_{n+\frac{1}{2}}(\cdot) \right] \end{aligned}$$

$P_{n|m|}(\cdot)$ is the associated Legendre function. $J_n(\cdot)$ and $j_n(\cdot)$ are the ordinary and spherical Bessel functions of the first kind. $N_n(\cdot)$ and $n_n(\cdot)$ are the ordinary and spherical Bessel functions of the second kind. $h_n^{(1)}(\cdot)$ and $h_n^{(2)}(\cdot)$ are the spherical Hankel functions of the first and second kind respectively.

The associated Legendre function exhibits the following orthogonality property [23, p. 117]

$$\int_{-1}^1 P_{n|m|}(\cos \theta) P_{n'|m'}(\cos \theta) d(\cos \theta) = \frac{2}{2n+1} \frac{(n+|m|)!}{(n-|m|)!} \delta_{nn'} \quad (2.6)$$

where δ is the Kronecker delta function. Similarly for exponentials

$$\int_0^{2\pi} e^{im\phi} e^{-im'\phi} d\phi = 2\pi \delta_{mm'}. \quad (2.7)$$

From (2.6) and (2.7) the spherical harmonics can be defined as

$$Y_{nm}(\theta, \phi) \triangleq \sqrt{\frac{2n+1}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} P_{n|m|}(\cos\theta) e^{im\phi}. \quad (2.8)$$

Using (2.6) and (2.7) it can be shown that

$$\int_0^{2\pi} \int_0^\pi Y_{nm}(\theta, \phi) Y_{pq}^*(\theta, \phi) d\theta d\phi = \delta_{np} \delta_{mq} \quad (2.9)$$

where * denotes the complex conjugate. This orthogonality property is used extensively in the solution to the reproduction problem.

The generalised modal expression for the classical wave equation is defined as

$$g(\mathbf{x}; k) = \sum_{n=0}^{\infty} \sum_{m=-n}^n G_{nm}(x; k) Y_{nm}(\hat{\mathbf{x}}) \quad (2.10)$$

where $G_{nm}(x; k)$ is a set of spherical harmonic coefficients for the spherical harmonics $Y_{nm}(\hat{\mathbf{x}})$. The subscript n , is referred to as the order of the spherical harmonic and m the mode. The modal solution to the wave equation takes on the form of a sum of coefficients multiplied by orthogonal functions. This form of equation is similar in nature to a Fourier series.

2.4 Sound Field Control

The derived sound field expression describes every point within a particular region of space. The Kirchhoff-Helmholtz theorem allows a more efficient description to be obtained. From the theorem it can be said that, ‘the sound field within a source-free volume is fully defined by the sound pressure and pressure gradient on the continuous surface enclosing the volume.’ [24, p. 282–284]

The sound field is defined by $g(\mathbf{x}; k)$ and the pressure gradient is calculated by taking the derivative of $g(\mathbf{x}; k)$ with respect to the distance x . From (2.5) the only term that depends on x is $h_n(x)$, ($h_n(x) = h_n^{(1)}(x)$ or $h_n^{(2)}(x)$). The derivative of this function is [22, p.197]

$$\frac{d}{dx} h_n(x) = h_{n-1}(x) - \frac{n+1}{x} h_n(x) \quad (2.11)$$

The derivative (2.11) only depends on lower order functions of the same kind. If a sound field is reproduced to the N^{th} order then the pressure gradient is also reproduced to the N^{th} order. Therefore reproducing a sound field to order N at radius x_0 ensures that the sound field is reproduced inside the sphere at the same order. A sound field description based on a surrounding surface can be efficiently utilised rather than considering an entire volume.

2.5 Sound Sources

The control of sound sources, in this case loudspeakers, is the key to reproducing a sound field. In the solution to the problem it will be assumed that the loudspeakers act as point sources. The equation used to describe a point source located at position \mathbf{y} is [3]

$$g(\mathbf{x}; k) = \frac{y e^{iky} e^{ik\|\mathbf{y}-\mathbf{x}\|}}{\|\mathbf{y}-\mathbf{x}\|} \quad \mathbf{x} \neq \mathbf{y}. \quad (2.12)$$

This formulation of the point source includes a normalisation term so that as $|\mathbf{y}_l| \rightarrow \infty$ the point source field becomes a plane-wave field,

$$g(\mathbf{x}; k) \rightarrow e^{ikx(\hat{\mathbf{y}}^T \cdot \hat{\mathbf{x}})}. \quad (2.13)$$

The ambisonics solution to the reproduction assumes that the loudspeakers are positioned far enough away that they can be considered plane-wave sources. Defining the loudspeakers as point sources is the key difference between the Ambisonics and the solution presented in this thesis.

Expressions for the point source and plane-wave fields using spherical harmonics are [19, p. 30]

$$\frac{y e^{iky} e^{ik\|\mathbf{y}-\mathbf{x}\|}}{\|\mathbf{y}-\mathbf{x}\|} = \sum_{n=0}^{\infty} \sum_{m=-n}^n 4\pi X_n(kx) R_n(ky) Y_{nm}(\hat{\mathbf{x}}) Y_{nm}^*(\hat{\mathbf{y}}) \quad x < y \quad (2.14)$$

and from [22, p. 227] as

$$e^{ikx(\hat{\mathbf{y}}^T \cdot \hat{\mathbf{x}})} = \sum_{n=0}^{\infty} \sum_{m=-n}^n 4\pi X_n(kx) Y_{nm}(\hat{\mathbf{x}}) Y_{nm}^*(\hat{\mathbf{y}}) \quad (2.15)$$

respectively, where

$$X_n(kx) = i^n j_n(kx)$$

and

$$R_n(ky) = ikye^{iky} i^{-n} h_n^{(1)}(ky).$$

2.6 Problem Formulation

The problem to be solved in this thesis is the reproduction of an arbitrary sound field using an array of loudspeakers. The basic solution to the problem involves obtaining mathematical descriptions of an arbitrary sound field and the sound field produced by loudspeakers. The reproduced sound field will be expressed in terms of loudspeaker signals. As it is desired for the two sound fields to be equal, the two descriptions are equated and the loudspeaker signals obtained. Playing these signals through the loudspeakers will reproduce the sound field therefore solving the problem.

Let $S(\mathbf{x}; k)$ be the sound field to be reproduced, where \mathbf{x} is an arbitrary observation point and k is the wave number. A sound field is generated by any number of farfield and nearfield point sources. An arbitrary sound field can also be represented by the superposition of plane-waves [3]. In this formulation the original sound field is defined in terms of Q plane-waves with arbitrary weights, phases and directions as

$$S(\mathbf{x}; k) = \sum_{q=1}^Q w_q e^{ikx(\hat{\mathbf{y}}_q^T \cdot \hat{\mathbf{x}})} \quad (2.16)$$

where w_q is a complex coefficient and $\hat{\mathbf{y}}_q$ is the direction of the q^{th} plane-wave.

Let L be the number of loudspeakers. From (2.12) the sound field produced by the l^{th} loudspeaker located at the position \mathbf{y}_l is

$$T_l(\mathbf{x}; k) = y_l e^{iky_l} \frac{e^{-ik\|\mathbf{y}_l - \mathbf{x}\|}}{\|\mathbf{y}_l - \mathbf{x}\|}. \quad (2.17)$$

Control of the produced sound field is gained by applying a complex-valued frequency dependant weighting function to each loudspeaker. The sound field produced by combining the L loudspeakers is

$$T(\mathbf{x}; k) = \sum_{l=1}^L a_l(k) T_l(\mathbf{x}; k). \quad (2.18)$$

The reproduction problem is solved when the arbitrary sound field and the reproduced sound field are the same, that is, $S(\mathbf{x}; k) = T(\mathbf{x}; k)$. In the following chapter, this problem is solved in 2D and 3D. In each case equations are developed to show the accuracy of the reproduced sound field as a function of the number of loudspeakers (L), the radial observation distance (x) and the wave number (k).

The key assumptions made when solving the problem were that each loudspeaker is positioned correctly and acts as a pure point source. Also effects such as reverberation in the listening room are ignored.

Sound Field Reconstruction

A solution to the reproduction problem described in Section 2.6 is presented in both two and three dimensions in this chapter. The key functional difference between the systems is their ability to recreate sound sources in either two or three dimensions. The two dimensional system therefore places sounds around the listener while the three dimensional system also places sounds above and below the listener. The three dimensional solution is verified by numerical simulation in Chapter 4 and the two dimensional solution verified by physical implementation in Chapter 5. A formal error analysis of the reproduced sound field is also presented.

3.1 Three Dimensional Reconstruction

3.1.1 Exact Reproduction

The spherical harmonic solution to the classical wave equation (2.10) allows an arbitrary sound field to be described by the set of spherical harmonic coefficients $G_{nm}(x; k)$. These coefficients are dependant on the radial position within a sound field. This is undesirable due to the dependance on the size of the sound field to the number of coefficients required. A more compact form of (2.10) is defined in Appendix A and given by

$$S(\mathbf{x}; k) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \alpha_{nm}(k) j_n(kx) Y_{nm}(\hat{\mathbf{x}}). \quad (3.1)$$

The newly defined sound field coefficients, $\alpha_{nm}(k)$, are a radially independent version of the $G_{nm}(x; k)$ coefficients given in (2.10). They provide a complete description of an arbitrary sound field. The other significant difference between (2.10) and (3.1) is the introduction of the spherical Bessel function $j_n(\cdot)$. A plot of this function for different orders of n is shown in Figure 3.1. The major observation from this figure is that as the order increases the range of x , for which the function is non-zero (or ‘turned on’), shifts along the x -axis. From this observation it is known that the spherical harmonics are only active in particular regions of the sound field depending on their order. More precisely the higher order spherical harmonics are active at increasing distances from the origin.

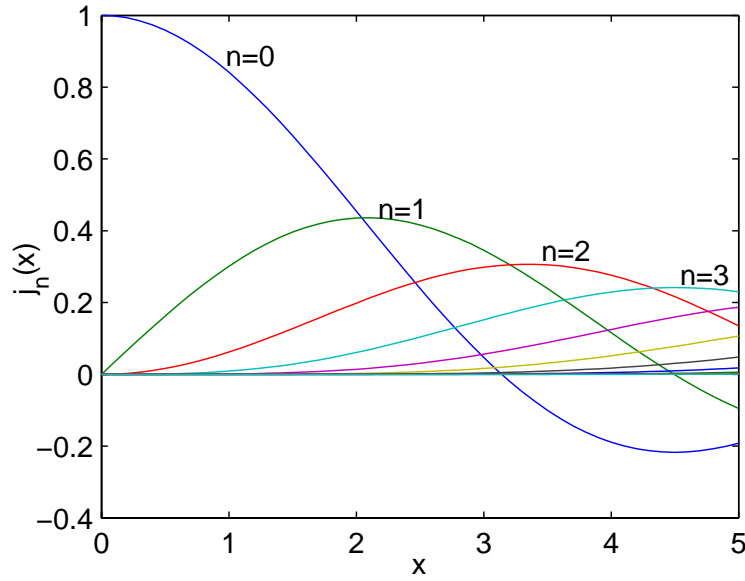


Figure 3.1: Spherical Bessel Function.

The sound field is to be reproduced using an array of loudspeakers. The sound field generated by an array of loudspeakers is obtained by combining (2.18) and (2.14) as

$$T(\mathbf{x}; k) = \sum_{n=0}^{\infty} \sum_{m=-n}^n 4\pi X_n(kx) \times \sum_{l=1}^L a_l(k) R_n(ky_l) Y_{nm}^*(\hat{\mathbf{y}}_l) Y_{nm}(\hat{\mathbf{x}}). \quad (3.2)$$

This equation is valid under the assumption that the observation point $\mathbf{x} = x\hat{\mathbf{x}}$ satisfies $x < \|\mathbf{y}_l\|$, $\forall l$ where \mathbf{y}_l is the location of the l^{th} loudspeaker. That is, all observation points are within the area enclosed by the loudspeakers.

The unknown in (3.2) is the loudspeaker weights (a_l), representing the sound emitted by each loudspeaker. Obtaining the loudspeaker weights for the original sound field solves the reproduction problem. To find the solution, the original sound field (3.1), is equated with the reproduced sound field (3.2) and the loudspeaker weights determined. This gives

$$\alpha_{nm}(k) j_n(kx) Y_{nm}(\hat{\mathbf{x}}) = 4\pi X_n(kx) \times \sum_{l=1}^L a_l(k) R_n(ky_l) Y_{nm}^*(\hat{\mathbf{y}}_l) Y_{nm}(\hat{\mathbf{x}})$$

$$\alpha_{nm}(k) = 4\pi i^n \times \sum_{l=1}^L a_l(k) R_n(ky_l) Y_{nm}^*(\hat{\mathbf{y}}_l) \quad (3.3)$$

for $n = 0, \dots, \infty$ and $|m| < n$.

If the loudspeaker weights are chosen to completely satisfy (3.3), the loudspeakers will exactly reproduce the arbitrary sound field. To completely satisfy (3.3) for all orders of n , requires an infinite set of sound field coefficients and thus an infinite number of loudspeakers. Obtaining loudspeaker weights that completely satisfy (3.3) is impractical. A practical system will consist of a finite number of loudspeakers and thus produce a sound field approximating the original.

3.1.2 Approximate Reproduction

A finite number of loudspeakers will always be used to reproduce a sound field. Consequently, to obtain the weights for the loudspeakers, a finite number of sound field coefficients are required. Using the sound field coefficients $n = 0, \dots, N$ results in an N^{th} order approximation to the reproduction problem. The approximate reproduction solution ensures that the first N spherical harmonics are reproduced exactly. To describe the error introduced by the approximate reproduction the normalised truncation error and the normalised expansion error are defined. These two errors are subsequently combined to give an expression for the total reproduction error in this approximate solution.

Normalised Truncation Error

By truncating the infinite set of spherical harmonics information describing the original sound field is lost. The error resulting from this discarded information is quantified by the normalised truncation error. The normalised truncation error is mathematically defined as

$$\epsilon_N(kx) \triangleq \frac{\int |S(\mathbf{x}; k) - \hat{S}_N(\mathbf{x}; k)|^2 d\hat{\mathbf{x}}}{\int |S(\mathbf{x}; k)|^2 d\hat{\mathbf{x}}} \quad (3.4)$$

where $S(\mathbf{x}; k)$ is the original sound field and $\hat{S}_N(\mathbf{x}; k)$ is the sound field truncated to the N^{th} spherical harmonic. Substituting (3.1) into (3.4) results in the following expression for the normalized truncation error

$$\epsilon_N(kx) = 1 - \frac{\sum_{n=0}^N \sum_{m=-n}^n |\alpha_{nm}(k) j_n(kx)|^2}{\int |S(\mathbf{x}; k)|^2 d\hat{\mathbf{x}}} \quad (3.5)$$

The derivation of (3.5) is given in Appendix A. From this equation it is observed that the truncation error is defined by the terms of the truncated series allowing exact evaluation. To relate the truncation error to the original sound field (3.5) is solved for (2.16) giving

$$\epsilon_N(kx) = 1 - \frac{4\pi \sum_{n=0}^N (j_n(kx))^2 \sum_{m=-n}^n \left| \sum_{q=1}^Q w_q Y_{nm}^*(\hat{\mathbf{y}}_q) \right|^2}{\sum_{q=1}^Q \sum_{p=1}^Q w_q w_p j_0(kx \|\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p\|)} \quad (3.6)$$

The derivation of (3.6) is given in Appendix A. From (3.6) it is observed that the normalised truncation error is dependant on the direction of the plane-waves present in the sound field. In the analysis of the 2D reproduction system an equation identical in nature to (3.6) is derived for the truncation error. As a plane-wave in 2D can be described using a single variable reducing the complexity, the dependence of this error to the nature of the sound field is considered in later discussion.

Normalised Expansion Error

The reproduced sound field is represented by an infinite series of spherical harmonics in (3.2). The entire range of harmonics are present due the physical realisation of the loudspeakers. The spherical harmonics in the reproduced sound field are weighted from the information contained in a limited series of spherical harmonic coefficients describing the original sound field. As discussed previously the first N spherical harmonic orders are perfectly recreated. The normalised expansion error describes the error in the reproduced spherical harmonics $n = N+1, \dots, \infty$. The normalised expansion error is mathematically defined as

$$E_N(kx) \triangleq \frac{\int |\tilde{T}_N(\mathbf{x}; k)|^2 d\hat{\mathbf{x}}}{\int |S(\mathbf{x}; k)|^2 d\hat{\mathbf{x}}} \quad (3.7)$$

where $\tilde{T}_N(\mathbf{x}; k)$ is the sound field produced by considering the reproduced spherical harmonics $n = N+1, \dots, \infty$. The normalised expansion error (3.7) is derived in Appendix A and given as

$$E_N(kx) = \frac{\sum_{n=N+1}^{\infty} \sum_{m=-n}^n |\kappa_{nm}(k) j_n(kx)|^2}{\sum_{n=0}^{\infty} \sum_{m=-n}^n |\alpha_{nm}(k) j_n(kx)|^2} \quad (3.8)$$

where

$$\kappa_{nm}(k) = 4\pi i^n \sum_{l=1}^L a_l(k) R_n(kr_l) Y_{nm}^*(\hat{\mathbf{y}}_l).$$

The expansion error is mathematically dependent on the loudspeaker weightings. Determining the relationship between the original sound field and the expansion error is non-trivial due to the complex relationship between the loudspeaker weights from the original sound field shown in (3.3). For the specific case of a sound field containing a single plane-wave (3.3) can be reduced to [3]

$$P_{n|m|}(\cos \vartheta) e^{-im\varphi} = \sum_{l=1}^L R_n(kr_l) a_l(k) P_{n|m|}(\cos \theta_l) e^{-im\phi_l} \quad (3.9)$$

where ϑ and φ are the elevation and azimuth of the plane-wave. From (3.9) it is observed that the calculation of the loudspeaker weights is dependant on the correlation between the direction of the plane-wave and the position of the loudspeakers.

In an initial attempt to demonstrate such a relationship the expansion error was calculated as the direction of a single plane-wave was varied. The fourth order expansion error at the radius $kx = 5.4$ is shown in Figure 3.2. The angular distances between the direction of the plane-wave and the three closest loudspeakers are also included. From the figure it is observed that the magnitude of the expansion error depends on the angular distance of the loudspeakers to the plane-wave. To characterise this relationship further the effect of the associated Legendre and the complex exponential to the angular distance needs to be considered.

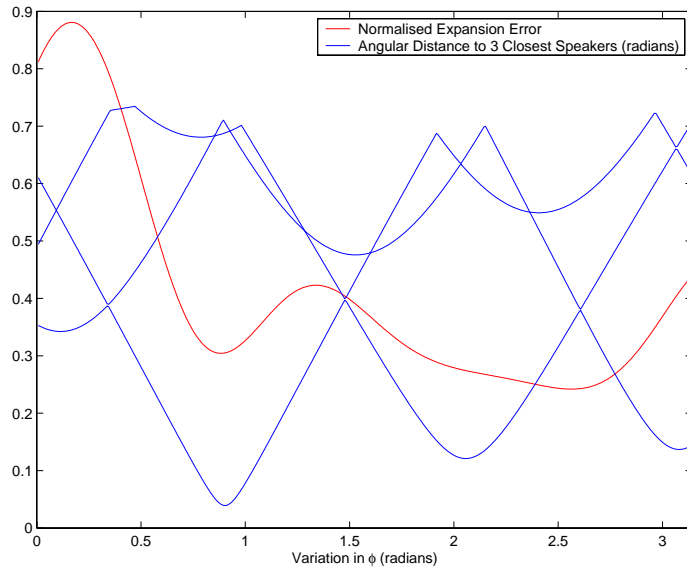


Figure 3.2: Normalised expansion error as plane-wave direction is varied.

Normalised Reproduction Error

An overall expression for the reproduction error can be obtained from the consideration of the normalised truncation and expansion errors. The normalised reproduction error is mathematically defined as

$$\Upsilon_N(kx) \triangleq \frac{\int |S(\mathbf{x}; k) - T(\mathbf{x}; k)|^2 d\hat{\mathbf{x}}}{\int |S(\mathbf{x}; k)|^2 d\hat{\mathbf{x}}} \quad (3.10)$$

where $T(\mathbf{x}; k) = \hat{S}_N(\mathbf{x}; k) + \tilde{T}_N(\mathbf{x}; k)$ represents the reproduced sound field. The normalised reproduction error expressed in spherical harmonics is

$$\Upsilon_N(kx) = \frac{\sum_{n=N+1}^{\infty} \sum_{m=-n}^n |[\alpha_{nm}(k) - \kappa_{nm}(k)] j_n(kx)|^2}{\sum_{n=0}^{\infty} \sum_{m=-n}^n |\alpha_{nm}(k) j_n(kx)|^2}. \quad (3.11)$$

The derivation of this equation is given in Appendix A. The error expressions derived for the approximate solution were evaluated over a thousand sound fields each containing fifty plane-waves with random amplitudes, directions and phases. The mean and standard deviation observed from the error equations is shown in Figure 3.3. The number of loudspeakers used to evaluate the expansion and reproduction errors was selected such that $L = (N + 1)^2 + N$. The selection of more loudspeakers than modes reduced the errors associated with uniform placement of the loudspeakers on the surface of a sphere (see Section 3.1.4). Each loudspeaker was placed at a distance of one metre from the origin.

The significant observation from Figure 3.3 is that each of the mean normalised error graphs are similar. That is, each reproduction order increases monotonically from the origin and that increasing the order N decreases the error.

Using the results from the normalised reproduction error it is observed that given the wave number k , the reproduction radius x and the reproduction order N equal to kx the expected reproduction is 8% with a standard deviation of 2%. Using this observation, it can be stated that for an arbitrary sound field given the maximum wave number k and the maximum reproduction radius x , if the reproduction order is selected such that

$$N = \lceil kx \rceil \quad (3.12)$$

then the expected reproduction error will be in the order of 8%. For a sound field containing a single plane-wave a similar observation was made in [3] where the expected error was 4%.

3.1.3 Calculating the Loudspeaker Weights

To satisfy (3.3) for the first N orders, the loudspeaker weights are calculated using the linear relationship

$$\mathbf{P}\mathbf{a} = \mathbf{b} \quad (3.13)$$

where

$$\mathbf{P} = 4\pi \begin{bmatrix} i^0 \mathbf{p}_0 \\ i^1 \mathbf{p}_1 \\ \vdots \\ i^n \mathbf{p}_n \end{bmatrix}, \quad \mathbf{p}_n = \begin{bmatrix} R_n(ky_1) Y_{n-n}^*(\hat{\mathbf{y}}_1) & \dots & R_n(ky_L) Y_{n-n}^*(\hat{\mathbf{y}}_L) \\ \vdots & & \vdots \\ R_n(ky_1) Y_{n0}^*(\hat{\mathbf{y}}_1) & \dots & R_n(ky_L) Y_{n0}^*(\hat{\mathbf{y}}_L) \\ \vdots & & \vdots \\ R_n(ky_1) Y_{nn}^*(\hat{\mathbf{y}}_1) & \dots & R_n(ky_L) Y_{nn}^*(\hat{\mathbf{y}}_L) \end{bmatrix} \quad (3.14)$$

is a $K \times L$ matrix

$$\mathbf{a} = \left[a_1(k) \quad a_2(k) \quad \dots \quad a_L(k) \right]^T \quad (3.15)$$

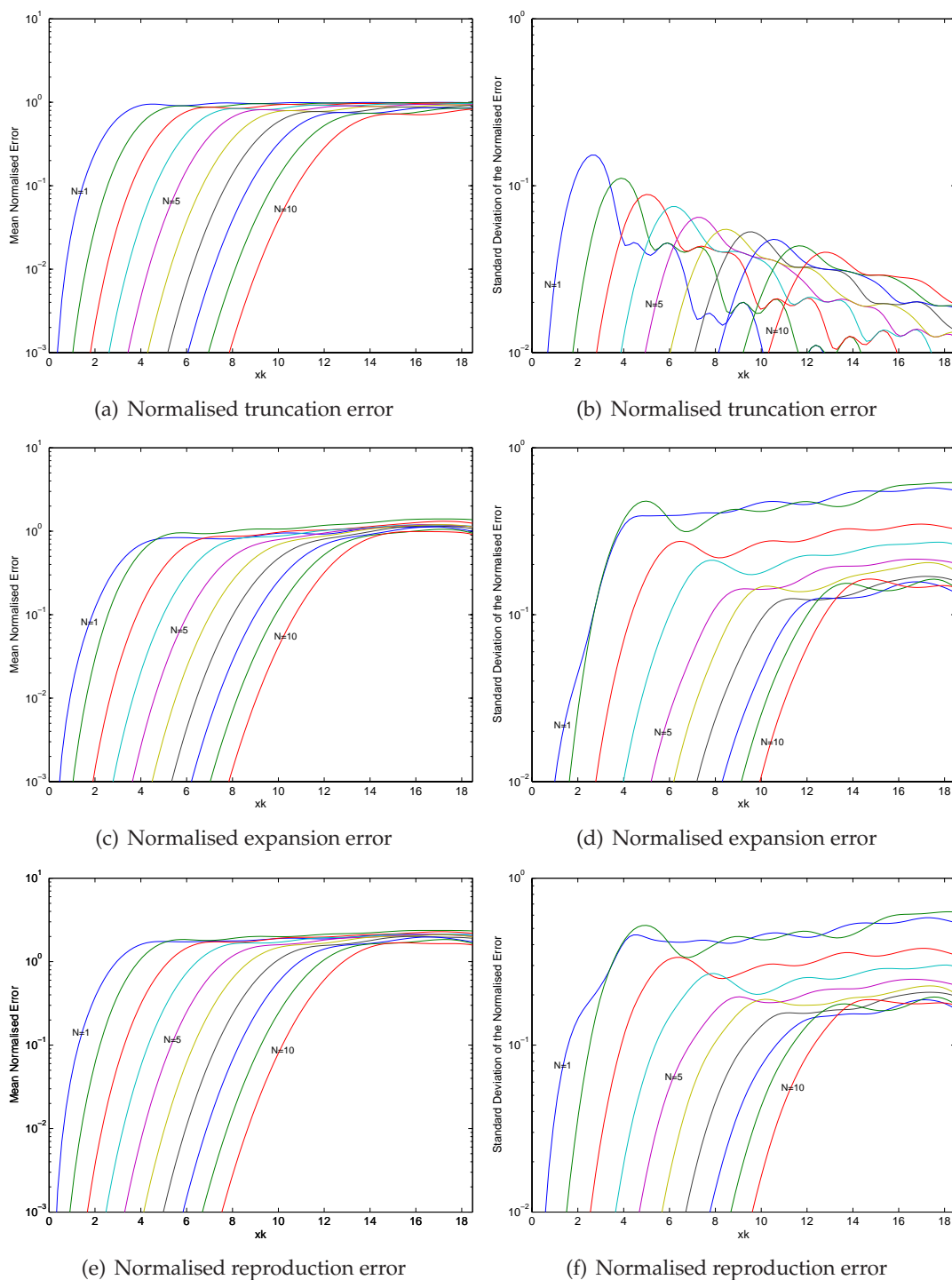


Figure 3.3: Plots of formulated 3D error equations for the approximate reproduction solution. (a), (c) and (e) contain the mean normalised error observed from a thousand random fifty plane-wave sound fields, (b), (d) and (f) shows the standard deviation of the normalised error from these sound fields. The plane-wave frequency for all sound fields was 1 kHz.

is a $L \times 1$ matrix representing the loudspeaker weights and

$$\mathbf{b} = \left[\alpha_{00}(k) \quad \alpha_{1-1}(k) \quad \alpha_{10}(k) \quad \alpha_{11}(k) \quad \dots \quad \alpha_{nm}(k) \right]^T \quad (3.16)$$

is a $K \times 1$ matrix corresponding to the sound field coefficients of the original sound field. (3.13) describes a system of K equations and L unknown loudspeaker weights. K is the number of sound field coefficients and is related to the order N by

$$K = \sum_{n=0}^N (2n+1) = (N+1)^2. \quad (3.17)$$

The loudspeaker weights can be calculated by solving the system of linear equations described by (3.13). If the number of loudspeakers L matches the number of sound field coefficients K the weights can be calculated exactly otherwise they are estimated.

If $K = L$ then the system (3.13) can be solved using

$$\mathbf{a} = \mathbf{P}^{-1}\mathbf{b}. \quad (3.18)$$

This is under the condition that \mathbf{P}^{-1} exists i.e. \mathbf{P} is non-singular. For \mathbf{P} to be non-singular $\det(\mathbf{P}) \neq 0$. To condition \mathbf{P} such that it is non-singular the position of the loudspeakers must be sparse.

If $K > L$ then (3.13) will form an over-determined linear system. An exact solution for the loudspeaker weights will not exist. The loudspeaker weights can be estimated by solving the least squares solution [25, p.236]

$$\min_{\mathbf{a}} \|\mathbf{P}\mathbf{a} - \mathbf{b}\|^2$$

Finally if $K < L$ the system is under-determined. In this case there is either no solution for the loudspeaker weights or an infinite number of solutions. In the case of infinite solutions the loudspeaker weights would be selected under the conditions of [25, p.271]

$$\min_{\mathbf{a}} \|\mathbf{a}\|^2 \text{ subject to } \mathbf{P}\mathbf{a} = \mathbf{b}$$

When solving (3.13) for the under and over determined cases the pseudo-inverse of \mathbf{P} is calculated giving $\mathbf{P}^\#$. The loudspeaker weights are then calculated as

$$\mathbf{a} = \mathbf{P}^\#\mathbf{b}. \quad (3.19)$$

All simulations presented in this thesis are calculated under the conditions of $K = L$ unless otherwise stated. The effect on the estimation of the loudspeaker weights is examined in Section 4.3.

3.1.4 Loudspeaker Placement

In the solution presented for the reproduction problem, the placement of the loudspeakers has not been specified. In the previous section it was stated that the loudspeakers must be placed sparsely so that the loudspeaker weights can be calculated. In [9] it is suggested that the loudspeakers be placed such that the angles between them and their distance to the centre of the listening area be equal. Positioning the loudspeakers to satisfy these conditions allows the reproduction system to reproduce sounds equally from all directions.

If the loudspeakers are placed on the surface of a sphere, the second condition of having equidistant loudspeakers to the centre of the listening area is satisfied. Unfortunately there is no known method of satisfying the first condition of equidistant angles between the loudspeakers. A solution to the *sphere packing problem* in mathematics provides approximate solutions to this problem. A library of 3D packings is available at [26] providing coordinates for up to 130 points, or in this application 130 loudspeakers. Figure 3.4 shows the positioning of 25 and 36 loudspeakers on a sphere as given by the library. These positions are used in the simulations provided in Chapter 4.

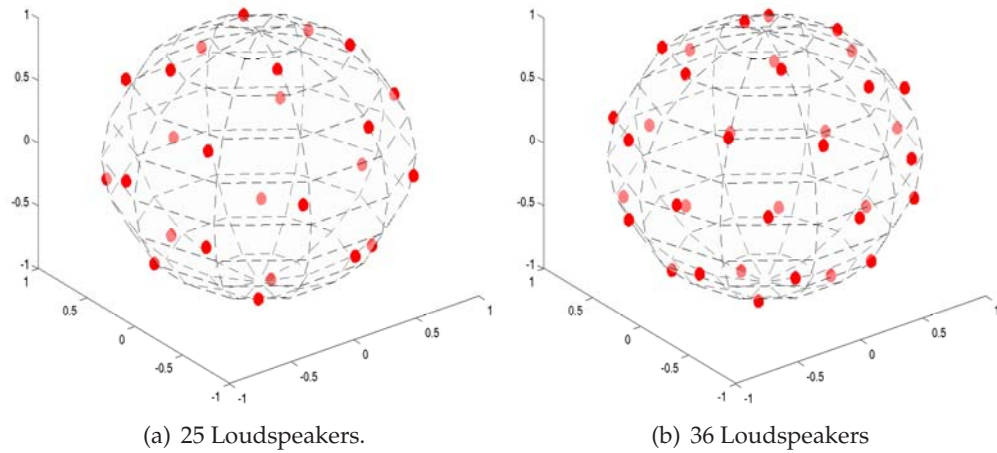


Figure 3.4: The two figures show the layout of twenty-five and thirty-six loudspeakers placed in an uniform packing around a unit sphere.

3.1.5 Alternative Loudspeaker Weight Calculation

In Section 2.4 it was shown that if a surface of a volume is reproduced to the N^{th} order, the sound field within the surface is also reproduced to the same order. To reproduce a sound field on the surface of a sphere, consider a theoretical continuous loudspeaker with radius r . To describe the sound field at the surface of the sphere a continuous weighting function for the theoretical loudspeaker is obtained.

The sound field produced by the continuous loudspeaker is given by

$$T(\mathbf{x}; k) = \sum_{n=0}^{\infty} \sum_{m=-n}^n 4\pi X_n(kx) R_n(kr) \times \int a(\hat{\mathbf{y}}; k) Y_{nm}^*(\hat{\mathbf{y}}) d\hat{\mathbf{y}} Y_{nm}(\hat{\mathbf{x}}). \quad (3.20)$$

The above equation is similar in nature to (3.2) where the summation has been replaced by an integral. The continuous loudspeaker weighting function $a(\hat{\mathbf{y}}; k)$ is only defined on the surface of a sphere. This arbitrary weighting function $a(\hat{\mathbf{y}}; k)$ can be expressed using spherical harmonics as

$$a(\hat{\mathbf{y}}; k) = \sum_{p=0}^{\infty} \sum_{q=-p}^p \gamma_{pq}(k) Y_{pq}(\hat{\mathbf{y}}) \quad (3.21)$$

where $\gamma_{pq}(k)$ is a set of coefficients that completely describe this function.

Substituting (3.21) back into (3.20) and equating with the original sound field (3.1) gives

$$\alpha_{nm}(k) = 4\pi i^n R_n(kr) \int \gamma_{pq}(k) Y_{pq}(\hat{\mathbf{y}}) Y_{nm}^*(\hat{\mathbf{y}}) d\hat{\mathbf{y}}.$$

Using the orthogonality property of spherical harmonics (2.9) the following relationship between the original sound field coefficients and the coefficients of the continuous loudspeaker is obtained

$$\gamma_{nm}(k) = \frac{\alpha_{nm}(k)}{4\pi R_n(kr)}. \quad (3.22)$$

To realise the theoretical loudspeaker physical loudspeakers are placed at the locations previously defined in Section 3.1.4. The loudspeaker weights are then obtained by sampling the continuous weighting function at the location of the physical loudspeakers. This alternative method places an additional constraint that all loudspeakers be equidistant from the origin.

For accurate reproduction the sampled loudspeaker weights need to be scaled by the surface area that they are acting over. The surface area is calculated by considering the surface area of the cap of a sphere [27]

$$SA_{cap} = 2\pi r h \quad (3.23)$$

where the height (h) is calculated from the minimum angle (v) between adjacent points in the sphere as

$$h = r - \cos\left(\frac{v}{2}\right) \times r. \quad (3.24)$$

Note that this calculation is only an approximation as v is not constant for all loudspeakers. This implies that the area that each loudspeaker is acting over is not symmetrical. A comparison of the two methods used to calculate the loudspeaker weights is given in Section 4.4.

3.2 Two Dimensional Reconstruction

3.2.1 Exact Reproduction

When solving the reproduction problem, the solution can be restricted to the recreation of sound sources located in a 2D plane. The key difference between the 2D and 3D solutions is the use of cylindrical harmonics instead of spherical harmonics. This solution is used in Chapter 5 where the sound system is physically implemented.

The arbitrary sound field given by (2.16) is defined in 2D using cylindrical coordinates in Appendix A as

$$S(\mathbf{x}; k) = \sum_{m=-\infty}^{\infty} \alpha_m(k) J_m(kx) e^{im\phi} \quad (3.25)$$

where ϕ is the azimuth of an arbitrary observation point. Similarly, the reproduced sound field (2.18) expressed in cylindrical harmonics is

$$T(\mathbf{x}; k) = \sum_{m=-\infty}^{\infty} J_m(kx) e^{im\phi} \sum_{l=1}^L a_l(k) B_m(kr_l) e^{-im\varphi_l} \quad (3.26)$$

where φ_l is the azimuth of the l^{th} loudspeaker,

$$B_m(kr) = \pi i k r e^{ikr} H_m^{(1)}(kr),$$

$$H_m^{(1)}(kr) = J_m(kr) + iY_m(kr)$$

and $H_m^{(1)}(\cdot)$ is commonly referred to as the ordinary Hankel function of the first kind.

As in the 3D case, the unknown in the 2D reproduction equation (3.26) is the loudspeaker weights, $a_l(k)$. A solution to the 2D reproduction problem is obtained by equating (3.25) with (3.26) and solving for the loudspeaker weights. This gives

$$\alpha_m(k) J_m(kx) e^{im\phi} = J_m(kx) e^{im\phi} \sum_{l=1}^L a_l(k) B_m(kr_l) e^{-im\varphi_l}$$

$$\alpha_m(k) = \sum_{l=1}^L a_l(k) B_m(kr_l) e^{-im\varphi_l} \quad (3.27)$$

for $m = 0, \dots, \infty$.

This result is identical in nature to that found previously. If the loudspeaker weights exactly satisfy (3.27) then the sound field will be reproduced exactly. To obtain loudspeaker weights that exactly satisfy this equation requires an infinite number of loudspeakers. This condition is not practically possible. As only a finite number of loudspeakers would ever be available, an approximation to the reproduction problem must be made.

3.2.2 Approximate Reproduction

In Section 3.1.2 the effect of limiting the sound field coefficients used to solve the reproduction problem was discussed. In the solution for 2D reproduction the same condition occurs. Using the error definitions given in Section 3.1.2, expressions for the errors present in the 2D solution are provided.

The 2D normalised truncation error is given as

$$\varepsilon_M(kx) = 1 - \frac{2\pi \sum_{m=-M}^M |\alpha_m(k) J_m(kx)|^2}{\int |S(\mathbf{x}; k)|^2 d\phi}. \quad (3.28)$$

For a sound field consisting of Q arbitrary plane-waves the truncation error is given as

$$\varepsilon_M(kx) = 1 - \frac{\sum_{m=-M}^M (J_m(kx))^2 \left| \sum_{q=1}^Q w_q e^{-im\varphi_q} \right|^2}{\sum_{q=1}^Q \sum_{p=1}^Q w_q w_p J_0(kx \|\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p\|)}. \quad (3.29)$$

The derivation of (3.28) and (3.29) is given in Appendix A. The form of (3.29) is similar to that given for the 3D reproduction solution. It is observed from the above equation, that the truncation error is dependent on the difference in direction of the plane-waves contained in the sound field. To demonstrate this dependance Figure 3.5 shows the fourth order normalised truncation error calculated at the position $xk = 5.4$ as the angular difference between two plane-waves was increased from 0 to $2\pi^c$.

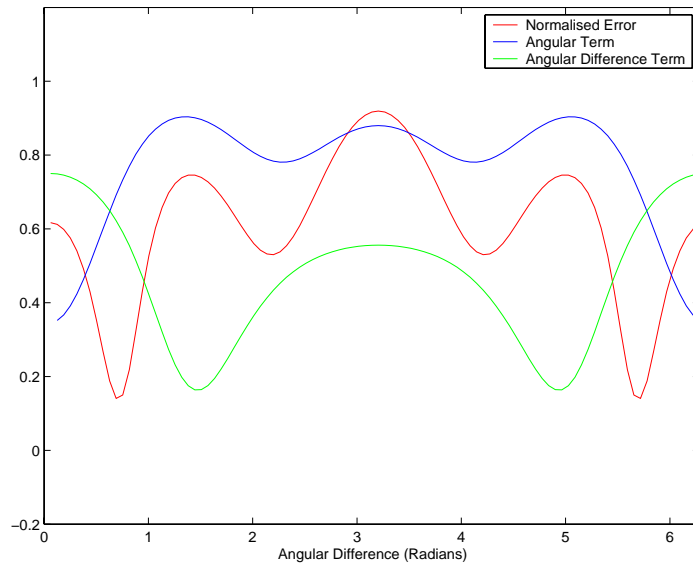


Figure 3.5: Normalised Truncation Error for two plane-wave sound field.

To illustrate the observed angular dependence in the derived equation, plots of the angular term $e^{-im\varphi_q}$ and the angular difference term $J_0(kx \|\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p\|)$ of (3.29) are included in the figure. It is observed from the figure that the sinusoidal nature of the truncation error is related to this angular term. It was observed from additional experiments that the frequency of the sinusoid is dependent on the mode of the reconstruction. It also appears that the angular difference component provides a form of modulation on the amplitude of the truncation error.

The 2D normalised expansion error is derived in Appendix A as

$$E_M(kx) = \frac{\sum_{m=-\infty}^{-M-1} |\kappa_m(k) J_m(kx)|^2 + \sum_{m=M+1}^{\infty} |\kappa_m(k) J_m(kx)|^2}{\sum_{m=-\infty}^{\infty} |\alpha_m(k) J_m(kx)|^2} \quad (3.30)$$

where

$$\kappa_m(k) = \sum_{l=1}^L a_l(k) B_m(kr_l) e^{-im\varphi_l}.$$

As in Section 3.1.2, the 2D normalised expansion error depends on the correlation of the directions of the plane-waves in the sound field to the location of the loudspeakers. The 2D normalised reproduction error is given as

$$\Upsilon_M(kx) = \frac{\sum_{m=-\infty}^{-M-1} |(\alpha_m(k) - \kappa_m(k)) J_m(kx)|^2 + \sum_{m=M+1}^{\infty} |(\alpha_m(k) - \kappa_m(k)) J_m(kx)|^2}{\sum_{m=-\infty}^{\infty} |\alpha_m(k) J_m(kx)|^2}. \quad (3.31)$$

The derivation of (3.31) is given in Appendix A. To obtain a measure of the performance of the 2D reproduction system for arbitrary sound fields, the mean and standard deviation of the normalised truncation, expansion and reproduction errors was calculated over a thousand sound fields consisting of fifty random plane-waves. The results of this calculation is shown in Figure 3.6. When calculating the expansion and truncation errors the number of loudspeakers was selected such that $L = 2M + 1$. The loudspeakers were uniformly distributed on a circle of radius one metre.

As the reproduction radius increases towards the loudspeaker positions the expansion and reproduction errors increase rapidly. This is due to the limited radially dependent information used to recreate the sound field. Using the results obtained for the total reproduction error it was found that given the reproduction radius x , wave number k and a reproduction order M equal to kx the expected error is 12% with a standard deviation of 7%. Thus, given a desired reproduction radius and maximum wave number kx , the reproduction order should be selected such that

$$M = \lceil kx \rceil. \quad (3.32)$$

This conclusion was also made in [3] for the reproduction of a single plane-wave although in that case the reproduction error was predicted to be 10%.

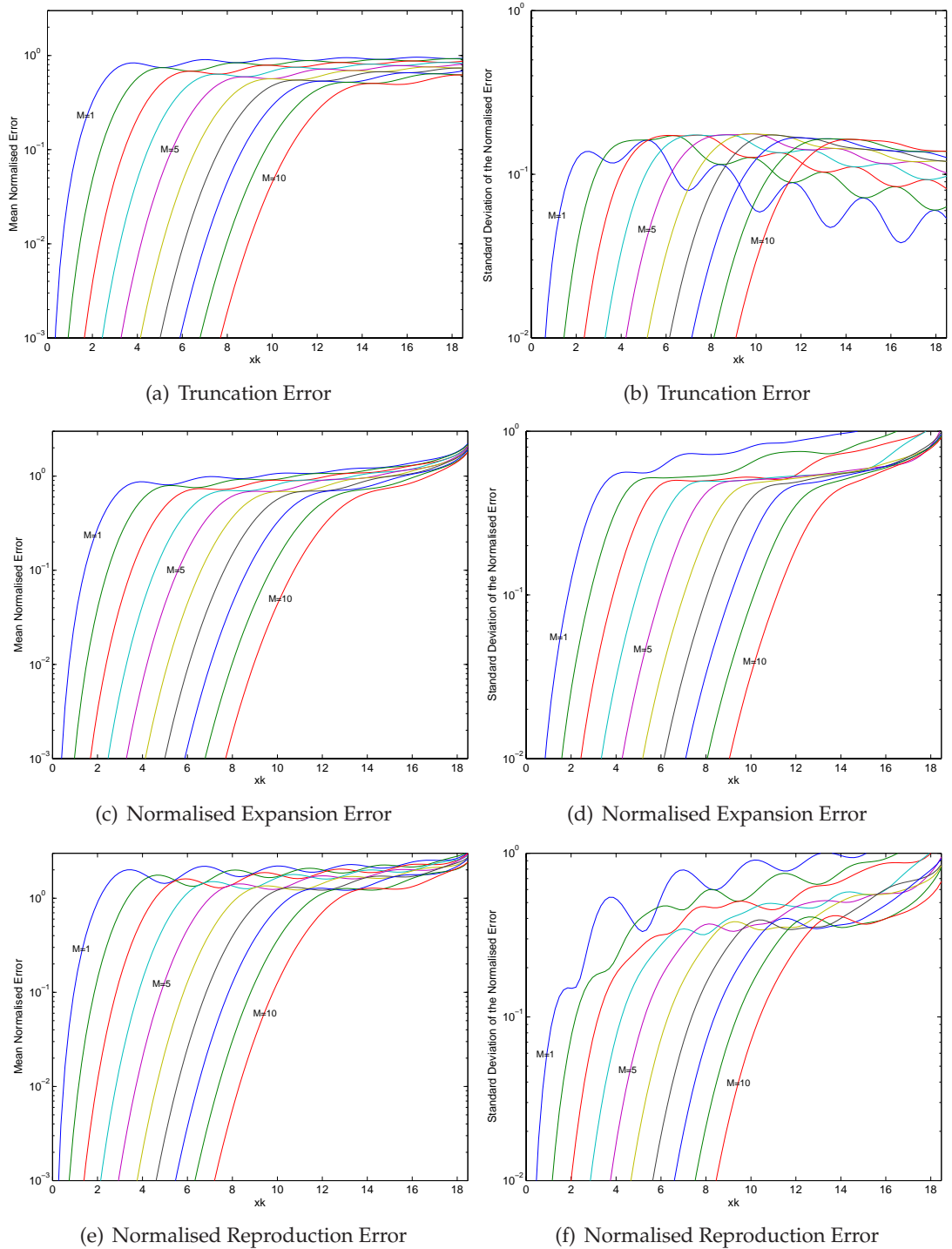


Figure 3.6: Plots of formulated 2D error equations for the approximate reproduction solution. (a), (c) and (e) contain the mean error observed from a thousand random fifty plane-wave sound fields, (b), (d) and (f) contain the standard deviation of the normalised error from these sound fields. The plane-wave frequency for all sound fields was 1 kHz.

3.2.3 Calculating the Loudspeaker Weights

To calculate the loudspeaker weights that satisfy (3.27) for $m = 0, \dots, M$ a system of linear equations is defined as

$$\mathbf{P}\mathbf{a} = \mathbf{b} \quad (3.33)$$

where

$$\mathbf{P} = \begin{bmatrix} B_{-m}(r_1k) e^{im\varphi_1} & \dots & B_{-m}(r_Lk) e^{im\varphi_L} \\ \vdots & & \vdots \\ B_0(r_1k) e^{-i0\varphi_1} & \dots & B_0(r_Lk) e^{i0\varphi_L} \\ \vdots & & \vdots \\ B_m(r_1k) e^{-im\varphi_1} & \dots & B_m(r_Lk) e^{-im\varphi_L} \end{bmatrix} \quad (3.34)$$

is a $K \times L$ matrix

$$\mathbf{a} = \begin{bmatrix} a_1(k) & a_2(k) & \dots & a_l(k) \end{bmatrix}^T \quad (3.35)$$

is a $L \times 1$ matrix representing the loudspeaker weights and

$$\mathbf{b} = \begin{bmatrix} \alpha_{-m}(k) & \dots & \alpha_0(k) & \dots & \alpha_m(k) \end{bmatrix}^T \quad (3.36)$$

is a $K \times 1$ matrix corresponding to the sound field coefficients of the original sound field. To calculate the loudspeaker weights the system of K equations is solved for the L unknown loudspeaker weights. K is the number of sound field coefficients and is related to the number of modes M by

$$K = 2 \times M + 1 \quad (3.37)$$

As discussed in Section 3.1.3, the loudspeaker weights for approximate reproduction are determined exactly if the number of loudspeakers and sound field coefficients match. If this is not the case the methods described in Section 3.1.3 are used to obtain estimates for the loudspeaker weights.

3.2.4 Loudspeaker Placement

To obtain the most accurate reconstruction for an arbitrary sound field each loudspeaker should be placed such that their distance from the origin and the angles between them are equal. Fortunately for the 2D case these positions can be calculated exactly. The position of the l^{th} loudspeaker on a circle of radius r is given by

$$y_l = \left[r \quad \frac{(l-1) \times 2\pi}{L} \right]$$

where L is the number of loudspeakers.

Numerical Simulation

To verify the solution to the 3D reproduction system described in Section 3.1 numerical simulations were performed. In particular, the derived error equations are compared with simulation results, the effect of estimating the loudspeaker weights is examined and a comparison of the two techniques derived for calculating the loudspeaker weights is made.

The simulations were developed using Mathworks MatLab. The bases of the scripts used to perform the simulations are included in Appendix D. All scripts used to obtain the results in this chapter have been included in the attached media.

4.1 Three Dimensional Sound Field Reconstruction

To simulate the reproduction system, a sound field that is to be reproduced is defined corresponding to (2.16). The sound field used in the following simulations consists of five 1 kHz plane-waves defined by $\mathbf{y}_q = [\theta \ \phi]$ and the complex coefficients w_q :

$$\begin{aligned} \mathbf{y}_1 &= [1.9551 \quad 2.7987] & w_1 &= 0.0563 \times e^{-1.5908ik} \\ \mathbf{y}_2 &= [1.6315 \quad 0.8934] & w_2 &= 0.1456 \times e^{3.1043ik} \\ \mathbf{y}_3 &= [2.0247 \quad -1.6075] & w_3 &= 0.5460 \times e^{2.5483ik} \\ \mathbf{y}_4 &= [2.3131 \quad -0.7491] & w_4 &= 0.1963 \times e^{0.9439ik} \\ \mathbf{y}_5 &= [1.9971 \quad 0.7717] & w_5 &= 0.3468 \times e^{-1.3941ik} \end{aligned}$$

The spherical radius for the listening area of the reproduction system was selected as $x_0 = 0.2\text{m}$. For the 1 kHz sound field the corresponding wave number is $k = 18.48$. Using (3.12) the reproduction order of $N = 4$ is selected. To ensure the loudspeaker weights are calculated exactly the number of loudspeakers is selected as $L = (N + 1)^2 = 25$. The loudspeakers were evenly distributed on a sphere with radius $r = 1\text{m}$.

The original and reproduced sound fields are shown in Figures 4.1 and 4.2. The figures show a slice of the 3D sound field along the $x - y$ plane at the heights $z=0.1$ and 0 metres respectively. The images along the first row of each figure illustrates the original sound field and the images along the second row illustrate the reproduced sound field. The images in the first column illustrate the real component of the sound field, and the second column illustrates the imaginary components. For each image, the horizontal axis is the x -axis and the vertical axis is the y -axis. The red circle indicates the previously selected listening area.

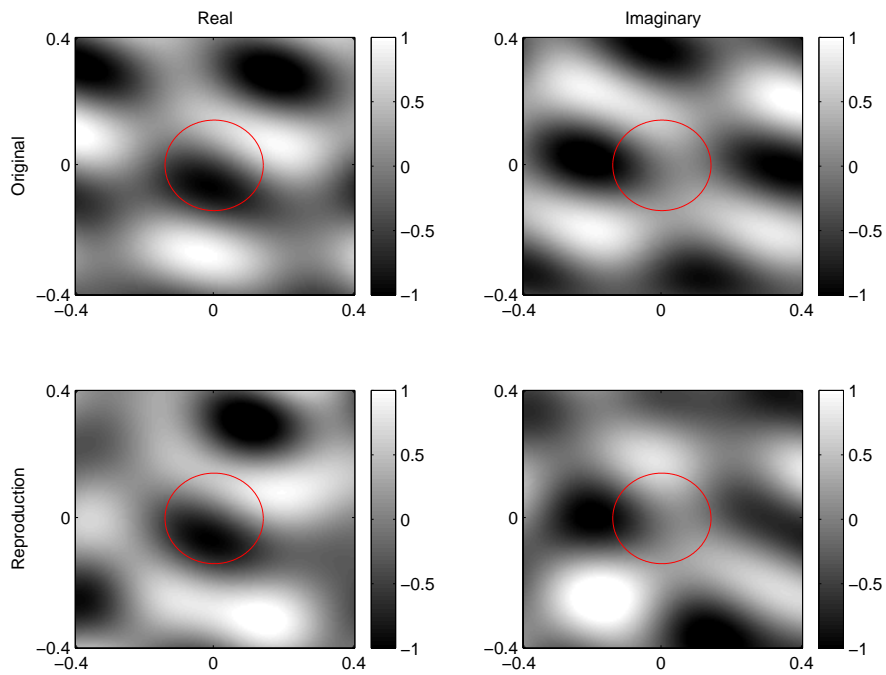


Figure 4.1: Sound field reproduction at height $z = 0.1\text{m}$

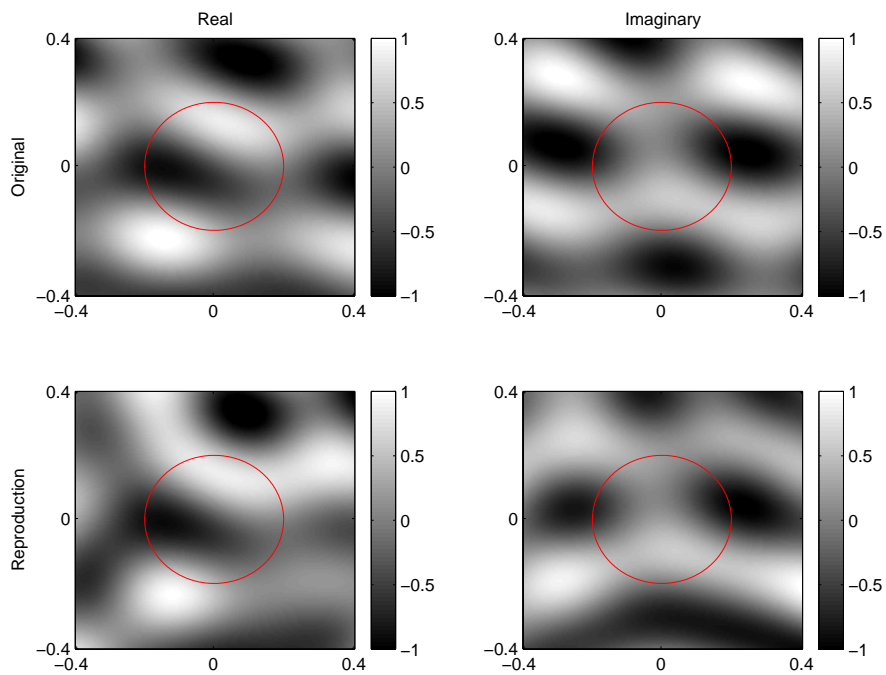


Figure 4.2: Sound field reproduction at height $z = 0\text{m}$

4.1.1 Error Measurement

To measure the accuracy of the 3D reproduction system using numerical simulation, the original and reproduced sound fields are compared at a specified spherical radius. This method allows direct verification of the equations given in Section 3.1.2. In sound field reproduction literature, this measure of error is referred to as the *Integrated Wavefront Error* [28, p. 16].

Alternatively, an error measurement could be calculated by comparing the volume enclosed by a sphere of a specified radius. Using Kirchhoff-Helmholtz theorem (see Section 2.4) additional knowledge about the reconstruction error is not gained from this measure. Also, the *volume error* can be calculated from the *surface error* by integration.

The normalised error of the reproduced sound field in the previous figures is 2.50%. From (3.11) the predicted error is 2.47%. The difference between the simulated and predicted error is likely to be caused by the size of the discrete blocks used to represent the volume on the sound field.

4.2 Error Equation Verification

Using the five plane-waves as defined in Section 4.1, Figure 4.3 illustrates the truncation and expansion components of the reproduced sound field. The original sound field is shown in Figure 4.3(a). The truncation of this sound field (the first four spherical harmonics) is shown in Figure 4.3(b). Figure 4.3(c) shows the spherical harmonics from $n = N + 1 \dots \infty$ of the reconstructed sound field. Finally Figure 4.3(d) shows the reconstructed sound field. Figure 4.3(d) is equal to the sum of Figures 4.3(b) and 4.3(c). From these figures, it is demonstrated that as the order of spherical harmonic increases, their impact on the sound field occurs further from the centre. This is as expected due to the spherical Bessel function contained in (3.1).

To verify that the equations presented in Section 3.1.2 are correct, Figure 4.4 shows a comparison of the errors obtained via simulation and from the derived equations. Figure 4.4(a) shows the errors obtained via simulation from a sound field containing 50 random plane-waves. Using the same sound field, the errors calculated from the derived equations is shown in Figure 4.4(b).

From the figure it can be seen that the two methods of obtaining a measure of error are practically identical. The maximum numerical difference between the simulated and calculated results was 0.5%.

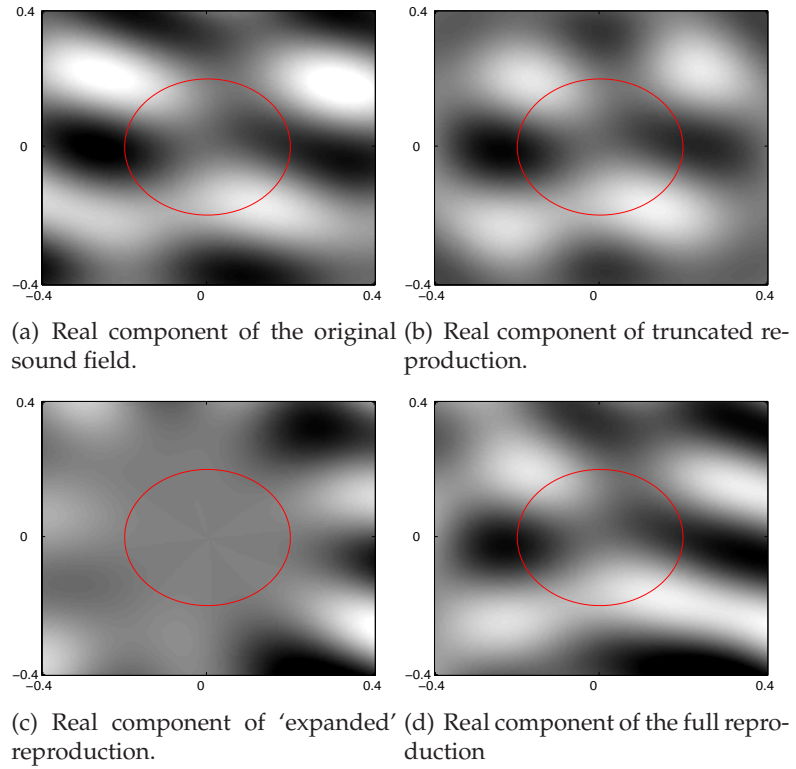


Figure 4.3: Illustration of the three errors defined in Section 3.1.2. The height of the sound field illustrated is $z = 0\text{m}$, the twenty-five loudspeakers were placed on a radius of $r = 1\text{m}$.

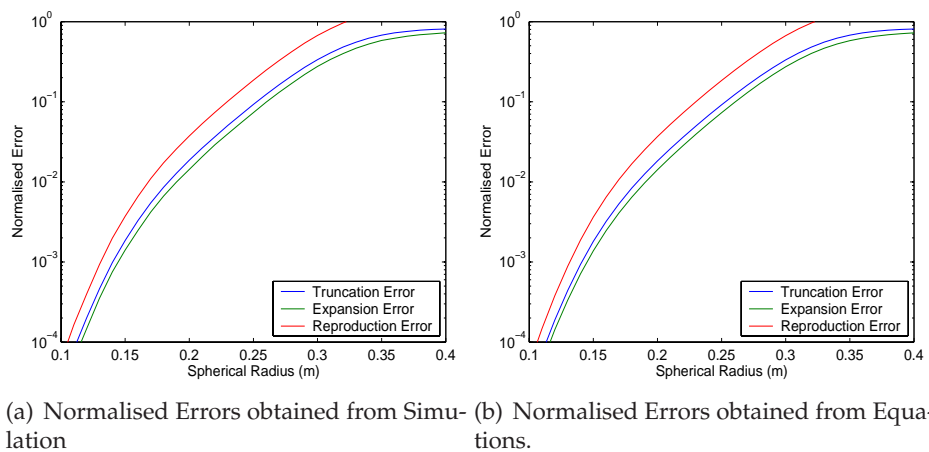


Figure 4.4: Comparison of the reproduction errors measured via numeric simulation and using the error formulas presented in Section 3.1.2. The sound field consisted of fifty random plane-waves at 1 kHz.

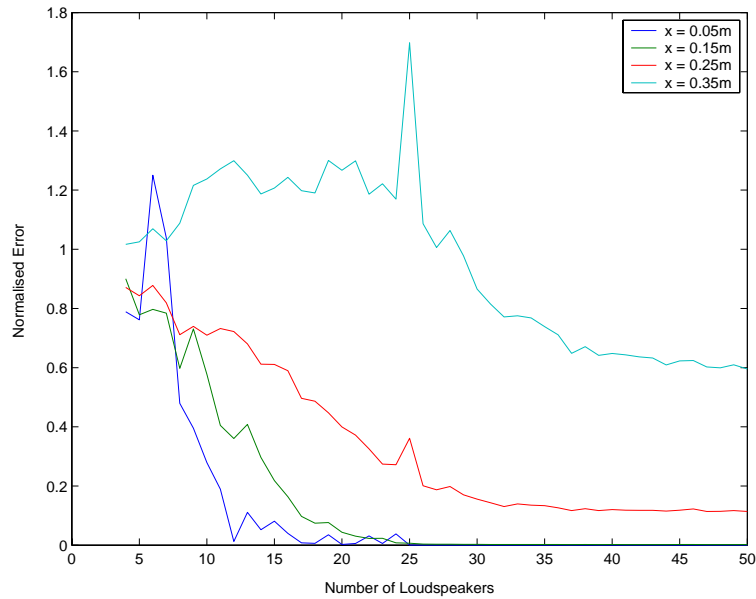


Figure 4.5: Effect of varying the number of loudspeakers for a fixed number of sound field coefficients on the estimation of the loudspeaker weights. The reproduction errors at the four spherical radii were averaged from ten random fifty plane-wave sound fields at 1 kHz. The loudspeaker radius was 1m.

4.3 Effect of Loudspeaker Weight Estimation

In Section 3.1.3 it was stated that if the number of sound field coefficients (K) and the number of loudspeakers (L) did not match, the loudspeaker weights would be estimated. In this section the effectiveness of this estimation is investigated.

Figure 4.5 shows the measured reproduction error when the sound field coefficients were held constant at $K = 25$, corresponding to a fourth order spherical harmonic reconstruction and the number of loudspeakers were varied from $L = 4$ to $L = 50$. The number of loudspeakers required for exact calculating of the loudspeaker weights is $L = 25$. The figure shows the variation in error at four spherical surfaces within the reconstructed sound field. From these results, the general observation is made that increasing the number of loudspeakers decreases the error in the reproduced sound field.

Upon closer examination it is observed that adding more loudspeakers does not give a continual decrease in reproduction error. This is particularly evident for the spherical radius $x = 0.25\text{m}$ and $L \geq 33$. This levelling is present due to the fixed number of sound field coefficients and thus the fixed amount of information available to describe the sound field. It would be expected that the error lines would ‘level-out’ at $L = 25$ corresponding to exact reproduction. As observed, this is not the case. The difference is caused by the approximations used to place the loudspeakers equally around the surface of the sphere (as discussed in Section 3.1.4). The spatial distribution on the loudspeakers increases as their number increases, allowing the complete description contained in the $K = 25$ coefficients to be reproduced.

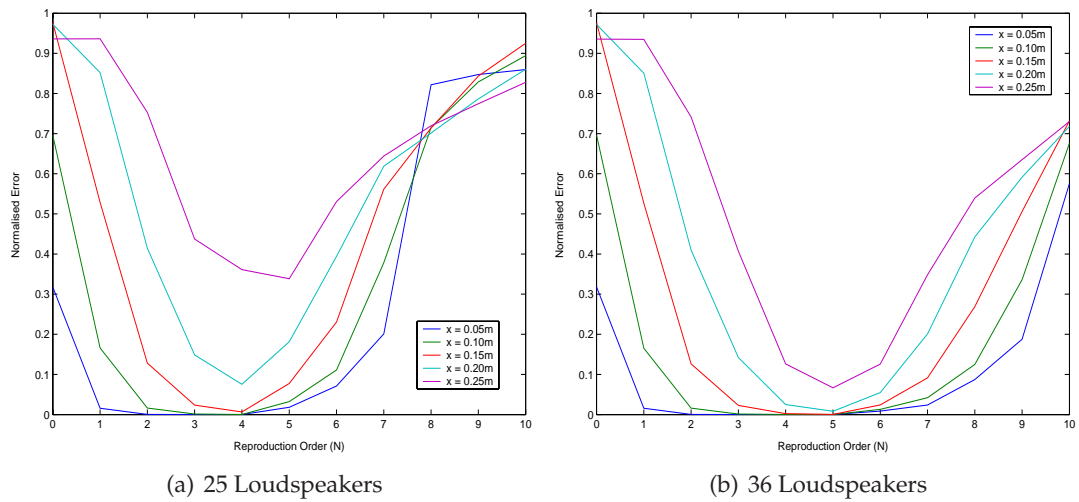


Figure 4.6: Reproduction error observed at different radii when the reproduction order (N) is varied for a fixed number of loudspeakers. Figure (a) shows the result for $L=25$ and Figure (b) shows the result for $L=36$. The results shown have been averaged from ten randomly selected sound fields containing fifty plane-waves at 1 kHz. The loudspeaker radius was 1m.

An alternative method of investigating the effect of loudspeaker weight estimation is to hold the number of loudspeakers constant and vary the number of sound field coefficients. The results of this experiment is shown in Figure 4.6.

From the figure, it is observed that increasing the number of sound field coefficients (K) to calculate the loudspeaker weights has a positive effect up until K becomes greater than L . When $K > L$, the linear system used to calculate the loudspeaker weights is over determined. As K is increased, the higher order spherical harmonic components of the sound field are used to generate the loudspeaker weights. The reproduction system is now attempting to reproduce more modes than loudspeakers. The effect of this, as observed in the figure, is to increase the reproduction error.

It can also be seen in Figure 4.6(a), that for the higher orders of sound field coefficients ($N=8, 9$ and 10), the reproduction error decreases with an increasing distance from the origin. Higher order spherical harmonics represent the sound field further from the origin (as seen by the spherical Bessel term in (3.1)) and contain more modes than the lower harmonics. When the over-determined system is solved, each mode is weighted equally, therefore the higher order spherical harmonics will have a greater presence in the reproduced sound field.

Using the results of Figure 4.5 and Figure 4.6 when calculating the loudspeaker weights, the following rules of thumb should be used:

1. If a sound field is represented by K sound field coefficients, increasing the number of loudspeakers above K will decrease the reproduction error until an asymptotic limit is reached;
2. For a fixed number of loudspeakers, when calculating the loudspeaker weights, K should be limited by L .

4.4 Comparison of Solutions

In Section 3.1, two methods for obtaining the loudspeaker weights were presented. The first was based on the formulation of matrices (Section 3.1.3) and the second based on the sampling of a weighting function for a continuous loudspeaker (Section 3.1.5). A comparison of the reproduction error in the produced sound fields for the two methods is shown in Figure 4.7. From the figure it can be seen that the matrix method produces a more accurate sound field at the centre, whereas the continuous method produces a more accurate sound field at a distance from the centre.

The radius of the ‘theoretical’ loudspeaker used to calculate the loudspeaker weights was $x = 0.2\text{m}$. The sound field at this radius will not contain particular spherical harmonics as determined by the spherical Bessel function. When the loudspeaker weights are calculated using this method they will not reproduce the spherical harmonic which are not present at the radius of the ‘theoretical’ loudspeaker. This is observed in Figure 4.7 where the continuous method performs poorly at the centre of the sound field. The listener of the reproduced sound field hears the sound field that is produced at the location of their ears, not in the center of their head (assuming they are positioned correctly). Therefore, the errors introduced by using the continuous loudspeaker method may be negligible in an implementation environment.

From this discussion it is concluded that using the ‘continuous loudspeaker’ method reproduces the spherical harmonic components of the sound field that exist in the sound field of the selected radius. In a practical implementation this method is likely to outperform the ‘matrix calculation’ method.

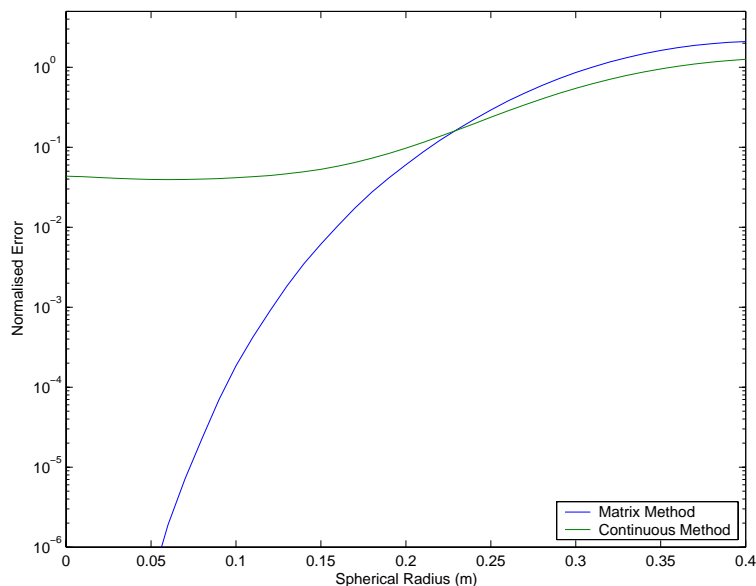


Figure 4.7: Comparison of the ‘matrix’ and ‘continuous’ methods for obtaining the loudspeaker weights. The results are averaged over twenty-five sound fields containing fifty random plane-waves at 1 kHz. The radius of the loudspeakers was 1m.

4.5 Summary of Results

In this chapter the 3D sound field reconstruction system has been verified using numerical simulation. The significant results from this chapter have been

- The derived error equations correspond closely to the results obtained from simulation.
- Increasing the number of loudspeakers used to reproduce the sound field has a limited benefit when the number of sound field coefficients remains constant.
- Increasing the number of sound field coefficients above the number of loudspeakers can have a negative effect on the quality of the reproduced sound field.
- The method of obtaining the loudspeaker weights by consideration of a continuous loudspeaker is likely to outperform the matrix based calculation method in a practical implementation.

Physical Implementation

In the SIGCOM Acoustics Room, located in the Engineering building at the Australian National University, a 2D reconstruction system was built in an echoic chamber.¹The system was built to demonstrate the already presented mathematical solution. Although the system was built from consideration of the 2D solution the same methodology applies to the 3D solution.

The system consists of eight loudspeakers uniformly spaced around a circle of radius 1.5 metres. Precalculated loudspeaker signals are passed through a digital to analogue converter, an audio amplifier and finally through to the loudspeakers. A full technical description of the constructed sound system is given in Appendix B.

5.1 Design

The sound field reconstruction system that was theoretically described in Chapter 3 is pictorially illustrated in Figure 5.1. The system takes as its inputs the sound field coefficients, and the coordinates of each loudspeaker. With these inputs, the system produces loudspeaker signals, recreating the sound field when played through the loudspeakers.

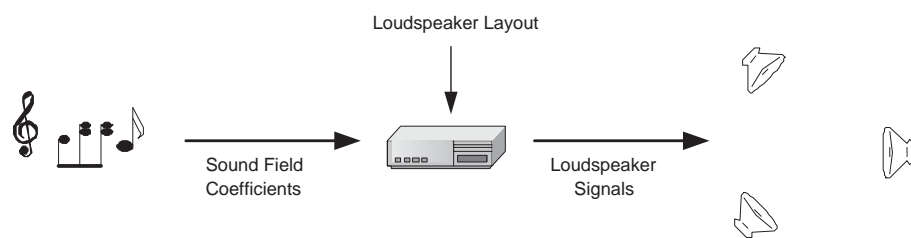


Figure 5.1: Basic System Implementation

The system is not limited by the order or modes of the sound field coefficients or the number or location of the loudspeakers. Given the inputs to the system, it dynamically adapts to produce the most accurate sound field reproduction.

The design of the listening environment was limited by the physical size of the echoic chamber and the available equipment. The following equipment was made available by

¹The echoic chamber is used to reduce un-modelled reverberation in the room. The echoic chamber also has a secondary function of isolating the listening environment from unwanted noise.

the Department of Engineering, FEIT and the Department of Telecommunications Engineering, RSISE;

- Echoic Chamber $3.5\text{m} \times 5.7\text{m} \times 2.3\text{m}$ (width \times breadth \times height),
- $11 \times$ loudspeakers,
- $4 \times$ stereo (2 channels) audio amplifiers,
- $16 \times$ channel output DSP and
- $8 \times$ loudspeaker stands $h = 1.2\text{m}$.

Using this equipment an eight loudspeaker system was produced. The number of loudspeakers was limited by the four stereo audio amplifiers. Due to the difficulty of placing loudspeakers accurately in 3D, a 2D system was constructed. The 2D system will produce a more accurate surface area reproduction than the corresponding 3D volume reproduction with the same radius. The loudspeakers were placed on loudspeaker stands, in a circle 45° apart, with a radius of 1.5 m. A detailed diagram of the layout of the system is provided in Appendix B, Figure B.4.

With the design and construction of the listening environment complete loudspeaker signals were required to begin testing. The first step in generating the signals is to identify the sound field to be reproduced. This process is described in Section 5.2. Using the description of the sound field and (3.33), the loudspeaker signals can be generated as required. This process is described in Section 5.3.

The system is to be implemented digitally, therefore the variables t and k represent *discrete* time and frequency in this chapter. The digital to analogue converter used as the interface between the computer and audio amplifiers supports the sampling rates of 44.1 kHz and 48.0 kHz. In this implementation, the lower sampling rate of 44.1 kHz was selected as the higher frequency representation was not required.

5.2 Sound Field Synthesis

In order to reproduce a sound field a sufficient description of that sound field is required. In Section 3.2, (3.25) allows a sound field to be described by a set of sound field coefficients. As the implemented system consists of eight loudspeakers, from the discussion in Section 3.2.3, nine sound field coefficients will be required, providing a 4th order cylindrical expansion of the sound field.

A microphone based system that records a sound field to produce the required sound field coefficients has not been built. The design and theory of such a system has been published in [29] and further researched in [30].

With no recording system, a method of synthesising a sound field was developed. The general methodology for synthesis is to obtain a sound sample and associate it with a point source. The sound sample can be placed in motion by changing the location of the point source.

Let $s(t)$ define the sample to be synthesised. The sound sample $s(t)$ is converted to the frequency domain using the Discrete Fourier Transform (DFT) to match the frequency

dependant sound field coefficients. Instead of transforming the sound sample into the frequency domain for all time (requiring an infinite number of samples), it is separated into frames.

When calculating the DFT over a finite number of samples, the samples must be multiplied by a windowing or shaping function, $w(t)$, to reduce spectral leakage. Spectral leakage causes the spectrum of frequency components in $s(t)$ to spread into adjacent frequencies. It occurs as the samples used in the DFT are assumed to be periodic for all time [31, p. 95–97].

To reduce spectral leakage, a windowing function is applied to the data reducing the order of discontinuity at the boundary of the periodic extension (start and end of the data frames). This is achieved by matching as many derivatives as possible at the boundary. This is generally achieved by setting the derivatives equal to zero [32, p. 52].

Applying the windowing function results in a significant portion of the data to be ignored due to the near zero values at the boundaries. To remove this negative effect, the window function is applied to overlapping partitions of the sound sample [32, p.56]. This implies that the total energy added to the signal by the windowing function should equal unity, that is

$$\sum_p |w(t - p(N_w - O_w))|^2 = 1 \quad (5.1)$$

where N_w is the length of $w(t)$ and therefore also the length of the data frame, O_w is the amount of overlap between data frames and p is the current frame number.

The windowing function selected for this implementation has been obtained from the standard for the AC-3 algorithm [33]. The AC-3 algorithm is principally used for the compression of digital audio. This windowing function exhibits the boundary properties required to reduce spectral leakage and is well suited for audio applications. The details of the windowing function are provided in [33, p. 105] and repeated in Table B.2. The length of the windowing function is $N_w = 512$ samples. The overlap required to satisfy (5.1) is $O_w = 256$ samples.

The process of generating a frame of data is illustrated in Figure 5.2. An example sound sample, $s(t)$ is shown in Figure 5.2(a). This sample is then multiplied by the shifted windowing functions in Figure 5.2(b) to produce the frames of data shown in Figure 5.2(c). Let $s_p(t)$ denote a frame of data. The frequency components of the data frame can be found using the DFT

$$s_p(t) \xrightarrow{\mathfrak{F}} s_p(k).$$

Using the frequency domain representation of the data frame, each frequency that has less power than a predefined threshold is removed. Completing this step allows the quantity of data to be reduced, a form of data compression. From the subset of frequencies, and the spherical harmonic representation of a 2D point source [19, p. 67], the $\alpha_m(k)$

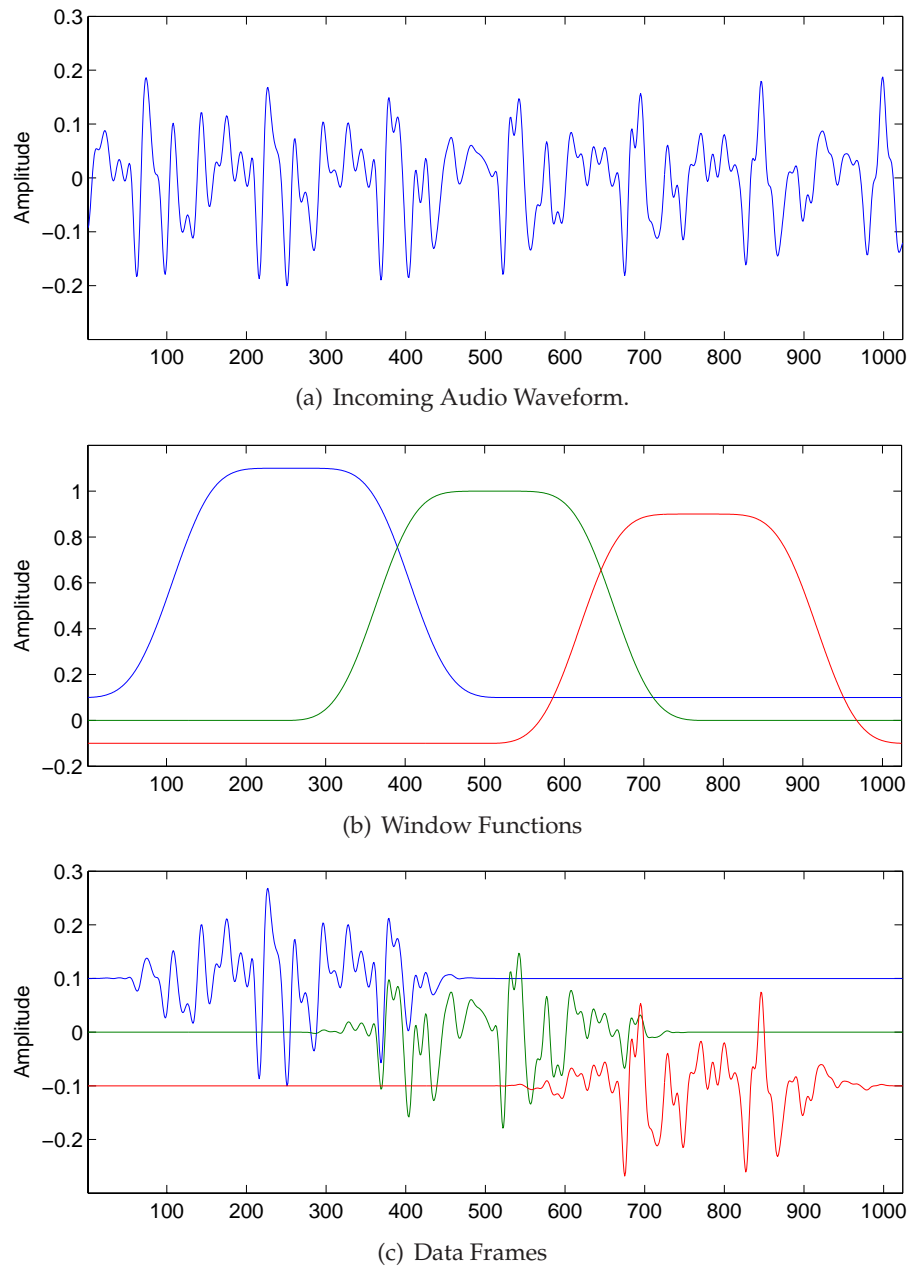


Figure 5.2: Illustration of the conversion of sampled audio to data frames. Note: In (b) and (c) the blue and red lines have been shifted in amplitude by +0.1 and -0.1 respectively to ease readability.

sound field coefficients are calculated as

$$\alpha_{pm}(k) = \sum_{m=-M}^M s_p(k) \pi i w(k) H_m^1(kr_p) e^{-im\phi_p} \quad (5.2)$$

where r_p and ϕ_p are the radius and azimuth of the point source for the duration of frame p of data and $w(k)$ is used to control the sound field intensity. To synthesise a sound field with multiple sound sources (5.2) is evaluated for each sound sample and the sound field coefficients summed.

If the sound source is located at a sufficient distance from the centre of the listening area there is a significant delay between the sound propagating from the source to when it is heard in the listening area. As the sound field coefficients record the sound field at the centre of the listening area these delays are observed. This effect is also present when the loudspeaker signals are calculated. For a sound to be present in the listen area at time t_n , it must have left the loudspeaker at time t_m where $t_m < t_n$. To manage these delays the frames of data is padded with zeros to avoid the unwanted effects of the cyclic nature of the DFT. The addition of these zeros allows the sample of data to move in time, rather than wrap within the current data frame [34, p548-560].

In the applications of sound field recording and the transmission of the sound fields (TV, radio) there is a requirement for the storage and broadcasting of the sound field coefficients. In the implementation process it was observed that the sound field coefficients were even ($\alpha_m(k) = \alpha_m^*(-k)$). It is therefore possible to produce real time varying sound field coefficient signals $\alpha_m(t)$. The number of orders or modes used to record the sound field would determine how many of these time varying signals would be required for the storage and broadcast.

Using the time varying sound field coefficient signals, existing audio compression algorithms (AC-3, MP3 etc.) could be used for compression. A moderate quality MP3 encoding compresses 1 minute of signal channel audio to approximately 0.5 megabytes (MB) of data. In the implemented system, nine sound field coefficient signals are present, and therefore a transmission rate of 4.5 MB per minute would be required. Alternatively the sound field coefficients could be stored for 60 hours on a double sided, dual layer DVD (with the storage capacity of 15.9 gigabytes (GB)). As a comparison, the AC-3 algorithm encodes '5.1' audio channels at a rate of 2.9 MB per minute [33, p. 15]. This allows recording for up to 90 hours on the same sized DVD.

5.3 Loudspeaker Signals

Using the nine sound field coefficients generated by the synthesis process, eight loudspeaker signals are required. The mismatch of coefficients to loudspeakers requires that the over-determined solution of (3.33) be used to produce the frequency dependant loudspeaker weights, that is

$$\mathbf{a}_p(k) = \mathbf{P}^\#(k) \mathbf{b}_p(k). \quad (5.3)$$

Note that $\mathbf{P}^\#(k)$ is not dependent on the current data frame and therefore can be calculated off line using the loudspeaker geometry for each discrete frequency k .

Using the frequency dependent loudspeaker weights, time based loudspeaker signals can be calculated for each frame with the inverse DFT

$$a_{lp}(k) \xrightarrow{\mathfrak{S}^{-1}} a_{lp}(t),$$

where $a_{lp}(t)$ is a time-based data frame for the l^{th} loudspeaker. Each of the loudspeaker data frames are multiplied by the window function and then added together to obtain the loudspeaker signals,

$$a_l(t) = \sum_p w(t - p(N_w - O_w) - d_p) \times a_{lp}(t - p(N_w - O_w))$$

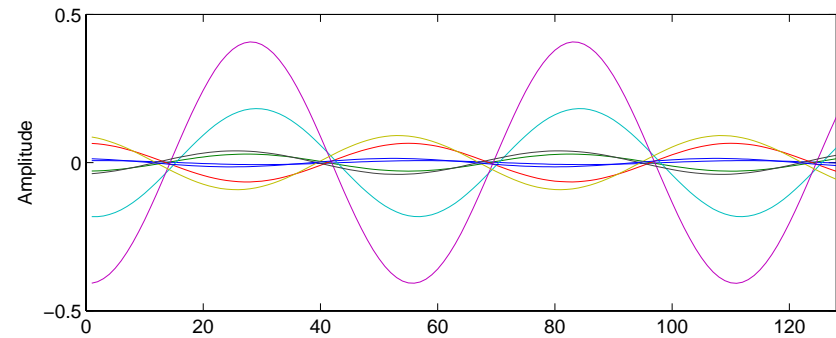
where d_p is a time shift applied to the windowing function. The shift depends on the radial position of the sound source and is required due to the natural delay of the time signals as discussed in the previous section.

The loudspeaker signals need to be transformed from an array of numbers to electrical signals used to excite the loudspeakers. To do this the loudspeaker signals are paired together and the data saved in four Microsoft Pulse Code Modulated (PCM) Wave format files. Each loudspeaker is assigned to either the left or right channel within a file. These files are then loaded into an audio application. The left and right channels of each audio file is associated with an output channel on the digital to analog converter. The eight channels of the digital to analog converter are then amplified by audio amplifiers and feed into the loudspeakers.

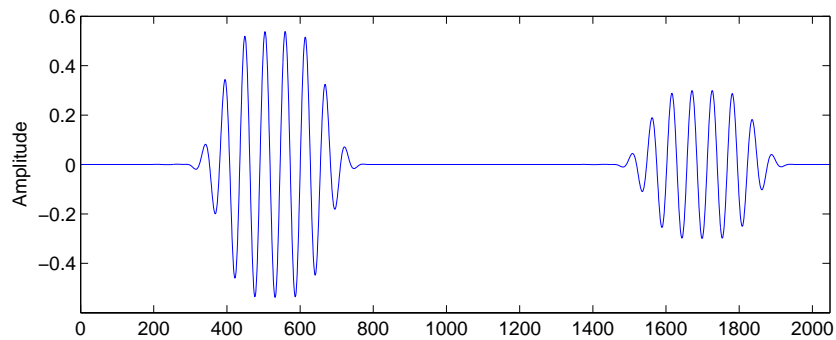
To illustrate that the reproduction system is not simply an amplitude panning algorithm the loudspeaker waveforms generated for a narrow band sound field is shown in Figure 5.3(a). From the figure it can be seen that there is a phase difference between the signals. Similarly, Figure 5.3(b) shows a loudspeaker data frame that consists of two sound samples placed at different depths. The delay introduced by the system to realise the depths is shown in this figure. Therefore the system introduces amplitude, phase and delay adjustments to reproduce the sound field.

5.4 Computational Cost

In this section the processing power required to generate the loudspeaker signals from the sound field coefficients is considered in relation to a real-time implementation of the system. The computational size and its relationship to the sound field and reproduction quality determines how much computational power is required to produce the loudspeaker signals. The process described in Section 5.3 is given below in pseudo-code for the conversion of a single data frame.



(a) Loudspeaker waveforms for a sound field containing 1 kHz point source located at $r = 2.5m$ and $\phi = 166^\circ$.



(b) A window of a loudspeaker waveform illustrating delay of two sound samples.

Figure 5.3: Loudspeaker waveform properties

```

loudspeaker_k = zeros(L, 512);

% Calculate the frequency dependent loudspeaker weights
for i=1:valid_frequencies
    % L * K by K x 1 matrix multiplication
    a = P_Inverse(frequency_index(i)) * b(i)
    % Map selected frequencies to complete frequency axis
    loudspeaker_k(:, frequency_index(i)) = a
end

% Calculate time based loudspeaker signals. The function
% loudspeaker_buffer implements the overlapping of data frames
% and sends the data to an audio amplifier
for i=1:L
    loudspeaker_signal = ifft(loudspeaker_k(i,:))
    loudspeaker_buffer(i, loudspeaker_signal)
end

```

From the pseudo-code, it can be seen that the time to produce the loudspeaker signals for one frame is linearly dependent on the number of sound field coefficients (K), the number of loudspeakers (L) and the frequency content of the frame.

Using a system implemented in MatLab on a 2.0 GHz Pentium IV (Mobile) CPU and 512 Megabytes of main memory the execution time to transform the sound field coefficient into loudspeaker signals for various audio files is presented in Table 5.1. The

number of loudspeakers and sound field coefficients was held constant to eight and nine respectively. The sample rate for each audio file was 44.1 kHz.

Audio Type	Length (m:s)	Freq Compression (%)	Execution Time (m:s)
Speech 1	0:60.0	95.2	0:24.18
Speech 2	0:60.0	94.8	0:35.3
Classical Music	0:20.4	75.9	0:24.72
Pop Music	1:30.4	66.8	2:15.32

Table 5.1: Timing Results

The results given in Table 5.1 demonstrate the linear timing relationship to the frequency content and the length of the audio samples. From the demonstrated linear relationships it is predicted that this system could be implemented to operate in ‘real-time’. To achieve this for data with low frequency compression more efficient programming languages and hardware would be required.

5.5 Reproduction Accuracy

The accuracy of the reproduced sound field, when compared to the original, will be influenced by the representation of the sound field and any un-modelled phenomenon. These influences are discussed below.

5.5.1 Modal Limitation

The system being built consists of eight loudspeakers for which the eight loudspeaker signals are being constructed from nine sound field coefficients, corresponding to a fourth order cylindrical harmonic reconstruction.

Using the error formulations presented in Section 3.2.2 the relative error at a specific position for a range of frequencies can be determined. A graph of the reproduction error is provided in Figure 5.4 for the listening radius of $x = 0.2\text{m}$. From the graph it is observed that an error of less than 10% will occur for frequencies below 1.15 kHz in the implemented system.

5.5.2 Loudspeakers

In the formation and solution to the reproduction problem, the sound sources used to reproduce the sound field have been modelled as point sources. To implement the system, the point sources have been replaced by loudspeakers. The loudspeakers will not behave as true point sources thus introducing errors in the reproduced sound field.

The solution to the reproduction problem also assumes that the exact position of the point sources is known. Positioning the loudspeakers in their precise locations is mechanically difficult. To quantify the error introduced by incorrect placement, a simulation was

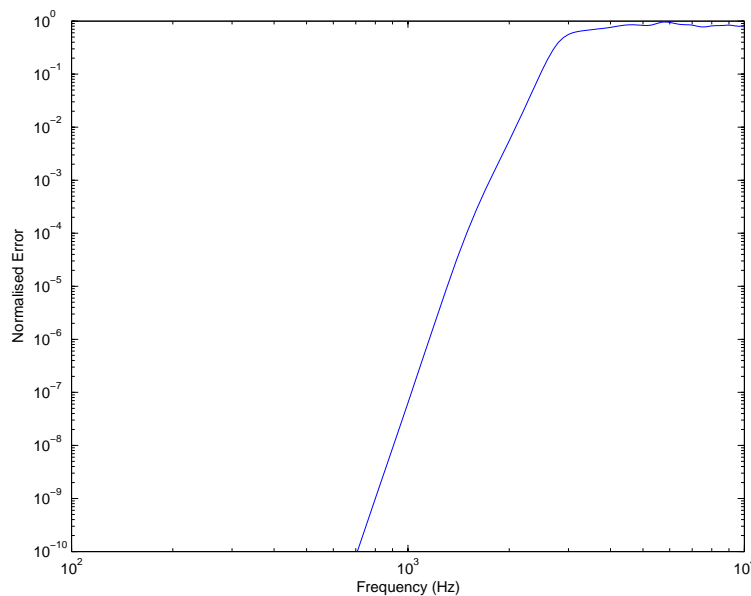


Figure 5.4: Expected modal error for the frequency range 0.1-10 kHz at the listening radius of $x_0 = 0.2\text{m}$.

constructed where the loudspeakers have been randomly shifted from their original position. The results of the simulation is shown in Figure 5.5 where the maximum error in the random placement is specified as a percentage of the loudspeaker radius along the horizontal axis. Using these results, a 2% maximum error (corresponding to 3cm) approximately doubles the error in the reproduced sound field.

5.5.3 Miscellaneous

Other minor effects that will introduce errors in the reproduced sound field are:

- Reverberation — The system was constructed in an anechoic chamber to reduce the effects of reverberation. The room insulation does not absorb 100% of the sound energy thus producing some reverberation. Also the loudspeakers, loudspeaker stands and the listener will add to the reverberation created within the room. As reverberation is not included in the system model, its effects will introduce errors to the sound field.
- Distortion — A number of components are used to transform the loudspeaker signals into pressure waves to generate the sound field. The signals pass through a digital to analogue converter, audio amplifier and loudspeaker, via copper cable. Each of these devices add amplitude and frequency distortion to the sound field. These effects have been minimised by using identical audio amplifiers and loudspeakers so that any distortion is directionally consistent.

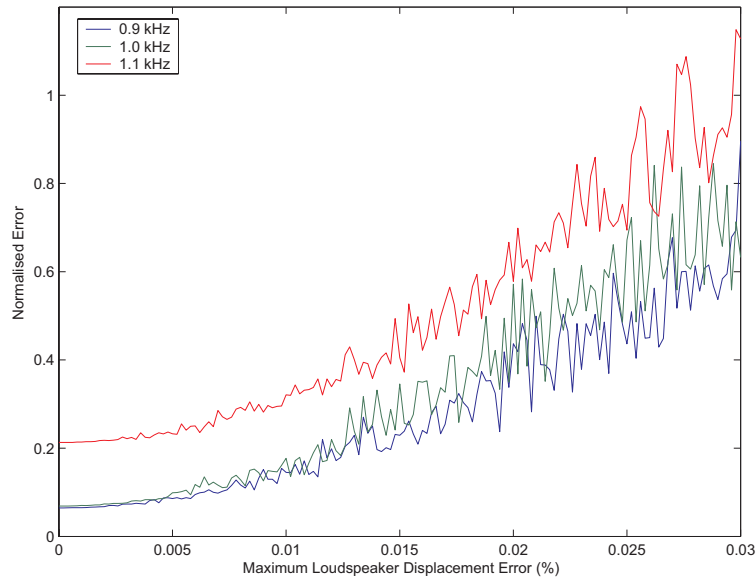


Figure 5.5: Expected reproduction error for the frequency range 0.1-10 kHz at the listening radius of $x = 0.2\text{m}$. The values shown have been averaged over twenty different sound fields consisting of twenty-five random plane-waves.

5.6 Listening Tests

A number of sound fields were synthesised and then listened to in the loudspeaker setup. The details of the listening tests are given in Appendix C. In each case the listeners were able to locate the origin of these sounds. This indicates the production of a sound field with *spatial* cues allowing sound localisation.

From the listening tests, the numerical accuracy of the reproduction can not be measured. This measurement requires the construction of a sufficient sound field recording system. When such a system exists the ability of the system to recreate a sound field will be known.

5.7 Summary of Results

In this chapter the 2D sound reproduction system has been successfully implemented. This implementation has been verified using the demonstrations described in Appendix C. The methodology used to implement the 2D system applies equally to the designed 3D system. The significant contributions and results obtained in this chapter are:

- The implementation of a spatial sound field synthesis system.
- The transformation of the narrowband system design to a broadband implementation.
- The demonstration that the reproduction system could be implemented in real-time. This is a necessary requirement for commercial adoption.

Sound Field Control

In this chapter two modifications to the presented sound reproduction system are introduced. The first modification uses non-uniform loudspeaker placement to increase the spatial reproduction accuracy for sounds originating in a predetermined range of directions. The second modification allows the listening area or ‘sweet spot’ to be moved from the centre of the listening area. This modification demonstrates the high level of control that the sound field reproduction system has over the reproduced sound field.

6.1 Non-uniform Loudspeaker Placement

The solution to the reproduction problem presented, requires a large number of loudspeakers to reproduce a large sound field required for multiple listeners. The current solution has only considered the placement of loudspeakers uniformly around the listening area. This uniform placement provides equal reproduction accuracy to sounds originating from all directions.

For most current applications of sound reproduction technology this uniform reproduction accuracy is not necessary. For example, in a movie theatre the listeners attention is focused towards the direction of the screen. It would therefore be desirable to have higher reproduction accuracy from sounds generated in the direction of the screen. Also in practical implementation it is difficult to reproduce sounds below a seated audience due to the insulating nature of the seats.

If the reproduction system is constructed with prior knowledge of its particular application, the positions of the speakers can be optimised for increased reproduction accuracy. To illustrate this solution, Figure 6.1 shows a comparison of a uniform and non-uniform loudspeaker placement scheme. The non-uniform loudspeaker placement was selected to optimise the reproduction accuracy in the direction of $\phi = 0^\circ$. Figure 6.1(b) shows the numerical accuracy at the radius $kx = 0.54$ as the azimuth was rotated in the range $\phi = [-\pi, \pi)^\circ$. From this figure it is observed that the reproduction accuracy increased in the region around $\phi = 0^\circ$ and decreased in the regions around $\phi = \pm\pi/2^\circ$. Unexpectedly the reproduction accuracy also increased in the region $\phi = \pi^\circ$. The accuracy is likely to have increased in this region due to the symmetry of the loudspeaker placement as observed from this direction.

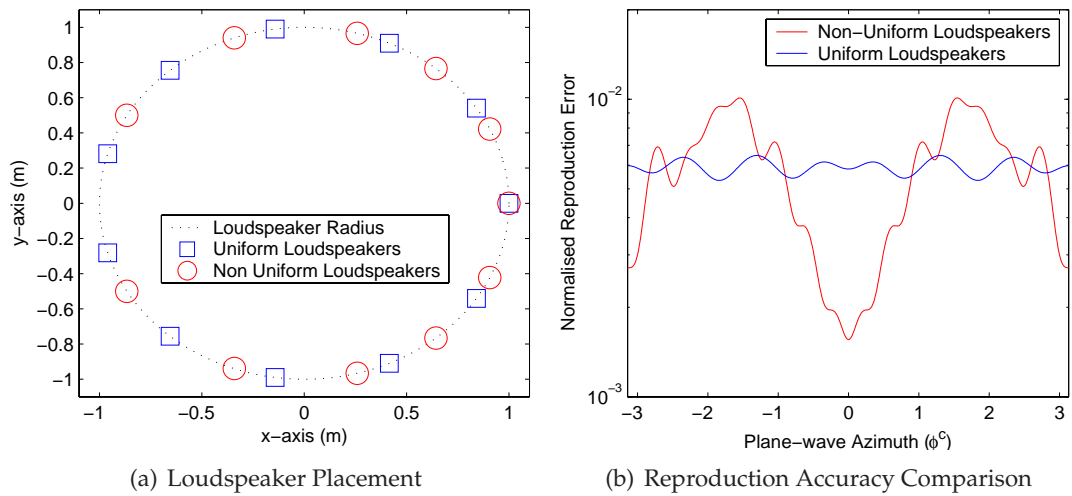


Figure 6.1: Effect of Non-uniform Loudspeaker Placement

6.2 Listening Area Repositioning and Duplication

The solution presented to the sound field reproduction system has only allowed the listening area to be placed in the centre of an array of loudspeakers. The ability to reposition the listening area may be desirable for many commercial and virtual reality applications. The current reproduction system has considered the loudspeakers to be placed at equal distances from the centre of the listening area as shown in Figure 6.2(a).

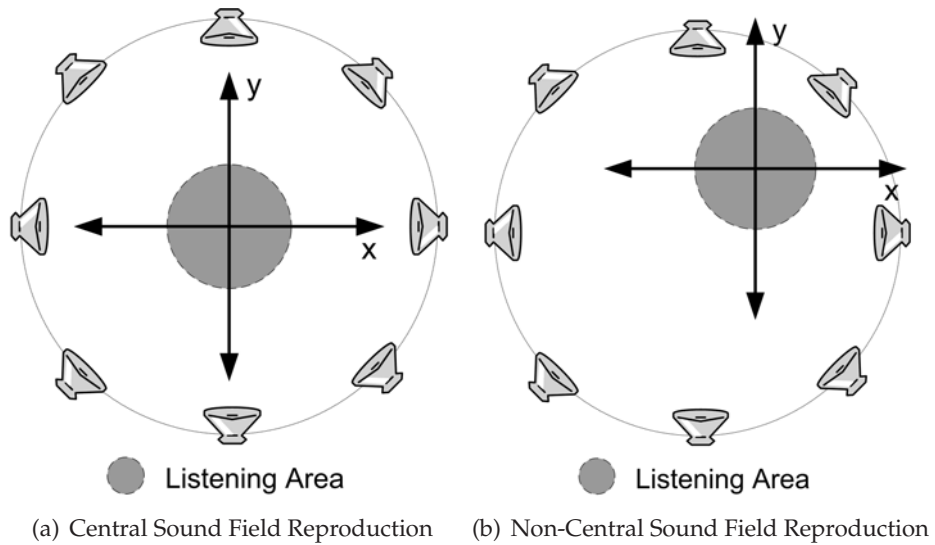


Figure 6.2: Loudspeaker Coordinate Reference

To reposition the listening area, the coordinate system of the reproduced sound field can be moved such that its origin corresponds to the centre of the desired listening area. This is shown in Figure 6.2. The modification required to reproduce the sound field

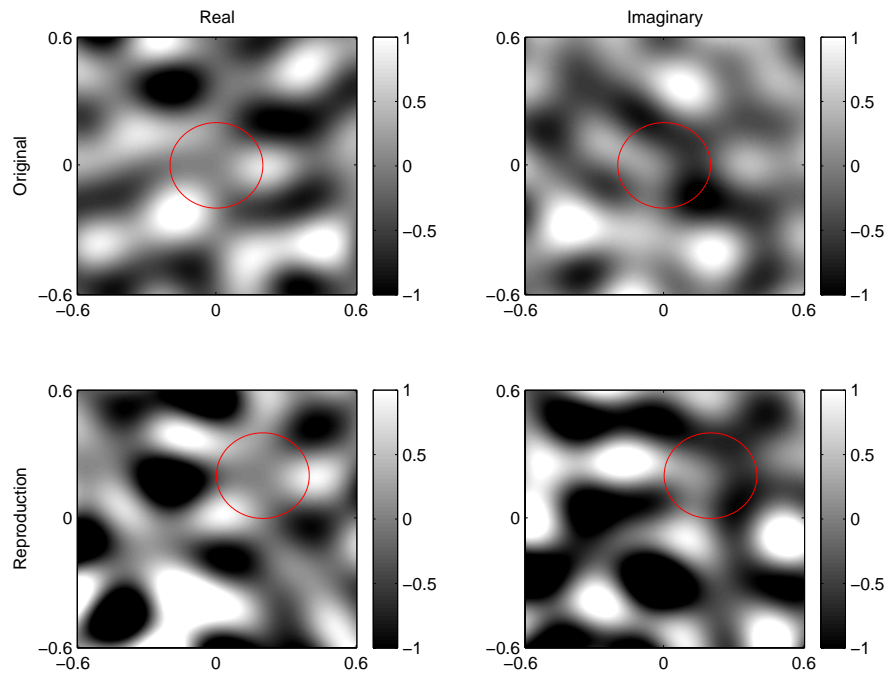


Figure 6.3: Sound field reproduction at height $z = 0\text{m}$.

with the repositioned listening area is to adjust the vectors describing the position of the loudspeakers. The loudspeaker positions are only adjusted mathematically and remain in their original physical position. The previously defined reproduction equations can then be applied with the adjusted loudspeaker positions to reproduce the sound field.

This method of repositioning the listening area is valid for the presented 2D and 3D reproduction systems. Figure 6.3 shows the result from a 3D simulation where the listening area was repositioned by 0.2m in both the x and y directions. The sound field used for this simulation consisted of 25 random plane-waves with a frequency of 1 kHz. The original spherical radius for the loudspeakers was 1m. The reproduction accuracy at the listening area was 7.36%. By simulation it was found that as the distance of the listening area increases from the origin the reproduction error increases. The result of the simulation is shown in Figure 6.4. The increased error is caused by the unequal distances of the loudspeakers to the centre of the new listening area.

With the ability of repositioning the listening area within in the reproduced sound field leads to the possibility of duplicating that listening area for more than one listener. A methodology to duplicate the listening area consists of calculating the loudspeaker weights for multiple listening areas and then combining the loudspeaker weights. The complication occurs in the combination of the loudspeaker weights. As observed in Figure 6.3 a sound field exists outside the designated listening area. When combining the loudspeaker weights for multiple listening areas, the residual sound field from the previously considered listening areas would need to be nullified.

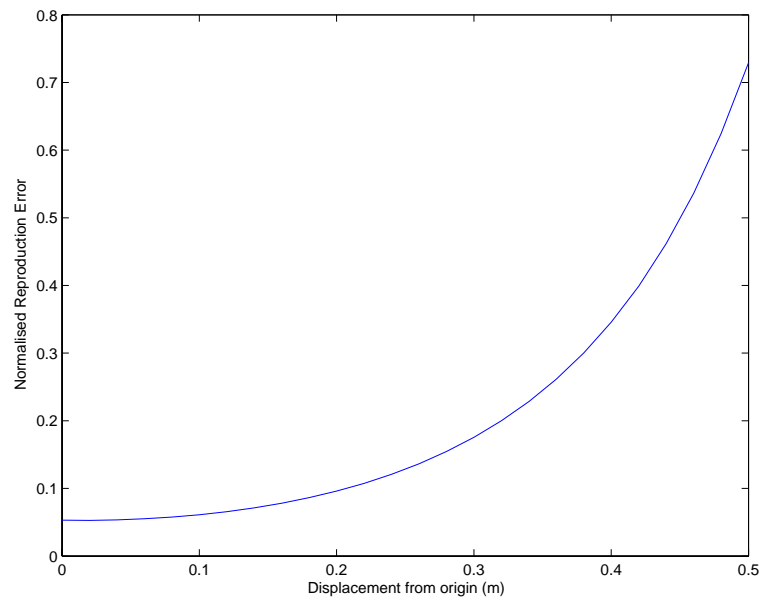


Figure 6.4: Reproduction error as displacement of the listening area is increased.

The duplication of this listening area allows the reproduction to be experienced by more than one listener without additional requirements on the spatial sound field recording system. It is anticipated that additional loudspeakers would be required to achieve the duplicate listening areas. It is expected that less loudspeakers would be required than simply increasing the size of the central listening area. This is an area for further enhancement of the reproduction system proposed in this thesis.

Conclusions and Further Work

7.1 Conclusions

The goal of this project was to develop an audio system capable of accurately reproducing arbitrary sound fields. The project succeeded in developing both a two and three dimensional reproduction system. The designed systems are presented in a general form allowing the quantity and placement of the loudspeakers to be specified in implementation. The research described in this thesis has made advances in the theoretical understanding of sound field reproduction systems and demonstrated commercial potential in addition to existing technologies.

The initial challenge in developing the systems was obtaining a mathematical description of an arbitrary sound field. Spherical harmonics and cylindrical harmonics were found to provide the mathematical tools allowing the development of the sound reproduction systems. The use of these orthogonal spatial functions allowed the complexity of a spatial system to be described using a series of coefficients.

In the theoretical design of the reproduction systems a measure of the spatial reproduction accuracy was derived. The reproduction accuracy was found to be dependent on the size of the listening area, the number of loudspeakers and the frequency content of the sound field. It was determined that to maintain the same reproduction accuracy, the number of loudspeakers has a 'squared' relationship with the size of the listening area in three dimensions. This relationship was found to be linear for the two dimensional system.

The two reproduction systems were verified using numerical simulation techniques. The results of the 3D simulation were included in this thesis. The 2D system was successfully implemented to show that the detailed mathematical analysis presented accurately described a practical system. This achieved the goal of this thesis. Through this implementation, it was shown that the system considers the amplitude, phase and delay of signals to individual sound sources to reproduce the sound field.

The flexibility of the designed system was demonstrated by showing that loudspeaker placement schemes can be used to increase the reproduction accuracy from sounds originating in predefined areas. The particular scheme presented promotes the designed system for use in commercial and home theatre systems. Also, the ability to reposition the listening area from the conventional central location was demonstrated.

In summary, the work presented in this thesis has demonstrated that the designed reproduction systems exhibit the following characteristics:

- Strong Theoretical Foundation – Mathematical equations have been derived stating the input, output and performance of the system.
- Flexibility – The designed reproduction system is capable of working with any number of loudspeakers placed in any location.
- Commercial Application – The developed reproduction and synthesis systems demonstrate capabilities above what is commercially available in the market today.

7.2 Further Research

In the process of executing the research for this project the following items have been identified as further research possibilities that would enhance the field of spatial sound reproduction

- Investigation of the impact and possible correction of reverberation within the listening environment to the quality of the reproduced sound field [35].
- The reproduction systems developed in this thesis assumed that the loudspeakers act as pure point sources. Practically this is not the case. The effect and correction of this non-ideal practicality would enhance the existing system design.
- In Chapter 3 the normalised truncation and expansion error were defined for arbitrary sound field. Further works needs to be completed to relate these errors to the sound field being reproduced.
- The proposed system was only implemented in two dimensions. A demonstration of the proposed system to reproduce 3D sound would demonstrate the full commercial opportunities of the theoretical work presented in this thesis.
- The audible sounds generated in Chapter 5 were processed off-line. A real-time implementation of this algorithm together with constructed system could allow the listener to vary the position of a sound source and hear the effects in real time. This would also demonstrate the commercial viability of the proposed system with current proposing power.
- In Section 6.2 a method of duplicating the listening areas for multiple listeners was suggested. There is still a large amount of work to implement this feature and test its effectiveness in a physical implementation.
- Throughout this thesis a numerical measure of the reproduction error has been used to quantify the performance of the designed system. It would be interesting to find a relationship between this performance measure and human spatial perception.

Mathematical Proofs

Proof of (3.1) – Simplified 3D Sound Field Representation

Consider a sound field derived from the sum of Q arbitrarily located sound sources (as shown in Figure A.1). Using (2.12) the sound field is described by,

$$S(\mathbf{x}; k) = \sum_{q=1}^Q \frac{w_q y_q e^{iky_q} e^{ik\|\mathbf{y}_q - \mathbf{x}\|}}{\|\mathbf{y}_q - \mathbf{x}\|} \quad (\text{A.1})$$

where $y_q \neq x$. Using (2.14) the sound field can be represented with spherical harmonics as

$$S(\mathbf{x}; k) = \sum_{q=1}^Q 4\pi w_q (-i) ky_q e^{iky_q} \sum_{n=0}^{\infty} \sum_{m=-n}^n h_n^{(2)}(ky_q) j_n(kx) Y_{nm}^*(\hat{\mathbf{y}}) Y_{nm}(\hat{\mathbf{x}})$$

where $x < y_q$. Rearranging the order of summation results in

$$S(\mathbf{x}; k) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \left[\sum_{q=1}^Q 4\pi w_q (-i) ky_q e^{iky_q} h_n^{(2)}(ky_q) Y_{nm}^*(\hat{\mathbf{y}}) \right] j_n(kx) Y_{nm}(\hat{\mathbf{x}}),$$

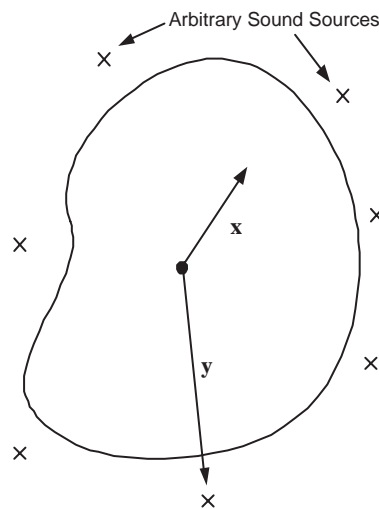


Figure A.1: Arbitrary Sound Field

which is equal to

$$S(\mathbf{x}; k) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \alpha_{nm}(k) j_n(kx) Y_{nm}(\hat{\mathbf{x}})$$

where

$$\alpha_{nm}(k) = \sum_{q=1}^Q 4\pi w_q(-i) k y_q e^{iky_q} h_n^{(2)}(k y_q) Y_{nm}^*(\hat{\mathbf{y}}),$$

thus completing the proof.

Proof of (3.5) – 3D Normalised Truncation Error

The 3D normalised truncation error is defined in (3.4) as

$$\epsilon_N(kx) \triangleq \frac{\int |S(\mathbf{x}; k) - \hat{S}_N(\mathbf{x}; k)|^2 d\hat{\mathbf{x}}}{\int |S(\mathbf{x}; k)|^2 d\hat{\mathbf{x}}} \quad (\text{A.2})$$

where $\hat{S}_N(\mathbf{x}; k)$ denotes the sound field described by the truncated series expansion of (3.1) to the order N and is equal to

$$\hat{S}_N(\mathbf{x}; k) = \sum_{n=0}^N \sum_{m=-n}^n \alpha_{nm}(k) j_n(kx) Y_{nm}(\hat{\mathbf{x}})$$

and

$$S(\mathbf{x}; k) - \hat{S}_N(\mathbf{x}; k) = \sum_{n=N+1}^{\infty} \sum_{m=-n}^n \alpha_{nm}(k) j_n(kx) Y_{nm}(\hat{\mathbf{x}}).$$

The numerator of (A.2) can be expressed as

$$\begin{aligned} \int |S(\mathbf{x}; k) - \hat{S}_N(\mathbf{x}; k)|^2 d\hat{\mathbf{x}} &= \int \sum_{n=N+1}^{\infty} \sum_{m=-n}^n \sum_{p=N+1}^{\infty} \sum_{q=-p}^p \alpha_{nm}(k) \alpha_{pq}^*(k) \times j_n(kx) j_p^*(kx) \\ &\quad \times Y_{nm}(\hat{\mathbf{x}}) Y_{pq}^*(\hat{\mathbf{x}}) d\hat{\mathbf{x}} \end{aligned}$$

which is simplified using the orthogonality property of spherical harmonics (2.9)

$$\int |S(\mathbf{x}; k) - \hat{S}_N(\mathbf{x}; k)|^2 d\hat{\mathbf{x}} = \sum_{n=N+1}^{\infty} \sum_{m=-n}^n |\alpha_{nm}(k) j_n(kx)|^2. \quad (\text{A.3})$$

Expanding the sound field energy gives

$$\int |S(\mathbf{x}; k)|^2 d\hat{\mathbf{x}} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \sum_{p=0}^{\infty} \sum_{q=-p}^p \alpha_{nm}(k) \alpha_{pq}^*(k) \times j_n(kx) j_p^*(kx) \times \int Y_{nm}(\hat{\mathbf{x}}) Y_{pq}^*(\hat{\mathbf{x}}) d\hat{\mathbf{x}}$$

which is again simplified using the orthogonality property of spherical harmonics (2.9)

$$\int |S(\mathbf{x}; k)|^2 d\hat{\mathbf{x}} = \sum_{n=0}^{\infty} \sum_{m=-n}^n |\alpha_{nm}(k) j_n(kx)|^2. \quad (\text{A.4})$$

From (A.3) and (A.4) the numerator of (A.2) can be expressed as

$$\int \left| S(\mathbf{x}; k) - \hat{S}(\mathbf{x}; k) \right|^2 d\hat{\mathbf{x}} = \int |S(\mathbf{x}; k)|^2 d\hat{\mathbf{x}} - \sum_{n=0}^N \sum_{m=-n}^n |\alpha_{nm}(k) j_n(kx)|^2. \quad (\text{A.5})$$

Substituting (A.5) into the definition of the normalised truncation error (A.2) gives

$$\epsilon_N(kx) = 1 - \frac{\sum_{n=0}^N \sum_{m=-n}^n |\alpha_{nm}(k) j_n(kx)|^2}{\int |S(\mathbf{x}; k)|^2 d\hat{\mathbf{x}}}$$

as required, thus completing the proof.

Proof of (3.6) – Plane-wave expansion of the 3D Normalised Truncation Error

The spherical harmonic expansion of a plane-wave is given in (2.15) as

$$e^{ikx(\hat{\mathbf{y}}^T \cdot \hat{\mathbf{x}})} = \sum_{n=0}^N \sum_{m=-n}^n 4\pi i^n j_n(kx) Y_{nm}^*(\hat{\mathbf{y}}) Y_{nm}(\hat{\mathbf{x}}). \quad (\text{A.6})$$

Equating (A.6) with the arbitrary sound field representation given in (3.1) it can be calculated that the sound field coefficients for a plane-wave sound field are

$$\alpha_{nm}(k) = 4\pi i^n Y_{nm}^*(\hat{\mathbf{y}}).$$

Using this expression, the numerator of (3.5) can be expressed as

$$\begin{aligned} \sum_{n=0}^N \sum_{m=-n}^n |\alpha_{nm}(k) j_n(kx)|^2 &= \sum_{n=0}^N \sum_{m=-n}^n \left| \sum_{q=1}^Q 4\pi w_q i^n Y_{nm}^*(\hat{\mathbf{y}}_q) j_n(kx) \right|^2 \\ &= (4\pi)^2 \sum_{n=0}^N \sum_{m=-n}^n \left| \sum_{q=1}^Q w_q Y_{nm}^*(\hat{\mathbf{y}}_q) j_n(kx) \right|^2 \end{aligned} \quad (\text{A.7})$$

where Q is the number of plane-waves and w_q is the complex weighting coefficient of the q^{th} plane-wave. For an arbitrary set of Q plane-waves the sound field is completely defined as (2.16)

$$S(x; k) = \sum_{q=1}^Q w_q e^{ikx(\hat{\mathbf{y}}_q^T \cdot \hat{\mathbf{x}})}$$

which is substituted into the denominator of (3.5) giving

$$\begin{aligned}
 \int |S(x; k)|^2 d\hat{\mathbf{x}} &= \int \sum_{q=1}^Q w_q e^{ikx(\hat{\mathbf{y}}_q^T \cdot \hat{\mathbf{x}})} \times \sum_{p=1}^Q w_p e^{-ikx(\hat{\mathbf{y}}_p^T \cdot \hat{\mathbf{x}})} d\hat{\mathbf{x}} \\
 &= \int \sum_{q=1}^Q \sum_{p=1}^Q w_q w_p e^{ikx((\hat{\mathbf{y}}_q^T \cdot \hat{\mathbf{x}}) - (\hat{\mathbf{y}}_p^T \cdot \hat{\mathbf{x}}))} d\hat{\mathbf{x}} \\
 &= \int \sum_{q=1}^Q \sum_{p=1}^Q w_q w_p e^{ikx \|\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p\| \left(\left(\frac{\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p}{\|\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p\|} \right)^T \cdot \hat{\mathbf{x}} \right)} d\hat{\mathbf{x}}.
 \end{aligned}$$

Using the spherical harmonic expansion of a plane-wave (A.6)

$$\begin{aligned}
 \int |S(x; k)|^2 d\hat{\mathbf{x}} &= \sum_{n=0}^N \sum_{m=-n}^n \sum_{q=1}^Q \sum_{p=1}^Q \int w_q w_p 4\pi i^n j_n(kx \|\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p\|) \\
 &\quad \times Y_{nm}^* \left(\frac{\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p}{\|\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p\|} \right) Y_{nm}(\hat{\mathbf{x}}) d\hat{\mathbf{x}} \\
 &= \sum_{n=0}^N \sum_{m=-n}^n \sum_{q=1}^Q \sum_{p=1}^Q w_q w_p 4\pi i^n j_n(kx \|\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p\|) \\
 &\quad \times Y_{nm}^* \left(\frac{\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p}{\|\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p\|} \right) \int Y_{nm}(\hat{\mathbf{x}}) d\hat{\mathbf{x}}.
 \end{aligned} \tag{A.8}$$

Using the spherical harmonics orthogonality property (2.9) and from the evaluation $Y_{00}(\hat{\mathbf{x}}) = (\sqrt{4\pi})^{-1} \forall \hat{\mathbf{x}}$

$$\begin{aligned}
 \int Y_{nm}(\hat{\mathbf{x}}) d\hat{\mathbf{x}} &= \sqrt{4\pi} \int Y_{nm}(\hat{\mathbf{x}}) Y_{00}^*(\hat{\mathbf{x}}) d\hat{\mathbf{x}} \\
 &= \begin{cases} \frac{\sqrt{4\pi}}{4\pi} \int d\hat{\mathbf{x}} = \sqrt{4\pi} & n=0, m=0 \\ 0 & n \neq 0. \end{cases}
 \end{aligned}$$

Therefore (A.8) can be reduced to

$$\begin{aligned}
 \int |S(x; k)|^2 d\hat{\mathbf{x}} &= \sqrt{4\pi} \sum_{q=1}^Q \sum_{p=1}^Q w_q w_p 4\pi i^0 j_0(kx \|\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p\|) Y_{00}^* \left(\frac{\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p}{\|\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p\|} \right) \\
 &= 4\pi \sum_{q=1}^Q \sum_{p=1}^Q w_q w_p j_0(kx \|\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p\|).
 \end{aligned} \tag{A.9}$$

Substituting (A.7) and (A.9) back into (3.5) gives

$$\varepsilon_N(kx) = 1 - \frac{4\pi \sum_{n=0}^N \sum_{m=-n}^n \left| \sum_{q=1}^Q w_q Y_{nm}^*(\hat{\mathbf{y}}) j_n(kx) \right|^2}{\sum_{q=1}^Q \sum_{p=1}^Q w_q w_p j_0(kx \|\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p\|)}$$

as required, thus completing the proof.

Proof of (3.8) – 3D Normalised Expansion Error

The 3D normalised truncation error is defined in (3.7) as

$$E_N(kx) \triangleq \frac{\int |\tilde{T}_N(\mathbf{x}; k)|^2 d\hat{\mathbf{x}}}{\int |S(\mathbf{x}; k)|^2 d\hat{\mathbf{x}}} \quad (\text{A.10})$$

where $\tilde{T}_N(\mathbf{x}; k)$ denotes the reproduced sound field created by considering the spherical harmonics from $n = N + 1, \dots, \infty$. Using (3.2) this sound field is expressed as

$$\tilde{T}_N(\mathbf{x}; k) = \sum_{n=N+1}^{\infty} \sum_{m=-n}^n 4\pi X_n(kx) \times \sum_{l=1}^L a_l(k) R_n(ky_l) Y_{nm}(\hat{\mathbf{x}}) Y_{nm}^*(\hat{\mathbf{y}}_l).$$

A simplification is made by defining the reproduction coefficient

$$\kappa_{nm}(k) = 4\pi i^n \sum_{l=1}^L a_l(k) R_n(ky_l) Y_{nm}^*(\hat{\mathbf{y}}_l)$$

giving

$$\tilde{T}_N(\mathbf{x}; k) = \sum_{n=N+1}^{\infty} \sum_{m=-n}^n \kappa_{nm}(k) j_n(kx) Y_{nm}(\hat{\mathbf{x}}).$$

The numerator of (A.10) is calculated as

$$\left| \tilde{T}_N(\mathbf{x}; k) \right|^2 = \sum_{n=N+1}^{\infty} \sum_{m=-n}^n \sum_{p=N+1}^{\infty} \sum_{q=-p}^p \kappa_{nm}(k) \kappa_{pq}^*(k) j_n(kx) j_p(kx) Y_{nm}(\hat{\mathbf{x}}) Y_{pq}^*(\hat{\mathbf{x}}),$$

giving

$$\int \left| \tilde{T}_N(\mathbf{x}; k) \right|^2 d\hat{\mathbf{x}} = \sum_{n=N+1}^{\infty} \sum_{m=-n}^n |\kappa_{nm}(k) j_n(kx)|^2, \quad (\text{A.11})$$

where the spherical harmonic orthogonality property (2.9) was used to simplify the equation. Substituting (A.11) and (A.4) into the definition for the normalised expansion error (A.10) gives

$$E_N(kx) = \frac{\sum_{n=N+1}^{\infty} \sum_{m=-n}^n |\kappa_{nm}(k) j_n(kx)|^2}{\sum_{n=0}^{\infty} \sum_{m=-n}^n |\alpha_{nm}(k) j_n(kx)|^2}$$

as required, thus completing the proof.

Proof of (3.11) – 3D Normalised Reproduction Error

The 3D normalised reproduction error is defined in (3.10) as

$$\Upsilon_N(kx) \triangleq \frac{\int |S(\mathbf{x}; k) - [\hat{S}_N(\mathbf{x}; k) + \tilde{T}_N(\mathbf{x}; k)]|^2 d\hat{\mathbf{x}}}{\int |S(\mathbf{x}; k)|^2 d\hat{\mathbf{x}}} \quad (\text{A.12})$$

where each of the sound fields are defined as

$$\begin{aligned} S(\mathbf{x}; k) &= \sum_{n=0}^{\infty} \sum_{m=-n}^n \alpha_{nm}(k) j_n(kx) Y_{nm}(\hat{\mathbf{x}}) \\ \hat{S}_N(\mathbf{x}; k) &= \sum_{n=0}^N \sum_{m=-n}^n \alpha_{nm}(k) j_n(kx) Y_{nm}(\hat{\mathbf{x}}) \\ \tilde{T}_N(\mathbf{x}; k) &= \sum_{n=N+1}^{\infty} \sum_{m=-n}^n \kappa_{nm}(k) j_n(kx) Y_{nm}(\hat{\mathbf{x}}). \end{aligned}$$

The numerator of (A.12) is calculated as

$$\begin{aligned} S(\mathbf{x}; k) - [\hat{S}_N(\mathbf{x}; k) + \tilde{T}_N(\mathbf{x}; k)] &= \sum_{n=N+1}^{\infty} \sum_{m=-n}^n (\alpha_{nm}(k) - \kappa_{nm}(k)) j_n(kx) Y_{nm}(\hat{\mathbf{x}}) \\ \left| S(\mathbf{x}; k) - [\hat{S}_N(\mathbf{x}; k) + \tilde{T}_N(\mathbf{x}; k)] \right|^2 &= \sum_{n=N+1}^{\infty} \sum_{m=-n}^n \sum_{p=N+1}^{\infty} \sum_{q=-p}^p (\alpha_{nm}(k) - \kappa_{nm}(k)) \times \\ &\quad (\alpha_{pq}^*(k) - \kappa_{pq}^*(k)) j_n(k) j_p^*(k) Y_{nm}(\hat{\mathbf{x}}) Y_{pq}^*(\hat{\mathbf{x}}) \\ \int \left| S(\mathbf{x}; k) - [\hat{S}_N(\mathbf{x}; k) + \tilde{T}_N(\mathbf{x}; k)] \right|^2 d\hat{\mathbf{x}} &= \sum_{n=N+1}^{\infty} \sum_{m=-n}^n |(\alpha_{nm}(k) - \kappa_{nm}(k)) j_n(kx)|^2, \end{aligned} \quad (\text{A.13})$$

where the orthogonality property of spherical harmonics (2.9) was used for simplification. Substituting (A.13) and the previously calculated (A.4) into (A.12) gives

$$\Upsilon_N(kx) = \frac{\sum_{n=N+1}^{\infty} \sum_{m=-n}^n |[\alpha_{nm}(k) - \kappa_{nm}(k)] J_n(kx)|^2}{\sum_{n=0}^{\infty} \sum_{m=-n}^n |\alpha_{nm}(k) J_n(kx)|^2}$$

as required, thus completing the proof.

Proof of (3.25) – 2D Sound Field Representation

Using the previously definition of an arbitrary sound field (A.1)

$$S(\mathbf{x}; k) = \sum_{q=1}^Q w_q y_q e^{iky_q} \frac{e^{-ik\|\mathbf{y}_q - \mathbf{x}\|}}{\|\mathbf{y}_q - \mathbf{x}\|}$$

along with [19, p 67 Eq (3.65)]

$$H_0^{(1)}(k\|\mathbf{y} - \mathbf{x}\|) = \sum_{m=-\infty}^{\infty} H_m^{(1)}(ky) J_m(kx) e^{im\phi} e^{-im\varphi}$$

where ϕ is the azimuth of \mathbf{x} and φ is the azimuth of \mathbf{y} , also using [19, Eq 2.1 and Eq 3.61]

$$\frac{e^{ik\|\mathbf{y} - \mathbf{x}\|}}{\|\mathbf{y} - \mathbf{x}\|} = \pi i H_0^{(1)}(k\|\mathbf{y} - \mathbf{x}\|)$$

gives

$$S(\mathbf{x}; k) = \pi i \sum_{q=1}^Q w_q y_q e^{iky_q} \sum_{m=-\infty}^{\infty} H_m^{(1)}(ky) J_m(kx) e^{im\phi} e^{-im\varphi_q}$$

where $\mathbf{y}_q \neq \mathbf{x}$. A 2D arbitrary sound field can therefore be defined as

$$S(\mathbf{x}; k) = \sum_{m=-\infty}^{\infty} \alpha_m(k) J_m(kx) e^{im\phi}$$

where

$$\alpha_m(k) = \pi i \sum_{q=1}^Q w_q y_q e^{iky_q} H_m^{(1)}(ky) e^{-im\varphi_q}$$

as required, thus completing the proof.

Proof of (3.28) – 2D Normalised Truncation Error

The 2D normalised truncation error is defined as

$$\varepsilon_M(kx) \triangleq \frac{\int |S(\mathbf{x}; k) - \hat{S}_M(\mathbf{x}; k)|^2 d\phi}{\int |S(\mathbf{x}; k)|^2 d\phi} \quad (\text{A.14})$$

where

$$\hat{S}_M(\mathbf{x}; k) = \sum_{m=-M}^M \alpha_m(k) J_m(kx) e^{im\phi}$$

and

$$S(\mathbf{x}; k) = \sum_{m=-\infty}^{\infty} \alpha_m(k) J_m(kx) e^{im\phi}.$$

The numerator of (A.14) is calculated as

$$\begin{aligned} S(\mathbf{x}; k) - \hat{S}_M(\mathbf{x}; k) &= \sum_{m=-\infty}^{-M-1} \alpha_m(k) J_m(kx) e^{im\phi} + \sum_{m=M+1}^{\infty} \alpha_m(k) J_m(kx) e^{im\phi} \\ \int |S(\mathbf{x}; k) - \hat{S}_M(\mathbf{x}; k)|^2 d\phi &= \int \sum_{m=-\infty}^{-M-1} \sum_{p=-\infty}^{-M-1} \alpha_m(k) \alpha_p(k) J_m(kx) J_p(kx) e^{im\phi} e^{-ip\phi} d\phi \\ &\quad + \int \sum_{m=M+1}^{\infty} \sum_{p=M+1}^{\infty} \alpha_m(k) \alpha_p(k) J_m(kx) J_p(kx) e^{im\phi} e^{-ip\phi} d\phi \\ &\quad + 2 \int \sum_{m=-\infty}^{-M-1} \sum_{p=M+1}^{\infty} \alpha_m(k) \alpha_p(k) J_m(kx) J_p(kx) e^{im\phi} e^{-ip\phi} d\phi \end{aligned}$$

Using the orthogonality property of exponentials (2.7)

$$\begin{aligned} \int |S(\mathbf{x}; k) - \hat{S}_M(\mathbf{x}; k)|^2 d\phi &= 2\pi \sum_{m=-\infty}^{-M-1} \alpha_m(k)^2 J_m(kx)^2 + 2\pi \sum_{m=M+1}^{\infty} \alpha_m(k)^2 J_m(kx)^2 \\ &= 2\pi \left(\sum_{m=-\infty}^{-M-1} |\alpha_m J_m(kx)|^2 + \sum_{m=M+1}^{\infty} |\alpha_m J_m(kx)|^2 \right). \end{aligned}$$

The total energy of the sound field is calculated as

$$\int |S(\mathbf{x}; k)|^2 d\phi = 2\pi \sum_{m=-\infty}^{\infty} |\alpha_m(k) J_m(kx)|^2 \quad (\text{A.15})$$

allowing

$$\int |S(\mathbf{x}; k) - \hat{S}_M(\mathbf{x}; k)|^2 d\phi = \int |S(\mathbf{x}; k)|^2 d\phi - 2\pi \sum_{m=-M}^M |\alpha_m(k) J_m(kx)|^2.$$

The 2D normalised truncation error can therefore be expressed as

$$\begin{aligned}\varepsilon_M(kx) &= \frac{\int |S(\mathbf{x}; k)|^2 d\phi - 2\pi \sum_{m=-M}^M |\alpha_m(k) J_m(kx)|^2}{\int |S(\mathbf{x}; k)|^2 d\phi} \\ &= 1 - \frac{2\pi \sum_{m=-M}^M |\alpha_m(k) J_m(kx)|^2}{\int |S(\mathbf{x}; k)|^2 d\phi},\end{aligned}$$

thus completing the proof.

Proof of (3.29) – Plane-wave expansion of the 2D Normalised Truncation Error

An arbitrary sound field described in terms of plane-waves is given by (2.16) as

$$S(\mathbf{x}; k) = \sum_{q=1}^Q w_q e^{ikx(\hat{\mathbf{y}}^T \cdot \hat{\mathbf{x}})}$$

A plane-wave sound field expressed in cylindrical harmonics is [19, p67]

$$e^{ikx(\hat{\mathbf{y}}^T \cdot \hat{\mathbf{x}})} = \sum_{m=-\infty}^{\infty} i^m J_m(kx) e^{im\phi} e^{-im\varphi}$$

giving

$$S(\mathbf{x}; k) = \sum_{m=-\infty}^{\infty} \sum_{q=1}^Q w_q i^m J_m(kx) e^{im\phi} e^{-im\varphi_q}$$

from which the sound field coefficients are

$$\alpha_m(k) = i^m \sum_{q=1}^Q w_q e^{-im\varphi_q}.$$

The numerator of (3.28) can be expressed as

$$2\pi \sum_{m=-M}^M |\alpha_m(k) J_m(kx)|^2 = 2\pi \sum_{m=-M}^M (J_m(kx))^2 \left| \sum_{q=1}^Q w_q e^{-im\varphi_q} \right|^2. \quad (\text{A.16})$$

The denominator of (3.28) can be expressed as

$$\begin{aligned} \int |S(\mathbf{x}; k)|^2 d\phi &= \sum_{q=1}^Q \sum_{p=1}^Q \int w_q w_p e^{ikx \|\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p\| \left(\left(\frac{\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p}{\|\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p\|} \right)^T \cdot \hat{\mathbf{x}} \right)} d\phi \\ &= \sum_{m=-\infty}^{\infty} \sum_{q=1}^Q \sum_{p=1}^Q w_q w_p i^m J_m(kx \|\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p\|) e^{im(\varphi_q - \varphi_p)} \int e^{-im\phi} d\phi \end{aligned}$$

The orthogonality property of exponentials gives (2.7)

$$\int e^{im\phi} e^{-i0\phi} d\phi = \begin{cases} 2\pi & m = 0 \\ 0 & m \neq 0 \end{cases}$$

allowing the simplification

$$\int |S(\mathbf{x}; k)|^2 d\phi = 2\pi \sum_{q=1}^Q \sum_{p=1}^Q w_q w_p J_0(kx \|\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p\|) \quad (\text{A.17})$$

(A.16) and (A.17) are substituted into (3.28) to give

$$\varepsilon_M(kx) = 1 - \frac{\sum_{m=-M}^M (J_m(kx))^2 \left| \sum_{q=1}^Q w_q e^{-im\varphi_q} \right|^2}{\sum_{q=1}^Q \sum_{p=1}^Q w_q w_p J_0(kx \|\hat{\mathbf{y}}_q - \hat{\mathbf{y}}_p\|)}$$

as required, thus completing the proof.

Proof of (3.30) – 2D Normalised Expansion Error

The 2D normalised expansion error is defined as

$$E_M(kx) = \frac{\int |\tilde{T}_M(\mathbf{x}; k)|^2 d\phi}{\int |S(\mathbf{x}; k)|^2 d\phi} \quad (\text{A.18})$$

where

$$\tilde{T}_M(\mathbf{x}; k) = \sum_{m=-\infty}^{-M-1} \kappa_m(k) J_m(kx) e^{im\phi} + \sum_{m=M+1}^{\infty} \kappa_m(k) J_m(kx) e^{im\phi}$$

and

$$\kappa_m(k) = \sum_{l=1}^L a_l(k) B_m(kr_l) e^{-im\varphi_l}.$$

The denominator can be calculated as

$$\begin{aligned}
 \left| \tilde{T}_M(\mathbf{x}; k) \right|^2 &= \sum_{m=-\infty}^{-M-1} \sum_{p=-\infty}^{-M-1} \kappa_m(k) \kappa_p^*(k) J_m(kx) J_p^*(kx) e^{im\phi} e^{-ip\phi} \\
 &+ \sum_{m=M+1}^{\infty} \sum_{p=M+1}^{\infty} \kappa_m(k) \kappa_p^*(k) J_m(kx) J_p^*(kx) e^{im\phi} e^{-ip\phi} \\
 &+ 2 \sum_{m=-\infty}^{-M-1} \sum_{p=M+1}^{\infty} \kappa_m(k) \kappa_p^*(k) J_m(kx) J_p^*(kx) e^{im\phi} e^{-ip\phi} \\
 \int \left| \tilde{T}_M(\mathbf{x}; k) \right|^2 d\phi &= 2\pi \sum_{m=-\infty}^{-M-1} |\kappa_m(k) J_m(kx)|^2 + 2\pi \sum_{m=M+1}^{\infty} |\kappa_m(k) J_m(kx)|^2 \quad (\text{A.19})
 \end{aligned}$$

Substituting (A.19) and the previously calculated (A.15) into (A.18) gives

$$E_M(kx) = \frac{\sum_{m=-\infty}^{-M-1} |\kappa_m(k) J_m(kx)|^2 + \sum_{m=M+1}^{\infty} |\kappa_m(k) J_m(kx)|^2}{\sum_{m=-\infty}^{\infty} |\alpha_m(k) J_m(kx)|^2}$$

as required, thus completing the proof.

Proof of (3.31) – 2D Normalised Reproduction Error

The 2D normalised reproduction error is defined as

$$\Upsilon_M(kx) \triangleq \frac{\int \left| S(\mathbf{x}; k) - \left(\hat{S}_M(\mathbf{x}; k) + \tilde{T}_M(\mathbf{x}; k) \right) \right|^2 d\phi}{\int |S(\mathbf{x}; k)|^2 d\phi} \quad (\text{A.20})$$

where

$$\begin{aligned}
 S(\mathbf{x}; k) &= \sum_{m=-\infty}^{\infty} \alpha_m(k) J_m(kx) e^{im\phi}, \\
 \hat{S}_M(\mathbf{x}; k) &= \sum_{m=-M}^M \alpha_m(k) J_m(kx) e^{im\phi}, \\
 \tilde{T}_M(\mathbf{x}; k) &= \sum_{m=-\infty}^{-M-1} \kappa_m(k) J_m(kx) e^{im\phi} + \sum_{m=M+1}^{\infty} \kappa_m(k) J_m(kx) e^{im\phi}
 \end{aligned}$$

and as previously defined

$$\kappa_m(k) = \sum_{l=1}^L a_l(k) B_m(kr_l) e^{-im\varphi_l}.$$

The denominator of (A.20) can be calculated as

$$\begin{aligned} S(\mathbf{x}; k) - \left(\hat{S}_M(\mathbf{x}; k) + \tilde{T}_M(\mathbf{x}; k) \right) &= \sum_{m=-\infty}^{-M-1} (\alpha_m(k) - \kappa_m(k)) J_m(kx) e^{im\phi} \\ &+ \sum_{m=-M}^M (\alpha_m(k) - \alpha_m(k)) J_m(kx) e^{im\phi} \\ &+ \sum_{m=M+1}^{\infty} (\alpha_m(k) - \kappa_m(k)) J_m(kx) e^{im\phi} \end{aligned}$$

$$\begin{aligned} S(\mathbf{x}; k) - \left(\hat{S}_M(\mathbf{x}; k) + \tilde{T}_M(\mathbf{x}; k) \right) &= \sum_{m=-\infty}^{-M-1} (\alpha_m(k) - \kappa_m(k)) J_m(kx) e^{im\phi} \\ &+ \sum_{m=M+1}^{\infty} (\alpha_m(k) - \kappa_m(k)) J_m(kx) e^{im\phi}. \end{aligned}$$

To simplify the following expression $\nu_m(k)$ is defined as

$$\nu_m(k) \triangleq \alpha_m(k) - \kappa_m(k)$$

giving

$$\begin{aligned} \int \left| S(\mathbf{x}; k) - \left(\hat{S}_M(\mathbf{x}; k) + \tilde{T}_M(\mathbf{x}; k) \right) \right|^2 d\phi &= \\ &\int \sum_{m=-\infty}^{-M-1} \sum_{p=-\infty}^{-M-1} \nu_m(k) \nu_p(k) J_m(kx) J_p(kx) e^{im\phi} e^{-ip\phi} d\phi \\ &+ \int \sum_{m=M+1}^{\infty} \sum_{p=M+1}^{\infty} \nu_m(k) \nu_p(k) J_m(kx) J_p(kx) e^{im\phi} e^{-ip\phi} d\phi \\ &+ 2 \int \sum_{m=-\infty}^{-M-1} \sum_{p=M+1}^{\infty} \nu_m(k) \nu_p(k) J_m(kx) J_p(kx) e^{im\phi} e^{-ip\phi} d\phi \end{aligned}$$

Using the exponential orthogonality property (2.7) gives

$$\begin{aligned} \int \left| S(\mathbf{x}; k) - \left(\hat{S}_M(\mathbf{x}; k) + \tilde{T}_M(\mathbf{x}; k) \right) \right|^2 d\phi &= 2\pi \sum_{m=-\infty}^{-M-1} |\nu_m(k) J_m(kx)|^2 \\ &+ 2\pi \sum_{m=M+1}^{\infty} |\nu_m(k) J_m(kx)|^2 \quad (\text{A.21}) \end{aligned}$$

Substituting (A.21) and the previously calculated (A.15) into (A.20) gives

$$\Upsilon_M(kx) = \frac{\sum_{m=-\infty}^{-M-1} |(\alpha_m(k) - \kappa_m(k)) J_m(kx)|^2 + \sum_{m=M+1}^{\infty} |(\alpha_m(k) - \kappa_m(k)) J_m(kx)|^2}{\sum_{m=-\infty}^{\infty} |\alpha_m(k) J_m(kx)|^2}$$

as required, thus completing the proof.

Detailed Implementation Design and Setup

B.1 Detailed Implementation Design and Setup

Table B.1 contains a detailed list of the equipment used to implement the reconstruction system. Photos of some of the major pieces of equipment used in the reproduction system are also given below.

Quantity	Description
1	Echoic Chamber, dimensions $3.5 \times 5.7 \times 2.3$ m (width \times breadth \times height)
1	Personal Computer with a 500 MHz Pentium III CPU and 512 MB of Main Memory
1	Creamware SCOPE Version 1.2 (Software)
1	Syntrillium Cool Edit Pro Version 1.2 (Software)
1	Creamware TDAT Analog 16 (used as a digital to analogue converter) – see Figure B.1
4	Yamaha Natural Sound Stereo Amplifier (AX-490) – see Figure B.2
8	Tannoy System 600 Loudspeakers – see Figure B.3
8	Loudspeaker Stands (height=1.2m)
8	Loudspeaker Cables
8	Audio Cables
16	Gold Loudspeaker Plugs
1	Optical Cable

Table B.1: Detailed Equipment List



Figure B.1: Creamware TDAT Analog 16



Figure B.2: Yamaha Natural Sound Stereo Amplifiers (AX-490)



Figure B.3: Tannoy System 600 Loudspeaker

From the specifications discussed in Section 5.1 the physical layout and cabling diagrams was constructed as shown in Figure B.4 and Figure B.5 respectively. A photo of the completed system is shown in Figure B.6 and Figure B.7. The procedure used to build the reproduction system is described below.

1. Speaker Placement – The first step in implementing the design was to accurately place the loudspeakers (see Figure B.4). This required
 - (a) Removing the floor insulation allowing for easier movement of equipment inside the room.
 - (b) Determining the origin of the listening area.
 - (c) Fastening a piece of paper at the origin with large arrows pointing to the location of each loudspeaker.
 - (d) Measuring a distance of 1.5m from the centre of the listening area in the direction of each loudspeaker. The loudspeaker stands were placed at these positions. Loudspeakers were then placed on each loudspeaker stand.
 - (e) Refitting the floor insulation inside the room. As the loudspeaker stands were now in position, the cut-outs of insulation that were previously used did not fit. To overcome this issue pieces of spare insulation were cut to fit.
2. Cabling – The second step was to cable the equipment together as per Figure B.5. This procedure was relatively straight-forward, except only four of the required eight audio cables connecting the digital to analogue converter to the audio amplifiers were present. Four additional cables were found and the appropriate connectors soldered to each end. The configuration for the audio amplifiers was such that the audio cables connected to the ‘CD’ input channel and the loudspeakers connected to the ‘A’ speaker output.
3. Amplifier Settings – The settings for each audio amplifier are identical, the input was selected to ‘CD’, the output set to Speaker ‘A’, and the volume set to 30dB.
4. Software Setup – The final stage in completing the implementation of the room was to configure the software. Two software packages were used:
 - (a) Scope – The scope software is used to configure the DSP hardware inside the computer. The setup for the hardware was relatively basic with four stereo wave file sources linked to four stereo output channels as shown in Figure B.8.
 - (b) Cool Edit Pro – Cool Edit Pro was used to play the four stereo wave files required for sound field reproduction. The output of each wave file was directed to one of the four stereo wave file sources in the Scope environment. A screen capture of this application is included as Figure B.9. The audio waveforms shown in this figure are those for Demonstration 5 described in Appendix C.

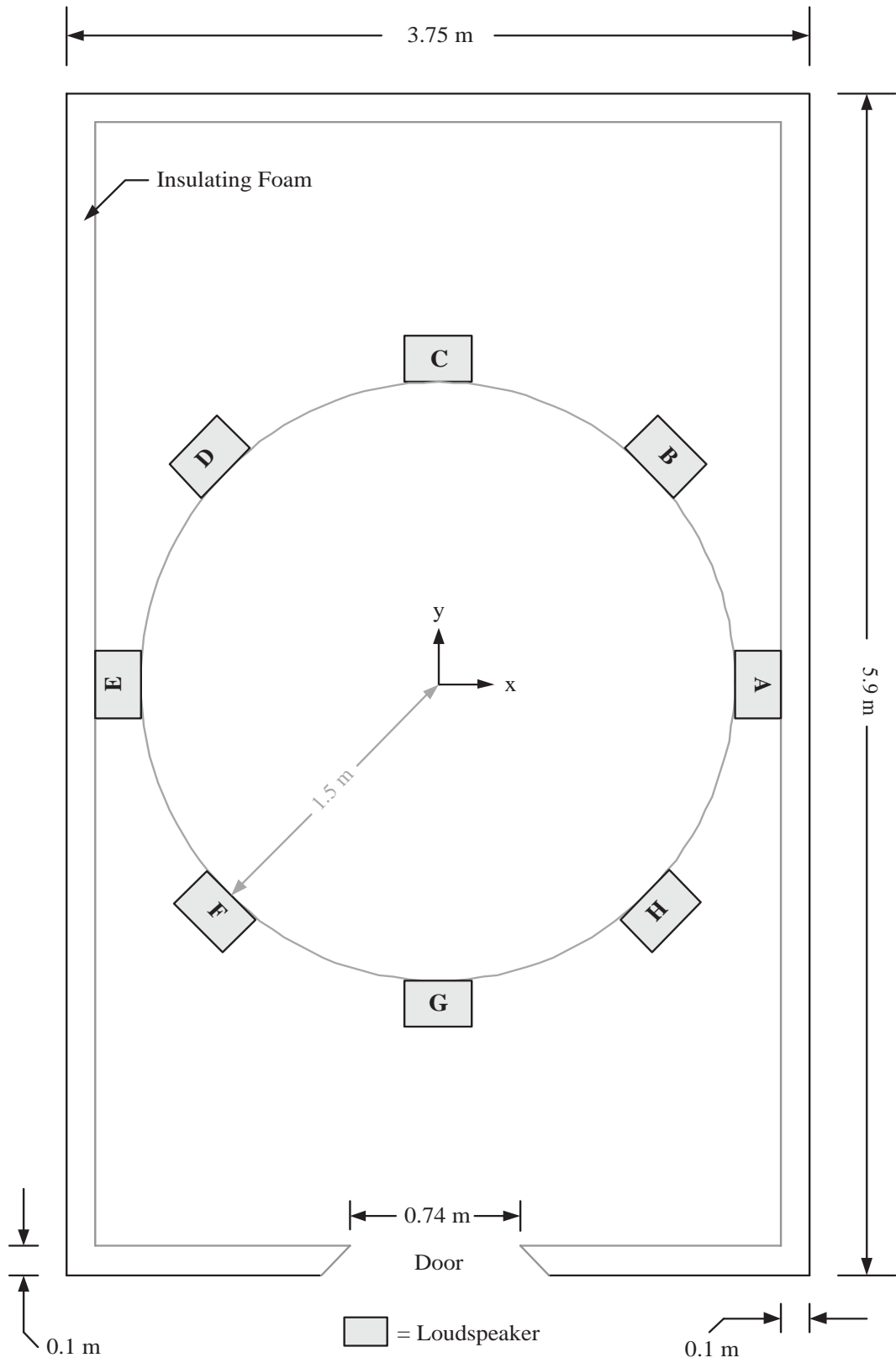


Figure B.4: Listening Environment Layout

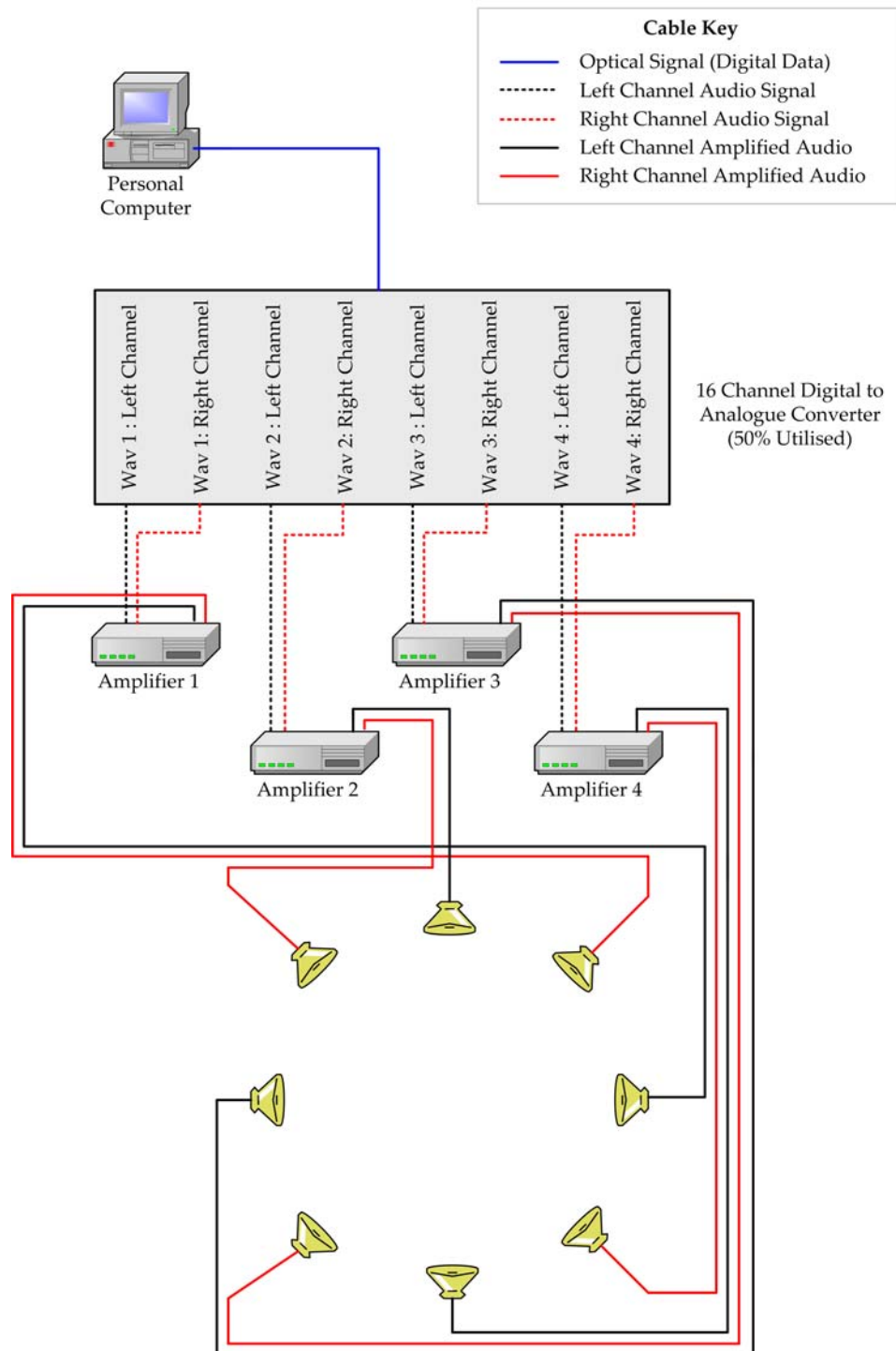


Figure B.5: Cabling Diagram for the sound field reproduction system.



Figure B.6: Photograph of the completed system.



Figure B.7: Photograph of the equipment used to drive the speakers.

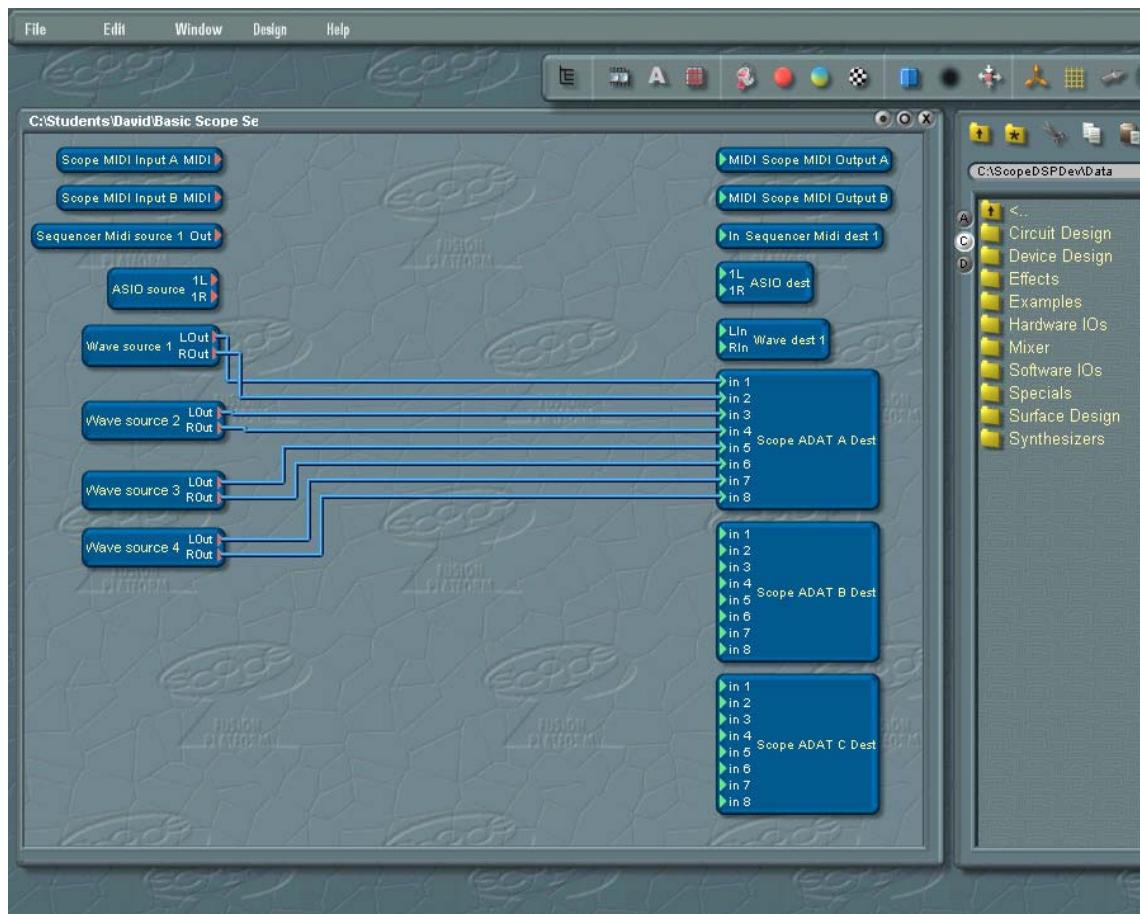


Figure B.8: Screen Capture of the Scope software package

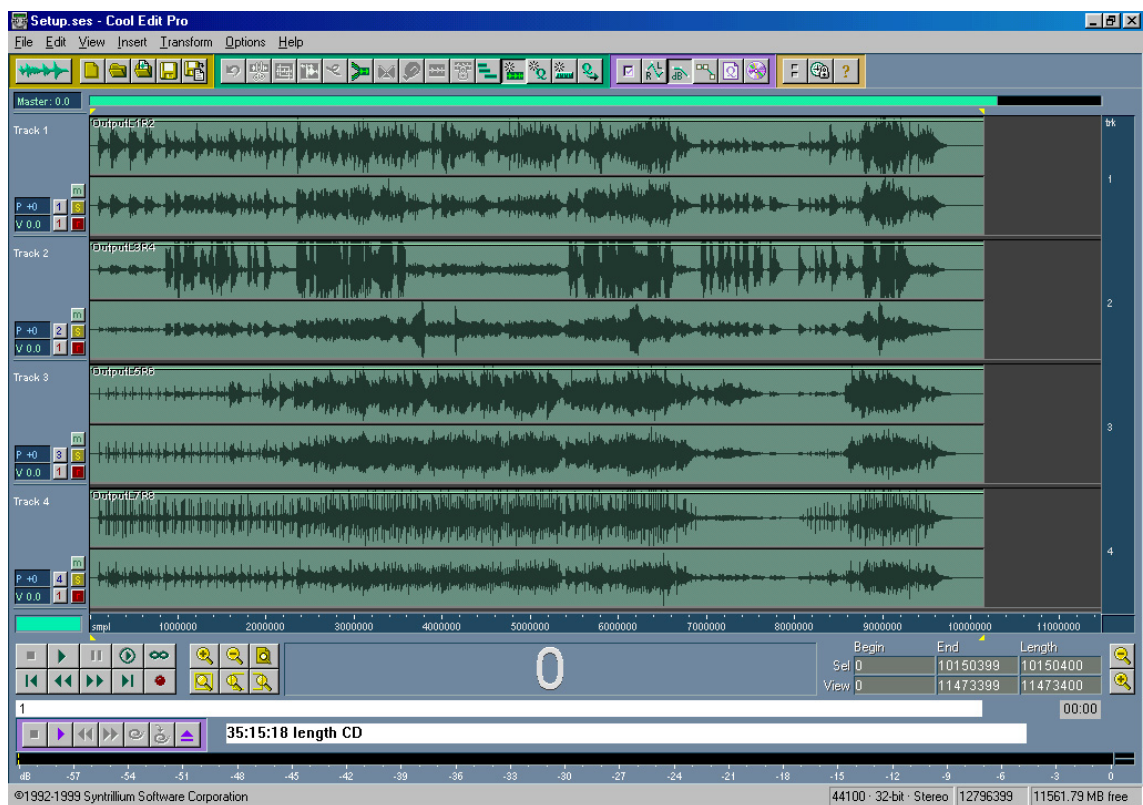


Figure B.9: Screen Capture of the Cool Edit Pro Software.

B.2 Windowing Function

The windowing function used in the implementation of the 2D sound field reproduction system is repeated here as Table B.2. Note that only the first 256 points of the 512 point window is shown. The second half of the window function is the mirror of the first i.e., $w[256] = w[257]$ and $w[1] = w[512]$. A plot of this function is shown in Figure B.10

	B=0	B=1	B=2	B=3	B=4	B=5	B=6	B=7	B=8	B=9
A=0	0.00014	0.00024	0.00037	0.00051	0.00067	0.00086	0.00107	0.0013	0.00157	0.00187
A=1	0.0022	0.00256	0.00297	0.00341	0.0039	0.00443	0.00501	0.00564	0.00632	0.00706
A=2	0.00785	0.00871	0.00962	0.01061	0.01166	0.01279	0.01399	0.01526	0.01662	0.01806
A=3	0.01959	0.02121	0.02292	0.02472	0.02662	0.02863	0.03073	0.03294	0.03527	0.0377
A=4	0.04025	0.04292	0.04571	0.04862	0.05165	0.05481	0.0581	0.06153	0.06508	0.06878
A=5	0.07261	0.07658	0.08069	0.08495	0.08935	0.09389	0.09859	0.10343	0.10842	0.11356
A=6	0.11885	0.12429	0.12988	0.13563	0.14152	0.14757	0.15376	0.16011	0.16661	0.17325
A=7	0.18005	0.18699	0.19407	0.2013	0.20867	0.21618	0.22382	0.23161	0.23952	0.24757
A=8	0.25574	0.26404	0.27246	0.281	0.28965	0.29841	0.30729	0.31626	0.32533	0.3345
A=9	0.34376	0.35311	0.36253	0.37204	0.38161	0.39126	0.40096	0.41072	0.42054	0.4304
A=10	0.4403	0.45023	0.4602	0.47019	0.4802	0.49022	0.50025	0.51028	0.52031	0.53033
A=11	0.54033	0.55031	0.56026	0.57019	0.58007	0.58991	0.5997	0.60944	0.61912	0.62873
A=12	0.63827	0.64774	0.65713	0.66643	0.67564	0.68476	0.69377	0.70269	0.7115	0.72019
A=13	0.72877	0.73723	0.74557	0.75378	0.76186	0.76981	0.77762	0.7853	0.79283	0.80022
A=14	0.80747	0.81457	0.82151	0.82831	0.83496	0.84145	0.84779	0.85398	0.86001	0.86588
A=15	0.8716	0.87716	0.88257	0.88782	0.89291	0.89785	0.90264	0.90728	0.91176	0.9161
A=16	0.92028	0.92432	0.92822	0.93197	0.93558	0.93906	0.9424	0.9456	0.94867	0.95162
A=17	0.95444	0.95713	0.95971	0.96217	0.96451	0.96674	0.96887	0.97089	0.97281	0.97463
A=18	0.97635	0.97799	0.97953	0.98099	0.98236	0.98366	0.98488	0.98602	0.9871	0.98811
A=19	0.98905	0.98994	0.99076	0.99153	0.99225	0.99291	0.99353	0.99411	0.99464	0.99513
A=20	0.99558	0.996	0.99639	0.99674	0.99706	0.99736	0.99763	0.99788	0.99811	0.99831
A=21	0.9985	0.99867	0.99882	0.99895	0.99908	0.99919	0.99929	0.99938	0.99946	0.99953
A=22	0.99959	0.99965	0.99969	0.99974	0.99978	0.99981	0.99984	0.99986	0.99988	0.9999
A=23	0.99992	0.99993	0.99994	0.99995	0.99996	0.99997	0.99998	0.99998	0.99998	0.99999
A=24	0.99999	0.99999	0.99999	1	1	1	1	1	1	1
A=25	1	1	1	1	1	1	0	0	0	0

Table B.2: Windowing Function $w[(10 \times A) + B]$, [33, p105]

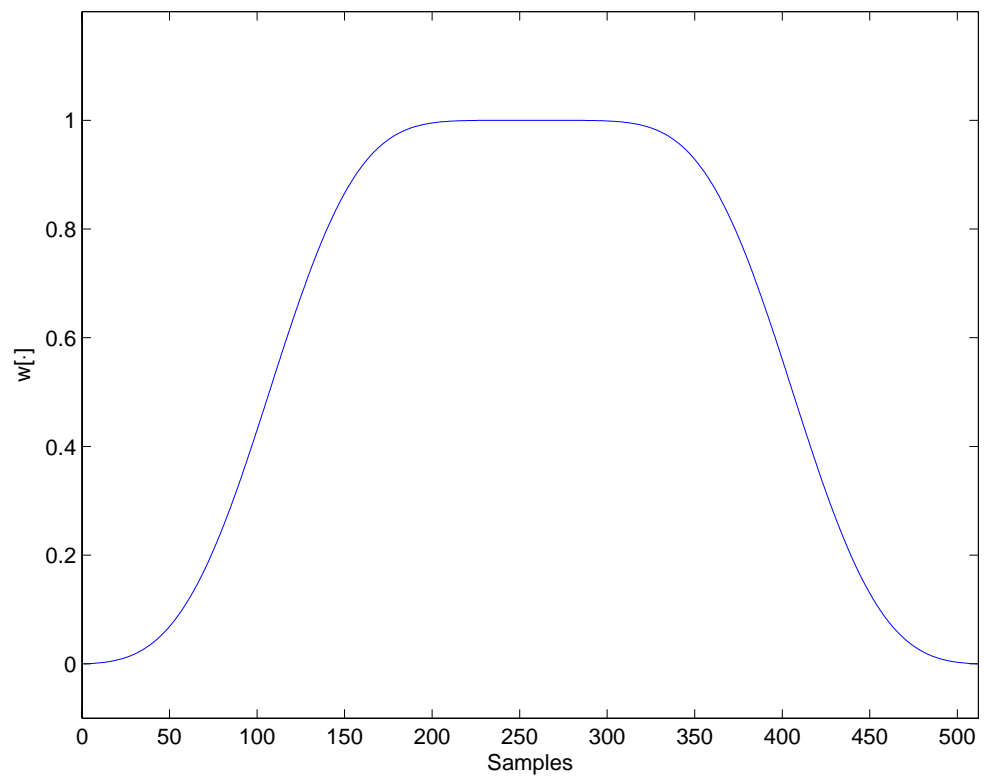


Figure B.10: AC-3 Window Function

Listening Test Summary

To demonstrate the 2D implementation of the sound field reproduction system five demonstrations were constructed. Each of the sound fields used in the demonstrations are described below. The sound files used for the demonstrations have been included in the compact disks attached to this thesis.

C.1 Single Stationary Narrowband Sound Source

In this demonstration, the sound field is constructed from a 800Hz stationary point source. The source is heard in bursts of 5 seconds. To illustrate the ability of positioning the point source it is moved to the locations $y_n = \begin{bmatrix} r & \phi \end{bmatrix}$

$$y_1 = [2.5\text{m} \quad 22.5^\circ]$$

$$y_2 = [2.5\text{m} \quad 212.25^\circ]$$

$$y_3 = [2.5\text{m} \quad 135^\circ]$$

$$y_4 = [2.5\text{m} \quad 315^\circ].$$

These point source locations are shown in Figure C.1.

C.2 Single Rotating Narrowband Sound Source

In the second demonstration a 800Hz point source is rotated around the room to construct the sound field. The sound source makes four rotations in an anti-clockwise direction. The rotations are completed in 30, 15, 7.5 and 3.25 seconds respectively. The point source maintains a distance of 2.5 meters from the origin of the listening area. The movement of the point sources is illustrated in Figure C.2.

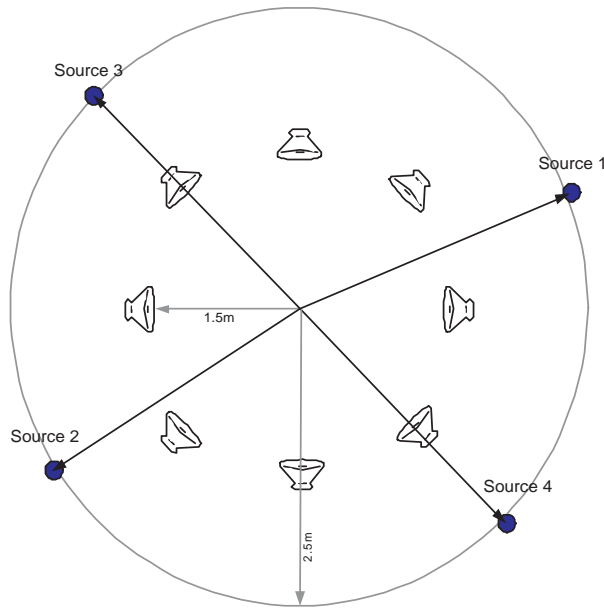


Figure C.1: Sound Field for Demonstration 1.

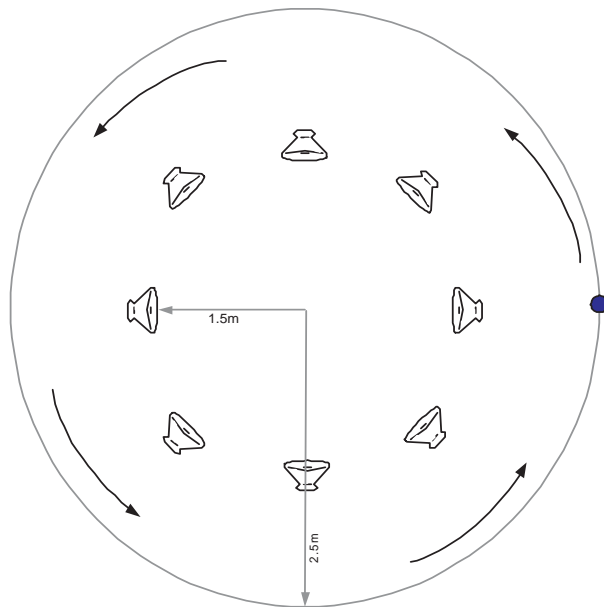


Figure C.2: Sound Field for Demonstrations 2 and 3.

C.3 Single Rotating Broadband Sound Source (Circle)

The third demonstration is similar to the second, but the narrowband sound source has been replaced with a broadband source. The sound source makes three anti-clockwise rotations of the sound field (as shown in Figure C.2). The rotations complete in 60, 30 and 15 seconds respectively. The sound source was extracted from the CD 'Paul Lyneham: A Memoir' [36]. This material is copyrighted to the Australian Broadcasting Corporation.

C.4 Single Rotating Broadband Sound Source (Square)

To illustrate the systems ability to recreate the depth of the original sound source, the fourth demonstration moves a broadband sound source in the path of a square. The path taken is shown in Figure C.3. A single clockwise rotation of the square is used for the duration of this demonstration. The sound source was extracted from the CD 'True History of the Kelly Gang' [37]. This material is copyrighted to the Australian Broadcasting Corporation.

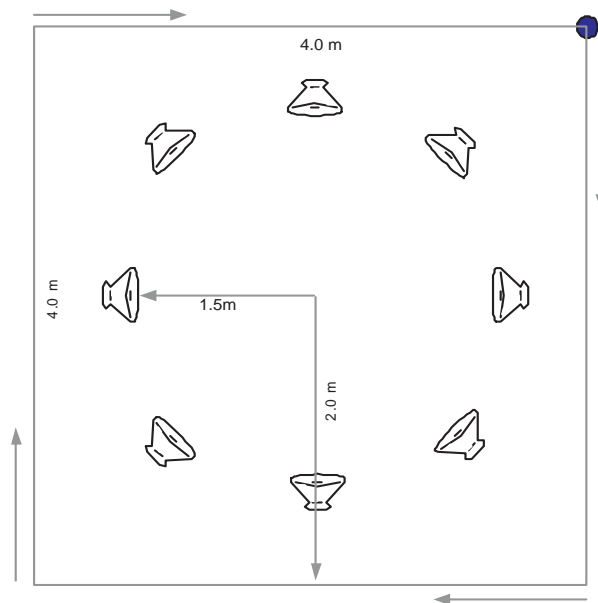


Figure C.3: Sound Field for Demonstration 4.

C.5 Multiple Stationary Broadband Sound Sources

The final demonstration consists of a sound field with multiple sound sources. The individual vocal and instrumental tracks from a song have been positioned around the listening area to give the impression of sitting in the middle of the band. The position of the sound sources is shown in Figure C.4. The song used for this demonstration is titled 'London Still' recorded by The Waifs. The individual tracks were made available by Triple J and the 'Noise' project at the Australian Broadcasting Corporation for a remix competition [38]. The copyright for the material used in this demonstration remains with the original artists.

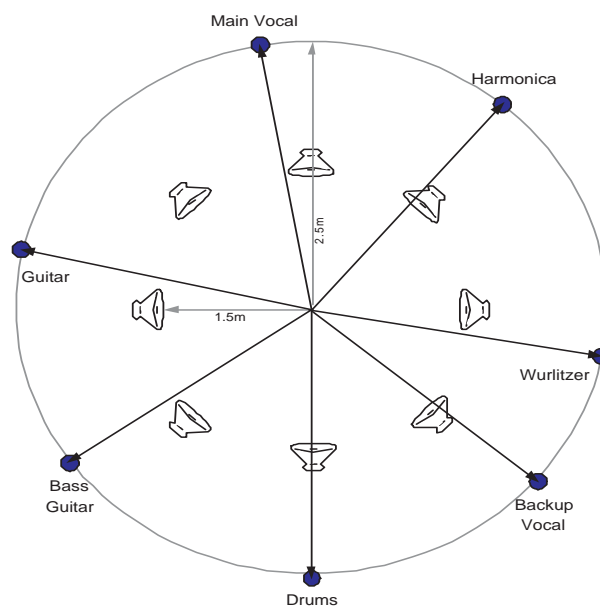


Figure C.4: Sound Field for Demonstration 5.

MatLab Code Listing

Throughout the project MatLab scripts were generated to verify and experiment with the design of the reproduction system. The code contained in this appendix has been divided into two sections. The first section contains the code used to simulate the 3D sound field reproduction system. This code forms the basis for the simulations provide in Chapter 4. The second section contains the code used to generate the loudspeaker audio files for the 2D implementation as discussed in Chapter 5.

The naming of functions referred to in the code listing is consistent with the definitions provided within this thesis. The remaining code developed during the execution of this project is included in the attached compact discs (as indicated in Appendix E).

D.1 Three Dimensional Simulation

simulation.m

```

%SIMULATION - 3D Sound Field Reconstruction Simulation
% This matlab function performs a simulation of the reproduction of a
% 3D sound field using an array of loudspeakers. The parameters used to
% formulate the problem are defined below in the section 'problem
% parameters'. The sound field that is reproduced is obtained from the
% function 'randomPlaneWaves(n)' where n is the number of plane-waves.
%
% The output of the function is two figures of the resulting
% reproduction. The first figure shows the real and imaginary components
% of the original and reproduced sound field. The second figure shows a
% plot of the normalised error as a function of the distance from the
% center of the sound field.
%
% Author David Excell
% Created 06-01-2003
% Modified 12-06-2003

function simulation

% Constants % meters/second
c = 340;

% Problem parameters
f = 1000 * (pi/180); % wavefield frequency of interest (radians/s)
xr = 0.2; % reproduction sphere radius (m)
rl = 1; % loudspeaker sphere radius (m)
z = 0; % image height in the sphere currently displaying (m)
imwh = 0.4; % range for display image x=y=-im:wh:im:wh (m)
scale = 0.005; % meters per image pixel
placement-error = 0; % maximum error for speaker placement (m)
type = 3; % 1 = truncation reproduction
% 2 = expansion reproduction
% 3 = total reproduction
method = 1; % 1 = Matrix based loudspeaker weight calculation
% 2 = Continuous Loudspeaker based loudspeaker weight
% calculation

% Obtain original sound field from random planewaves generated from the script
% 'randomPlaneWaves.m'. The description of each plane-wave is stored in
% the format [radius theta phi amplitude phase]. Note that the radius is

```

```

% not needed for the a plane-wave but is included in this formation to
% allow the specification of point sources.
randomPlaneWaves(5)
load planewaves.mat ypt

% Calculations based on parameters
k = (360*f)/c; % calculate the wavenumber
pix_wh=im_wh*2/scale+1;

% Optimum number of recorded spherical harmonic coefficients and loud
% speakers for accurate reproduction within the radius 'xr'
Ns = ceil(k*xr); % Order of reproduction required for accuracy within xr sphere
L = (Ns+1)^2; % Number of loudspeakers required

% calculate the co-ordinates for each image pixel
pt = zeros(pix_wh^2, 3);
idx = 1;
for x=-im_wh:scale:im_wh
for y=flipdim(-im_wh:scale:im_wh,2)
% Obtain the value using spherical harmonics
pt(idx, 1) = x;
pt(idx, 2) = y;
pt(idx, 3) = z;
idx=idx+1;
end
end

% Convert image points to spherical co-ordinates
[r theta phi] = c2s(pt); pt = [r theta phi];

% Obtain the sound field coefficients for the sound field
alpha = getSphericalCoefficients(k, Ns, ypt, 1);

% obtain 'L' loudspeaker positions in spherical co-ordinates
spkr_pt = speakerpositions(L);
[r theta phi] = c2s(spkr_pt); spkr_pt = [r*rl theta phi];

% calculate loudspeaker weights using the sound field coefficients and
% the loudspeaker positions
switch method
case 1
a = getLoudSpeakerWeights(k, alpha, spkr_pt, Ns);
case 2
a = getContinuousLoudSpeakerWeights(alpha, rl, k, Ns, spkr_pt);
end

```

```

figure(1); subplot(2,2,1)
imagesc(real(field)); setaxes;
title('Real'); colormap('gray'); caxis([-1 1]); colorbar;
ylabel('Original')
hold on; plot(circle_x, circle_y, 'r'); hold off

subplot(2,2,2);
imagesc(imag(field)); setaxes;
title('Imaginary'); colormap('gray'); caxis([-1 1]); colorbar;
hold on; plot(circle_x, circle_y, 'r'); hold off

subplot(2,2,3)
imagesc(real(reproduction_field)); setaxes
title(''); colormap('gray'); caxis([-1 1]); colorbar;
ylabel('Reproduction')
hold on; plot(circle_x, circle_y, 'r'); hold off

subplot(2,2,4);
imagesc(imag(reproduction_field)); setaxes
title(''); colormap('gray'); caxis([-1 1]); colorbar;
hold on; plot(circle_x, circle_y, 'r'); hold off

% Display the calculated reproduction accuracy
figure(2);
semilogy(0:scale:im_wh, errors_r)
xlabel('Radius (m)')
ylabel('Normalised Error')

% Some text about the simulation
disp(' ');
disp(sprintf('Number of loudspeakers used: %d', I))
disp(sprintf('Reproduction Error (radius): %0.3g percent', 100 * errors_r( xr/scale + 1) ))

% -----
function field = CalculateActualFieldExp(pt, k, ypt)

% CalculateActualFieldExp is used to obtain an image of the original
% sound field that the system is attempting to reproduce. The equation
% implemented in this function is (2.16) given in Section 2.6
% of this thesis.
%
% Inputs:
% pt - spherical coordinates of the points to evaluate the
% sound field for.
% k - frequency of the sound field.
% ypt - plane-waves making up the original sound field

% generate an 'image' of the original and reproduced sound fields.
field = CalculateActualFieldExp(pt, k, ypt);
switch type
case 1 % truncated reproduction
reproduction_field = CalculateReproducedFieldSph(pt, spkr_pt, Ns, k, a);
case 2 % 'expansion' reproduction
reproduction_field = CalculateReproducedFieldExp(pt, spkr_pt, k, a) - ...
CalculateReproducedFieldSph(pt, spkr_pt, Ns, k, a);
case 3 % total reproduction
reproduction_field = CalculateReproducedFieldExp(pt, spkr_pt, k, a);
end
% calculate the points surrounding the listening area so that it can be
% indicated in the figure.
t = [0:0.01:2*pi 0];
r = r*cos(2*pi/(2*xr));
circle_x = r*cos(t)/(scale) + pix_xwh/2+scale/2;
circle_y = r*sin(t)/(scale) + pix_xwh/2+scale/2;

% Remove NaN's from the field (caused by divide by zero) and replace with zero
field(find(isnan(field))==1)=0;
reproduction_field(find(isnan(reproduction_field))==1)=0;

% Calculate the normalised error as a function of the spherical radius.
errors_r = zeros(im_wh/scale+1,1);
count = 1;
for nr=0:scale:im_wh
I = find(nr-(scale/2) <= pt(:,1) & pt(:,1) <= nr+(scale/2));
if type == 2
errors_r(count) = sum(abs(reproduction_field(I)).^2) / sum( abs(field(I)).^2 );
else
errors_r(count) = sum(abs(field(I)-reproduction_field(I)).^2) / sum( abs(field(I)).^2 );
end
count = count +1;
end

% Reformat the image from a (1 x pix_xwh^2) array to a (pix_xwh x pix_xwh)
% array
field = reshape(field, pix_xwh, pix_xwh);
reproduction_field = reshape(reproduction_field, pix_xwh, pix_xwh);

% Display the real and imaginary components of the original and
% reproduced sound fields.

```

```

% coordinate system to the Cartesian coordinate system.
[x y z] = s2c(sprk_pt(:,1),sprk_pt(:,2),sprk_pt(:,3));
sprk_ptc = [x y z];
[x y z] = s2c(pt(:,1), pt(:,2), pt(:,3));
cpt = [x y z];

% Calculate the reproduced sound field as per (2.18). The reproduced sound
% field is calculated by considering one loudspeaker at a time and
% summing the individual sound fields together.
field = zeros(size(pt,1),1);
for p=1:size(sprk_pt,1)
    % Calculate the distance between the loudspeaker and each sound
    % field image point.
    diff = repmat(sprk_ptc(p,:), size(pt,1), 1);
    diff = diff - cpt;
    dis = distance(diff);

    % Evaluate the loudspeaker sound field equation
    sprk = sprk_pt(p,1) * exp(i*k* sprk_pt(p,1) );
    spr = sprk * (exp(-i*k*dis)/dis);

    % Apply the weight to the sound field and sum to the total
    % reproduced sound field
    field = field + a(p) * spr;
end

% -----
function y = CalculateReproducedFieldSph(pt, sprk_pt, N, k, a)

% CalculateReproducedFieldSph is used to obtain an image of the sound
% field produced by the reproduction system using spherical harmonics
% The primary role of this function is to evaluate the truncation
% and expansion error which are defined in terms of spherical harmonics.
% The equation implemented in this function is (3.2) which is
% described in Section 3.2 of this thesis.
%
% Inputs:
% pt - spherical coordinates of the points to evaluate the
% sound field for.
% sprk_pt - spherical coordinates for the speaker positions
% N - Order to which the spherical harmonics should be
% evaluated to.
% k - frequency of the sound field.
% a - Loudspeaker weights

```

```

% determine the size of the sound field image
num_pts = size(pt,1);

% store the distance from origin to each point
dis = pt(:,1);

% Calculate the unit vectors (x hat and y hat) for each image
% point in Cartesian coordinates
pt(:,1) = ones(num_pts,1);
[x y z] = s2c(pt(:,1), pt(:,2), pt(:,3));
cpt = [x y z];

[x y z] = s2c(ypt(:,1), ypt(:,2), ypt(:,3));
cypt = [x';y';z'];

% Calculate the dot product of the two unit vectors
dotprod = cpt * cypt;

% calculate the plane-wave weights
w = ypt(:,5) .* exp(i*k*ypt(:,4));

% Format arrays so that sound field can be calculated using the .*
% operator
w = repmat(w',num_pts,1);
dis = repmat(dis,1, size(ypt,1));

% Calculate the contents of the original sound field as given by (2.16)
field = sum(w .* exp(i.*k.*dis.*dotprod),2);

% -----
function field = CalculateReproducedFieldExp(pt, sprk_pt, k, a)

% CalculateReproducedFieldExp is used to obtain an image of the
% sound field produced by the system in an attempt to reproduce the
% original. The basic equation implemented in this function is (2.18)
% given in Section 2.6 of this thesis.
%
% Inputs:
% pt - spherical coordinates of the points to evaluate the
% sound field for.
% sprk_pt - spherical coordinates for the speaker positions
% k - frequency of the sound field.
% a - loudspeaker weights

% Convert the loudspeaker and sound field coordinates from the Spherical

```

getSphericalCoefficients.m

```

H = waitbar(0, 'Calculating Reproduced Soundfield');
% Initialise sound field
y = zeros(size(pt,1),1);
% Loop through each order of spherical harmonic
for n=0:N
    waitbar(n/N,H)
    % Get each of the modes for the current order 'n'
    m=-n:n;
    % Repeat the loudspeaker weightings so that they can be by point
    % multiplied with each of the speaker spherical harmonics
    am = repmat(R(n,k*spkr_pt(:,1)) .* a, [1, size(m,2)]);
    % Calculate the loudspeaker spherical harmonics
    spk = sum( am .* conj(spharmonic(spkr_pt,n,m)) );
    % Repeat the loudspeaker spherical harmonics so they can be point
    % multiplied with the spherical harmonics for each of the sound
    % field image points.
    spk = repmat(spk, size(pt,1), size(m,1));
    % Calculate the sound field spherical harmonic and multiply with
    % the loudspeaker spherical harmonics.
    z = sum(spk .* spharmonic(pt,n,m), 2);
    % Apply the spherical bessel weighting function to the harmonics.
    y = y + X(n,k*pt(:,1)) .* z;
end

y = y.*(4*pi);
close(H);

% -----
function setaxes
% setaxes is used to format the axes on figure 1 in the output of the
% simulation function

a = gca;
set(a, 'XTick', [1 80 161]); set(a, 'YTickLabel', [-0.4 0 0.4])
set(a, 'YTick', [1 80 161]); set(a, 'YTickLabel', [0.4 0 -0.4])

```

```

%GETSPHERICALCOEFFICIENTS - Obtain 3D Sound Field Coefficients
% GETSPHERICALCOEFFICIENTS(K,N,YPT,TYPE) generates the sound field
% coefficients for the sources located at YPT with frequency K up to the
% order N.
%
% TYPE is used to specify the source model to use. TYPE==1 corresponds to
% a plane-wave sources and TYPE==2 corresponds to point sources.
%
% Author David Excell
% Created 09-01-2003
% Modified 12-06-2003
function alpha = getSphericalCoefficients(k, N, ypt, type)
% Initialise memory for coefficients
alpha = zeros((N+1)^2, 1);

switch type
case 1 % Plane-wave Sources
    for i=1:size(ypt,1)
        alpha = alpha + ypt(i,5)*getPlaneWaveSphericalCoefficients(k, N, ypt(i,:));
    end
case 2 % Point Sources
    for i=1:size(ypt,1)
        alpha = alpha + ypt(i,5)*getPointSourceSphericalCoefficients(k, N, ypt(i,:));
    end
end
% -----
function y = getPlaneWaveSphericalCoefficients(k, N, ypt)
% getPlaneWaveSphericalCoefficients calculates the sound field
% coefficients for a single plane-wave (ypt). Note that ypt(4)
% corresponds to the phase of the plane-wave. The bases of the
% equation implemented in this function is given by (2.15)

y = zeros((N+1)^2,1);
low = 1;
for n=0:N
    up = low+(2*n);
    y(low:up) = exp(-i*k*ypt(4)) * (i^n) * (conj(spharmonic(ypt, n, -n:n)))';
    low = up+1;
end

```

```

end
y = (4*pi) * y;
% -----
function y = getPointSourceSphericalCoefficients(k, N, Ypt)
% getPointSourceSphericalCoefficients calculates the sound field
% coefficients for a single point source (Ypt). The bases of the
% equation implemented by this function is given by (2.14).
y = zeros((N+1)^2,1);
low = 1;
for n=0:N
    up = low+(2*n);
    y(low:up) = R(n, k*Ypt(1)) * (i^n) * (conj(spharmonic(Ypt, n, -n:n)));
    low = up+1;
end
y = (4*pi) * y;

```

getLoudspeakerWeights.m

```

%GETLOUDSPEAKERWEIGHTS - Obtain solution to reproduction equation
% GETLOUDSPEAKERWEIGHTS(K,ALPHA,SPKR_PT,NS) calculates the loudspeaker
% weights for the loudspeakers specified by SPKR_PT for the sound field
% specified by ALPHA at the frequency K. NS specifies the order of sound
% field coefficients to consider.
%
% The solution to the reproduction problem in this function is based on
% the linear relationships that are described in Section 3.1.3 of
% this thesis.
%
% Author David Excell
% Created 09-01-2003
% Modified 12-06-2003
function a = getLoudspeakerWeights(k, alpha, spkr_pt, NS)
L = size(spkr_pt,1); % Number of speakers in reproduction
K = (NS+1)^2; % Number of alpha coefficients available
% Make sure the correct range of sound field coefficients is being
% used.
alpha = alpha(1:(NS+1)^2);

```

```

% Calculate the 'P' matrix as described by equation (3.13)

```

```

P = zeros((NS+1)^2,L);
low=1;
for n = 0:NS
    up = 2*n+low;
    spkr_vals = repmat(R(n,k*spkr_pt(:,1)'), 2*n+1, 1) .* conj(spharmonic(spkr_pt, n, -n:n));
    P(low:up,:) = 4*pi*(i^n)*spkr_vals;
    low = up+1;
end
% Solve (3.13) for the loudspeaker weights using (3.18) if L==K and
% (3.19) otherwise
if L == K
    a = inv(P) * alpha;
else
    a = pinv(P) * alpha;
end

```

getContinuousLoudspeakerWeights.m

```

%GETCONTINUOUSLOUDSPEAKERWEIGHTS - Obtain solution to reproduction equation
% GETCONTINUOUSLOUDSPEAKERWEIGHTS(ALPHA,RL,K,NS,SPKR_PT) calculates the
% loudspeaker weights for the loudspeakers specified by SPKR_PT on the
% spherical radius RL for the sound field specified by ALPHA at the
% frequency K. NS specifies the order of sound field coefficients to
% consider.
%
% The solution to the reproduction problem used in this function is based on
% the continuous loudspeaker method that is described in Section 3.1.5
% of this Thesis.
%
% Author David Excell
% Created 11-02-2003
% Modified 12-06-2003
function a = getContinuousLoudspeakerWeights(alpha, rl, k, NS, spkr_pt)
% Calculate the gamma coefficients for the continuous loudspeaker
% function as in (3.22)
gamma = ones((NS+1)^2, 1);
low=1;
for n=0:NS

```

```

up=2*n+low;
gamma(low:up) = alpha(low:up) / (4*pi*(i^n)*R(n,k*r1));
low = up+1;
end

% Sample the continuous loudspeaker function at the point 'spkr_pt' and
% obtain the loudspeaker weights
a = zeros(size(spkr_pt,1),1);
low=1;
for n=0:Ns
    up = low+2*n;
    a = a + (gamma(low:up) * spharmonic(spkr_pt,n, -i*(-n:n)))';
    low = up+1;
end

% Get the mean angle between the loudspeakers
theta = mean_angle(size(spkr_pt,1));

% Calculate the unit surface area of the cap that the loudspeaker
% is acting over.
h = 1 * (1-cos(cheta/2));
SA = 2 * pi * h;

% Scale the loudspeaker weights by the surface area
a = a * SA;

function theta = mean_angle(L)

% get the loudspeaker positions (in cartesian co-ordinates)
pt = speakerpositions(L);

% Convert to spherical co-ordinates
[theta phi rpt] = cart2sph(pt(:,1), pt(:,2), pt(:,3));

% Calculate the minimum angular distance between each of the
% loudspeakers
dist = Inf*ones(L,1);
for n=1:L
    for m=1:L
        if not(n==m)
            d = sphere_dist(theta(n),phi(n), theta(m), phi(m));
            if d < dist(n)
                dist(n) = d;
            end
        end
    end
end
end

end

theta = mean(dist);

end

%SPHARMONIC - Spherical Harmonic
% SPHARMONIC(PT,N,M) calculates the spherical harmonics for the PT given
% the order N and modes M. PT and M may be vectors but N must be scalar.
% PT must be specified using spherical coordinates in the form
%
% PT = [radius theta phi]
%
% The equation implemented in this function is given by (2.8) in this
% thesis.
%
% Author David Excell
% Created 16-12-2002
% Modified 12-06-2003

function y = spharmonic(pt, n, m)

% Calculate the associate Legendre component of the spherical harmonic.
% The output is returned for each of the modes m=0:n
P = legendre(n, cos(pt(:,2)))';

Pm = zeros(size(pt,1), size(m,2));
Am = zeros(size(m));

for a=1:size(m,2)
    % Factorial function only excepts scalars that why Am is calculated
    % in a loop
    Am(a) = sqrt( ( (2*n+1)/(4*pi) ) * ( factorial(n-abs(m(a)))/factorial(n+abs(m(a))) ) );

    % Need to format the sasoicate Legendre component for m=-n:n
    % so the .* operator can be used.
    Pm(:,a) = P(:, abs(m(a))+1);
end

% Reformat Am co-efficient matrix so it can be point multiplied with
% the spherical coordinates.
Am = repmat(Am, size(pt,1), size(m,1));

```

```

% Problem constants
% -----
c = 340;           % speed of sound (meters/second)
M = 4;            % order of reproduction
L = 8;            % number of loudspeaker
r1 = 1.5;         % loudspeaker radius
samples = 512;    % Number of samples in each window
overlap = 256;    % Overlap between each window
windows_per_loop = 512; % How many windows to operate on before
                    % saving data to temporary file.
padding = 3*samples; % allow time delay of data
sp = samples + padding; % total window data length

% Start timing
t0 = clock;

% Load the sound file that is to be synthesised
[y, Fs, Nbits] = wavread(filename);

% Only consider one channel of the sound file as only point source is
% being considered
y = y(1:44100,1);

% If the original file is not sampled at the correct rate, resample the
% file so this is the case.
if not(Fs==Fs_Output)
    y = resample(y, Fs_Output, Fs);
end

Fs = Fs_Output;
len = size(y,1);
disp(sprintf('\nLoaded sound file %s (%0.3gs)', filename, len/Fs));

% Calculate index for the frequency bins that are generated when the
% Fourier transform is taken over the window size. Note that the
% negative frequencies will be discarded as the input signal is real.
k = (0:(sp/2)) / (sp/2);
k = (2*Fs*pi/c) * k';

% Get the window function that will be used to shape the input data
load window.mat w

% Pad the sound with zeros so that there are zeros for half the
% window size at the beginning and end of the sound field and also so
% that is an integer number of windows within the sound file. Note
% that one window is 0.04 seconds so that this should have little

```

```

% Multiply all the components of (2.8) together to complete the
% operation
y = Am.*Pm.*exp(i*pi*(pt(:,3)*m));

```

D.2 Two Dimensional Implementation

GenerateSoundfield.m

```

% GENERATESOUNDFIELD is used to transform a given sound sample into a sound
% field. The sound sample is modelled as a point source located in two
% dimensional space.
%
% The script will first resample the file to 44100Hz. Then in 512 sample
% windows, it will transform the sound sample into a sound field. Each
% window overlaps the next by 256 samples. The output of the function is
% four stereo wave files that can be played through an array of eight
% loudspeakers producing the required sound field.
%
% The script will work on sound samples of any length. If the sound sample
% is longer than 6.4 seconds, every 6.4 seconds of data will be
% written to temporary files to stay within main memory limits. After
% completion the temporary files are converted into loudspeaker wave files.
%
% David Excell
% Created 18-04-2003
% Modified 23-04-2003

```

```

function generateSoundField

global c M L r1 power_threshold

% User parameters ...
% -----
filename = 'mono800Hz.wav'; % Target sample wavefile
Nbit_Output = 32;          % Number of bits per wavefile sample
Fs_Output = 44100;         % Output file sample rate.
power_threshold = 1e-3;    % Frequency Power Threshold
tmpFileName = 'TempSpeaker%d.bin'; % Temp file mask
wavFileName = 'OutputL%dR%d.wav'; % Wave file name mask

```

```

% impact on the generated sound field.
s = mod(size(y,1), samples);
s = sp-s;
s_up = floor(s/2) + overlap + padding;
s_low = ceil(s/2) + overlap;

y = [zeros(s_low,1); y; zeros(s_up,1)];
len = size(y,1)-padding;

% Generate the inverse loudspeaker matrix for each discrete frequency
% that will be returned from the Fourier transform
disp('Calculating loudspeaker inverse matrix')
global P_inverse
P_inverse = loudspeakermatrix(k);

% Determine how many windows the sound sample will be divided into
% (Note: the windows overlap).
windows = (len / ( samples - overlap )) - 1;

% Generate the location information for each windows. This is used to
% give the impression that the sound source is moving
type = 1;
switch type
case 1
    % Sound Sample is stationary:
    ypt = [2.5 * ones(windows,1) (0*pi/180) * ones(windows,1)];
case 2
    % Circular Rotation:
    ypt = [3 * ones(windows,1) (2*pi*((0:(windows-1))/windows))];
case 3
    % Square Rotation:
    % Calculate the length the sound moves along each edge
    ws1 = ceil(0.25*windows);
    ws2 = windows - 3*ws1;
    ws1ha = ceil(0.5*ws1);
    ws1hb = ws1-ws1ha-1;
    ws2ha = ceil(0.5*ws2);
    ws2hb = ws2 - ws2ha-1;
    % Calculate the positions alongh each edge (cartesian coords)
    yptA = [2 * ones(ws1,1) flipud(2*((-ws1ha):(ws1hb)))/ws1ha)];
    yptB = [flipud(2*((-ws1ha):(ws1hb)))/ws1ha) -2 * ones(ws1,1)];
    yptC = [-2 * ones(ws1,1) 2*(((-ws1ha):(ws1hb)))/ws1ha)];
    yptD = [2*(((-ws2ha):(ws2hb)))/ws2ha 2 * ones(ws2,1)];
end

% Join the edges together and convert to cylindrical coords.
ypt = [yptA; yptB; yptC; yptD];
ypt = cart2cyl(ypt);

end

% Setup some variables and indexes for use in the loop
w_single = w;
% Input window function
w_loudspeaker = repmat(w, 1, L)';
% Output window function
alpha = struct('data', {}, 'I', {});
% Storage of soundfield values
alpha = repmat(alpha, 1, windows_per_loop);
% Array of storage for each windows
low_a = 1; up_a = samples;
% Indexes of the current data range
% for the sound sample
% Index for the sound sample position
% Timing info.
yptidx=1;
tg = 0; ts = 0;

loops = ceil(windows/windows_per_loop);

% Check if the temporary files are required, if so, make sure they do
% not already exist, and if they do, remove them. A handle to each file
% is also created so that it can be easily written to within the loop.
if loops > 1
    fid=zeros(8,1);
    for i=1:8
        outfile = sprintf(tmpFileName, i);
        if exist(outfile, 'file') > 0
            disp(sprintf('Deleting: %s.', outfile));
            delete(outfile);
        end
        fid(i) = fopen(sprintf(tmpFileName, i), 'a+');
    end
end

% This loop does the most significant work of generating the sound
% field. The general process is to create the alpha coefficients to
% describe the sound field, and then use these coefficients to
% generate the waveforms to be played by each loudspeaker.
compress = 0; count=0;
tloop = clock;
for j = 1:loops
    % Calculate wall time until the loop terminates
    if j>=1
        s = etime(clock, tloop);
        time_per_loop = s/j;
    end
end

```

```

time_left = ceil((loops-j)*time_per_loop);
m = floor(time_left/60);
time_left = mod(time_left,60);
s = sprintf('(Time Left = %d:%d [m:s])', m, time_left);
else
s='';
end
disp(sprintf('Starting Step %d of %d %s', j, loops, s));

% If this is the last loop, there may not be enough windows left
% to evaluate as per the prespecified 'windows_per_loop'
if j*windows_per_loop > windows
windows_per_loop = windows - (j-1)*windows_per_loop;
end

% Create array to store the loudspeaker waveforms, also if we have
% been through the loop before, we need to get the values that will
% be overlapped.
waveform = zeros(L, (samples-overlap) * (windows_per_loop+1)*padding);
if j > 1
waveform(:, 1:(overlap*padding)) = values;
end

% Generate the sound field coefficients for each window of data of the
% original sound field in the index range low_a:up_a. (Note: up_a -
% low_a = samples).
tx = clock;
for i=1:windows_per_loop
% Apply windowing function
data = y(low_a:up_a) .* w_single;
data = [zeros(1*samples,1); data; zeros(2*samples,1)];

% Generate and store sound field coefficients
[alpha(i),data_alpha(i),I c] ...
= generateAlpha(data, k, ypt(yptIdx, :));
compress = compress + c;
count = count +1;

% Progress indexes
low_a = low_a + (samples-overlap); up_a = up_a + (samples-overlap);
yptIdx = yptIdx+1;
end
ts = ts + etime(clock,tx);
yptIdx = yptIdx - windows_per_loop;

% Now transform the sound field coefficients into loudspeaker

```

```

% waveforms.
tx = clock;
low_b = 1; up_b = sp;
for i = 1:windows_per_loop

% Get the loudspeaker waveform for this window
data = generateLoudspeaker(alpha(i), sp, w, ypt(yptIdx, :));
yptIdx = yptIdx + 1;

% Slot the window into the waveform, noting that it overlaps
% with previous data, and thus needs to be added.
if not(up_b > size(waveform,2))
waveform(:, low_b:up_b) = waveform(:, low_b:up_b) + data;
end

% Progress indexes.
low_b = low_b + (samples-overlap); up_b = up_b + (samples-overlap);
end

tg = tg + etime(clock,tx);

% If there is more than one loop need to store current waveform in
% an out of main memory location (i.e. file on the hard disk).
if not(loops==1)
% Determine how much data we can store. If this is not the last
% loop, the last 'overlap' values can not be stored as they
% still have to overlap with the next window.
ePoint = size(waveform,2);
if not(j==loops)
ePoint = ePoint - overlap - padding;
end

% Write the data to each of the open files.
for i=1:8
fwrite(fid(i), waveform(i, 1:ePoint), 'double');
end

% Store the data that is to be overlapped by the next window.
if not(j==loops)
ePoint = ePoint+1;
values = waveform(:, ePoint:size(waveform,2));
end
end
compress = compress / count;

% Clear up some memory

```

```

if not(loops==1)
    clear waveform alpha
end

% Transform the arrays of data into the wave files to be played
% through each loudspeaker
for i = 1:(L/2)
    L1 = 2*i-1;
    L2 = 2*i;

    % Determine where to get the waveforms from, either memory or
    % temporary storage files.
    if (loops==1)
        left = waveform(L1,:);
        right = waveform(L2,:);
    else
        fseek(fid(L1), 0, -1);
        left = fread(fid(L1), inf, 'double');
        fseek(fid(L2), 0, -1);
        right = fread(fid(L2), inf, 'double');
    end

    % Generate output filename and check that it does not already
    % exist. If it does exist delete it.
    outfileName = sprintf(wavFileName, L1, L2);
    if exist(outfileName, 'file') > 0
        disp(sprintf('Deleting: %s', outfileName));
        delete(outfileName);
    end

    disp(sprintf('Creating: %s (Levels L: %0.4g R: %0.4g)', ...
        outfileName, mean(abs(left)), mean(abs(right))))

    % Write the waveform to the wav file.
    wavwrite([left right], Fs_Output, Nbit_Output, outfileName);
end

% Clean up the temporary storage files if nessecary.
if loops > 1
    for i=1:8
        fclose(fid(i));
        outfile = sprintf(tmpFileName, i);
        disp(sprintf('Deleting: %s', outfile));
        delete(outfile);
    end
end

% End timer and report on elapsed time,
s = etime(clock, t0);
m = floor(s/60); s = s-(60*m);
disp(sprintf(''))
disp(sprintf('Elapsed Time %d:%0.4g', m, s));

m = floor(ts/60); ts = ts-(60*m);
disp(sprintf('Synthesis Time %d:%0.4g', m, ts));

m = floor(tg/60); tg = tg-(60*m);
disp(sprintf('Generation Time %d:%0.4g', m, tg));

disp(sprintf('Average Comprison %0.3g', compress));

function [alpha, I, compress] = generateAlpha(Y, k, Ypt)

% generateAlpha determines the sound field (alpha) coefficients for a
% sound field with a point source (described by the sound sample 'y')
% at the location 'ypt'. As this function transforms data from the time
% domain to the frequency domain, any frequencies with power below
% 'power.threshold' are removed. The index of the frequencies maintained
% are returned in 'I'.

global M power.threshold

% Get the Fourier transform of the signal and remove the negative
% frequencies (don't require them as the input signal is real)
Y = fft(y);
Y = Y(1:(0.5*size(Y,1)+1));

% Find the dominant frequencies above the power threshold.
Pyy = Y.* conj(Y);
I = find(Pyy >= power.threshold);
Y = Y(I); k = k(I);

% Work out the compression rate
compress = 1-size(I,1) / size(Pyy,1);

% Define variables to simplify the following expressions
x = Ypt(1); phi = Ypt(2);

% Calculate the sound field coefficient from the pure
% point source
m = (-M:M);
k_part = repmat(i .* pi .* k .* 2.5 ,1,size(m,2));

```

```

m_part = repmat(exp(-i*m*pi), size(k,1),1);
alpha = k_part .* besseli(m, k*r) .* m_part ;

% Weight the point source coefficients with the sound sample 'Y'.
Y = repmat(Y, 1, 2*M+1);
alpha = Y .* alpha;

function P_inv = loudspeakermatrix(k)

% Generate the inverse of the 'P' matrix. The 'P' matrix is dependant
% on the sound field frequency and the loudspeaker positions.

global M L r l

% Know the speakers are evenly distributed around a circle, therefore
% their angles can be easily calculated.
spkr_angles = (0:L-1).*(2*pi/L);
spkr_angles = -spkr_angles+pi;

% Obtain memory for storing the result.
P_inv = zeros(size(k,1), L, 2*M+1);

% Values not dependant on frequency.
m = (-M:M)';
P_spkr = exp(-i*m*spkr_angles);

% Calculate of the 'P' matrix as given in Equation 3.34
for kidr=1:size(k,1)
    P = repmat(B(m, k(kidr)*r), 1, L) .* P_spkr;
    P(find(isnan(P)==1))=0;
    P_inv(kidr, :, :) = pinv(P);
end

function waveform = generateLoudspeaker(alpha, samples, w, ypt)

% GenerateLoudSpeaker transforms the sound field described by the sound
% field coefficients (alpha) into 'L' loudspeaker waveforms. These
% waveforms can then be amplified through loudspeakers to reproduce
% the sound field.

global M L P_inv

% Get some memory to store the result.
loudspeaker_weights = zeros(L, samples/2+1);

% For each frequency dependent set of coefficients calculate the
m_part = repmat(exp(-i*m*pi), size(k,1),1);
alpha = k_part .* besseli(m, k*r) .* m_part ;

% Weight the point source coefficients with the sound sample 'Y'.
Y = repmat(Y, 1, 2*M+1);
alpha = Y .* alpha;

function P_inv = loudspeakermatrix(k)

% Generate the inverse of the 'P' matrix. The 'P' matrix is dependant
% on the sound field frequency and the loudspeaker positions.

global M L r l

% Know the speakers are evenly distributed around a circle, therefore
% their angles can be easily calculated.
spkr_angles = (0:L-1).*(2*pi/L);
spkr_angles = -spkr_angles+pi;

% Obtain memory for storing the result.
P_inv = zeros(size(k,1), L, 2*M+1);

% Values not dependant on frequency.
m = (-M:M)';
P_spkr = exp(-i*m*spkr_angles);

% Calculate of the 'P' matrix as given in Equation 3.34
for kidr=1:size(k,1)
    P = repmat(B(m, k(kidr)*r), 1, L) .* P_spkr;
    P(find(isnan(P)==1))=0;
    P_inv(kidr, :, :) = pinv(P);
end

function waveform = generateLoudspeaker(alpha, samples, w, ypt)

% GenerateLoudSpeaker transforms the sound field described by the sound
% field coefficients (alpha) into 'L' loudspeaker waveforms. These
% waveforms can then be amplified through loudspeakers to reproduce
% the sound field.

global M L P_inv

% Get some memory to store the result.
loudspeaker_weights = zeros(L, samples/2+1);

% For each frequency dependent set of coefficients calculate the
loudspeaker_weights = zeros(L, samples/2+1);

% Remove any NaN's caused by 'divide by zero' errors
loudspeaker_weights(find(isnan(loudspeaker_weights)==1)) = 0;

% Recreate the negative frequencies (so we get a real output signal)
X = conj(fliplr(loudspeaker_weights));
loudspeaker_weights = [loudspeaker_weights X(:,2:(samples/2))];

% Calculate the inverse Fourier transform of the loudspeaker
% 'frequency' signals. Need to extract the real component so that
% small imaginary parts are discarded.
waveform = real(real(ifft(loudspeaker_weights, [], 2)));

% Calculate the position of the data so that the windowing function
% can be applied. The values used to calculate the position where
% obtained experimentally.
pos = ceil(261*(ypt(1)-1.6)+424);
s = pos-256;
e = samples-pos-256;
wq = [zeros(1,s) w' zeros(1,e)];

% Apply the window function. The fliplr function is needed as it
% appears the transformation process reverses the data and makes
% the audio sound distorted.
waveform = waveform .* repmat(wq,8,1);
waveform = fliplr(waveform);

```

Summary of Content on Attached Media

Two compact disks (CDs) are provided with this Thesis. These CDs contain an electronic version of this document, the MatLab source code generated during the execution of the project, configuration files for the 'Scope' software and the sound files used for the demonstrations. The organisation of the data stored on the CDs is shown in Figure E.1 below.

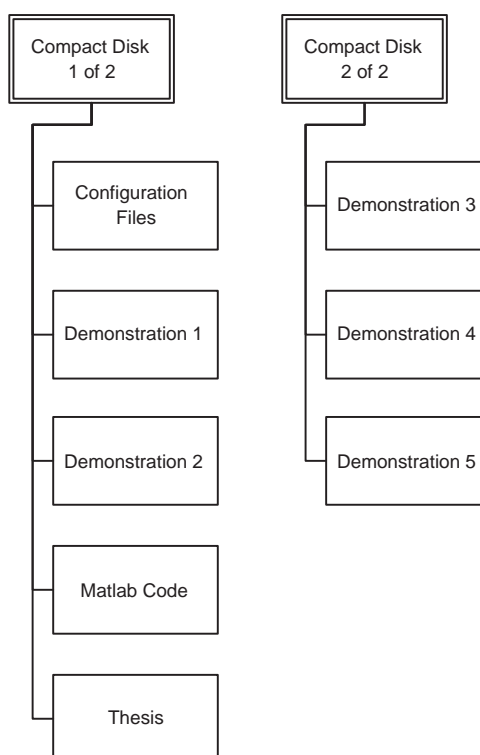


Figure E.1: Contents of the attached compact disks.

Bibliography

- [1] E. Skudrzyk, *The Foundations of Acoustics*. New York: Springer – Verlag, 1971.
- [2] Chris Kyriakakia, Panagiotis Tsakalides and Tomlinson Holman, “Surrounded by Sound,” *IEEE Signal Processing Magazine*, pp. 55–66, January 1999.
- [3] Darren B. Ward and Thushara D. Abhayapala, “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, September 2001.
- [4] Brian Bowers, “Scanning Our Past from London: Bell and the Telephone – the 125th Anniversary,” *Proceedings of the IEEE*, vol. 89, no. 6, pp. 984–986, June 2001.
- [5] J. Hull, “Surround Sound Past, Present and Future,” Dolby Laboratories Inc, Tech. Rep., 1999.
- [6] G. Dickins, “Soundfield representation, reconstruction and perception,” Master’s thesis, Department of Telecommunications Engineering, Research School of Information Science and Engineering, March 2003.
- [7] M. Gerzon, “With-height sound reproduction,” *J. Audio Eng. Soc.*, vol. 21, no. 1, pp. 2–10, Jan/Feb 1972.
- [8] M. J. Leese. (1998) Ambisonic Surround Sound FAQ. [Online]. Available: http://members.tripod.com/martin.leese/Ambisonic/faq_latest.html
- [9] J. S. Bamford, “An analysis of Ambisonic sound systems of first and second order,” Master’s thesis, University of Waterloo, 1995.
- [10] R. H. A. Sontacchi, “Enhanced 3D sound field synthesis and reproduction system by compensating interfering reflections,” in *Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-00)*, December 2000.
- [11] A. J. Berkhout D. de Vries and J. J. Sonke, “Array technology for acoustic wave field analysis in enclosures,” *J. Acoust. Soc. Am.*, vol. 102, no. 5, pp. 2757–2770, November 1997.
- [12] Diemer de Vries and Jan Baan, “Auralization of sound fields by wave field synthesis,” Laboratory of Acoustic Imaging and Sound Control, Delft University of Technology, Tech. Rep., 1999.

-
- [13] R. A. Kennedy, T. D. Abhayapala, and D. B. Ward, "Broadband nearfield beamforming using a radial beampattern transformation," *IEEE Transactions on Signal Processing*, vol. 46, no. 8, pp. 2147–2156, August 1998.
- [14] R. Green, "Spherical Harmonic Lighting: The Gritty Details," Sony Computer Entertainment Americas, Tech. Rep., January 2003.
- [15] Dolby Laboratories, "The Evolution of Dolby Film Sound," Dolby Laboratories, Tech. Rep.
- [16] R. Loyola. (2002, January) DVD audio overview. [Online]. Available: <http://www.techtv.com/screensavers/howto/story/0,24330,2334015,00.html>
- [17] Anders Torger, "Performance space recreation," *SurSound mailing list*, 16 January 2003.
- [18] NXT. (2002) NXT Technology Review. [Online]. Available: <http://www.nxtsound.com>
- [19] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, 2nd ed. New York: Springer, 1998.
- [20] Thushara D. Abhayapala, "Modal analysis and synthesis of broadband nearfield beamforming arrays," Ph.D. dissertation, Australian National University, 1999.
- [21] G. Baldcock and T. Bridgeman, *The Mathematical Theory of Wave Motion*. Ellis Horwood Limited, 1981.
- [22] E. G. Williams, *Fourier Acoustics*. Academic Press, 1999.
- [23] T. MacRobert, *Spherical Harmonics*, 3rd ed. Pergamon Press, 1967.
- [24] P. Nelson and S. Elliot, *Active Control of Sound*. New York: Academic Press, 1992.
- [25] Gene H. Golub and Charles F. Van Loan, *Matrix Computations*, 3rd ed. John Hopkins University Press, 1996.
- [26] N. Sloane and R. Hardin, and W.D. Smith et al.. Tables of spherical codes. [Online]. Available: <http://www.research.att.com/~njas/packings>
- [27] Drexel University. (2003) Math FAQ: Sphere Formulas. [Online]. Available: <http://mathforum.org/dr.math/faq/formulas/faq.sphere.html>
- [28] J. Vanderkooy and S. P. Lipshitz, "Anomalies of wavefront reconstruction in stereo and surround-sound reconstruction," in *Proceedings of the 83rd Convention of the Audio Engineering Society*, October 1987, preprint 2554.
- [29] T. Abhayapala and D. Ward, "Theory and design of high order sound field microphone using spherical microphone array," in *Proc. IEEE Int. Conf. Acoust. Speech Sig. Process. ICASSP02*, May 2002.

-
- [30] M. Chan, "Theory and design of higher order sound field recording," Department of Engineering, FEIT, ANU," Honours Thesis, June 2003.
- [31] L. W. Couch, *Digital and Analog Communication Systems*, 6th ed. Prentice Hall, 2001.
- [32] Fredric J. Harris, "On the use of Windows for Harmonic Analysis with the Discrete Fourier Transform," *Proceedings of the IEEE*, vol. 66, no. 1, pp. 51–83, Jan 1978.
- [33] A. T. S. Committee, "ATSC Standard: Digital audio compression (AC-3), Revision A," Tech. Rep., August 2001.
- [34] A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*, international ed. United States of America: Prentice Hall, 1989.
- [35] Darren B. Ward, "On the performance of acoustic cross talk cancellation in a reverberant environment," *Journal of the Acoustic Society of America*, vol. 110, no. 2, pp. 1195–1198, August 2001.
- [36] Paul Lyneham (Read by Max Gillies and Heather Steen), *Paul Lyneham: A Memoir [Audio CD]*. Australian Broadcasting Corporation, 2002.
- [37] Peter Carey (Read by Paul English and Robert Grubb), *True History of the Kelly Gang [Audio CD]*. Australian Broadcasting Corporation, 2001.
- [38] Australian Broadcasting Corporation. (2002, May) noise. [Online]. Available: <http://www.abc.net.au/noise/>