



2009-10-06

# Development and Comparison of Methods for Measuring Directional Sound Arrivals in Rooms

Brian Trevor Thornock  
*Brigham Young University - Provo*

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>

 Part of the [Astrophysics and Astronomy Commons](#), and the [Physics Commons](#)

---

## BYU ScholarsArchive Citation

Thornock, Brian Trevor, "Development and Comparison of Methods for Measuring Directional Sound Arrivals in Rooms" (2009). *All Theses and Dissertations*. 1925.

<https://scholarsarchive.byu.edu/etd/1925>

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact [scholarsarchive@byu.edu](mailto:scholarsarchive@byu.edu), [ellen\\_amatangelo@byu.edu](mailto:ellen_amatangelo@byu.edu).

DEVELOPMENT AND COMPARISON OF METHODS FOR THE MEASUREMENT  
OF DIRECTIONAL SOUND ARRIVALS IN ROOMS

by

Brian T. Thornock

A thesis submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Physics and Astronomy

Brigham Young University

December 2009



Copyright © 2009 Brian T. Thornock

All Rights Reserved



BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Brian T. Thornock

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

\_\_\_\_\_  
Date

\_\_\_\_\_  
Timothy W. Leishman, Chair

\_\_\_\_\_  
Date

\_\_\_\_\_  
Kent L. Gee

\_\_\_\_\_  
Date

\_\_\_\_\_  
Brian D. Jeffs



BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the thesis of Brian T. Thornock in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

---

Date

---

Timothy W. Leishman  
Chair, Graduate Committee

Accepted for the Department

---

Ross L. Spencer, Chair  
Department of Physics and Astronomy

Accepted for the College

---

Thomas W. Sederberg, Associate Dean  
College of Physical and Mathematical Sciences





## ABSTRACT

# DEVELOPMENT AND COMPARISON OF METHODS FOR THE MEASUREMENT OF DIRECTIONAL SOUND ARRIVALS IN ROOMS

Brian T. Thornock

Department of Physics and Astronomy

Master of Science

In room acoustics, the directional information of sound arrivals at a listening location can be used to diagnose the origins of problematic reflections so offending surfaces or other features can be properly treated. It can also be used for other purposes, including the study of psychoacoustic indicators. Many methods have been developed in the past to derive directional information, but despite their benefits, each has had significant drawbacks that have necessitated further research into their properties and development of an improved method. This thesis presents a review of past methods, their benefits and shortcomings. It discusses many theoretical and practical issues pertaining to the Polar ETC method and methods using the cross-correlation function. It also presents a new short-time correlation-based method (STCM) for gathering directional information of sound arrivals in rooms. Computer programs were developed for the implementation of the theory. Numerical simulations



and experimental measurements are shown and the results are compared to those obtained by the Polar Energy Time Curve (ETC) method. The STCM is shown to be an improvement over past methods in terms of its ability to distinguish between simultaneous arrivals, its accuracy, its computational efficiency and its equipment requirements. Limitations of the method are also discussed.



## ACKNOWLEDGMENTS

I would like to thank Dr. Tim Leishman for his guidance and many insightful comments and questions. He has been my mentor for several years and without his encouragement and support, I would not have arrived at this point.

I would also like to thank Drs. Gee and Jeffs for their expertise in their respective fields and for their time and effort to support me in this project.

Many thanks to James Esplin for his hard work and insightful comments and questions. Many of the items discussed in the practical considerations chapter were investigated by him.

My thanks also go to Buye Xu for his superior intellectual skills and everyone else with whom I have shared an office for allowing me to intrude to ask random questions and for facilitating the thinking process.

I would especially like to thank my wife for her support and patience with me while I executed this project. Thank you, Rosie, for all you have done. I could not have done this without you there.



# Contents

<b>Table of Contents</b>	<b>xv</b>
<b>List of Figures</b>	<b>xix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Previous Work . . . . .	2
1.2 Motivations for Research . . . . .	4
1.3 Objectives and Scope of Work . . . . .	4
1.4 Plan of Development . . . . .	5
<b>2 Past Methods</b>	<b>7</b>
2.1 Warble Tone Detection . . . . .	8
2.2 Polar ETC . . . . .	9
2.3 Correlation-Based Cartesian Technique . . . . .	27
2.4 Tetrahedral Array . . . . .	28
2.5 Modified Tetrahedral Array . . . . .	29
2.6 Cross Correlation in a Steady-State Sound Field . . . . .	31
2.7 Beamforming Techniques . . . . .	32
2.8 Intensity Measurements . . . . .	33
2.9 Ambisonics and Spherical Harmonic Decomposition . . . . .	33
2.10 Chapter Conclusions . . . . .	34
<b>3 Theoretical Developments</b>	<b>37</b>
3.1 Preliminary Considerations . . . . .	37
3.2 Generalized System Setup . . . . .	39
3.3 Time Delay Space . . . . .	42
3.4 The Cross Correlation Function . . . . .	44
3.5 Special Considerations for the Cross Correlation Function . . . . .	46
3.6 Alternate Cross Correlation Pairs . . . . .	48
3.7 Procedure for Directional Impulse Responses . . . . .	49
3.8 Chapter Summary . . . . .	50



<b>4</b>	<b>Practical Issues</b>	<b>51</b>
4.1	Introduction . . . . .	51
4.2	Sampling Rate and Accuracy . . . . .	51
4.3	Filtering Effects of Equipment . . . . .	52
4.4	Impulse Response Generation Programs . . . . .	54
4.4.1	Discrete Sources in Free Space . . . . .	55
4.4.2	Enhanced Image Source Method . . . . .	55
4.4.3	Modified Modal Expansion . . . . .	55
4.5	The Directional Impulse Response Program . . . . .	65
4.5.1	Time Delays and Allowable Error . . . . .	66
4.5.2	Selection of Peaks in the Cross Correlation . . . . .	67
4.5.3	Relative Amplitude and Dominant Reflections . . . . .	69
4.5.4	Display of the Results . . . . .	70
4.6	Matching of Microphones . . . . .	71
4.6.1	Phase Matching . . . . .	71
4.6.2	Magnitude Matching . . . . .	74
4.7	Array Dimensions and Geometry . . . . .	74
4.7.1	Array Dimensions . . . . .	75
4.7.2	Array Geometry . . . . .	78
4.8	Microphone Directivity . . . . .	79
4.9	Subsequent versus Simultaneous Measurements . . . . .	80
4.10	Maximum Usable Time Length . . . . .	81
4.10.1	Acoustic Arrival Density . . . . .	83
4.10.2	Empirical Results . . . . .	84
4.10.3	Strong Late Arrivals and Other Exceptions . . . . .	87
4.11	Chapter Conclusions . . . . .	88
<b>5</b>	<b>Numerical Verification, Experiments, and Results</b>	<b>91</b>
5.1	A Single Source in Free Space . . . . .	92
5.2	Two Asymmetrically-Placed Sources in Free Space . . . . .	93
5.3	Two Symmetric Sources in Free Space . . . . .	94
5.4	Modeled Rectangular Room with Uniform Absorption . . . . .	97
5.5	Modeled Rectangular Room with Nonuniform Absorption . . . . .	101
5.6	Low-Pass Filtered Room Models . . . . .	104
5.7	Experimental Results . . . . .	108
5.7.1	Margin of Error . . . . .	109
5.7.2	Single Reflector Experiment . . . . .	110
5.7.3	Single Small Reflector Experiment . . . . .	111
5.7.4	Dual Symmetric Reflector Experiment . . . . .	111
5.7.5	Dual Asymmetric Reflector Experiment . . . . .	113
5.7.6	Variable Acoustics Chamber Experiment . . . . .	114
5.7.7	Summary of Results . . . . .	115
5.7.8	Off-Axis Experiments . . . . .	117

---

5.8	Discussion . . . . .	119
5.8.1	EASERA . . . . .	119
5.8.2	Erroneous Solutions and Cross Correlation Peak Selection . . .	120
5.8.3	High-Frequency Content . . . . .	120
5.8.4	Frequency Magnitude Compensation . . . . .	121
5.8.5	Noise Gating . . . . .	124
5.9	Chapter Conclusions . . . . .	126
<b>6</b>	<b>Comparison of Methods</b>	<b>131</b>
6.1	Polar ETC . . . . .	131
6.1.1	Polar ETC Program . . . . .	131
6.2	Comparison Results . . . . .	133
6.2.1	Numerical Comparisons . . . . .	134
6.2.2	Experimental Comparisons . . . . .	135
6.3	Discussion . . . . .	141
6.4	Chapter Conclusions . . . . .	141
<b>7</b>	<b>Conclusions</b>	<b>143</b>
7.1	Summary of Findings . . . . .	144
7.2	Contributions . . . . .	145
7.3	Future Work . . . . .	146
	<b>Bibliography</b>	<b>148</b>
<b>A</b>		<b>155</b>
A.1	Sources in Free Space IR Generation . . . . .	155
A.2	Allen-Berkley IR Generation . . . . .	157
A.3	Modified Modal Expansion . . . . .	162
A.4	STCM Program . . . . .	166
A.5	Polar ETC Programs . . . . .	179
A.5.1	Cardioid IR Generation . . . . .	179
A.5.2	Subcardioid IR Generation . . . . .	185
A.5.3	Hypercardioid IR Generation . . . . .	191
A.5.4	Polar ETC Matlab Program . . . . .	196
<b>B</b>		<b>203</b>
B.1	Spherical Coordinate Rotations . . . . .	203
B.2	Rotation of a Cardioid Solid . . . . .	206



# List of Figures

2.1	Example of a source and receiver in a simple two-dimensional “room.”	11
2.2	IR for the two-dimensional “room” in Fig. 2.1.	12
2.3	IR and histogram of a numerically-simulated long, narrow room.	19
2.4	Two-dimensional polar pattern for an ideal cardioid.	20
2.5	Two-dimensional directivity plots for (a) an ideal subcardioid pattern and (b) an ideal hypercardioid pattern.	21
2.6	Two-dimensional polar pattern for a baffled piston in the far field for $ka = 10$ .	24
2.7	Plot of the modifying factor for a baffled-piston directivity with (a) fixed $ka$ values and $\phi_A$ varying from 0 to $90^\circ$ and (b) $\phi_A = 20^\circ$ and $ka$ varying from 5 to 25.	26
2.8	Four-microphone Cartesian array used for a correlation-based technique.	28
2.9	Five-microphone tetrahedral array with microphones located at the apexes and at the center of gravity.	30
3.1	Cartesian microphone arrangements for (a) a seven-microphone array and (b) a four-microphone array where a single microphone is located at the origin and the others are located a distance $d$ away along the axes.	40
3.2	Physical setup for theoretical development including a seven-microphone Cartesian array and radiating point source in free space.	41
3.3	Graphical representation of a source (shaded circle) and two receivers in free space with different source-receiver path lengths for the demonstration of time delays between receivers.	41
3.4	Graphical representation of a sound arrival in time delay space.	43
3.5	Cross correlation in $x$ (top), $y$ (middle), and $z$ for a microphone array centered at (0 m, 0m, 0 m) and two sources located at (20 m, 4 m, 6 m) and (20 m, -4 m, 6 m).	47
4.1	Comparison of cross correlation along the $x$ -axis for two closely-located sources in free space when sampled at (a) $f_s = 48$ kHz and (b) $f_s = 192$ kHz.	53

4.2	Overlay of magnitude classical (solid) and modified modal expansions at (a) low frequencies and (b) high frequencies for an inclusion window of $20\Delta f$ . . . . .	59
4.3	Magnitude error between classical and modified modal expansions for differing inclusion window widths. . . . .	60
4.4	Spectral standard deviation of the magnitude error in percent versus inclusion window width. . . . .	60
4.5	Overlay of phase of classical (solid) and modified modal expansions at (a) low frequencies and (b) high frequencies for an inclusion window width of $20\Delta f$ . . . . .	61
4.6	Phase error between classical and modified modal expansions for differing inclusion window widths. . . . .	62
4.7	Spectral standard deviation of phase error in percent versus inclusion window width. . . . .	62
4.8	Overlay of IRs generated using classic (solid) and modified modal expansions early in the response. . . . .	63
4.9	IR error between classical and modified modal expansions for differing frequency window widths. . . . .	64
4.10	Standard deviation of IR Amplitude error in percent versus inclusion window width. . . . .	64
4.11	Overlay of IRs generated using modified modal response, method of images, and low-pass filtered method of images. . . . .	65
4.12	Definition of angles as computed by the DIR. . . . .	68
4.13	Scatter plot showing arrival angles for a long, narrow room for $t = 90$ ms. . . . .	71
4.14	Plot of maximum inherent angular error (degrees) and window width (samples) versus array diameter for $f_s=192$ kHz. . . . .	77
4.15	Cross correlation using simultaneous and subsequent IR measurements for (a) the full time-record length and (b) zoomed in to show detail . . . . .	82
4.16	Plot showing measured and predicted arrival densities for a rectangular room. . . . .	85
4.17	Impulse response for a numerically modeled room 30 x 8 x 4.5 m with nonuniform $\alpha$ . . . . .	85
4.18	Impulse response for a numerically modeled room 30 x 8 x 4.5 m with nonuniform $\alpha$ and an artificially-added, strong late arrival. . . . .	87
4.19	Cross correlation along the $x$ -axis for the case of (a) the normal IR and (b) the IR with an artificially added large spike at $t = 130$ ms . . . . .	89
5.1	Illustration of a single source in free space located near a Cartesian microphone array. . . . .	93
5.2	Impulse response for a single sound arrival at a measurement location. . . . .	94

5.3	Illustration of two asymmetrically-placed sources in free space located near a Cartesian microphone array for (a) the $x$ - $y$ plane and (b) the $x$ - $z$ plane. . . . .	95
5.4	Impulse response for two asymmetrically placed sources producing simultaneous sound arrivals at a measurement location. . . . .	96
5.5	Illustration of two symmetrically-placed sources in free space located near a Cartesian microphone array. . . . .	97
5.6	Impulse response for two symmetrically placed sources producing simultaneous sound arrivals at a measurement location. . . . .	98
5.7	Impulse response (top) and histogram (bottom) for a numerically modeled 10 x 8 x 7 m rectangular room with $\alpha = 0.1$ and arrows indicating peaks analyzed. . . . .	99
5.8	Impulse response (top) and histogram (bottom) for a numerically modeled room 30 x 8 x 4.5 m with nonuniform $\alpha$ and arrows indicating peaks analyzed. . . . .	102
5.9	Image source IR with a low-pass filtered version of it superposed. . .	104
5.10	Original cross correlation along the $x$ -axis with a low-pass filtered version of it superposed. . . . .	105
5.11	Two pieces of custom equipment used for the experimental setup; (a) microphone positioner and (b) altazimuth-mounted laser pointer. . .	109
5.12	Experimental setup used for a single reflector case in an anechoic chamber where dotted lines represent rays traced by reflections. . . . .	110
5.13	Experimental setup used for a single, small reflector case in an anechoic chamber where dotted lines represent rays traced by reflections. . . .	111
5.14	Experimental setup used for a dual symmetric reflector case in an anechoic chamber where dotted lines represent rays traced by reflections. . . . .	112
5.15	Experimental setup used for a dual asymmetric reflector case in an anechoic chamber where dotted lines represent rays traced by reflections. . . . .	113
5.16	Experimental setup used in a variable acoustics chamber where dotted lines represent rays traced by reflections. . . . .	115
5.17	Cross correlation function for simultaneous arrivals in the dual asymmetric reflector experiment with the loudspeaker oriented $180^\circ$ from its original position. . . . .	118
5.18	Smoothed frequency-response curve of a Mackie HR-824 loudspeaker used for frequency magnitude compensation. . . . .	122
5.19	Section of an IR (a) before and (b) after the frequency magnitude compensation process. . . . .	123
5.20	Example cross correlation function (a) before and (b) after the magnitude compensation process. . . . .	124
5.21	An IR shown (a) with a superposed noise gate threshold and (b) the modified IR after a noise gate has been applied. . . . .	125
5.22	Cross Correlation functions for the dual symmetric reflector experiment using original IRs. . . . .	127

5.23	Cross Correlation functions for the dual symmetric reflector experiment using noise gated IRs. . . . .	128
5.24	Cross Correlation functions for the dual symmetric reflector experiment using noise-gated and magnitude-compensated IRs. . . . .	129
B.1	A cardioid solid facing in the $+z$ direction with views of (a) the entire solid and (b) the solid cut in half to show its cardioid profile. . . . .	207
B.2	Cartesian coordinate system and angular definitions for (a) the original coordinate system and (b) the rotate coordinate system. . . . .	209
B.3	Cardioid solids facing in the (a) $+x$ direction and (b) $+y$ direction. . . . .	212

# Chapter 1

## Introduction

In the realm of room acoustics, one of the most common measurements is the impulse response (IR). The IR ideally describes how the room responds temporally at one location to a pressure impulse originating at another location, where both the source and receiver are omnidirectional. The measurement may be used to derive standard room acoustic parameters including, but not limited to, reverberation time, clarity, strength, definition, speech transmission index, etc. The sequence of reflections arriving at a receiving location after the direct sound arrival are apparent and, as evidenced by the parameters, are known to have a large impact on the subjective perception of the space [1–3].

Despite the usefulness of the common IR, it often fails to provide enough information about how sound arrives at the receiver location. For example, it may show the amplitude and delay of a strong late arrival corresponding to a perceived echo. However, it does not provide any information about the direction from which the arrival has come. Many methods for determining the direction of sound arrivals in rooms have been explored, but they have lacked in practicality, accuracy, and reliability. This thesis presents a new approach to directional IR receiving measurements



that overcomes these limitations and provides acousticians with a valuable new tool for diagnostics and troubleshooting in listening environments.

## 1.1 Previous Work

One of the earliest and most laborious methods of determining directional information of sound arrivals was to simply use the speed of sound and the arrival time values in an IR to determine overall path lengths traversed by a given ray. Then, using the geometry of the room, one would predict the most likely path that the ray traversed to arrive at the receiver location. As one could imagine, this method was often inaccurate and time consuming.

Another common method was to shield the omnidirectional receiving microphone from (mid-to-high frequency) arrivals of different directions using some type of portable barrier. Through multiple measurements and trial and error, an acoustician could ascertain the general direction of which a specific arrival came. Yet another common method used involved the use of highly directional sources and receiving microphones to explore paths of sound transmission and reflection.

One of the first major efforts to acquire directional measurements in rooms was pioneered by Thiele in 1953. He used a directional microphone at the focal point of a parabolic reflector oriented in different directions to detect warble tones in rooms [4]. Little additional work was done on measurement methods until 1987, when Becker introduced his Polar ETC method for determining directional arrivals [5, 6]. The method utilized a set of six Energy-Time Curve (ETC) measurements made with a cardioid microphone sequentially oriented along the Cartesian axes. One year later, Yamasaki and Itow developed a method using four microphones, one at the origin and one located a small positive distance away along each the Cartesian axes. With this

configuration, they implemented a basic correlation technique to extract the location of “virtual image sources” [7]. Following Yamasaki and Itow, Sekiguchi *et al.* used a similar method but employed an array with the microphones set at the vertices of a regular tetrahedron [8]. Choi *et al.* further modified the technique by placing a fifth microphone at the tetrahedron’s center of gravity [9].

In 2002, Gover introduced a more complicated method that utilized two different-sized spherical arrays of 32 microphones each, with a beamforming algorithm to determine the directions of sound arrivals [10–12]. This method has not been used for measuring arrivals of reflections directly, but has been used for identifying areas of high transmission in panels and walls and might be of some use in determining directions of arrivals. Another recent method that merits mention is one that utilizes special microphone arrays, sometimes called Soundfield<sup>®</sup> or ambisonic microphones. These arrays use decoding algorithms to allow the user to effectively change microphone placement and polar pattern after the fact. This method has been used by both Essert and Farina and involve processing the signals such that an intensity measurement is extracted and directions of arrival are derived from that [13–17].

Other work by Noël *et al.* has used a cross-correlation technique with alternative array arrangements for finding directional sound arrivals from multiple primary sources that radiate continuously in a reverberant environment [18]. Abdou and Guy also used intensity measurements to characterize the directional aspects of a room response [19]. Each of the methods mentioned here will be discussed in greater detail in Chapter 2.

## 1.2 Motivations for Research

The directional information of sound in rooms is far more useful than for just the identification of surfaces that produce unwanted reflections, though that is a major concern [20, 21]. Other uses for this directional information include the validation of diffuse field assumptions commonly used in room acoustics, identification of the other sound field characteristics, and determination of psychoacoustic indicators [22–25].

Many of the methods mentioned above can provide useful (though perhaps not complete) directional IR results, but the results either take too much time, require the use of overly expensive equipment, or are not always inaccurate. In some cases, to be discussed later, the results may be completely erroneous due to invalid assumptions.

## 1.3 Objectives and Scope of Work

The objectives of the research for this thesis were several-fold. The first was to develop a method for extracting complete directional information from a small three-dimensional array of microphones at a location of interest in a room. This method was to discriminate between all sequentially and simultaneously arriving wave fronts. In addition, the method needed to be accurate in its calculation of directions of sound arrivals and computationally efficient so as to be usable on even a modest computer system. Furthermore, it needed to be executable with a modest amount of equipment typical of that available to an acoustical consultant or dedicated consumer.

This research will review work previously done in the area of directional IRs. The review, however, may not be completely comprehensive. There may be methods of which the author is unaware or details about known methods of which he is unaware. Furthermore, the new method presented herein should not be construed to be the only possible solution to the problem.

## **1.4 Plan of Development**

The following chapter will present an in-depth look at past methods including their benefits and shortcomings. Chapter 3 will present a new theory for determining the complete directional information of a room response at a single point using a cross correlation technique. Chapter 4 will discuss the computer programs created to implement the theory introduced in Chapter 3 with special emphasis on their practical considerations. Chapter 5 will present the results of numerical room response simulations to verify the theory. In addition, Chapter 5 will present the results of experiments conducted to determine practical limitations and real-world accuracy. Chapter 6 will contain the results from comparisons of the new method to previous methods. Finally, Chapter 7 will summarize the results, draw conclusions, and discuss any future work that may need to be done in the area.



# Chapter 2

## Past Methods

As mentioned in Ch. 1, many methods have been developed to extract the directional information of sound arrivals in rooms. The various methods have been introduced because their predecessors have been flawed in some way, rendering them inaccurate or invalid in certain situations. The specialized methods mentioned in Chapter 1 will be discussed in greater detail in this chapter, with their governing principles, fundamental assumptions, and their shortcomings. Where possible, enhancements will be suggested for some of the methods, though the author does not claim a comprehensive working knowledge of all the methods discussed herein and thus is not able to make adequate suggestions for optimizing every technique.

The purpose here is to identify the shortcomings of the methods to properly introduce the reader to the field and explain why there is still a need for an improved method.

## 2.1 Warble Tone Detection

One of the first specialized methods for taking directional measurements was implemented by Thiele in 1953 [4]. A parabolic reflector was attached to a cardioid microphone with the opening of the reflector facing the microphone diaphragm. The rear of the microphone (and thus the opening of the reflector) was aimed in different directions in an attempt to detect warble tones in rooms. This method was successful at some frequencies, but upon consideration of the physics and design principles of the equipment used, it is obvious that it would not be able to adequately function over a broad frequency range.

The first problem resides in the cardioid microphone. The difference between a cardioid microphone and a typical pressure microphone is that there is an opening in the microphone capsule allowing for the introduction of sound pressure on the back of the diaphragm as well as on the front, creating a pressure gradient across the diaphragm. This opening is designed to produce an ideal cardioid pattern, but only at a single design frequency. As soon as the frequency shifts away from this design frequency, the directivity of the microphone becomes nonideal. At low frequencies, the directivity often resembles either a subcardioid pattern, or an omnidirectional pattern with the back slightly flattened. At very low frequencies, the response actually becomes omnidirectional due to the diffraction effects of very long wavelengths. At high frequencies, the response becomes that of a hypercardioid, which no longer has a response null at 180 degrees off axis. This being the case, any broadband excitation signal of the room (and therefore the microphone) would result in a microphone response that would pick up acoustic signals from the rear (facing into the room) as well as from the front (facing into the reflector).

The second problem has to do with the parabolic reflector itself. It can be thought

of as a concave acoustic mirror, obeying the same basic laws as a concave optical mirror. For any given parabola, a set focal point and the strength of focusing at that focal point depend on the distance from the parabola to the source. For a room, the distance from the surfaces (sources) will not always be equal, so the focusing effect of the parabola will not always be the same. In addition, if one changes the frequency of the sound incident on the parabola, its reflection effects can also be altered. When a parabolic reflector becomes on the order of or small relative to a wavelength of incident sound scattering and diffraction effects become more important, altering the ideal focusing at the microphone. The directivity of the arrangement is again frequency-dependent.

Furthermore, detection of warble tones does not necessarily correspond to the more useful detection of the IR and is therefore not a very practical measure in the context of measuring directional sound arrivals over time. This method was largely abandoned because of the obstacles it poses.

## 2.2 Polar ETC

One of the most widely used methods for determining directional sound arrivals in rooms is the Polar ETC [5, 6]. It is based on the concept of the energy time curve (ETC) developed by Heyser [26–28]. The ETC was originally thought to show the total energy density at a specific location as a function of time. It is computed by combining the original IR with a Hilbert-transformed version of the IR, the imaginary part of the complex time signal (which also functions to approximate the kinetic energy) and then the total is expressed as a level in dB. The Polar ETC measurement is carried out at a specified in the room by sequentially orienting a cardioid microphone along each of the Cartesian axes in the positive and negative directions. The ETCs

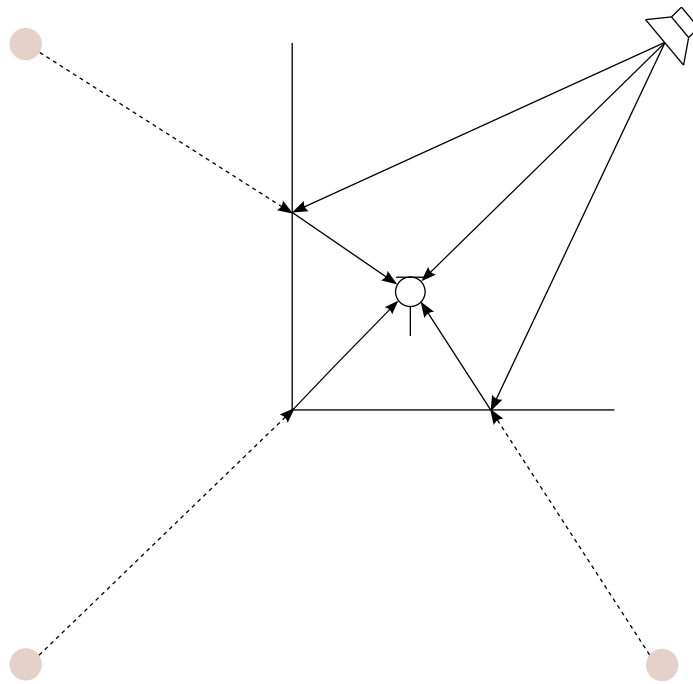


are then processed using the principle of direction cosines and normalized using the supposed “total energy density” to extract out the spherical angles of arrival relative to the Cartesian axes chosen for the measurement. It is important to note that the term “total energy density” as used by D’Antonio *et al.* [5] is not equal to the total energy density as typically defined in acoustics.

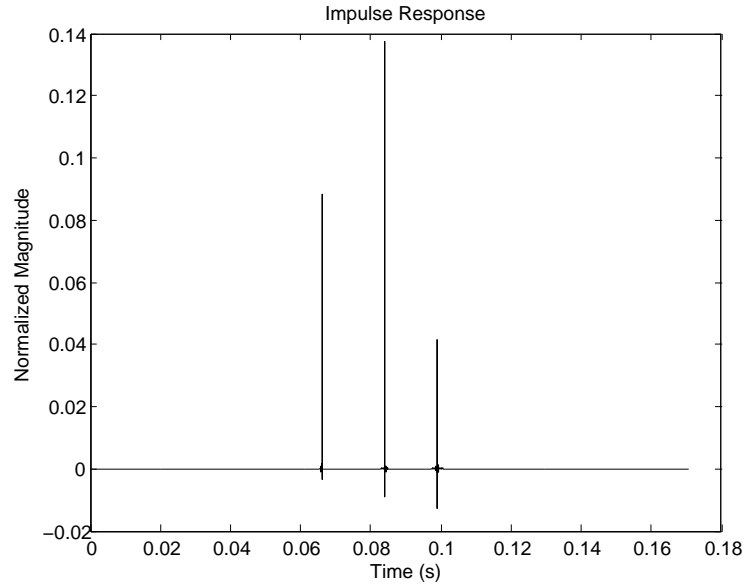
### Simultaneous Arrivals

The most significant problem with the Polar ETC is the fundamental assumption that only one sound arrival reaches the measurement position at any one time. To illustrate why this is such a problem, consider the simple case of a two dimensional “room” with perfectly reflecting walls (see Fig. 2.1). The source and receiver lie along the diagonal from the corner and the image sources due to the two walls are shown as shaded circles. The IR would look like that shown in Fig. 2.2. The IR was computed by placing discrete sources in free space using the method of images to simulate the room shown. The direct sound arrival is followed by a high peak (stronger than the direct arrival) caused by simultaneously arriving reflections from the side walls. This peak is followed by a final smaller peak due to reflection from the corner.

With the assumption of only a single arrival at any given time, the Polar ETC would generate a direction of arrival of either 0 or 180 degrees (facing directly toward or away from the primary source) for the peak involving two simultaneously arriving reflections. To see why this is the case, consider an ideal cardioid oriented along the two Cartesian axes for the room in Fig. 2.1, in both the positive and negative directions. According to the Polar ETC method, the measured quantity  $E$  is called the “energy density.” This terminology is confusing and incorrect. For the sake of correctness, the quantity  $E$  discussed here is the squared magnitude. The squared



**Figure 2.1** Example of a source and receiver in a simple two-dimensional “room.”



**Figure 2.2** IR for the two-dimensional “room” in Fig. 2.1.

magnitudes for each microphone orientation as defined by the Polar ETC are

$$E_F = E[1 + \cos(\phi)], \quad (2.1)$$

$$E_B = E[1 - \cos(\phi)], \quad (2.2)$$

$$E_L = E[1 + \sin(\phi)], \quad (2.3)$$

$$E_R = E[1 - \sin(\phi)], \quad (2.4)$$

In the above equations, the subscripts indicate the direction in which the microphone is pointing ( $F$  for front,  $B$  for back,  $L$  for left, and  $R$  for right). The quantity  $E_F$  is then the squared magnitude for the microphone facing in the front direction, and so

on. If one assumes  $N$  arrivals within a given sample, the various squared magnitudes can then be represented in an expanded form as

$$E_F = E_1[1 + \cos(\phi_1)] + E_2[1 + \cos(\phi_2)] + \dots + E_N[1 + \cos(\phi_N)], \quad (2.5)$$

$$E_B = E_1[1 - \cos(\phi_1)] + E_2[1 - \cos(\phi_2)] + \dots + E_N[1 - \cos(\phi_N)], \quad (2.6)$$

$$E_L = E_1[1 + \sin(\phi_1)] + E_2[1 + \sin(\phi_2)] + \dots + E_N[1 + \sin(\phi_N)], \quad (2.7)$$

$$E_R = E_1[1 - \sin(\phi_1)] + E_2[1 - \sin(\phi_2)] + \dots + E_N[1 - \sin(\phi_N)], \quad (2.8)$$

For the case in Fig. 2.1 there are only two simultaneous arrivals. The equations above then become

$$E_F = E_1[1 + \cos(\phi_1)] + E_2[1 + \cos(\phi_2)], \quad (2.9)$$

$$E_B = E_1[1 - \cos(\phi_1)] + E_2[1 - \cos(\phi_2)], \quad (2.10)$$

$$E_L = E_1[1 + \sin(\phi_1)] + E_2[1 + \sin(\phi_2)], \quad (2.11)$$

$$E_R = E_1[1 - \sin(\phi_1)] + E_2[1 - \sin(\phi_2)]. \quad (2.12)$$

The formulation for the angle of arrival  $\phi$  is derived from these expressions by subtracting the  $-y$ -axis (right) squared magnitude from the  $+y$ -axis squared magnitude (left) to get the  $y$ -axis component and subtracting the  $-x$ -axis (back) squared

magnitude from the  $+x$ -axis squared magnitude to get the  $x$ -axis component. The angle is calculated as the inverse tangent of their ratio. The result is

$$\phi = \tan^{-1} \left( \frac{E_L - E_R}{E_F - E_B} \right). \quad (2.13)$$

For the case at hand, the two arrivals are symmetric about the line from source to receiver, i.e.,  $\phi_1 = -\phi_2$ . We can then reduce the equations for the squared magnitudes to the following forms:

$$E_F = E[2 + 2\cos(\phi)], \quad (2.14)$$

$$E_B = E[2 - 2\cos(\phi)], \quad (2.15)$$

$$E_L = E(2), \quad (2.16)$$

$$E_R = E(2). \quad (2.17)$$

Substitution of these values into Eq. (2.13) yields

$$\phi = \tan^{-1} \left( \frac{0}{4E\cos(\phi)} \right). \quad (2.18)$$

As seen,  $\phi$  in Eq. (2.18) could be either 0 or 180. Determination of which value is correct is done by examining the sign of the denominator.

As a more generalized result of the above formulation, for any number of arrivals within one sample of each other in a digitized signal, the method will produce a direction of arrival for some intermediate location. This represents a single equivalent

source that could be used to produce the same ETC result in the absence of the various reflections that actually generate a single peak.

As a proof that the Polar ETC does produce directions of arrival that for a single equivalent source, we may adapt the discussion above to a three-dimensional case. The six different squared magnitudes become

$$E_F = E[1 + \cos(\phi)\cos(\theta)], \quad (2.19)$$

$$E_B = E[1 - \cos(\phi)\cos(\theta)], \quad (2.20)$$

$$E_L = E[1 + \sin(\phi)\cos(\theta)], \quad (2.21)$$

$$E_R = E[1 - \sin(\phi)\cos(\theta)], \quad (2.22)$$

$$E_U = E[1 + \sin(\theta)], \quad (2.23)$$

$$E_D = E[1 - \sin(\theta)], \quad (2.24)$$

where  $\phi$  is the azimuthal angle and  $\theta$  is the elevation from the horizontal plane and the  $U$  and  $D$  subscripts denote the up and down directions, respectively. For  $N$  simultaneous arrivals, these quantities become

$$E_F = E_1[1 + \cos(\phi_1)\cos(\theta_1)] + E_2[1 + \cos(\phi_2)\cos(\theta_2)] + \dots + E_N[1 + \cos(\phi_N)\cos(\theta_N)], \quad (2.25)$$

$$E_B = E_1[1 - \cos(\phi_1)\cos(\theta_1)] + E_2[1 - \cos(\phi_2)\cos(\theta_2)] + \dots + E_N[1 - \cos(\phi_N)\cos(\theta_N)], \quad (2.26)$$

$$E_L = E_1[1 + \sin(\phi_1)\cos(\theta_1)] + E_2[1 + \sin(\phi_2)\cos(\theta_2)] + \dots + E_N[1 + \sin(\phi_N)\cos(\theta_N)], \quad (2.27)$$

$$E_R = E_1[1 - \sin(\phi_1)\cos(\theta_1)] + E_2[1 - \sin(\phi_2)\cos(\theta_2)] + \dots + E_N[1 - \sin(\phi_N)\cos(\theta_N)], \quad (2.28)$$

$$E_U = E_1[1 + \sin(\theta_1)] + E_2[1 + \sin(\theta_2)] + \dots + E_N[1 + \sin(\theta_N)], \quad (2.29)$$

$$E_D = E_1[1 - \sin(\theta_1)] + E_2[1 - \sin(\theta_2)] + \dots + E_N[1 - \sin(\theta_N)], \quad (2.30)$$

The solutions for the angles of arrival  $\phi_A$  and  $\theta_A$  are

$$\phi_A = \tan^{-1} \left( \frac{E_L - E_R}{E_F - E_B} \right), \quad (2.31)$$

$$\theta_A = \sin^{-1} \left( \frac{E_U - E_D}{2E} \right). \quad (2.32)$$

If we then substitute the squared magnitude values for each direction into Eqs.( (2.31)) and( (2.32)), we obtain

$$\phi_A = \tan^{-1} \left( \frac{\sum_{n=1}^N E_n \sin(\phi_n)}{\sum_{n=1}^N E_n \cos(\phi_n)} \right) \quad (2.33)$$

$$\theta_A = \sin^{-1} \left( \frac{\sum_{n=1}^N E_n \sin(\theta_n)}{\sum_{n=1}^N E_n} \right) \quad (2.34)$$

For equistrength simultaneous arrivals, the equations reduce to

$$\phi_A = \tan^{-1} \left( \frac{\sum_{n=1}^N \sin(\phi_n)}{\sum_{n=1}^N \cos(\phi_n)} \right) \quad (2.35)$$

$$\theta_A = \sin^{-1} \left( \frac{\sum_{n=1}^N \sin(\theta_n)}{N} \right) \quad (2.36)$$

In the case of a dominant arrival along with weaker arrivals, the angles collapse to approximately the arrival angles for the dominant reflection.

Equations (2.34) and (2.36) show that, depending on the relative elevation angles of the arrivals, the resulting elevation angle will not always be at the mean value of the elevation angles. Consider, for example, the case with arrivals symmetric about the horizontal plane, i.e.,  $\theta_1 = -\theta_2$ . This will result in a summation value of 0, and therefore the resulting elevation angle will be  $0^\circ$ . If we were to change the elevation angles to, say,  $\theta_1 = 70^\circ$  and  $\theta_2 = -20^\circ$ , the resulting elevation angle will be  $\theta_A = 17.38^\circ$  as opposed to the mean value of  $25^\circ$ . This shift gets even more extreme as one angle approaches  $\pm 90^\circ$ . If the angles are  $\theta_1 = 90^\circ$  and  $\theta_2 = 0^\circ$ , then the resulting value is  $\theta_A = 30^\circ$ . Equations (2.33) and (2.35), however, show that the resulting angle  $\phi_A$  is simply the average of the angles  $\phi_1$  and  $\phi_2$ .

Consider the IR shown in Fig. 2.3. It shows the IR for a long, narrow room that is described in greater detail in Sec. 5.5. The ‘‘histogram’’ below the IR shows the

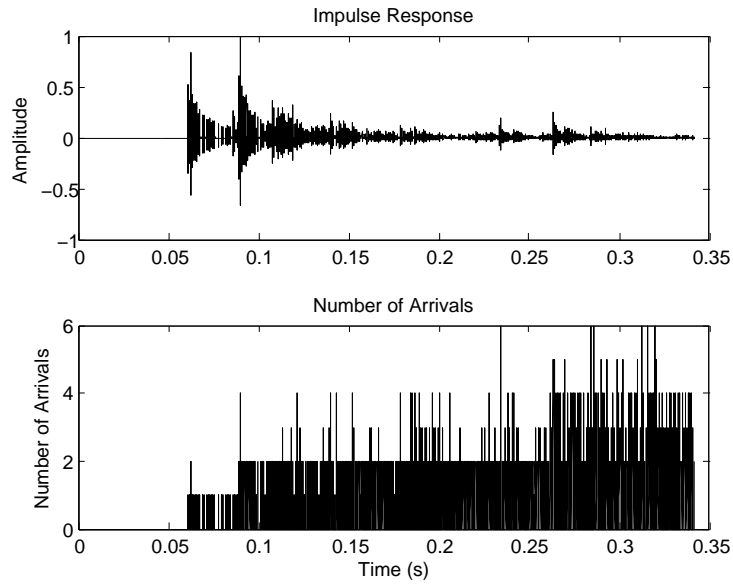


number of reflections present within one sample. Both the IR and histogram were created using the Berkley-Allen method of images adapted for use in Matlab [29]. The sampling frequency was 192 kHz, which corresponds to a sample interval (sampling period) of  $5.2 \mu\text{s}$ . Note how during the early part of the IR, the assumption of a single arrival within a given sample is not a bad assumption, but beginning at approximately 90 ms, the average number of reflections per sample is no longer 1 and therefore the assumption of a single arrival will result in erroneous results (the error of which is determined by number and relative strength of the simultaneous arrivals). These simultaneous arrivals can present serious problems. Chapter 6 some experimental results that show just how serious the discrepancies can be. One very important thing to note is that this phenomenon of multiple arrivals within a given sample is very common. Virtually every room will have acoustical path lengths that are nearly equal, and these will result in arrivals within a single sample. This problem only increases in severity as the time into the IR increases.

### Microphone Directivity

Another problem with this method lies in the use of a typical cardioid microphone. The derivation assumes that the directivity pattern of the microphone is a perfect cardioid over all frequencies (see Fig. 2.4). However, as discussed previously, this is clearly not the case; as frequency changes from the design frequency (e.g., 1 kHz) this fundamental assumption is violated. While it may seem that any deviation from a cardioid pattern will produce calculated values that are erroneous, consider the case of idealized hypercardioid and subcardioid microphone directivity patterns (see Fig. 2.5) [30].

To determine the arrival angles, the Polar ETC uses the “energy densities” measured along each Cartesian direction. The derivations in previous section can be



**Figure 2.3** IR and histogram of a numerically-simulated long, narrow room.

altered to use generalized directivity functions for the cardioid family. These functions are of the form

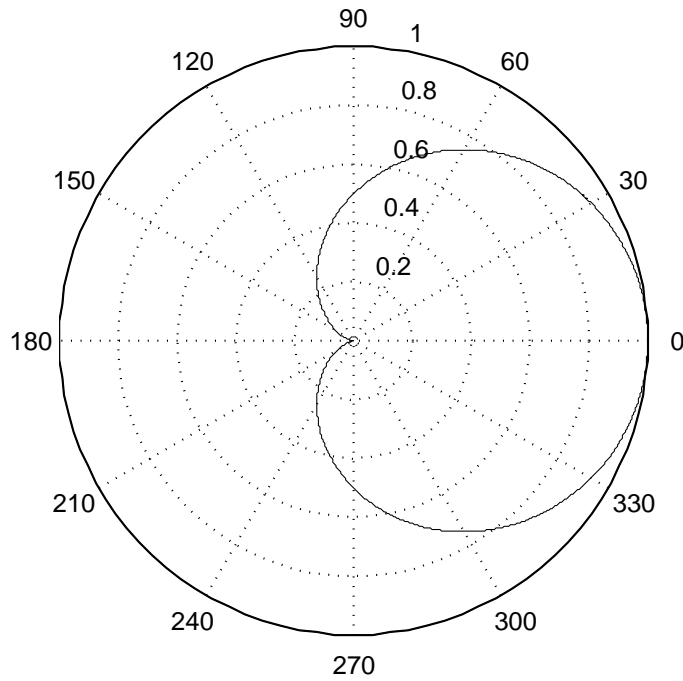
$$W = A + B\cos(\theta), \quad (2.37)$$

where the ratio of  $A$  to  $B$  is what determines the directivity. Table 2.1 shows what the  $A$  and  $B$  values are for some common directivity patterns [30].

If we define the  $+z$  direction to be the positive-facing direction and  $\theta$  to be the elevation angle from the  $x$ - $y$  plane, the polar pattern expression in Eq. (2.37) becomes

$$W_z = A + B\sin(\theta), \quad (2.38)$$

where the subscript denotes the direction in which the microphone is oriented. By



**Figure 2.4** Two-dimensional polar pattern for an ideal cardioid.

extension, the generalized polar pattern for a microphone facing in the  $-z$  direction is

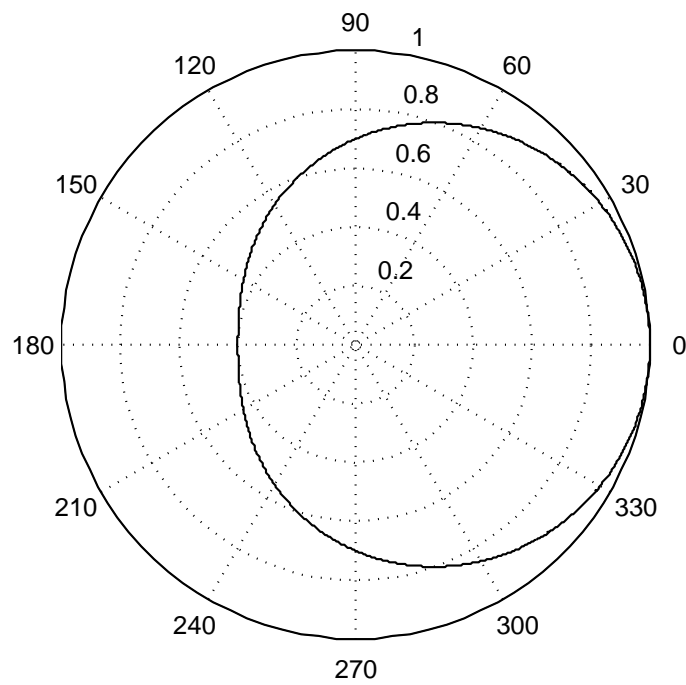
$$W_{-z} = A - B\sin(\theta). \quad (2.39)$$

If we now perform a rotation of the function so that it now points in the  $+x$  direction, we obtain (see Appendix B)

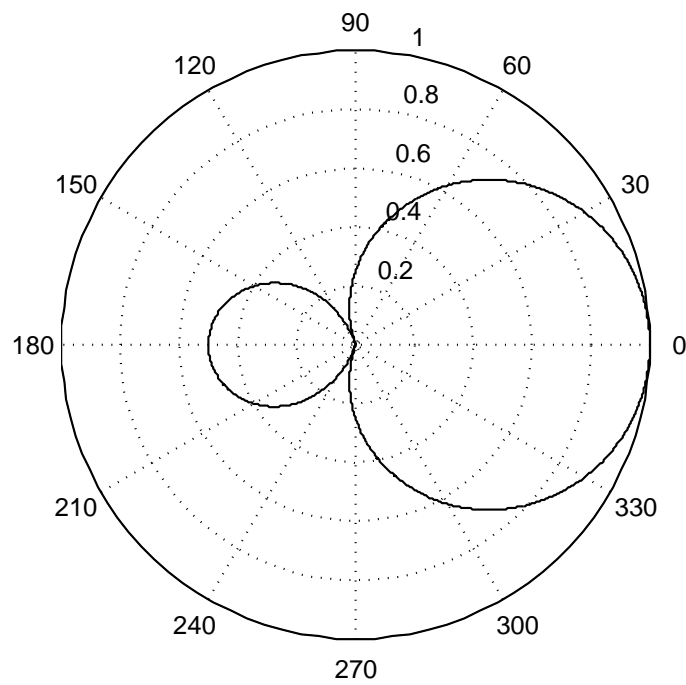
$$W_x = A + B\cos(\phi)\cos(\theta). \quad (2.40)$$

For a microphone facing in the  $-x$  direction, the directivity is

$$W_{-x} = A - B\cos(\phi)\cos(\theta). \quad (2.41)$$



(a)



(b)

**Figure 2.5** Two-dimensional directivity plots for (a) an ideal subcardioid pattern and (b) an ideal hypercardioid pattern.

**Table 2.1**  $A$  and  $B$  ratio factors for common directivity patterns.

	$A$	$B$
Omnidirectional	1	0
Subcardioid	0.7	0.3
Cardioid	0.5	0.5
Supercardioid	0.37	0.63
Hypercardioid	0.25	0.75
Figure 8 (bipolar)	0	1

Finally, the rotation of the directivity so that it faces in the  $+y$  direction yields

$$W_y = A + B \sin(\phi) \cos(\theta), \quad (2.42)$$

while the directivity for the  $-y$  facing microphone is

$$W_{-y} = A - B \sin(\phi) \cos(\theta). \quad (2.43)$$

If we now use these generalized expressions for microphone directivity in the equations for angles in the Polar ETC [see Eqs. (2.31) and (2.32)], we obtain

$$\begin{aligned}
\phi &= \tan^{-1} \left( \frac{E_L - E_R}{E_F - E_B} \right) \\
&= \tan^{-1} \left( \frac{[A + B \sin(\phi_A) \cos(\theta_A)] - [A - B \sin(\phi_A) \cos(\theta_A)]}{[A + B \cos(\phi_A) \cos(\theta_A)] - [A - B \cos(\phi_A) \cos(\theta_A)]} \right) \\
&= \tan^{-1} \left( \frac{2B \sin(\phi_A) \cos(\theta_A)}{2B \cos(\phi_A) \cos(\theta_A)} \right) \\
&= \tan^{-1} \left( \frac{\sin(\phi_A)}{\cos(\phi_A)} \right) \\
&= \phi_A,
\end{aligned} \quad (2.44)$$

$$\begin{aligned}
\theta &= \sin^{-1} \left( \frac{E_U - E_D}{2E} \right) \\
&= \sin^{-1} \left( \frac{[A + B \sin(\theta_A)] - [A - B \sin(\theta_A)]}{2BE} \right) \\
&= \sin^{-1} \left( \frac{2B \sin(\theta_A)}{2BE} \right) \\
&= \sin^{-1} (\sin(\theta_A)) \\
&= \theta_A,
\end{aligned} \tag{2.45}$$

where the subscript  $A$  denotes the actual value of the arrival angle. Thus, with any cardioid-family microphone directivity pattern, the results are the same. One should note, however, that when either  $A$  or  $B$  is 0, the equations for  $\phi$  and  $\theta$  are undefined due to an indeterminate ( $0/0$ ) form. While this is readily apparent for the case when  $B = 0$ , it is not so obvious in the case when  $A = 0$ . The reason why  $A = 0$  is a problem is that, for a bipolar pattern, both lobes are in phase, and thus the  $+x$ -facing and  $-x$ -facing patterns are equal and cancel each other out, resulting in the indeterminate form.

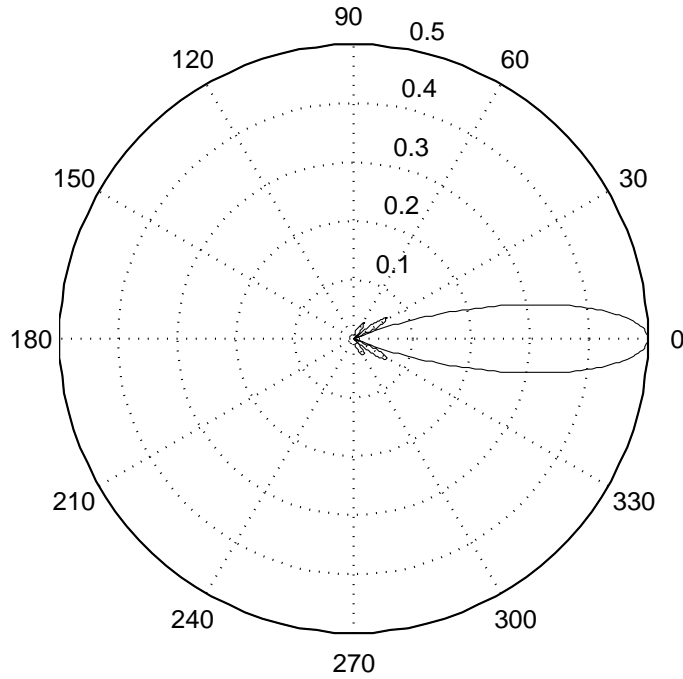
Let us now consider the case of a microphone directivity that is not a member of the ideal cardioid family. If we consider a hypothetical microphone with the polar response of a baffled piston (see Fig. 2.6), the directivity can be represented as

$$W(\phi) = \begin{cases} \frac{J_1(k \sin(\phi))}{k \sin(\phi)} & -\frac{\pi}{2} < \phi < \frac{\pi}{2} \\ 0 & \frac{\pi}{2} < \phi < \frac{3\pi}{2} \end{cases} \tag{2.46}$$

Substituting the directivity of a baffled piston into Eq. (2.13) produces

$$\phi = \tan^{-1} \left( \frac{\frac{J_1(k \cos(\phi_A))}{k \cos(\phi_A)} - \frac{J_1(k \cos(\phi_A - \pi))}{k \cos(\phi_A - \pi)}}{\frac{J_1(k \sin(\phi_A))}{k \sin(\phi_A)} - \frac{J_1(k \sin(\phi_A - \pi))}{k \sin(\phi_A - \pi)}} \right) \tag{2.47}$$

$$= \tan^{-1} \left( \frac{J_1(k \cos(\phi_A)) + J_1(-k \cos(\phi_A))}{J_1(k \sin(\phi_A)) + J_1(-k \sin(\phi_A))} \tan(\phi_A) \right) \tag{2.48}$$



**Figure 2.6** Two-dimensional polar pattern for a baffled piston in the far field for  $ka = 10$ .

This expression is clearly quadrant-dependent, due to the directivity pattern only being nonzero for half of the polar plane. If we assume an arrival in the first quadrant, it reduces to

$$\phi = \tan^{-1} \left( \frac{J_1(k a \cos(\phi_A))}{J_1(k a \sin(\phi_A))} \tan(\phi_A) \right). \quad (2.49)$$

The effect of the directivity pattern on the result thus depends on the factor

$$\frac{J_1(k a \cos(\phi_A))}{J_1(k a \sin(\phi_A))}, \quad (2.50)$$

which in turn depends on  $\phi_A$  and  $ka$ . A plot of this factor is shown in Fig. 2.7 for both fixed  $ka$  with  $\phi_A$  varying and fixed  $\phi_A$  with  $ka$  varying. The factor can thus vary between negligible and very large values, depending on the the combination of

$\phi_A$  and  $ka$ . In reality, many caridoid-family microphones do not have ideal directivity functions, so effects similar to those of a baffled-piston directivity could possibly be introduced. The extent of those effects are typically not determinable analytically, but a numerical implementation of actual directivity functions may be able to determine how much of an impact they have on the results.

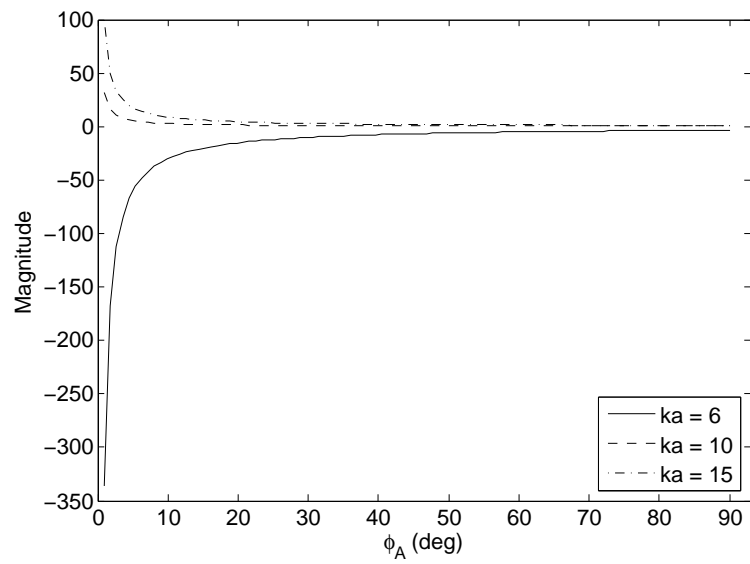
### The Energy Time Curve (ETC)

One final problem has to do with the ETC itself. When Heyser introduced it, he believed it was a more correct and useful measure than the IR because it involved both real and imaginary pressure signals in the time domain. He believed that the two components related to the potential energy density (the pressure signal) and the kinetic energy density (a  $90^\circ$  phase shifted and integrated version of the pressure signal which is not necessarily proportional to the acoustic particle velocity). It was later shown, however, that this is typically not the case and the ETC is not always a valid measure due to the acausal nature of the doublet response created when using a Hilbert transform to derive the imaginary part of the time signal [31,32]. In addition to the use of the acausal doublet response, the ETC is an envelope function for the decay of energy in a room, which does not lend itself well to exploring situations containing fine detail.

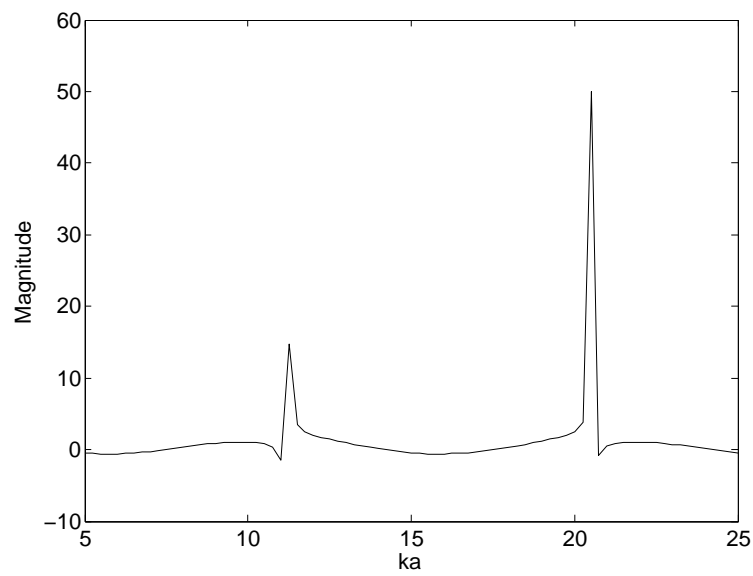
Despite its drawbacks, one should not dismiss the Polar ETC. There are many cases in which it performs adequately because there are single instantaneous sound arrivals of interest in the IR or some of the simultaneous arrivals are of low enough level that they do not significantly affect the localization of a single, much stronger arrival.

To improve the method, one might first use squared IRs in place of ETCs to eliminate signal processing ambiguities. This would allow for a more accurate portrayal





(a)



(b)

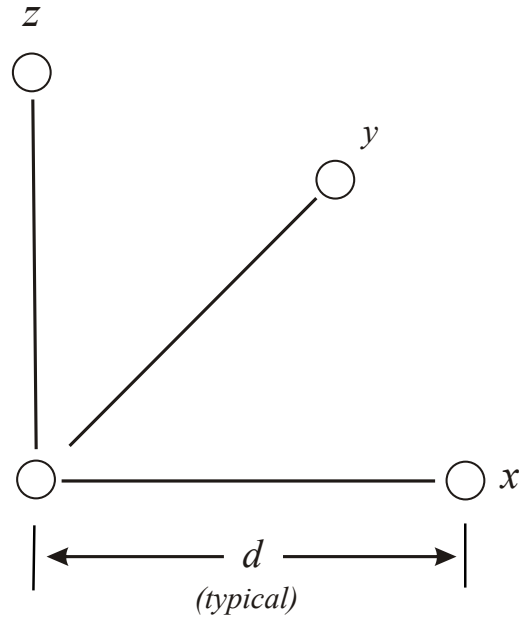
**Figure 2.7** Plot of the modifying factor for a baffled-piston directivity with (a) fixed  $ka$  values and  $\phi_A$  varying from 0 to  $90^\circ$  and (b)  $\phi_A = 20^\circ$  and  $ka$  varying from 5 to 25.

of the system and would result in finer detail in the time record. Actual directivity data for a particular microphone, or an array of microphones might also be used to improve the results over those based on an ideal cardioid assumption. As for multiple arrivals within one sample, the only things that can really be done is to increase the sampling rate as high as possible and thus make the sampling period as short as possible or increase frequency content to make the width of the arrival (in samples) as short as possible. This would force more arrivals into adjacent samples to increase the feasibility of the assumption that there is only one sound arrival in each sample.

## 2.3 Correlation-Based Cartesian Technique

In 1987, a correlation technique was introduced in Japan by Yamasaki and Itow [7]. The technique utilized four omnidirectional microphones comprising a Cartesian array. One microphone was located at the origin (or the measurement point of interest) and the other three microphones were located a small, known distance away from the origin along each of the three Cartesian axes (see Fig. 3.2). The cross correlation was found between each of the signals and all peaks with an amplitude above a certain threshold were selected to compute the direction of arrival for those peaks.

The technique first takes the cross correlation between the origin microphone and each of the three other microphones over the course of the entire signal and then identifies all peaks in the cross correlation that exceed the threshold. The associated time delays are then used to calculate the direction of arrival for each peak by finding the Cartesian coordinates of a “virtual image source.” These Cartesian coordinates are then transformed into spherical coordinates using the standard equations. The method not only assumes a single arrival within a single peak in the IR, but it also seems to assume that there is no spreading in time of the IR peaks. If the IR at the



**Figure 2.8** Four-microphone Cartesian array used for a correlation-based technique.

point of interest contains a large peak, the method will find similar peaks at the other microphones and use just the time delay between the microphones for that one peak for the direction of arrival.

The use of a correlation technique like this prevents any beneficial modification to the method. The use of a Cartesian array, however, proves useful due to its simple geometry, as will be discussed further in Chapter 3.

## 2.4 Tetrahedral Array

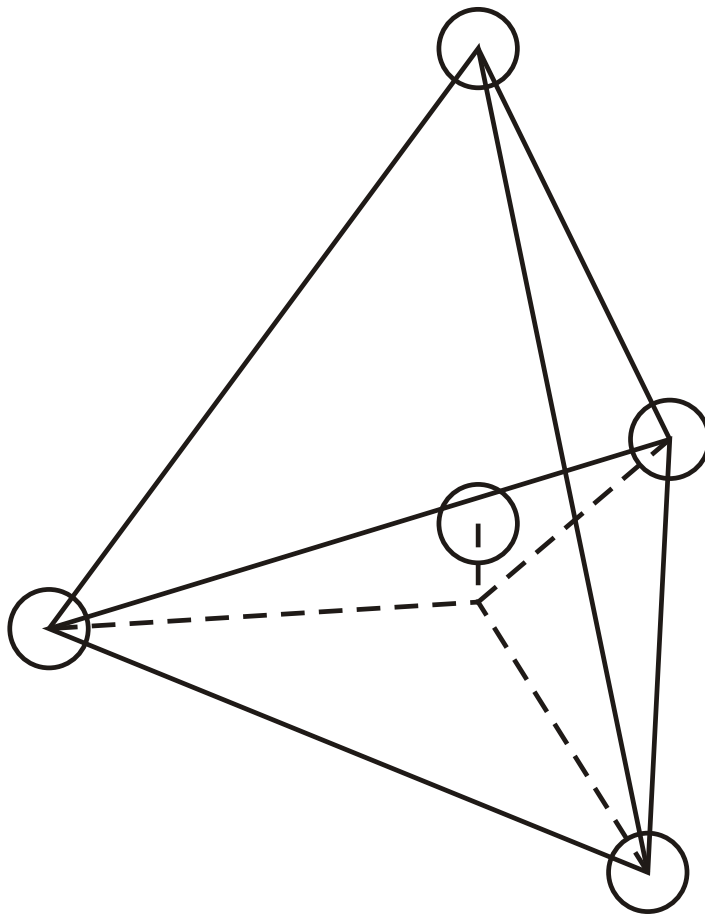
The method developed by Sekiguchi *et al.* in 1992 employed the use of an array of four microphones mounted at the apexes of an imaginary regular tetrahedron,

with one vertex pointing straight up and the other three forming a triangle in the horizontal plane relative to the room [8]. The primary reason for this geometry was the reduced computing power it required. When the method was developed in the early 1990's, this concern was more important than it is today. The method used a correlation technique like that used by Yamasaki and Itow, but they were able to improve measurement accuracy by utilizing a deconvolution method that eliminated non-ideal characteristics of the primary sound source. The problems with the method are essentially the same as those addressed in the preceding section.

## 2.5 Modified Tetrahedral Array

In 2003, Choi *et al.* developed a technique similar to that described above with modifications [9]. This array consisted of the four microphones described in the previous method, plus a fifth microphone located at the geometric center (see Fig. 2.9). The authors claimed to detect more sound arrivals by sequentially changing sets of four microphones in the calculations. This meant they would actually form five different groups of four microphones, calculate out the directions of arrival for each of the five sets, then discard all overlapping arrivals. The result was more detected arrivals. They apparently used the same basic correlation technique used by the other authors, although the use of alternate microphone sets did allow for improvement over past methods.

To further improve the method, they used a peak detection algorithm to eliminate most of the IR except for the peaks that exceeded a specified threshold. They interpolated the resulting data using a spline interpolation, effectively increasing their sample rate to help with reducing the sampling period. It is unknown whether the discarding of information in the peak detection algorithm improved the measurements



**Figure 2.9** Five-microphone tetrahedral array with microphones located at the apexes and at the center of gravity.

to any significant degree, but it is a point that should be mentioned.

Although the fifth microphone did allow them to extract more information (even some overlapping reflections in one Cartesian direction), the authors only calculated virtual image source locations and not actual directions of reflection origination.

## 2.6 Cross Correlation in a Steady-State Sound Field

A method published in 2006 uses a cross correlation function and matrix calculations developed by Noël *et al.* [18]. It was developed to localize a single sound source in a reverberant room, where the sound source cannot be turned on and off. Specifically, the method was developed for the localization of a dominant sound source in a large factory and the associated reflections from walls. This would allow one to determine where acoustic treatment would be most beneficial. One of the main drawbacks to the method for use in room acoustics diagnostics is that it was designed for an accuracy of  $\pm 5$  degrees in both the azimuthal and polar angles. An average error this large can result in identification of surfaces that actually do not play a role in a sound arrival of interest. In addition, the method requires extensive calculations (reduced by using infinite series of Legendre polynomials) and several microphones (up to 15 in practice, but up to more than a hundred in theory) in very complex geometrical arrangements. In some cases, computations could take several hours depending on the computational power available. In addition, one of the requirements for their matrix calculations is knowledge of a ratio that cannot be assessed until after the measurements have been made. This could result in the taking of an erroneous measurement the error of which is not known until after the fact. In the room acoustics problem considered in this thesis, the method should not only be accurate, but also computationally efficient, requiring only a modest amount of equipment to produce rapid results.

One should note that the authors were trying to solve for directions of arrival in a sound field that was most likely nearly diffuse. This is most likely what led to such a complex solution.

## 2.7 Beamforming Techniques

Beamforming is a method that has been used to great extent in other fields, such as SONAR and radio astronomy, and has existed for decades. The application of it to directional room acoustics was a method extensively developed in 2002 by Gover [10–12]. This method was not originally intended for DIR measurements, but for determining directions of sound arrivals resulting from sound transmission through partitions between rooms. He utilized two different sized open spherical microphone arrays with 32 microphones each and a beamforming algorithm to find the direction of sound arrivals in rooms. The first array was placed at the location of interest and the room was excited using a maximal length sequence (MLS). The same thing was done with the smaller array. The different sizes allowed the usable frequency of the measurements to range from approximately 300 Hz to about 3 kHz. With this setup, he was able to create 60 beams in his calculations into which the sound arrivals were divided.

The first problem is more of a logistical one. The sheer number of sensors used and the number of channels of data to be acquired make the method difficult for the average acoustician or dedicated consumer to implement. Another problem is that of the limited bandwidth. Another problem is that of accuracy. The method divides the three dimensional space up into 60 regions or beams. Each beam has its own associated width (usually expressed in degrees at the 3-dB down point) and it can be steered over the surface of an imaginary sphere. The beam is unable to resolve

multiple arrivals that fall within the same beam.

## 2.8 Intensity Measurements

In the mid-1990's Abdou and Guy explored the use of intensity measurements to determine the directional characteristics of sound in rooms [19]. They claimed that intensity could potentially offer more information about the spatial aspects of the sound due to the particle velocity content, but they also affirmed that intensity is a difficult measurement to work with. They used an MLS signal to excite the room and measured the pressure impulse response and the intensity impulse response along each axis. One of the main concerns is that in order to obtain a particle velocity measurement (and thus the intensity impulse response), a finite-difference method is typically used along each of the Cartesian axes. However, bias error is inherent in the finite difference method. Despite these difficulties, they achieved results having average errors in the range of  $2.3^\circ$  to approximately  $15^\circ$ . In addition, even single arrivals are smeared over space due to the inaccuracies in the intensity measurement. They did not specify whether simultaneous arrivals were resolvable.

## 2.9 Ambisonics and Spherical Harmonic Decomposition

In the late 1970's, Fellgett and Gerzon pioneered a new way of recording and playing back sound that has become known as Ambisonics [33]. The basis of Ambisonics lies in a spherical harmonic decomposition of the sound field of interest. These data are then stored and, using a known playback loudspeaker configuration, reproduced upon playback to reconstruct the recorded sound field to first-order accuracy. The probe



used to record the sound field consists of a special arrangement of four microphones arranged on a regular tetrahedron. However, instead of being arranged on the apexes of the tetrahedron, like two of the methods previously mentioned, they are centered on each face of the tetrahedron and face out normally. The signals are recorded as regular audio signals, but are then converted from “A” format (regular pressure signals) to “B” format (the zeroth and three first-order spherical harmonics). The benefit of having these four lowest-order spherical harmonics is that after the sound field has been recorded once, the signals can be combined in different ratios to produce virtually “steered,” coincident microphone sets. In practice, the sets are not quite perfectly coincident due to the the finite displacement of each microphone from a single point.

Using the idea of perfectly coincident microphone sets, Essert (1996) and Farina (2007) separately worked on extracting directional information from room IRs using different methods [13–17]. The methods entail using the B-format microphone signals to measure the acoustic intensity. It is unknown how similar the intensity method used is to that used by Abdou and Guy [19]. The intensity measurements are then used to extract the directional information in the room. One source of inaccuracy in these measurements comes from the fact that the derived B-format microphone responses for a typical first-order ambisonics probe are not very uniform over frequency. This introduces errors similar to those induced by assuming a cardioid microphone has a truly cardioid directivity function over all frequencies in the Polar ETC.

## 2.10 Chapter Conclusions

As suggested by the number of methods discussed in this chapter, directional impulse response information is extremely important in room acoustics. However, each of the

methods has had its own set of shortcomings or drawbacks. These problems range from using extremely complicated setups, extensive equipment, and long measurement times to flawed assumptions in the approach used to solve the problem. For this reason, this area of research still warrants further investigation. A new method for extracting directional information in room responses is presented in Ch. 3.



# Chapter 3

## Theoretical Developments

### 3.1 Preliminary Considerations

One of the main pitfalls of former DIR methods was their dependence on certain questionable assumptions. Perhaps the most unreliable is the assumption that only a single sound arrival occurs within a given measurement sample, made primarily to simplify expressions and allow for the use of more compact analytic forms. This is usually not an accurate depiction of what occurs in practice. In most any room, multiple sound-travel paths of the same length will eventually produce multiple arrivals in a given sample after a certain amount of time has elapsed in the IR. While the required length of time depends on the nature of the room, it invariably occurs, meaning the assumption of only individual arrivals will eventually be violated. These equal-length paths could be thought of as a result of a combination of equipartition of energy and Fermat's principle. Fermat's principle was developed for geometrical optics in order to bypass many of the difficulties in working with wave-based optics. Acoustics and optics share many similarities, and thus we can bypass many of the difficulties in wave-based acoustics by considering geometric acoustics and using a

similar approach as geometric optics [34].

To show that there are multiple path lengths that a sound wave can travel that are exactly the same length, let us consider an omnidirectional source emitting acoustic rays. If the radiation impedance over the surface of the source is approximately uniform, then the amount of acoustic energy propagating in any one ray is approximately equal to that of any other ray. Fermat's principle states that a ray will travel the path of shortest length. If the rays are already propagating in a certain direction (which has been established by the normal surface velocity of the source), then they will take the shortest path to the receiver based on the direction they are traveling. When a ray strikes a boundary, its direction has been changed and the path it takes to the receiver is then determined by the direction it is traveling after reflection. This process occurs as many times as necessary until the ray arrives at the receiver location. If we then assume that the surfaces the rays strike are all continuous, there is an infinite continuum of possible path lengths. The point in time where the statistical probability of there being multiple equal-length paths is high (cutoff time) depends on the room geometry and source and receiver placement. Despite the variance in the value of the cutoff time from room to room, it always occurs. If the room boundaries are sufficiently reflective, the amount of time in the IR where there are multiple simultaneous arrivals will far outweigh the amount of time in the IR where single arrival paths dominate.

These paths that are exactly the same length create arrivals that arrive at *exactly* the same time, no matter how frequently the signal is sampled or how carefully the measurement is taken. Having considered this, an "average" room, or a room where the walls, floor and ceiling are not perfectly smooth nor perfectly reflecting, is bound to have at least some discrete paths that are of the same length. With this in mind, the assumption that there is only one sound arrival at one time (though this may be

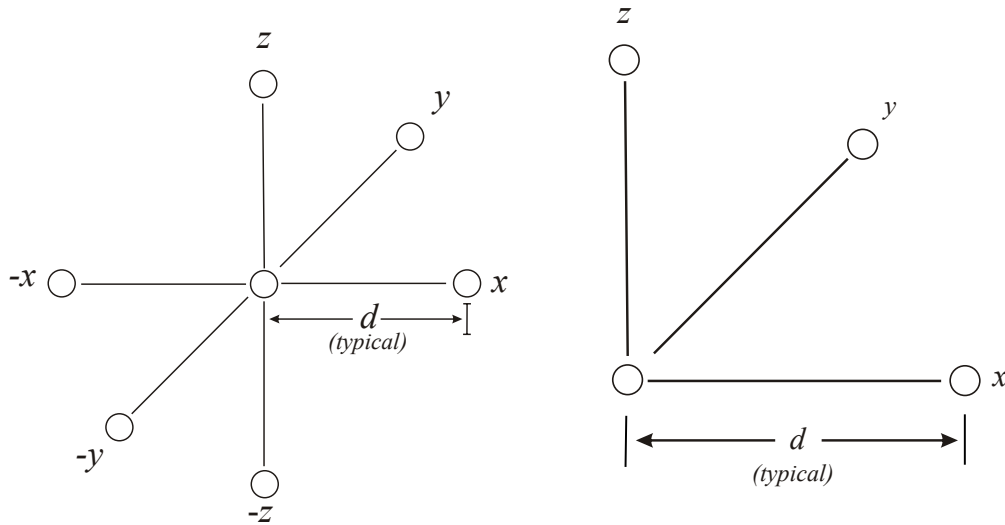
valid early in the IR) is not a safe assumption, especially when strong late arrivals are present, as these are often a result of multiple arrivals as opposed to a single later arrival. If we consider the two-dimensional “room” first introduced in Ch. 2 (see Fig. 2.1), we can see just how detrimental this assumption can be; there is nothing that can be done to salvage the measurement when there are simultaneously arriving wave fronts.

## 3.2 Generalized System Setup

In order to adequately resolve the issues discussed in Chs. 1 and 2 regarding the complete and efficient extraction of directional information in room acoustics, one may benefit from first considering the problem in the simplest of forms. Suppose a point source is radiating into free space. The source is located, relative to an arbitrary origin, at the Cartesian coordinates  $(x_s, y_s, z_s)$ . In order to extract three-dimensional directional information about its sound field, one must have at least four receivers in the field and at least one of these must not be in the same plane as the other three. If there is not at least one microphone out of plane with the other three, angles of arrival may only be determined in the single plane. (This would be much like trying to construct a vector with  $x$ ,  $y$ , and  $z$  components using only an incomplete set of  $x$  and  $y$  components.)

For the method developed in this work, the receivers comprised a seven-microphone array mounted in the Cartesian arrangement as depicted in Fig. 3.2. One microphone was located at the origin and the remaining six were each located a fixed distance  $d$  along a Cartesian axis from the origin. If the system is considered to be time invariant, the same results would follow by using a single microphone placed sequentially at each of the seven locations. Any other arrangement of at least four microphones with

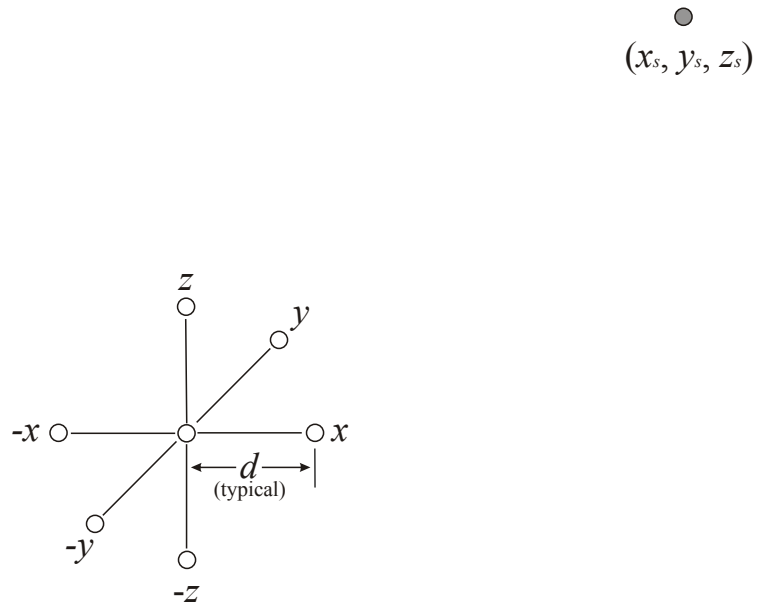
one microphone out of plane with the other three could also be used if the geometry of the measurement points is known.



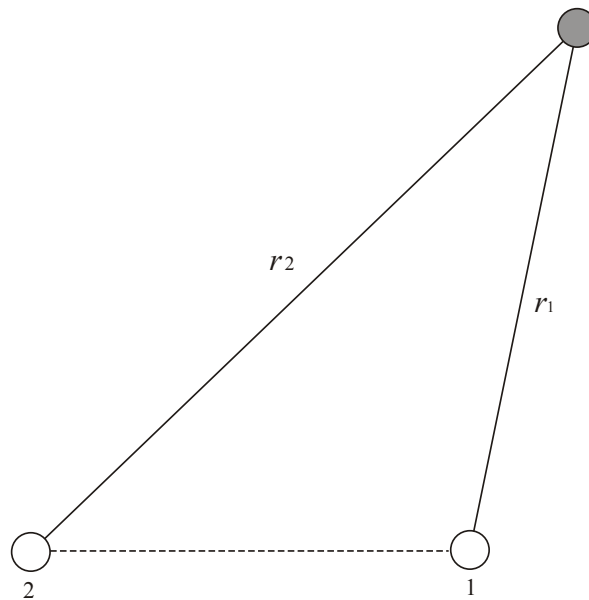
**Figure 3.1** Cartesian microphone arrangements for (a) a seven-microphone array and (b) a four-microphone array where a single microphone is located at the origin and the others are located a distance  $d$  away along the axes.

Now that the microphone array has been defined and the source placed, let us consider the geometry of the situation without any assumptions except that the speed of sound  $c$  is known along with the separation distance  $d$  of the microphones from the origin. With a source radiating into free space, the path length of the sound arrival from the source to the receivers can be called  $r_1$  and  $r_2$  (see Fig. 3.3). The time delay is then  $(r_1 - r_2)/c$  and represents the time delay of the arrival at position 1 relative to position 2 (a negative time delay indicates that an arrival arrives at position 1 prior to position 2).

Since the source coordinates were expressed relative to the origin, one can express the distance of the source from each of the six microphones as  $(x_s - d, y_s, z_s)$ ,  $(x_s + d,$



**Figure 3.2** Physical setup for theoretical development including a seven-microphone Cartesian array and radiating point source in free space.



**Figure 3.3** Graphical representation of a source (shaded circle) and two receivers in free space with different source-receiver path lengths for the demonstration of time delays between receivers.



$y_s, z_s$ ),  $(x_s, y_s-d, z_s)$ ,  $(x_s, y_s+d, z_s)$ ,  $(x_s, y_s, z_s-d)$  and  $(x_s, y_s, z_s+d)$  for the  $+x_s, -x_s, +y_s, -y_s, +z_s$ , and  $-z_s$  microphones, respectively. By grouping the vectors based on the Cartesian axis on which the pair of microphones resides, one can also express the coordinates as a function of the time delay along each axis between the wave fronts arriving at one microphone relative to the other. These time delays can be expressed as follows:

$$\Delta t_x = \frac{\sqrt{(x_s - d)^2 + y_s^2 + z_s^2} - \sqrt{(x_s + d)^2 + y_s^2 + z_s^2}}{c}, \quad (3.1)$$

$$\Delta t_y = \frac{\sqrt{x_s^2 + (y_s - d)^2 + z_s^2} - \sqrt{x_s^2 + (y_s + d)^2 + z_s^2}}{c}, \quad (3.2)$$

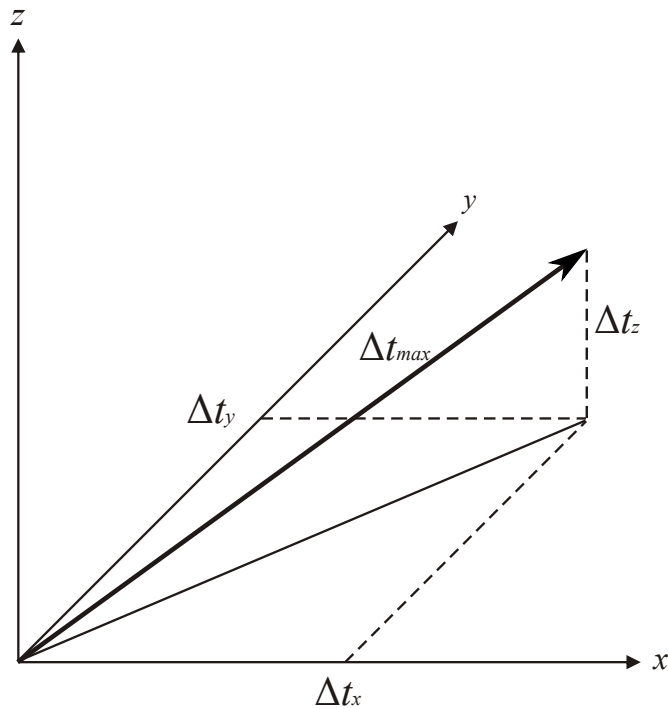
$$\Delta t_z = \frac{\sqrt{x_s^2 + y_s^2 + (z_s - d)^2} - \sqrt{x_s^2 + y_s^2 + (z_s + d)^2}}{c}, \quad (3.3)$$

where  $c$  is the speed of sound in air. The subscript for each  $\Delta t$  denotes the axis along which the time delay is computed and all time delay values are given in seconds. These equations are simply a three-dimensional extension of the situation considered in Fig. 3.3 where the receivers are equidistant from the origin. These arguments can be extended to a system with multiple radiating sources by simply identifying the time delays along each axis and correctly grouping the  $\Delta t_x$ ,  $\Delta t_y$ , and  $\Delta t_z$  for each arrival. To ensure correct grouping of time delays, a constraint is required.

### 3.3 Time Delay Space

To further elucidate the situation, one might consider using a representation that could be called “time delay space” (see Fig. 3.4). The conceptualization again involves Cartesian axes, but instead of distances along the  $x$ ,  $y$  and  $z$  axes, the time delays between pairs of microphones become the coordinates. Thus, a plot becomes that of

$\Delta t_x$ ,  $\Delta t_y$  and  $\Delta t_z$ . The maximum time delay along any one axis occurs when the time delays in the other two dimensions are zero. This corresponds to an arrival along just that one axis. In this case, the delay is simply the distance between the two microphones divided by the speed of sound, or  $2d/c$ . Indeed, it is true that the time delays for any sound arrival originating from outside the array will lie on the surface of a sphere in time delay space with a radius of  $2d/c$  for a seven microphone Cartesian arrangement and  $d/c$  for a four microphone Cartesian arrangement (see Fig. 3.1(b)). This is due to the fact that regardless of the physical distance from the origination of the arrival to the array origin, a time delay for a microphone pair a distance  $2d$  apart can never exceed  $2d/c$ . In other words



**Figure 3.4** Graphical representation of a sound arrival in time delay space.

$$\Delta t_{max} = \sqrt{(\Delta t_x)^2 + (\Delta t_y)^2 + (\Delta t_z)^2} = \frac{2d}{c}. \quad (3.4)$$

In terms of grouping the correct time delay values together to compute a direction of sound arrival, this equation may be used as a constraint. Only groupings of time delays that meet the requirement should be considered in the computation of solutions. In practice, the constraining is done by computing the time-delay-space magnitude of every possible combination of  $\Delta t_x$ ,  $\Delta t_y$ , and  $\Delta t_z$ . Only those sets of  $\Delta t_x$ ,  $\Delta t_y$ , and  $\Delta t_z$  that meet the criterion within a certain tolerance are considered. The advantage of working in time delay space over working in spatial coordinates is that the time delay arrival angles simplify calculations while being identical to the spatial arrival angles (as long as the time delays have been defined properly). There is no need for scaling of any sort. The spherical angles of arrival can be calculated using a simple coordinate transformation from Cartesian to spherical coordinates. If the  $+x$  direction in Fig. 3.4 is chosen to be  $\phi = 0^\circ$  and  $\theta$  is chosen to be the elevation angle from the horizontal plane, then the angles of arrival can be calculated using the following equations:

$$\phi = \tan^{-1} \frac{\Delta t_y}{\Delta t_x} \quad (3.5)$$

$$\theta = \sin^{-1} \frac{\Delta t_z}{\sqrt{\Delta t_x^2 + \Delta t_y^2 + \Delta t_z^2}} \quad (3.6)$$

where the inverse tangent is quadrant specific.

### 3.4 The Cross Correlation Function

Before the angles of arrival for any arbitrary number of arrivals can be found using Eqs. 3.5 and 3.6, it is necessary to determine the time delay along each axis. The

method that best suits this situation is the cross correlation function.

The cross correlation function can be thought of as a tool that produces the time delay between the arrivals of a single acoustic signal at two different locations. It can be computed in the time domain, but is often computed as the inverse Fourier transform of the cross power spectrum. This method was comprehensively described by Knapp and Carter [35]. Briefly, let us first consider the complex frequency-domain spectra of the two received signals,  $S_1$  and  $S_2$ . The complex cross spectrum is then defined as

$$S_{12} = S_1 S_2^*, \quad (3.7)$$

where the  $*$  denotes the complex conjugate. Then if we transform this function into the time domain, we obtain the cross correlation function, represented as

$$R_{12} = \mathcal{F}^{-1}\{S_{12}\}, \quad (3.8)$$

where  $\mathcal{F}^{-1}$  denotes the inverse Fourier Transform. This can also be expressed in integral form as

$$R(\tau)_{12} = \frac{1}{T} \int_{\tau}^{T+\tau} x_1(t) x_2(t - \tau) dt, \quad (3.9)$$

where  $\tau$  is the time delay value that maximizes the expected value of the integrand and  $T$  is the measurement period.

The peaks in the cross-correlation function correspond to a stronger relationship between the two signals at that specific time delay. The time-axis values of the peaks can therefore be interpreted as the time delays corresponding to the different sound arrivals at one location relative to another. It is common to have both positive and negative time delay values which simply represent wave fronts traveling along the axis of measurement in either direction. It is a relatively simple practice to obtain the

cross correlation from any two measured signals.

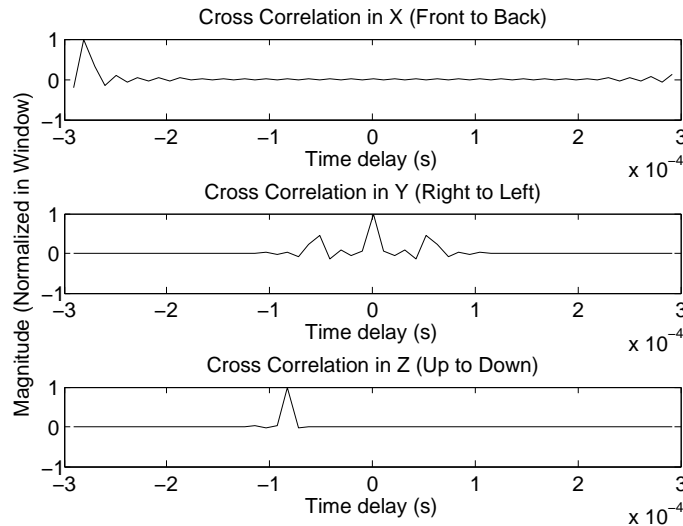
### 3.5 Special Considerations for the Cross Correlation Function

Some things must be taken into consideration when searching for the arrival time delay values. In practice, the number of sound arrivals comprising any given peak in the IR is unknown, as is their correct pairing. The distinguishing characteristics between single and multiple arrival scenarios should be discussed.

For a single arrival, only one peak in the cross correlation would be expected. Though this certainly is the case, one may wonder what happens if two sources lie at different distances along the same radial line out from the origin. They will produce identical time delays and therefore produce identical directions of arrival. However, if one wishes to individually assess a peak in the impulse response, the IR can be time gated by truncating the data that lies outside the range of  $\pm \Delta t_{max}$  of the peak of interest (see Sec. 3.7). The sources then only need to be physically separated by the distance  $2d$ . Any greater separation will mean the signal will get time filtered completely and therefore is of no concern.

A scenario involving multiple arrivals is much more complex and warrants greater discussion. Due to the fact that the cross correlation simply produces the time delays that correspond to a maximum relationship (or maximum expected value as shown in [35]), then the maximum peak in the cross correlation graph corresponds to the time delay that is “most likely”. In the case of multiple arrivals, however, this maximum will tend to appear at the time delay that does not necessarily correspond to an actual arrival. Therefore, when considering which peaks correspond to physical arrivals and which correspond to nonphysical arrivals, great care is required. It is possible, for

example, to have a set of three cross correlation graphs (one each for  $x$ ,  $y$  and  $z$ ) where two of the graphs have only a single large peak and the third has three peaks (see Fig. 3.5). The coordinates for the two sources used to generate the figure below are (20 m, 4 m, 6 m) and (20 m, -4 m, 6 m).



**Figure 3.5** Cross correlation in  $x$  (top),  $y$  (middle), and  $z$  for a microphone array centered at (0 m, 0m, 0 m) and two sources located at (20 m, 4 m, 6 m) and (20 m, -4 m, 6 m).

In this case demonstrated in Fig. 3.5, one should note first that there are multiple arrivals. Secondly, it should be realized that in the graph with multiple peaks, the tallest peak may correspond to a nonphysical arrival and should therefore be treated with care. The shorter peaks in this case correspond to physical arrivals, but one should note that any peaks corresponding to physical arrivals need to be significantly above the noise floor of the cross correlation function to avoid being classified as “noise.”

Multiple simultaneous arrivals also create the possibility of degenerate time delays. The term degenerate time delay as used here means that a single peak in the cross correlation graph may correspond to more than one sound arrival. In Fig. 3.5, the

time delay values in both  $x$  and  $z$  are degenerate, meaning the time delays correspond to more than one arrival, while the time delay values in  $y$  are distinct.

A final consideration is the averaging method used for extracting the cross correlation. When the cross spectrum is calculated in the frequency domain, its noise immunity is increased by the use of averaging. In many cases, IRs are measured as the average of several IRs taken sequentially. The cross spectrum, however, has maximum noise immunity when several cross spectra are averaged together as opposed to taking the cross spectrum of an average of several signals. In other words, the cross correlation will contain less noise if the cross spectrum it is derived from is the average of the cross spectra of several sequential measurements as opposed to being the cross spectrum of the average of the measurements.

### 3.6 Alternate Cross Correlation Pairs

Even after all the previous steps have been taken, an alternate “view” of the directional information may be desired. This could be potentially helpful if there are suspected degenerate time delays. It would involve simply taking the signals already measured and computing the cross correlation between different microphone pairs. The number of different combinations can be quite large, depending on the number of microphones in the array. All the possible equations governing the extraction of directional information using this method will not be presented. They are easily calculated using the geometry of the array in conjunction with generalized source coordinates and the law of cosines. Again, this method provides a different “viewpoint” of the directional information, though it does not introduce any new information. It may also be possible to implement the DIMUS [36] or CLEAN-SC [37] methods to deconvolve the spatial IR of the microphone array.

The use of alternate cross correlation pairs produces a new maximum time delay ( $\Delta t_{max}$ ) that is dependent on the array geometry and the cross correlation pairs chosen. In addition, the time delays in the cross-correlation graph do not relate to the same axes as the microphone array axes and need to be adjusted accordingly. This, too, can be done using relatively simple geometry and trigonometry.

### 3.7 Procedure for Directional Impulse Responses

Now that a set of equations has been developed which are dependent on only the physical constants  $c$  and  $d$ , and the measured time delay values  $\Delta t_x$ ,  $\Delta t_y$ , and  $\Delta t_z$ , we can consider applications of this method to IRs. However, due to the fact that the cross correlation yields the time delays for maximum relationship between two signals, its use on an *entire* IR is not optimal for identifying simultaneous arrivals. It would produce a graph whose peaks would relate to all sound arrivals over the entire course of the IR.

When a set of IRs is gathered from an array of microphones, the IRs will have a series of peaks, not all of which will be of interest to the user. In order to ensure that only the peak(s) of interest is (are) considered when calculating angles of arrival, one may window, or select only a small portion of the signal. The maximum time delay between microphone pairs defined in Eq. (3.4) may be used as the time window length. It is implemented using acquired IRs as follows:

1. Define the peak of interest in the origin IR.
2. Note the time,  $t$ , at which the peak arrives.
3. Window each IR, leaving only the response between  $t - \Delta t_{max}$  and  $t + \Delta t_{max}$  with non-zero values.



4. Fourier transform the windowed signals into the frequency domain.
5. Take the cross spectrum between the signals for each pair of microphones.
6. Transform the cross spectra into the time domain via Fourier transform to obtain the cross correlation functions.
7. Find the time delay values (time-axis values of the peaks) of the different sound arrivals.

Items 4 - 6 may be replaced by computing the cross correlation directly in the time domain. It should be noted that after the time windowing, the  $\Delta t = 0$  point in the cross correlation graphs will correspond to the time  $t$ . However, Eqs. (3.5) and (3.6) depend only on the time delay values and not on the actual arrival time  $t$ . In any case, it is possible to automate measurements such that all peaks in the cross correlation exceeding a threshold will be used to calculate arrival angles.

## 3.8 Chapter Summary

The theory for a new method of extracting complete directional information from IRs has been developed using only a few assumptions, all of which are very valid in practice. The cross correlation function has been implemented for estimating the time delay between pairs of microphones along the Cartesian axes and the angles of arrivals have been extracted from this information. Constraints have also been introduced to facilitate the determination of valid solution sets. Now that the theory has been introduced, we will consider some practical issues that must be addressed before a useful implementation of the theory is possible.

# Chapter 4

## Practical Issues

### 4.1 Introduction

Before the method introduced in Ch. 3 (hereafter called the STCM, for Short-Time Correlation Method) can be fully utilized, several practical issues must be addressed. These affects the accuracy, operability, and procedure in some way. Each of these issues will be addressed in order to ensure the best measurement approach for both numerical simulations and experimental setups.

### 4.2 Sampling Rate and Accuracy

One of the first things to consider when utilizing the STCM as described in Ch. 3 is the effect of sampling a signal. When a continuous-time signal is sampled, there is a loss of information. This loss can greatly impact the foundational cross-correlation, especially when working with very small time windows. For example, a 7.62 cm array diameter results in a time window of approximately 444  $\mu\text{s}$  (222  $\mu\text{s}$  in the positive and negative directions). With a sampling rate of 192 kHz, this equates to approximately

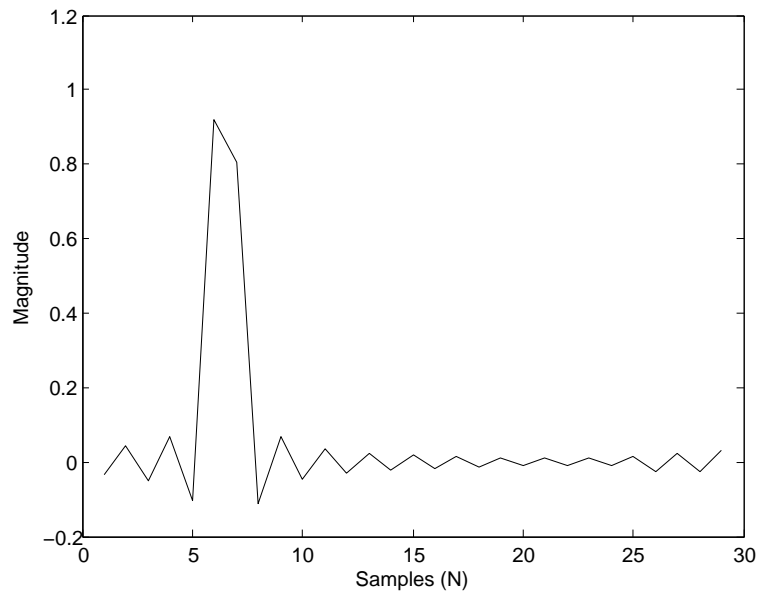
85 samples. This means only 85 distinct time delays are possible along each axis, resulting in a total of 614,125 ( $85^3$ ) possible results to represent a continuous scale of results. If the sampling rate is reduced to 96 kHz, the number of possible results is decreased to 76,766. If it is reduced to 48 kHz, the number decreases to 9,956. Thus, in principle, the sampling rate should be chosen to be as high as is possible.

With decreased sampling rates come increasingly erroneous results due to the lack of information about what is happening between samples. In these cases, the actual peaks of the cross correlation are usually estimated using some form of interpolation. Conversely, a high sample rate allows for finer resolution of both single peaks and multiple peaks located closely in the cross-correlation function. For example, Fig. 4.1(a) shows a cross correlation plot using a sample rate of 48 kHz. Fig. 4.1(b) shows a cross correlation plot using a sample rate of 192 kHz. In each case, the signal has frequency content up to the Nyquist frequency. Note how the single peak in Fig. 4.1(a) becomes resolved into three separate peaks when sampled at 192 kHz.

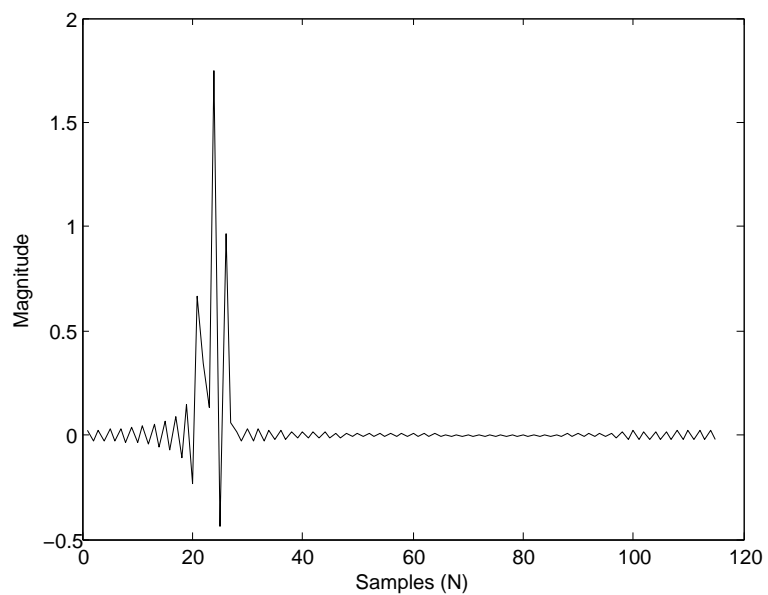
As can be seen in Fig. 4.1(a), the two different arrivals are not separated well enough to determine if there is a single arrival or two distinct arrivals. An inspection of Fig. 4.1(b) shows that there are indeed multiple arrivals, with the separation in azimuth being  $3^\circ$ .

### 4.3 Filtering Effects of Equipment

The IR is, in practice, derived from the frequency response function between the excitation signal and the signal measured at the microphone location. A perfect impulse in the time domain corresponds to a constant in frequency that extends to an infinite frequency. However, an infinite frequency-response bandwidth is impossible in practice. Because aliasing due to sampling will corrupt the measurement, filters are



(a)



(b)

**Figure 4.1** Comparison of cross correlation along the  $x$ -axis for two closely-located sources in free space when sampled at (a)  $f_s = 48$  kHz and (b)  $f_s = 192$  kHz.

intentionally employed to limit the bandwidth. As soon as the frequency content is truncated, the corresponding time response begins to widen from a perfect impulse to a peak of finite width. This in turn affects the cross-correlation function by widening the peaks.

The effect of the limited frequency response on the time response is observed regardless of the cause. Because a measurement system inherently acts as a bandlimited filter, equipment choice is important to the accuracy of the DIR. For example, a loudspeaker used to excite a room acts as a low-pass filter because of its high frequency rolloff. Microphones also roll off at high frequencies, but many precision mics respond better at higher frequencies than typical loudspeakers. The greatest accuracy can be obtained when wide-bandwidth sources and receivers are used along with high sampling rates.

In addition to the typical loudspeaker low-pass filtering effect, high frequency content is further reduced when the loudspeaker is oriented in a direction other than towards the receiver. Because most loudspeakers do not exhibit uniform directivity over frequency due to diffraction, high-frequency radiation in directions other than on-axis is dramatically reduced, which results in further widening of the peaks in the IR and reduced time resolution in the cross-correlation function. Experimental results for off-axis measurements are presented in Sec. 5.7.8.

## 4.4 Impulse Response Generation Programs

As a first step to validate the theory derived in Ch. 3, numerical simulations were utilized to generate IRs of hypothetical rooms. Three different programs were used in the effort, each of which are discussed in Appendix A.

#### 4.4.1 Discrete Sources in Free Space

The first program (Program 1) used was for simpler cases. It was developed originally by Ryan Chester [38] and modified by the author to suit the needs of the current research. It placed point sources and receivers in free space using Cartesian coordinates to simulate simple image source problems. Actual source and receiver coordinates were arbitrary; relative source and receiver coordinates were of primary importance as they were in free space. Sources in this case were equistrength, which represent perfectly reflective surfaces. Program 1 was used for the first four case studies described in Ch. 5.

#### 4.4.2 Enhanced Image Source Method

In addition to discretely placed sources, a second program (Program 2) was needed for creating more densely populated IRs. An updated version of the method developed by Berkley and Allen [29] for modeling rectangular rooms was created for Matlab. The program allowed for the inclusion of real pressure reflection coefficients for each of the six room surfaces individually, thus allowing the user to more closely approximate a real room. One limitation is that the simulated room must be rectangular, but this was sufficient for testing the DIR's ability to operate in more densely populated portions of an impulse response.

#### 4.4.3 Modified Modal Expansion

The first two programs generated IRs using Kronecker delta functions low-pass filtered at the Nyquist frequency placed in specific sample bins whose amplitude was determined by the surface reflection coefficients. While low-pass filtering at the Nyquist frequency will spread out the peaks in the IR slightly, it is much more likely that

peaks in IRs will have a greater finite width to them, i.e., they may be spread out over two or more samples. A third program was developed to more closely model measured IRs. A classical modal expansion for rectangular rooms is common and gives a reasonable representation of the room frequency-response function, but the process requires long computation times due to the millions of room modes required for broadband calculations.

The IR can be found by taking the inverse Fourier transform of the frequency-response function resulting from a modal expansion. The windowing function used when computing the IR from the frequency-response function is very important in getting an accurate result. A rectangular window (truncation) can be used, but tends to generate artifacts in the tail of the IR due to its time-domain characteristics. For the case of the DIR, this is not so important, because there is a cutoff in the time domain after which the likelihood of obtaining meaningful results is dramatically decreased (see Sec. 4.10). If a more accurate response is desired, many common windowing functions exist that can improve the response in the time domain.

In order to approximate a classical modal expansion with a reasonable computation time, a third program was written using an adaptation of a modified modal expansion developed by David Nutter [39]. It utilizes the streamlined method to reduce computation time by only summing modes over a limited frequency range near a frequency of interest. This band-limited summation eliminates the need to keep information from very low-order modes at higher frequencies and vice versa. The result is a faster computation using only dominant subsets of modes. One drawback is the amount of memory needed to compute necessary preliminary data tables. A 16-core supercomputer with 128 GB of RAM was used to generate and process the data tables, resulting in a wide-frequency expansion of the rooms being considered.

As a result, the equations for the modal expansion change from

$$\hat{p}(\vec{r}, \vec{r}_0, k) = -4\pi\hat{A} \sum_{N=0}^{\infty} \frac{\Psi_N(\vec{r}_0)\Psi_N(\vec{r})}{V\Lambda_N \left[ k^2 - k_N^2 - jk \left( \frac{\langle \alpha \rangle_s S}{4V} \right) \right]} \quad (4.1)$$

where the summation over  $N$  indicates a three-dimensional summation over  $n_x$ ,  $n_y$ , and  $n_z$ , to

$$\hat{p}(\vec{r}, \vec{r}_0, k) = -4\pi\hat{A} \sum_{N=N_L}^{N_U} \frac{\Psi_N(\vec{r}_0)\Psi_N(\vec{r})}{V\Lambda_N \left[ k^2 - k_N^2 - jk \left( \frac{\langle \alpha \rangle_s S}{4V} \right) \right]} \quad (4.2)$$

where where for each value of  $k$ ,

$$N_L \leq N \leq N_U \quad (4.3)$$

are chosen such that

$$k - \rho\Delta k \leq k_N \leq k + \rho\Delta k \quad (4.4)$$

where  $\Delta k = \frac{2\pi\Delta f}{c}$  and  $\rho$  is the inclusion-window width factor. The summations, being only over a limited frequency band, are then concatenated in frequency to yield the frequency-response function over the entire frequency band of interest.

Because this method excludes many modes at most frequencies, the width of the inclusion frequency window is vital. If not chosen properly, very large errors between the classical and modified modal expansions result in erroneous IRs in the time domain. In order to quantify these errors, a classical modal expansion of a 3 x 5 x 4 m room with a uniform absorption coefficient  $\alpha = 0.1$  was generated to be fully populated up to 5 kHz. The term fully populated means that every mode whose natural frequency is equal to or less than the specified frequency is included. This is done by finding the axial mode numbers for the specified frequency and performing the

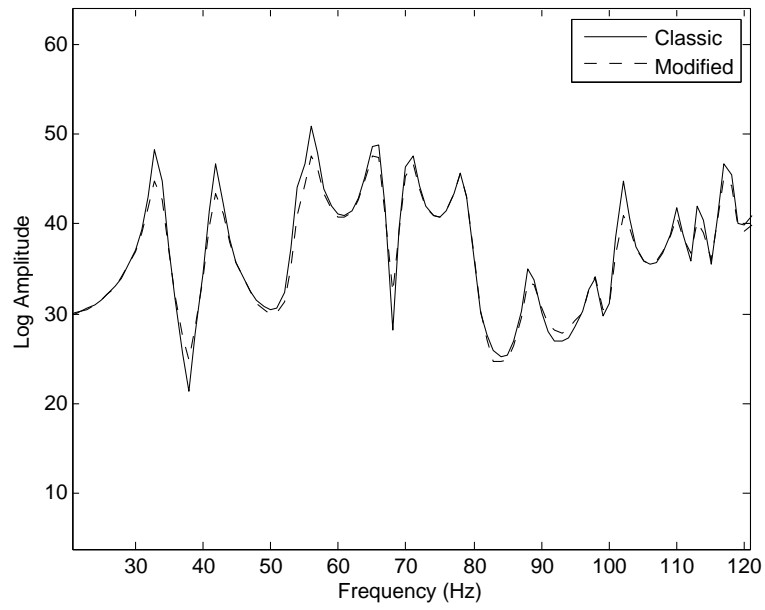


summation up to each of those mode numbers along their respective axis. The modal bandwidth  $\Delta f$  of this room is approximately 2 Hz. In order to determine the errors, the inclusion frequency window in the modified modal expansion was determined as a constant factor times the modal bandwidth. Window widths of 1, 2, 5, 10, 20, 50, and 100 times the modal bandwidth were used in comparative calculations. Figure 4.2 shows overlays of the the frequency-response magnitudes from the classical and modified expansions for 100 Hz bandwidths at both low and high frequencies where an inclusion window width of  $20\Delta f$ . The frequency resolution is 1 Hz, the source strength is unity, and the levels are shown in log scale. Figure 4.3 shows the difference between the classical modal expansion and the modified expansion for the different frequency window widths listed above. Figure 4.4 shows the standard deviation of the error as a function of window width.

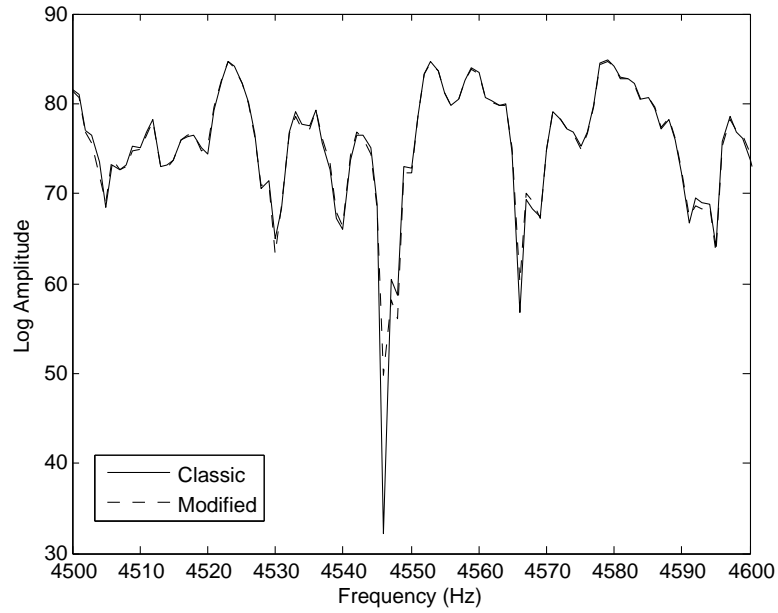
As seen in Figs. 4.2 through 4.4, the modified modal expansion begins to reasonably approximate the magnitude response of the classic modal expansion with frequency window widths of approximately  $20\Delta f$ . In general, the dips in the frequency response are where the two methods differ the most. As seen in Figs. 4.3 and 4.4, the error between the modified modal expansion and the classical modal expansion is reduced as the frequency window width is increased. In fact, if the frequency window width is increased to the Nyquist frequency, the two would match exactly.

In order to further check the accuracy of the modified modal expansion, the same plots as in Figs. 4.2 through 4.4 were generated for the phase of the frequency response functions and are shown in Figs. 4.5 through 4.7.

Figures 4.5 through 4.7 illustrate that phase errors are more significant than magnitude errors, but on the whole, the modified modal expansion gives a very good approximation of the classic modal expansion starting at inclusion window widths of about  $20\Delta f$ .

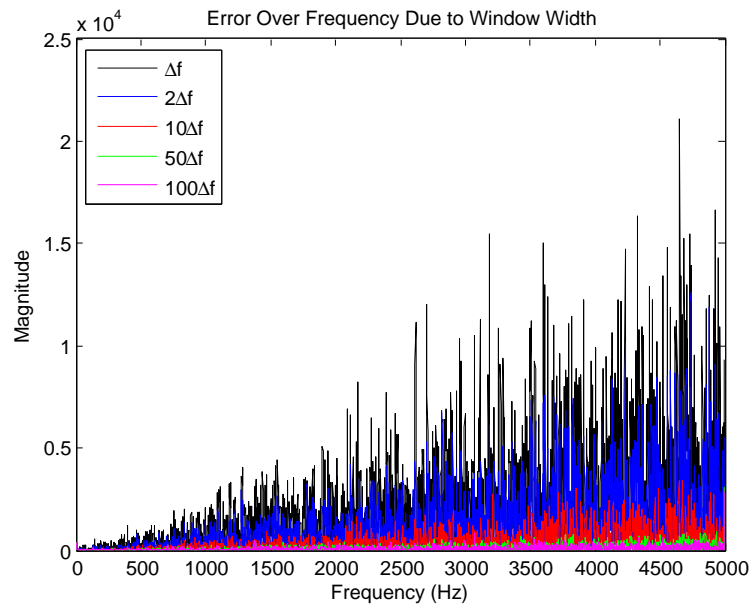


(a)

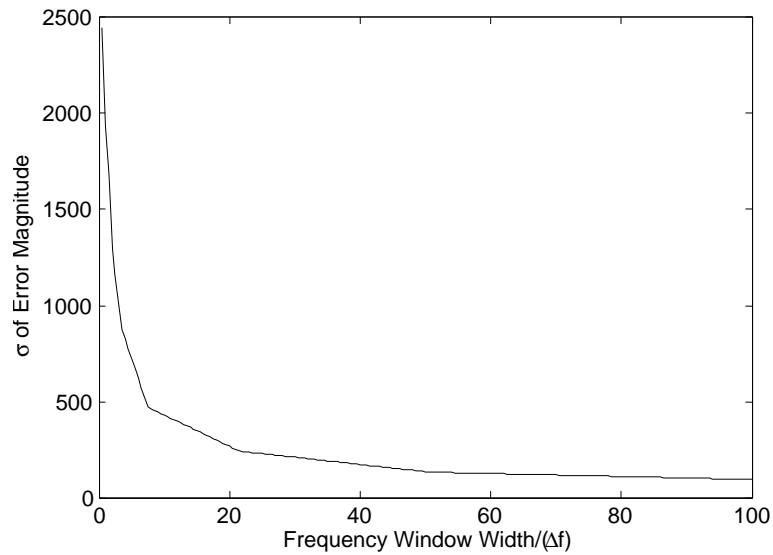


(b)

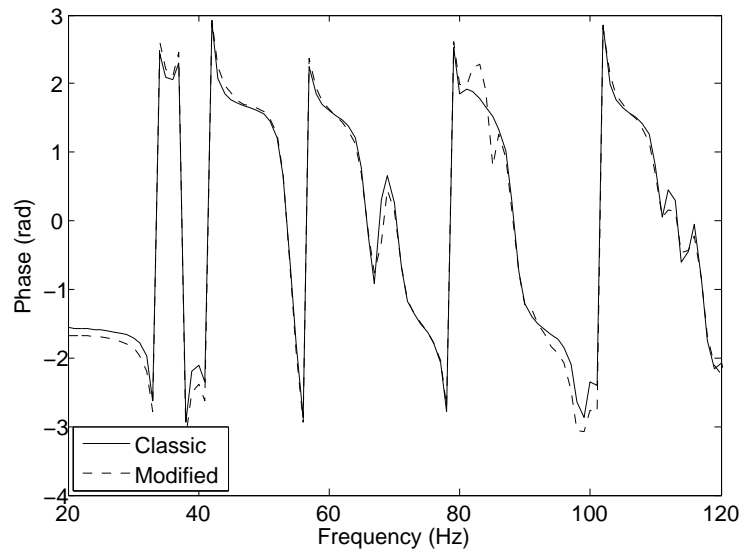
**Figure 4.2** Overlay of magnitude classical (solid) and modified modal expansions at (a) low frequencies and (b) high frequencies for an inclusion window of  $20\Delta f$ .



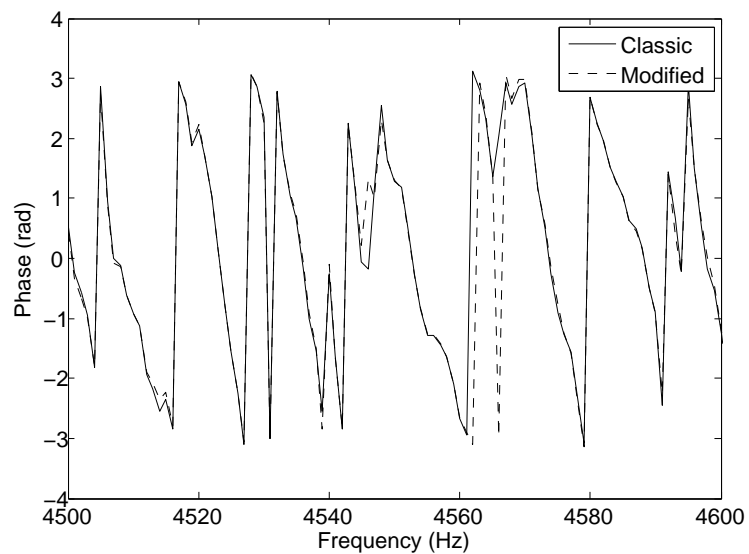
**Figure 4.3** Magnitude error between classical and modified modal expansions for differing inclusion window widths.



**Figure 4.4** Spectral standard deviation of the magnitude error in percent versus inclusion window width.

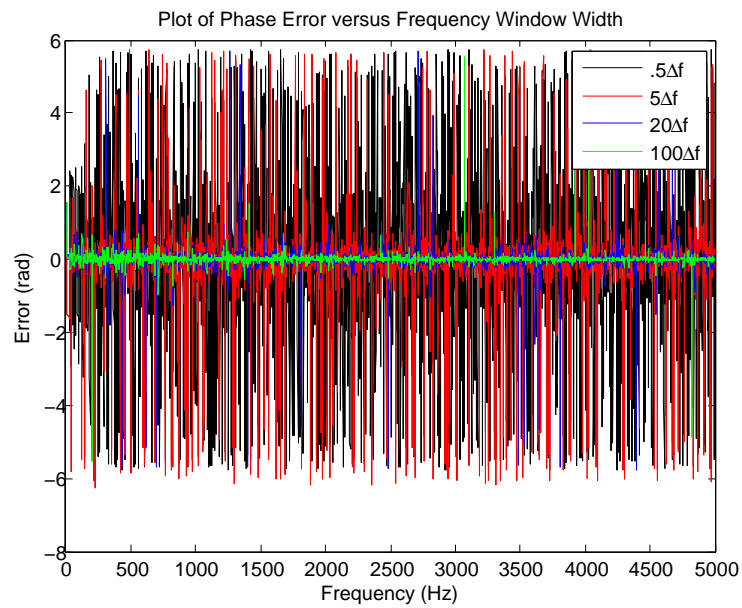


(a)

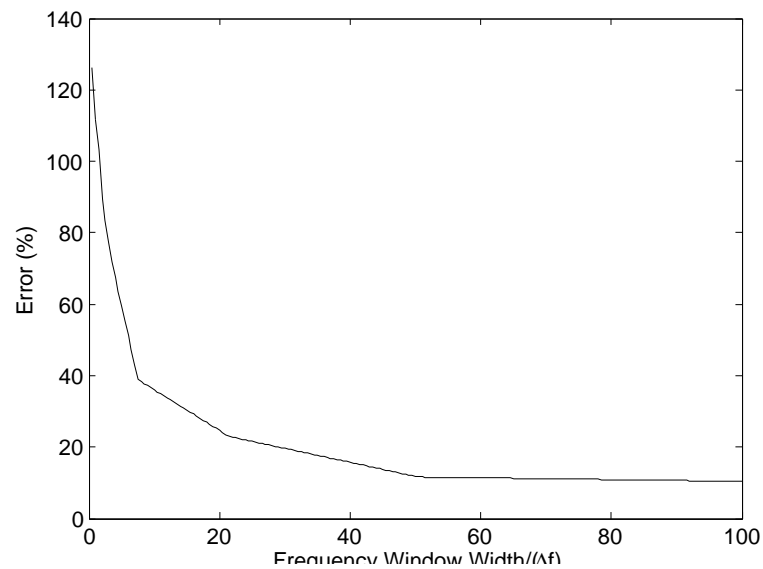


(b)

**Figure 4.5** Overlay of phase of classical (solid) and modified modal expansions at (a) low frequencies and (b) high frequencies for an inclusion window width of  $20\Delta f$ .

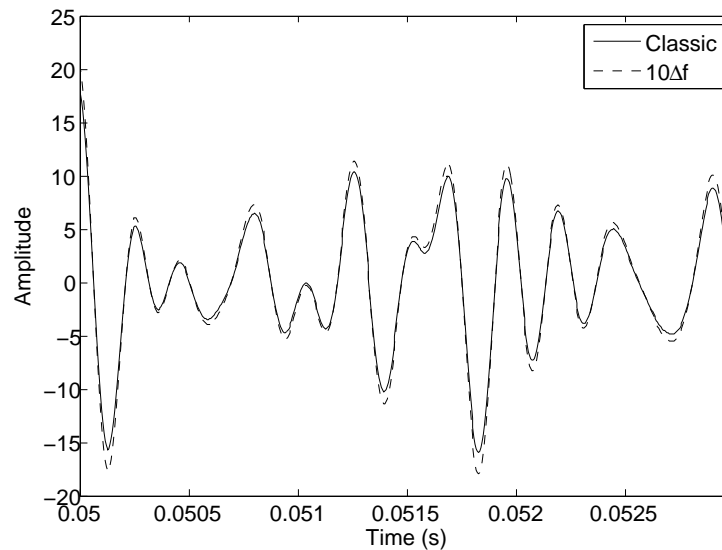


**Figure 4.6** Phase error between classical and modified modal expansions for differing inclusion window widths.



**Figure 4.7** Spectral standard deviation of phase error in percent versus inclusion window width.

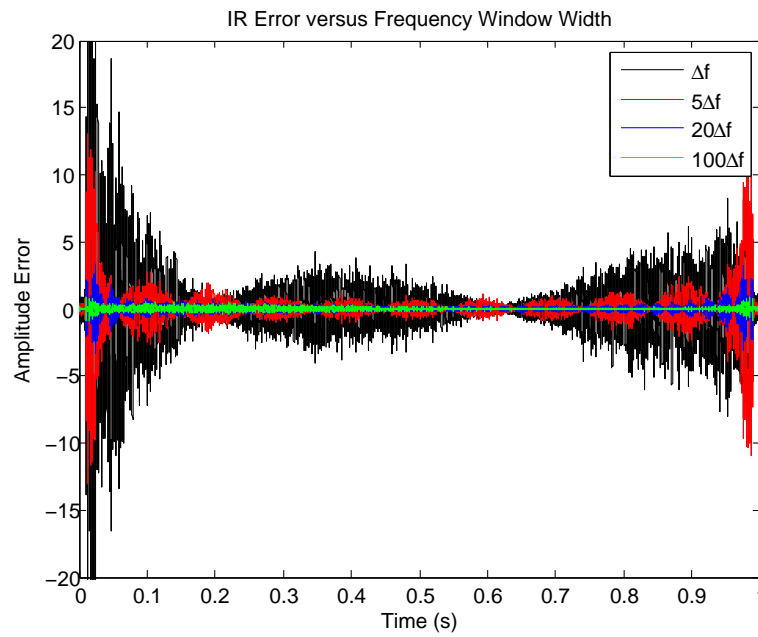
As a final check of the modified modal expansion's accuracy, IRs were generated by both methods for comparison. Figure 4.8 shows good agreement between IRs at relatively early times in the IR. The errors at the end of the tail are artifacts of the rectangular windowing function used in this instance. In reality, they rarely detract from the DIR due to the maximum usable time limit (see Sec. 4.10).



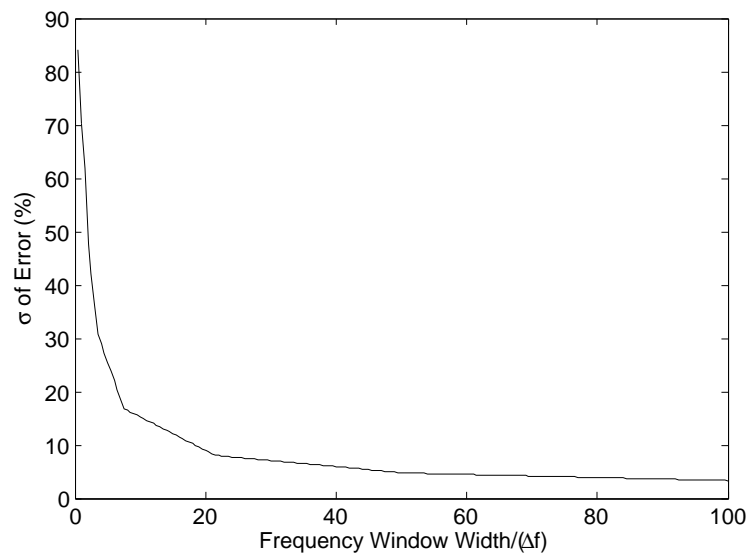
**Figure 4.8** Overlay of IRs generated using classic (solid) and modified modal expansions early in the response.

In short, the modified modal expansion was used to successfully reduce the computational load of producing wide-bandwidth frequency response functions. Depending on room size and population frequency, the modified modal expansion reduced computation time by a factor as great as 300.

For additional comparison, results from the image source method and the modified modal expansion for the simulated room are shown in Fig. 4.11. In addition to the full-bandwidth image source IR, the image source IR after passing through a 15th-order Bessel low-pass filter with a cutoff frequency of 5.8 kHz is also shown in the figure. The cutoff frequency was chosen to be 5.8 kHz to ensure that the response of

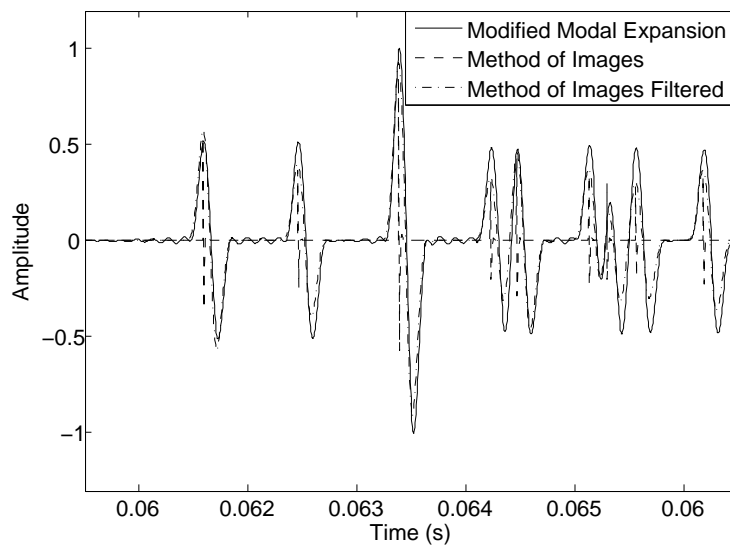


**Figure 4.9** IR error between classical and modified modal expansions for differing frequency window widths.



**Figure 4.10** Standard deviation of IR Amplitude error in percent versus inclusion window width.

the filter was flat up to 5 kHz. It shows that the filtered image source IR and modified modal expansion IR match very well. The only difference is that the amplitude of the filtered image source IR matches the unfiltered image source IR more closely than the modified modal expansion IR. This being the case, it appears to be more beneficial to use the filtered image source IR than the modified modal expansion IR for the purpose considered here.



**Figure 4.11** Overlay of IRs generated using modified modal response, method of images, and low-pass filtered method of images.

With all of these programs available, it is possible to generate the realistic IRs needed to implement the STCM and verify its accuracy and stability. The numerical computations provide efficient controls for preliminary validation.

## 4.5 The Directional Impulse Response Program

A program was created in Matlab to implement the theoretical aspects of the STCM described in Ch. 3. To allow for a clear understanding of the discussion in the following



chapter, several features of the program will be discussed here. A copy of the program can be found in Appendix A.

### 4.5.1 Time Delays and Allowable Error

As discussed in Ch. 3, the time delays between the three axial microphone pairs in the microphone array lie on the surface of a sphere in time delay space. In the continuous-time domain, one could filter out any result that differed, even slightly, from the maximum time delay value and would thus be left with only the reflections that arrived *exactly* at the time of interest. However, in the discrete-time domain there will be some ambiguity. An allowable deviation from the composite time delay value can account for most errors that occur during digitization. The deviation  $\delta$  from the time delay is given as a fraction of the radius of the ideal sphere in time delay space. This can be expressed as

$$\delta = \frac{2dc}{\beta}, \quad (4.5)$$

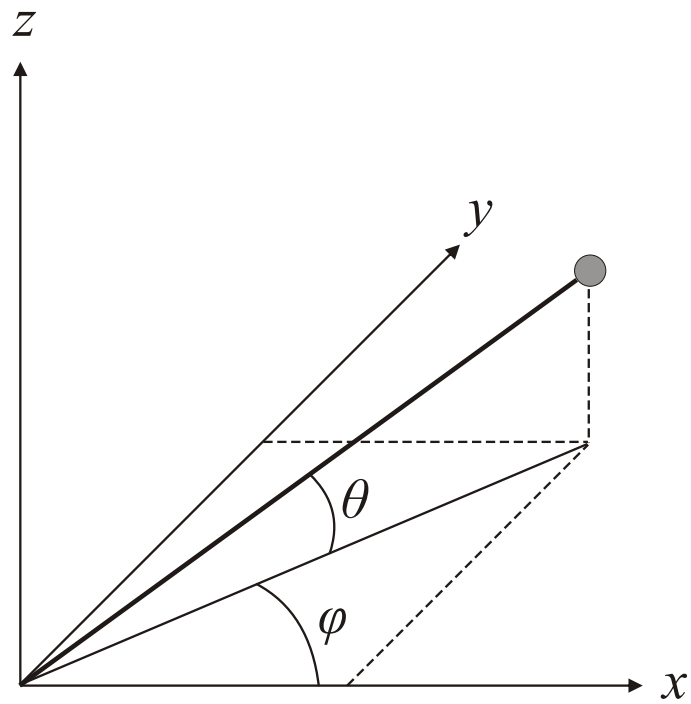
where  $d$  is the distance from the origin to each microphone,  $c$  is the speed of sound, and  $\beta$  is the fraction of the original time delay. For example, if  $\beta = 100$ , the program would allow an error of +/- 1 % of the maximum time delay value.

Choosing an appropriate allowable error is very important in this program. If the allowable error is too large, the program will begin including arrivals that come either earlier or later than the reflections that are to be singled out and arrivals that don't physically exist. If the allowable error is too small, the program will begin excluding peaks that may have arrived at the appropriate time, but due to the sampling-induced error, were assigned values that deviated from the desired time delay value by more than the specified error.

A general guideline for selection of an appropriate time delay error is given as follows. If the peak to be inspected is relatively early in the IR, the user should specify a greater allowable error. If the peak of interest is later in the IR, a smaller allowable error should be specified. The selection of a larger allowable error for early peaks is due to fewer peaks in the cross correlation caused by the (presumably) lower reflection density of the IR at that point. A larger allowable error allows the program to extract the reflections that arrive at the same time while not excluding any valid reflections. Conversely, when a peak is later in the impulse response, the reflection density is much higher and there will be many peaks in the cross correlation which, with the large allowable error, will result in a very high likelihood of many erroneous solution sets that meet the given constraint. Tighter time constraints should be specified for this reason. The selection of the best allowable error is a matter of experimentation, as each new room and source/receiver configuration is unique.

### 4.5.2 Selection of Peaks in the Cross Correlation

Another issue critical to the successful use of the Matlab program is the selection of peaks in the cross correlation function. After a peak of interest in the IR is selected for further examination, the program displays the cross correlation graphs for each of the three Cartesian axes. All angles calculated by the program are relative to these axes and their assignment to the physical realm is arbitrary. It is recommended, however, that the user choose and consistently use an intuitive configuration (e.g., microphone 1 is in the  $+x$  direction or the ceiling of the room is in the  $+z$  direction). It should also be noted that the program operates under the assumption of a right-handed coordinate system where  $\theta$  is the elevation from the horizontal plane and  $\phi$  is the azimuth measured in the counterclockwise direction from the  $+x$  direction (see Fig. 4.12).



**Figure 4.12** Definition of angles as computed by the DIR.

Once the three cross correlation graphs are displayed, the program will prompt the user to select a number of peaks in each cross correlation graph. To ensure the highest likelihood of including all valid arrivals, all peaks that are strong relative to the noise floor should be chosen. In the case of flattened peaks, one must decide if the sampling rate is sufficient to indicate a single arrival or not. If the peak is flattened and the sampling rate is high, then the peak should be chosen by interpolation, which is typically done by the user, though an automated peak detection algorithm could also be programmed to do the interpolation.

One feature that may be added in the future is automation of peak selection. An algorithm would need to be developed that would select all peaks significantly above the noise floor. However, it may prove difficult for the program to automatically determine what the noise floor is, especially when the program is operating in regions of high reflection density long after the direct sound arrival.

### **4.5.3 Relative Amplitude and Dominant Reflections**

Once all the cross correlation peaks have been selected, the program uses the corresponding time delay values to calculate the time delay vector magnitude of every possible combination. These magnitudes are filtered using the allowable time delay error, leaving only the values within the specified error. These valid sets of time delay values are then used to calculate the spherical angles of arrival relative to the coordinate system in use.

Because the STCM is to be used as a diagnostic tool, it would be very useful to determine the relative strengths of the sound arrivals. This information would allow one to determine which surfaces need to be treated most. One possible solution is beamforming. A beamforming scan over an entire solid angle would show areas with the greatest energy concentration. The drawback, however, is that the limited

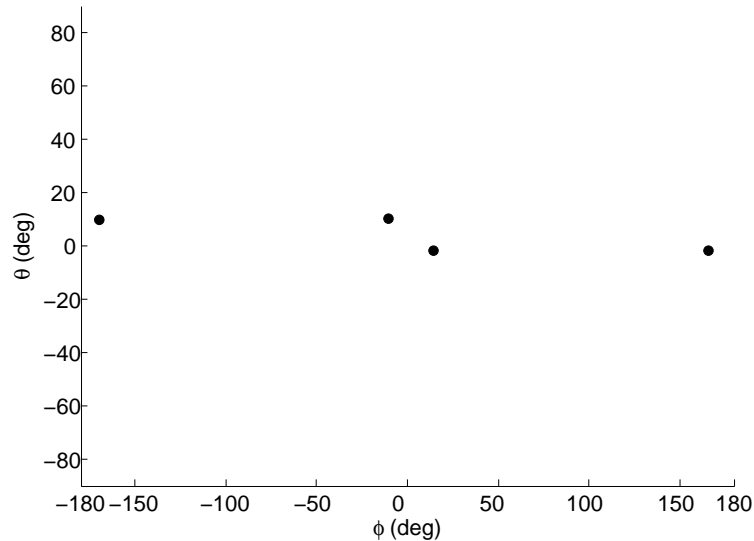
number of microphones in the array would result in beams that are rather wide and thus unable to identify which of two closely spaced arrivals is stronger.

Another possible method for determining the relative strength of arrivals is to use the amplitudes of the peaks in the cross correlation function. Peak amplitudes are related to the cross power, or the power common to both signals for a given time delay. One can assume that the greater the combined cross power from the three axes, the stronger the particular arrival pertaining to those three time delays. The problem here is that any erroneous results may show a very strong relative amplitude even though they are not physically present.

#### 4.5.4 Display of the Results

After the relative amplitudes are calculated, the program results are displayed in two ways. First, the program displays in tabular form the values of  $\theta$  and  $\phi$  for each sound arrival. The second display is a scatter plot of the arrivals where the axes represent  $\theta$  and  $\phi$  (see Fig. 4.13). The size of the circle displayed can be assigned such that it is directly proportional to the relative amplitude of the arrival (i.e., a larger circle indicates a stronger sound arrival) using one of the methods discussed above, though that has not been done in this case.

Once the results are displayed, the program is finished and can be run again to examine another peak in the IR. The process provides a quick and in-depth analysis of a room's reflection profile at the array location and can be done accurately over the course of just a few minutes.



**Figure 4.13** Scatter plot showing arrival angles for a long, narrow room for  $t = 90$  ms.

## 4.6 Matching of Microphones

The matching of microphones (primarily phase matching) is critical when working in an experimental setting. The effects of both phase and magnitude matching will be discussed briefly.

### 4.6.1 Phase Matching

Matching microphones for phase is critical when using the STCM, as it is based on the cross correlation function. If we consider two signals with unity magnitudes but with one signal phase shifted relative to the other, it is easy to see how the complex cross spectrum will differ. The phase difference translates directly into a time delay when the cross spectrum is transformed into the time domain. If not corrected, this time delay is incorporated into the cross correlation, resulting in erroneous values. The amount of phase difference between the microphones thus impacts the amount

of error in the resulting angles of arrival. To determine the amount of error, let us consider two signals between which we wish to estimate the time delay. If the signals are transformed to the frequency domain, we can represent the complex amplitudes as

$$\hat{p}_1 = p_{pk,1} e^{j\gamma_1} \quad (4.6)$$

and

$$\hat{p}_2 = p_{pk,2} e^{j\gamma_2} \quad (4.7)$$

where the  $pk$  subscript denotes the peak value of the signal. The relative phase between the two is

$$\gamma = \gamma_1 - \gamma_2. \quad (4.8)$$

This phase difference translates into a frequency-dependent relative time shift in the time domain according to the equation

$$\gamma = \omega\tau \quad (4.9)$$

where  $\tau$  is the relative time shift and  $\omega$  is the frequency in rad/s. The error generated by this relative time shift is dependent not only on frequency, but also on the arrival angle, since the important factor is the percent time shift caused by the phase mismatch. Rearranging Eq. (4.9) to solve for the relative time shift we obtain

$$\begin{aligned} \tau &= \frac{\gamma}{\omega} \\ &= \frac{\gamma}{2\pi f} \end{aligned} \quad (4.10)$$

$$= \frac{\delta}{360f}, \quad (4.11)$$

where  $\delta$  is the phase shift in degrees and  $f$  is the frequency in Hz. For the case of a microphone array with a diameter of 7.62 cm, the maximum time delay is 222  $\mu\text{s}$ . If we arbitrarily choose a uniform  $10^\circ$  phase shift between microphones, Eq. (4.11) can be simplified to

$$\tau = \frac{1}{36f}. \quad (4.12)$$

As seen, the relative time shift is reduced as the frequency is increased. However, for the 7.62 cm-diameter array, there is a time shift of more than 10% the maximum time delay value up to frequencies around 1.25 kHz. Exactly what error this will create in a measurement setting depends on the arrival angles and frequency content in the system, as a system with less high-frequency content will be even more susceptible to the phase shift errors. In addition, the phase difference between microphones is rarely constant over frequency and the errors will thus be governed by the frequency-dependent phase difference relationship. As a simple example, if an arrival has arrival angles of  $\phi = 30^\circ$ ,  $\theta = 0^\circ$  and the  $+x$  microphone has a  $10^\circ$  phase shift relative to the  $-x$  microphone, at 2.5 kHz the relative time shift will be 11.1  $\mu\text{s}$ . This will produce an arrival angle of  $\phi = 28.63^\circ$ . While this may not seem like much of an error, a  $10^\circ$  phase shift is a relatively mild difference and frequencies below 2.5 kHz will produce greater relative time delays. Because of the large number of variables that affect how great the errors due to phase mismatch are, a comprehensive quantitative investigation of the errors is outside the scope of this work.

In order to ensure accurate results, the phase differences between microphones must be accounted for. This can be done in a variety of ways. One is to use a single microphone with sequential measurements. Assuming no drastic change in microphone or room conditions (i.e., time invariance), the microphone is perfectly matched to itself from one measurement to another (see Sec. 4.9). This solution,



however, requires more time and much care in measurement setup to ensure correct spacings, etc.

A second way to compensate for phase differences is to use a matched set of microphones. However, matched microphones are typically much more expensive than unmatched microphones.

A third method is to use microphones that may not be matched but that are relatively calibrated (e.g., using the switching technique [40]). The calibration correction can be done as a post processing operation after the IRs are taken but prior to the calculation of the cross spectrum and cross correlation.

### 4.6.2 Magnitude Matching

The matching of magnitudes is less important when working with the DIR. To understand why this is true, consider two frequency-domain signals that are matched in phase, but have differing magnitudes,  $|\hat{A}_1|$  and  $|\hat{A}_2|$ . When the signals are multiplied for the cross spectrum, the resulting magnitude will be  $|\hat{A}_1||\hat{A}_2|$ . Because the Fourier transform is a linear operation, the cross correlation will also have a resulting magnitude  $|\hat{A}_1||\hat{A}_2|$ , as opposed to the magnitude of  $|\hat{A}_1|^2$  for perfectly matched signals. This is simply the difference of a constant in the cross correlation amplitude, which is present for the cross correlation in each direction and thus does not have an impact on the angular results. The cross power measurement may be off, but the relative cross power will still be intact.

## 4.7 Array Dimensions and Geometry

Another issue to consider is the array itself. Determining the size and geometry of the array is a delicate balancing act requiring one to identify which aspects are most

important. The two aspects of the array to be discussed here are its size and geometry.

### 4.7.1 Array Dimensions

The dimensions of the array have a noticeable impact on the DIR's performance and are a balance between two conflicting consequences: the maximum inherent angular error due to time window size and the inclusion of samples that correspond to arrivals in adjacent IR peaks when analyzing a peak of interest.

The maximum inherent error is simply the maximum error due to sampling and is a result of the number of samples in a given array time window (cross correlation time window). In practice, it is simply the error generated by having a one-half sample error in the time-domain quantization. The maximum value occurs when there is a half sample error about a zero time delay value in the cross correlation function. It can be quantified by taking the arcsin of one half of a sample divided by the number of samples per cross correlation time window,

$$E_0 = \sin^{-1} \frac{1}{2M}, \quad (4.13)$$

where M is the number of samples in a given cross correlation time window. The number of samples in a given cross correlation time window is simply the the product of the time it takes for sound to propagate along the axis and the sampling rate. Mathematically, M can be expressed as

$$M = f_s \Delta t_{max} \quad (4.14)$$

Substituting the expression for M in Eq. (4.14) into Eq. (4.13) yields

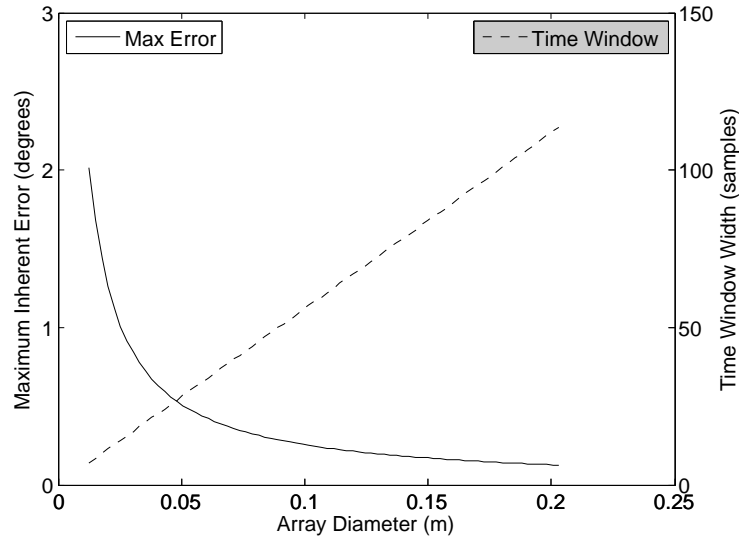
$$E_0 = \sin^{-1} \frac{1}{2f_s \Delta t_{max}} \quad (4.15)$$

From this, one can see that more samples should be included in a given time window to reduce the error. This can be done by either increasing the sampling frequency or increasing the array size. However, if we assume that the sampling rate is always the highest possible, the only way to further reduce this error is to increase the array size to include more samples for a cross correlation time window.

Though more samples reduces the error inherent in using a digital signal, it can also create more serious problems. When more samples are included in a time-windowed IR, the number of adjacent samples about the peak of interest is increased. This means that more data is included in the cross spectrum and therefore the cross correlation function. The increased number of samples means an increased likelihood of including information pertaining to peaks that do not compose the peak of interest. This is particularly problematic when using the STCM in a portion of the IR long after the direct sound. When the peak density is very high, it is beneficial to reduce the number of samples in the time window, excluding information that does not belong to the arrivals composing the peak of interest. This reduces the statistical likelihood of producing erroneous solutions that meet the constraints and therefore improves the reliability and usefulness of the STCM at these later times.

Figure 4.14 illustrates the conflicting nature of these two quantities. Here a comparison of both the inherent angular error and the number of samples in a time window as a function of array size (in inches) for a seven microphone array.

For the purpose of the numerical simulations performed herein, an array diameter of .075 m was chosen to reduce the number of samples in the time window (relative to the 0.1 m array used to develop the code) while still keeping the inherent angular error to an acceptably low level. The curves in Fig. 4.14 appear to intersect at an array diameter of approximately .05 m, but in numerical simulations run with that diameter, there appeared to be an increase in angular error while an increase in



**Figure 4.14** Plot of maximum inherent angular error (degrees) and window width (samples) versus array diameter for  $f_s=192$  kHz.

performance with erroneous sources was not found.

Another aspect of the array diameter that can affect the functionality of the STCM is how the array diameter relates to the distance between the array origin and reflecting surfaces. To increase the accuracy and robustness of the DIR, results from several different microphone pairings are compared. The correlation centers, or the geometric locations between two microphones, are shifted slightly from the geometric origin of the array when this is done. To visualize this, consider a pair of microphones with Cartesian coordinates  $(d, 0, 0)$  and  $(-d, 0, 0)$ , respectively. The geometric center is simply the average of each Cartesian component. In this case the geometric center is  $(0, 0, 0)$ . Now consider two microphones with Cartesian coordinates  $(d, 0, 0)$  and  $(0, d, 0)$ . The geometric center in this case is  $(\frac{d}{\sqrt{2}}, \frac{d}{\sqrt{2}}, 0)$ , which is clearly not the origin.

To accurately compare the results from different correlation pairings of microphones, we assume that the distance from reflecting surfaces to the array must be

much greater than the array diameter. This is to ensure that the error due to different cross correlation centers is negligible and thus the arrival angles relative to the different correlation axes are equal. If the spacing between microphones is  $d$  and the surface-to-array distance is  $r$ , we can represent this in equation form as

$$r \gg 2d, \quad (4.16)$$

for a six or seven microphone array and as

$$r \gg d, \quad (4.17)$$

for a four microphone array.

### 4.7.2 Array Geometry

The geometry of the microphone array is a property that allows for even more precise operation of the STCM if properly selected. The seven microphone array used here was chosen for very specific reasons. One of its main advantages over other array geometries is that the origin microphone allows for the calculation of arrival directions using multiple independent microphone sets. When the angles of arrival are found using these independent sets, the results from each set can be compared to the other sets' results. Doing this enables further reduction of erroneous solution sets because only the solutions common to all microphone configurations are kept. While it reduces the number of erroneous sets, it does not eliminate erroneous sets resulting from time delays that meet the constraints as discussed earlier. For the purpose of this research, three different cross correlation microphone pairings were used. The first consisted of the  $+/-x$ ,  $+/-y$  and  $+/-z$  microphones. The second pairing used the origin microphone along with the  $+x$ ,  $+y$  and  $+z$  microphones to produce the cross correlation along each

axis while the third pairing used the origin microphone in conjunction with the  $-x$ ,  $-y$  and  $-z$  microphones. Other pairings are possible. For example, the  $+x$  microphone could be paired with the  $+z$  microphone, but these other pairings are beyond the scope of this research.

While the research concerned here uses a Cartesian microphone array, it is not required. The only requirement is that the array geometry be known and a fixed coordinate system be defined relative to the array. If another array geometry is chosen, the equations for arrival angles (see Eqs. (3.5) and (3.6)) must simply be reworked to represent the new physical situation. The one problem when using an alternate array geometry is that alternate constraint equations applying to the geometry being used must be found.

## 4.8 Microphone Directivity

When using the DIR, the microphones are assumed to be omnidirectional, which is not always the case. Even type 1 precision microphones begin exhibiting nonuniform directivity at high enough frequencies. These nonuniformities are primarily due to diaphragm sizes and diffraction/scattering effects of the body when the wavelength of the impinging sound is on the order of the smallest of the microphone dimensions.

If the effects of these nonuniformities are determined to be appreciable, post processing using the microphone's measured directivity may not be possible since it is unknown which direction specific arrivals came from. The only real solution is to low-pass filter the signal so that frequencies higher than the uniform directivity are not used in computation of the cross spectrum and thus the cross correlation itself. Low-pass filtering, however, introduces its own problems (see Sec. 5.8.3).

## 4.9 Subsequent versus Simultaneous Measurements

As mentioned earlier, measurements at the microphone positions can be taken either simultaneously or subsequently, depending on whether the user requires short measurement time or a minimum of equipment. Of course, this assumes that the system being measured is time-invariant, i.e., there is no significant change in vital parameters or the noise floor between measurements. Noise sources that can produce these changes include HVAC systems, noisy lighting, strong mechanical vibrations, opening and closing of doors, people in the room, etc. Some vital parameters include the speed of sound (which is temperature and humidity dependent) and the reverberation time (which can be affected by people or objects entering or leaving the room between measurements) as well as source and/or receiver locations.

As a measure of this effect, two microphones were positioned arbitrarily in an anechoic chamber (for clarity, the microphones are labeled microphone A and microphone B, where the labels are arbitrary, but do not change from one set of measurements to the next). The IR was measured at each of the microphone locations using a single loudspeaker excitation. The IR was then measured again a short time later using the same source and receiver locations. The cross correlation was measured between microphones A and B using simultaneous IRs and then measured using the IR from microphone A at the early measurement time and the IR from microphone B at the later measurement time. The results are shown in Fig. 4.15. The cross correlation using IRs measured at two different instances in time is so similar to the cross correlation measured at the same instance that it had to be zoomed in to a very small window in order to see any difference at all. While the anechoic chamber is a very controlled environment, it is clearly acceptable to use subsequent measurements from a practical standpoint as long as the assumptions of system time invariance are not

violated.

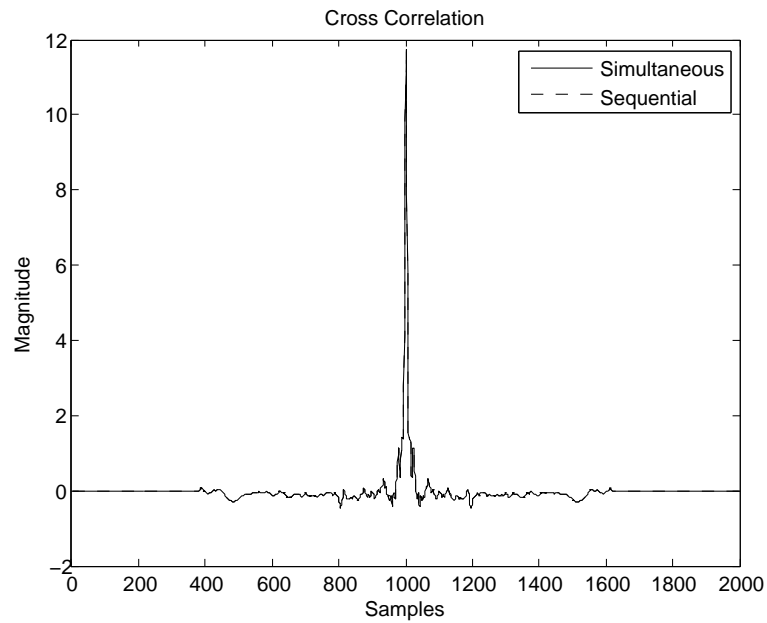
One particular advantage of the seven-microphone Cartesian array is in the use of simultaneous measurements. A large portion of multichannel audio interfaces on the market today have eight inputs. This would allow for recording all seven microphone signals and the excitation signal at once. The addition of the excitation signal allows for the use of any type of excitation signal provided it contains all frequencies of interest and allows for a post processing deconvolution. This means that a dedicated acoustical measurement package is not required, though it is beneficial. The ability to record all measurements at once further increases the flexibility of the STCM by requiring only minimal equipment but allowing for rapid measurements with more equipment if time is a factor.

## 4.10 Maximum Usable Time Length

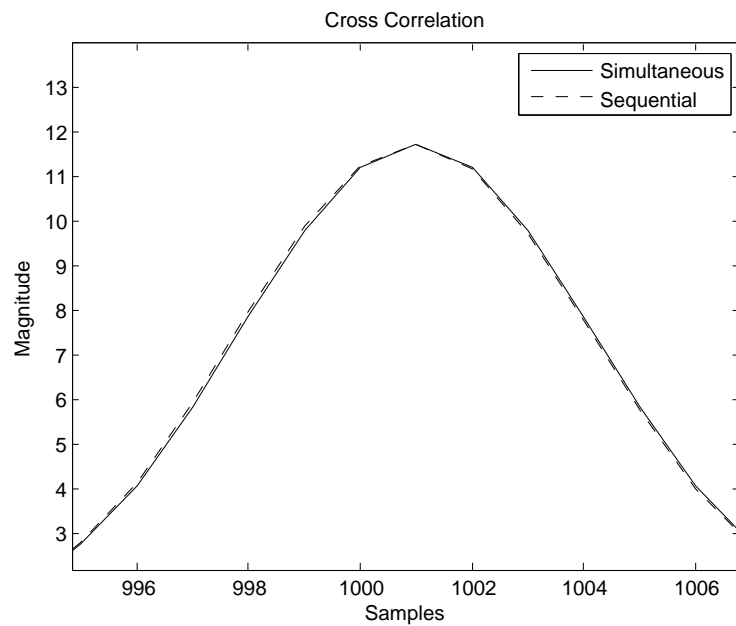
One final consideration is the maximum usable time length of the STCM after the direct sound arrival. As previously discussed, high reflection-density portions of IRs long after the direct sound arrival can create problems for the STCM. The primary source of these problems is that the cross correlation functions become much less well defined and the number of peaks in the cross correlation increases dramatically with high reflection density. This results in a large number of possible solutions and a high statistical probability of combinations not relating to actual arrivals matching the constraints, which results in predicted arrivals that don't exist.

In order to mitigate these possible issues, several things can be done. First, some arrivals can be eliminated by simply observing their relative amplitudes in comparison to those of the other calculated arrivals. Another solution is to compare the results from one microphone correlation configuration to those produced by another. For





(a)



(b)

**Figure 4.15** Cross correlation using simultaneous and subsequent IR measurements for (a) the full time-record length and (b) zoomed in to show detail

example, consider the seven microphone array. The results of the  $+/-x$ ,  $+/-y$  and  $+/-z$  microphone correlation pairs can be compared to those produced by the pairings of the  $+x$ ,  $+y$  and  $+z$  microphones with the origin. If more robustness is desired, additional cross correlation pairings can be used and compared to ensure a high degree of confidence in the results.

The final solution for mitigating the issues is to avoid using the STCM too far out in the IR. Many factors determine the maximum usable time, and that time is unique to each room. It depends on room volume, mean path length, and reverberation time. A more in-depth treatment of this issue follows.

#### 4.10.1 Acoustic Arrival Density

Because the arrival density is the primary cause of the STCM being limited to an early portion of the IR, its value is important to predict. If we consider a rectangular room and an IR created using the image source method, the number of reflections within a given time window can be represented mathematically. A sphere with radius  $r_1$ , where  $r_1 = c(t + \Delta t)$ , contains all of the room reflections up to the time  $t + \Delta t$ , where  $t$  is a given time after  $t = 0$  in the impulse response,  $\Delta t$  is one half the time window width (the signal is windowed from  $t - \Delta t$  to  $t + \Delta t$ ) and  $c$  is the speed of sound. We can also state that

$$nV = \frac{4}{3}\pi r_1^3, \quad (4.18)$$

where  $n$  is the number of reflections up to the specified time in the IR and  $V$  is the volume of the room. If we then consider a smaller sphere of radius  $r_2 = c(t - \Delta t)$ , we

can state that

$$(n - N)V = \frac{4}{3}\pi r_2^3, \quad (4.19)$$

where  $N$  is the number of reflections contained between the times  $t - \Delta t$  and  $t + \Delta t$ . Solving Eqs. (4.18) and (4.19) for  $N$  yields

$$N = \frac{4\pi}{3V} (r_1^3 - r_2^3). \quad (4.20)$$

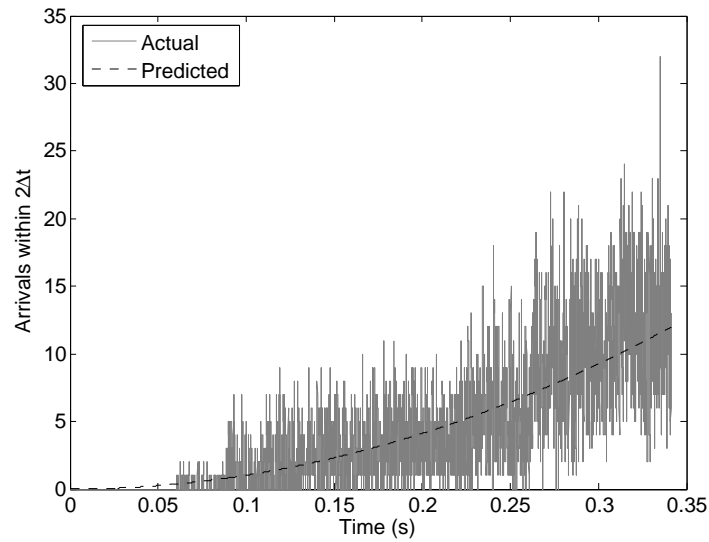
If we substitute for the values of  $r_1$  and  $r_2$  in terms of  $c$ ,  $t$ , and  $\Delta t$ , and factor out  $c$ , we obtain

$$N = \frac{4\pi c^3}{3V} [(t + \Delta t)^3 - (t - \Delta t)^3]. \quad (4.21)$$

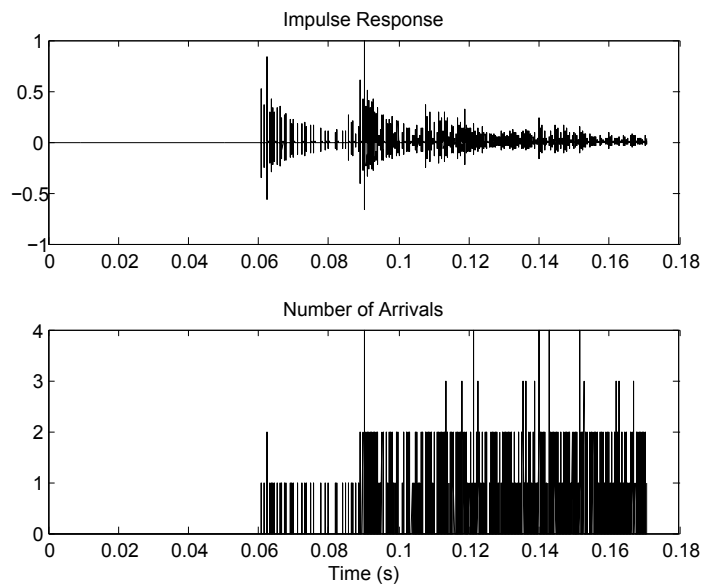
To validate the final equation, the image source code mentioned earlier in this chapter was modified to compute the actual arrival density for a window width corresponding to a 7.62 cm array diameter for a modeled room with dimensions of 30 x 8 x 4.5 m and absorption coefficients of  $\alpha_{walls} = 0.2775$ ,  $\alpha_{floor} = 0.51$ , and  $\alpha_{ceiling} = 0.37$ . A plot showing the result, along with the predicted arrival density from Eq. (4.21), are shown in Fig. 4.16. The actual arrival density varies between values much greater and much less than the predicted value, but on average, the predicted value follows the actual value trend. The IR for the room, along with a histogram of arrivals per sample are shown in Fig. 4.17.

### 4.10.2 Empirical Results

While the results from the predicted arrival density are promising, it is difficult to obtain an expression for the maximum usable time length by simply manipulating the



**Figure 4.16** Plot showing measured and predicted arrival densities for a rectangular room.



**Figure 4.17** Impulse response for a numerically modeled room 30 x 8 x 4.5 m with nonuniform  $\alpha$ .

equation. To determine a maximum usable time that is simple to calculate, simulations were run to approximate the usable time length for several rooms, including the room presented in the previous section and those presented in the following chapter. From these examples, the arrival density at the maximum usable time length was found to be within the range of  $1.5 \leq N \leq 2.5$ . If a nominal value of  $N = 2$  is chosen and if  $t$  in Eq. (4.21) is replaced with  $T_{max}$ , we obtain the following expression for the maximum usable time which is

$$T_{max} \approx \frac{\sqrt{3\pi\Delta tc(3V - 4\pi\Delta t^3 c^3)}}{6\pi\Delta tc^2}. \quad (4.22)$$

This is the approximate time in seconds to which the STCM is valid before too many erroneous solution sets are produced. For the seven-microphone array, the quantity  $\Delta tc$  is  $2d$ , where  $d$  is the distance from the origin to each microphone. Equation (4.22) can then be rewritten as

$$T_{max} \approx \frac{\sqrt{6\pi d(3V - 32\pi d^3)}}{12\pi dc}. \quad (4.23)$$

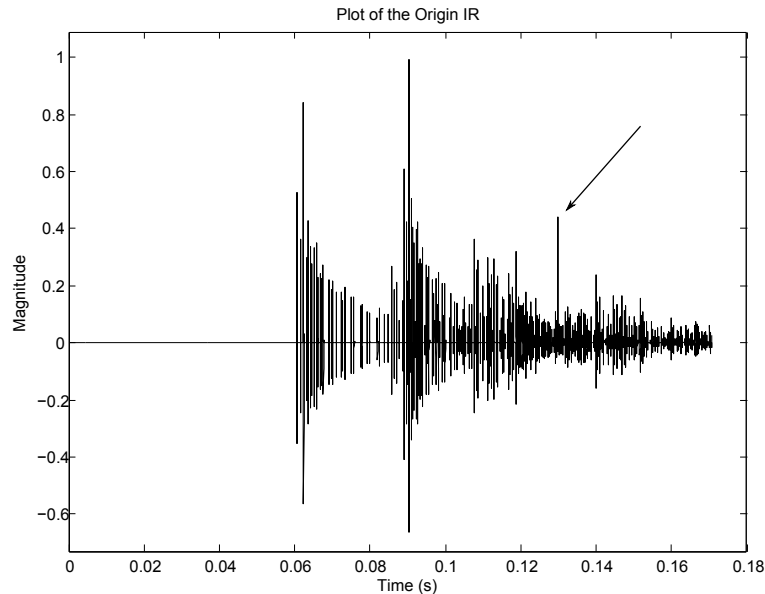
This equation is based solely on the arrival density in the room, which is an estimate of how many arrivals are present in a given array time window for a given room size. It does not take into account the fact that many of those arrivals may have reduced amplitudes due to absorption of room surfaces. When this is the case, it may be possible for the STCM to function in regions for which  $N > 2$ .

Again, it should be noted that Eqs. (4.22) and (4.23) do not represent an exact temporal cutoff, but rather a general guideline for the maximum time in the impulse response for which the STCM may be valid. It is possible that the STCM could be used for much longer time ranges depending on unique room characteristics. The final empirical formulation may not be as valid for differing array sizes or geometries

as it is for the current array size of 7.62 cm and its Cartesian geometry. For other cases, one should determine an appropriate  $N$  value through experimentation.

### 4.10.3 Strong Late Arrivals and Other Exceptions

One exception to the maximum usable time length given above should be noted: the case of a very strong, late arrival. As an example, consider the IR shown in Figure 4.18. It is identical to the IR shown in Fig. 4.17 with the addition of a strong late arrival at  $t = 130$  ms. Using Eq. (4.22) to find  $T_{max}$  produces a maximum usable time of only 110 ms, which is significantly earlier (relative to the direct sound arrival).



**Figure 4.18** Impulse response for a numerically modeled room 30 x 8 x 4.5 m with nonuniform  $\alpha$  and an artificially-added, strong late arrival.

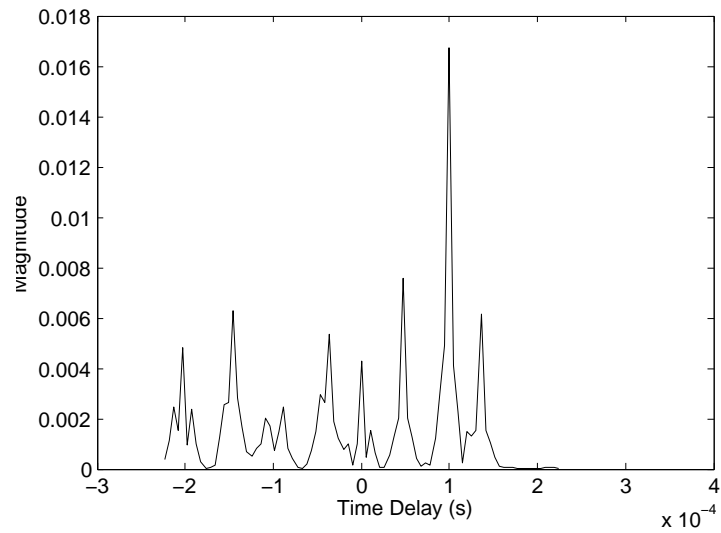
If we now consider the cross correlation in only one direction for the IR with and without the large spike at  $t = 130$  ms, we get the plots shown in Fig. 4.19. This clearly shows that the presence of a large peak in the IR results in a large peak in the

cross correlation because there is a greater amount of cross power associated with the stronger arrival than for the other peaks. Because this peak is so pronounced, it is easy to extract the correct angles of arrival. In general, the presence of a few strong peaks late in the IR allows for the continued use of the STCM in these situations.

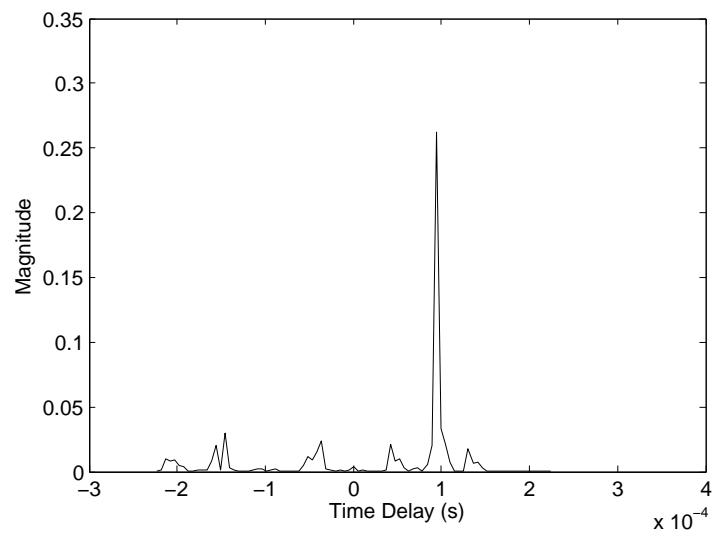
In addition to a strong late arrival, there may be portions of the IR that are not well populated, even long after the direct sound arrival. In these cases, the reflection density is low enough for the STCM to be used effectively as well. This is another exception to the rule of thumb for the maximum usable time stated in the previous subsection.

## 4.11 Chapter Conclusions

Many of the practical issues for using the STCM have been discussed. Despite the large number of considerations, once the parameters have been established, the STCM requires very little setup and maintenance. Many of the factors discussed here are a matter of experimentation for any specific case. For example, the allowable error in the constraint equations will typically be a matter of tweaking the constraint factors for any given room and any given time after the direct sound arrival. Even with all of these considerations, the following chapter will show that the STCM is still accurate and useful.



(a)



(b)

**Figure 4.19** Cross correlation along the  $x$ -axis for the case of (a) the normal IR and (b) the IR with an artificially added large spike at  $t = 130$  ms





# Chapter 5

## Numerical Verification, Experiments, and Results

Using the developments of Chs. 3 and 4, a Matlab script was written to implement the STCM using numerical IRs (see Appendix A). A seven-microphone Cartesian array was used in all cases. The IRs were generated using the programs described in Ch. 4 for several different scenarios ranging from a single source in free space to more complicated, numerically simulated rooms. All of the IRs were bandlimited at the Nyquist frequency (96 kHz) to eliminate aliasing. The speed of sound was assumed to be 343 m/s and all receivers and sources were ideal point receivers and sources. In addition, all FFTs were the same length as the truncated (time-windowed) IRs used for calculating the cross spectrum and cross correlation.

In addition to the numerical verification of the STCM, some elementary experiments were conducted to assess its usability in more realistic settings. Each IR was recorded at a sample rate of 192 kHz and bandlimited to the Nyquist frequency to eliminate aliasing. The IRs measured experimentally were put into the same program, adjusting only parameters such as the speed of sound, array spacing, and factors in

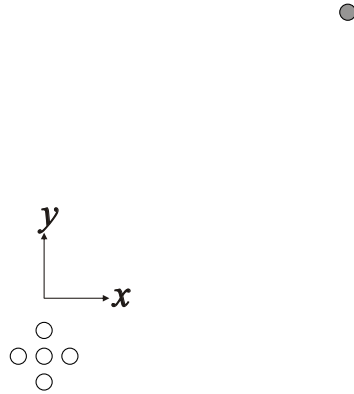
the constraint equations.

In each case, the errors in arrival angles produced by the STCM were averaged for composite inspection and comparison. In addition, erroneous solutions (solutions that matched the constraints but did not pertain to a known physical arrival) were not considered because methods have been developed that should reduce the number of erroneous solutions while increasing accuracy (see Sec. 5.8). The optimal constraint factors were chosen by trial and error to give the most accurate results and the time delay values were chosen manually from the cross correlation functions. This manual time delay selection is not the most desirable or accurate approach to time delay selection, but it made for the simplest verification of the method operability. It consisted of visually interpolating any peaks that were flattened and estimating the time delay value from the visual inspection. This could be automated by using spline or zero-padding interpolation or a combination thereof (see Sec. 5.8.3).

## 5.1 A Single Source in Free Space

The simplest case possible is a single sound arrival at a single instant in time (see Fig. 5.1). The coordinates of the single source here were chosen to be irrational numbers to make the angles interesting. The source was located at (9.42 m, 10.87 m, 6.93m)(( $3\pi$  m,  $4e$  m,  $4\sqrt{3}$  m)). This location results in arrival angles of  $\phi = 49.08^\circ$ ,  $\theta = 25.72^\circ$ . The IR of this case at the origin is shown as Fig. 5.2. As seen in the figure, the IR consists of a single peak and is zero at all other times (except for some small artifacts caused by the anti-aliasing filter).

The arrival angles produced by the program were  $\phi = 48.13^\circ$ ,  $\theta = 26.27^\circ$ . These values result in an average error of  $0.75^\circ$ . If this source were an actual sound source located in free space, the angular error would produce a spatial error of the evaluated

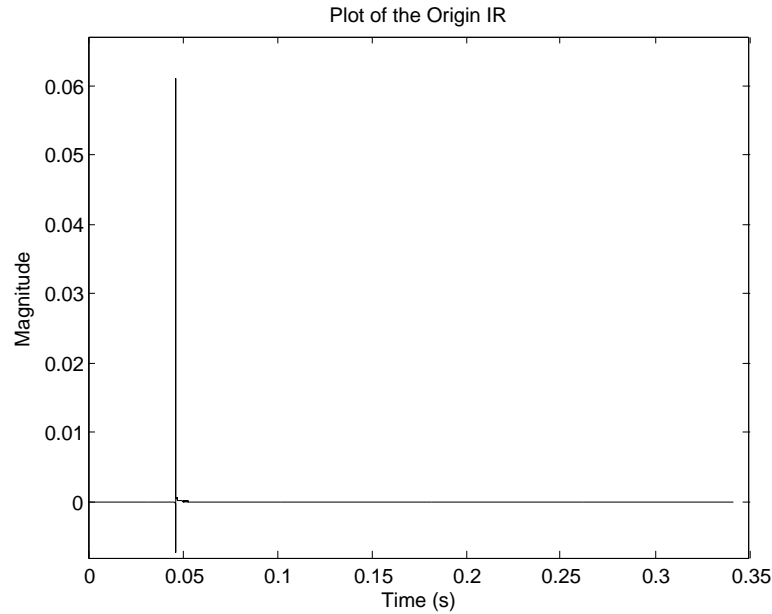


**Figure 5.1** Illustration of a single source in free space located near a Cartesian microphone array.

source location that is less than 21 cm from the actual location. The spatial error would be even smaller if the sound source were actually a reflection located closer to the microphone array than the source located in free space.

## 5.2 Two Asymmetrically-Placed Sources in Free Space

As a second simple case, consider two equistrength sources, equidistant from the origin, but with different coordinates: (9 m, 2 m, 15 m) and (15 m, 9 m, 2 m) (see Figs. 5.3). These values result in arrival angles of  $\phi = 12.53^\circ$ ,  $\theta = 58.42^\circ$  and  $\phi = 30.96^\circ$ ,  $\theta = 6.52^\circ$ , respectively. The IR at the origin for this situation is shown in Fig. 5.4. The arrival angles calculated by the STCM were  $\phi = 11.58^\circ$ ,  $\theta = 58.95^\circ$  and



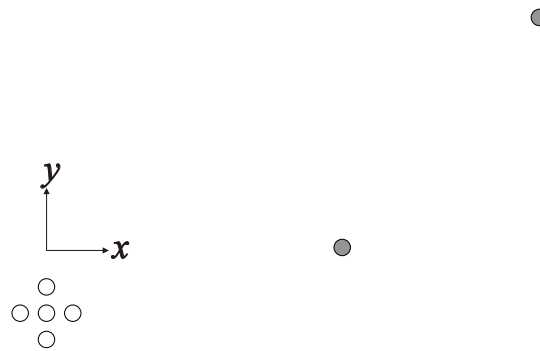
**Figure 5.2** Impulse response for a single sound arrival at a measurement location.

$\phi = 31.49^\circ$ ,  $\theta = 5.56^\circ$ , respectively. The average error was approximately  $0.72^\circ$ .

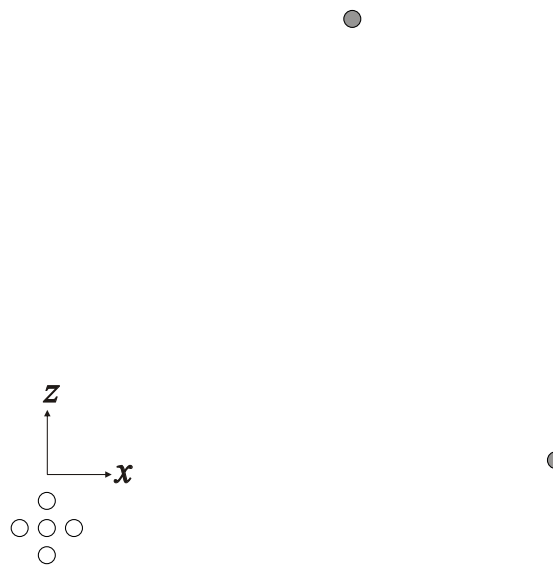
As noted in Ch. 2, the case of two simultaneous arrivals is one of the situations that most previous methods have been unable to handle. This is due to the assumption that only one arrival is present within any given sample. If one arrival was assumed for the scenario in Fig. 5.4, it would have resulted in arrival angles of  $\phi = -25.57^\circ$ ,  $\theta = 49.98^\circ$ , with an average error of  $17.6^\circ$ .

### 5.3 Two Symmetric Sources in Free Space

Next, we consider another situation involving two equistrength, equidistant sources, but in this case the sources are located symmetrically about the  $y$  axis (see Fig. ??). The coordinates for the two sources are (9.42 m, 10.87 m, 6.93m) and (9.42 m, -10.87 m, 6.93m) ( $(3\pi$  m,  $4e$  m,  $4\sqrt{3}$  m) and  $(3\pi$  m,  $-4e$  m,  $4\sqrt{3}$  m), respectively). These

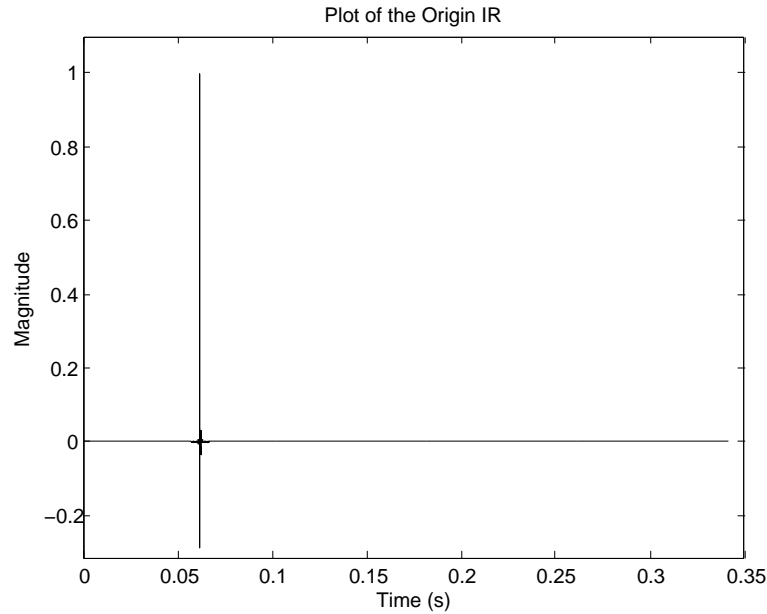


(a)



(b)

**Figure 5.3** Illustration of two asymmetrically-placed sources in free space located near a Cartesian microphone array for (a) the  $x-y$  plane and (b) the  $x-z$  plane.

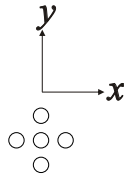


**Figure 5.4** Impulse response for two asymmetrically placed sources producing simultaneous sound arrivals at a measurement location.

correspond to arrival angles of  $\phi = 49.08^\circ$ ,  $\theta = 25.72^\circ$  and  $\phi = -49.08^\circ$ ,  $\theta = 25.72^\circ$ , respectively. The IR at the origin for this case is shown in Fig. 5.6.

A symmetric arrangement demonstrates another issue: now there are eight total solutions sets that will be returned, as every possible combination of time delays meets the constraining criteria established (see Sec. 3.3). However, there only exist two unique solution sets and the remainder are degenerate. Eliminating the degenerate sets will result in only the unique, correct solutions. This was done by searching all solution sets and eliminating any sets whose arrival angles were within  $\pm 1^\circ$  for both  $\phi$  and  $\theta$  of any other set.

The resulting solution sets from the STCM algorithm are  $\phi = 48.13^\circ$ ,  $\theta = 26.27^\circ$  and  $\phi = -48.13^\circ$ ,  $\theta = 26.27^\circ$ , respectively, with an average error of approximately  $0.75^\circ$ . A similar arrangement with the sources located at (20 m, -4 m, 6 m) and (20 m, 4 m, 6 m) was considered and the average error for the solutions returned was



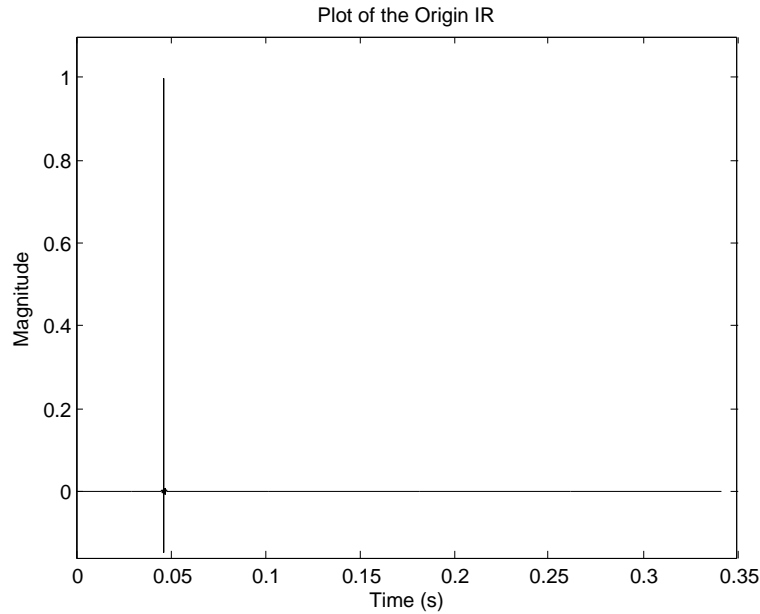
**Figure 5.5** Illustration of two symmetrically-placed sources in free space located near a Cartesian microphone array.

0.55°.

## 5.4 Modeled Rectangular Room with Uniform Absorption

In order to further investigate the utility of the program, a more realistic case was created. The IRs were created to simulate a rectangular room with dimensions 10 x 8 x 7 m and a uniform absorption coefficient of  $\alpha = 0.1$ . The source was placed at (5 m, 5 m, 5 m) and the receiver at (1 m, 1 m, 1 m). These locations were intentionally chosen to provide many instances of simultaneous arrivals. The sampling frequency was 192 kHz and the array diameter was 7.62 cm. The direct sound arrival occurred





**Figure 5.6** Impulse response for two symmetrically placed sources producing simultaneous sound arrivals at a measurement location.

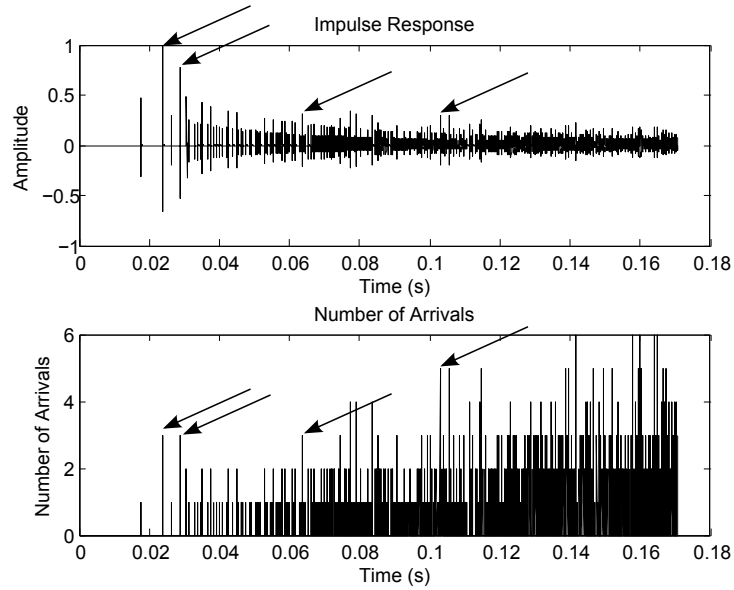
at  $t_0 = 18$  ms.

The image source method was used to generate the IRs using the Allen-Berkley method discussed in Ch. 4 and the code found in Appendix A. The normalized IR of the origin microphone and a histogram of the number of reflections per sample are shown in Fig. 5.7.

The average reflection density quickly exceeds one arrival per sample. The IR also becomes quite dense in a relatively short period of time. The IR was truncated in order to increase the IR generation speed as the number of image sources calculated goes as the number of samples cubed [29].

Four peaks in the IR were chosen for analysis at approximately 25 ms, 29 ms, 64 ms, and 103 ms. The number of arrivals for each of these samples was 3, 3, 3, and 5, respectively.

The first peak requiring analysis is the largest peak in the impulse response, which



**Figure 5.7** Impulse response (top) and histogram (bottom) for a numerically modeled 10 x 8 x 7 m rectangular room with  $\alpha = 0.1$  and arrows indicating peaks analyzed.

occurs immediately after the direct sound arrival. As seen in Fig. 5.7, the amplitude is approximately twice that of the direct sound. This is because three first-order reflections arrive within this single sample. The actual angles of arrival, calculated angles of arrival, differences and average errors are presented in Table 5.1.

**Table 5.1** Results for Peak 1,  $t = 25$  ms, for a numerically modeled room.

	Actual			Measured			Difference			Avg. Error
$\phi$	45.0°	-61.7°	151.7°	45.0°	-61.05°	151.05°	0.0°	0.65°	0.65°	0.43°
$\theta$	-52.7°	25.36°	25.36°	-51.96°	25.83°	25.83°	0.75°	0.45°	0.45°	0.55°

The final column of Table 5.1 shows that the average error is quite small, on the order of  $0.5^\circ$ . The peak is very well separated from other arrivals and thus the results

were expected to be very good. This separation ensures that only the time-delay components that compose the peak of interest are included in the cross correlation, thus reducing the chance of errors and erroneous solutions.

The next peak chosen for analysis was the second largest peak in the IR. This peak also has significantly greater amplitude than the direct sound and is likewise the result of three simultaneous arrivals. The results for the analysis of the peak are shown in Table 5.2.

**Table 5.2** Results for Peak 2,  $t = 29$  ms, for a numerically modeled room.

	Actual			Measured			Difference			Avg. Error
$\phi$	-135.0°	-61.7°	151.7°	-135.0°	-62.23°	152.23°	0.0°	0.53°	0.53°	0.35°
$\theta$	20.8°	-41.36°	-41.36°	20.4°	-41.5°	-41.5°	0.4°	0.14°	0.14°	0.23°

In this instance, the average error is even smaller than for the first peak. Again, it is very well separated from other arrivals, so the performance of the STCM would be similar to that shown for discrete sources in free space.

The next peak at 64 ms was chosen for three reasons: (1) it consists of multiple arrivals, (2) it is in a more densely populated portion of the IR, and (3) it has greater amplitude than the surrounding arrivals. The results for this analysis are presented in Table 5.3.

We find for this case that the average errors increase. This may be due primarily to the increased presence of other arrivals around the peak of interest within the time window.

For the final peak at 103 ms, the STCM was unable to provide any useful results. The number of erroneous sources was greater than the number of actual arrivals shown

**Table 5.3** Results for Peak 3,  $t = 64$  ms, for a numerically modeled room.

	Actual			Measured			Difference			Avg. Error
$\phi$	-74.36°	-61.7°	151.7°	-73.17°	-60.2°	152.7°	1.19°	1.51°	1.18°	1.32°
$\theta$	53.41°	-70.2°	-70.2°	50.70°	-68.6°	-71.3°	2.7°	1.6°	1.1°	1.81°

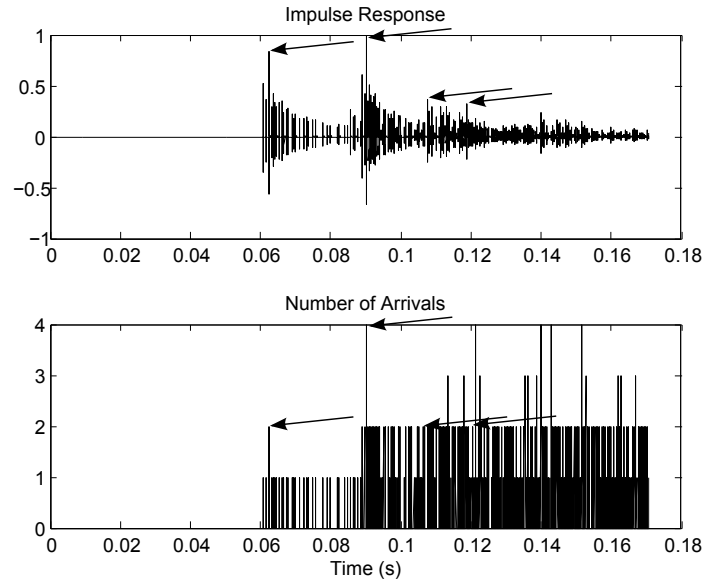
in the histogram (see Fig. ??). This was most likely due to the dense population of this portion of the IR, which results from being so long after the direct sound ( $t = 103$  ms, a difference of 85 ms between direct sound and peak of interest).

## 5.5 Modeled Rectangular Room with Nonuniform Absorption

Due to the positive results in the previous case, a more complex modeled room was constructed. An elongated room was chosen due to the late first-order reflections typical of such a room. It had dimensions (30 m, 8 m, 4.5 m) and had absorption coefficients of  $\alpha_{walls} = 0.2775$ ,  $\alpha_{floor} = 0.51$  and  $\alpha_{ceiling} = 0.37$ . The source and receiver were arbitrarily placed at (25 m, 1.5 m, 1.25 m) and (5 m, 7 m, 2.5 m), respectively. The array diameter was 3" and the sampling rate was again 192 kHz. The direct sound arrival occurred at  $t_0 = 60$  ms.

The origin IR and histogram of arrivals are shown in Fig. 5.8. As seen in the figure, there are multiple peaks that would be of great interest if this were a real room. Most notable is the peak at  $t = 90$  ms resulting from four simultaneous arrivals. This is not only 30 ms after the direct sound arrival, but is also nearly double the amplitude of the direct sound. Three other peaks were chosen for analysis, with arrival times

at 62 ms, 108 ms, and 119 ms. The number of arrivals for each peak are 2, 2, and 2, respectively. The results for the peak at 62 ms are shown in Table 5.4.



**Figure 5.8** Impulse response (top) and histogram (bottom) for a numerically modeled room 30 x 8 x 4.5 m with nonuniform  $\alpha$  and arrows indicating peaks analyzed.

**Table 5.4** Results for Peak 1,  $t = 62$  ms, for a numerically modeled, long, narrow room.

	Actual		Measured		Difference		Avg. Error
$\phi$	$-15.37^\circ$	$20.55^\circ$	$-14.62^\circ$	$20.70^\circ$	$0.75^\circ$	$0.15^\circ$	$0.45^\circ$
$\theta$	$14.21^\circ$	$-3.35^\circ$	$14.18^\circ$	$-4.23^\circ$	$0.03^\circ$	$0.88^\circ$	$0.45^\circ$

The results were very close to the actual values. This particular peak was again well isolated from other arrivals in the impulse response, which likely ensured greater accuracy both in number of arrivals and angles of arrival.

The second peak of interest is at  $t = 90$  ms. This peak is still reasonably well

separated from other arrivals in the impulse response. The results for this case are presented in Table 5.5 and are very close to the actual values.

**Table 5.5** Results for Peak 2,  $t = 90$  ms, for a numerically modeled, long, narrow room.

	Actual				Measured				Difference				Avg. Error
$\phi$	-169.61°	-10.38°	14.03°	165.96°	-169.85°	-10.20°	14.15°	165.93°	0.24°	0.18°	0.12°	0.03°	0.14°
$\theta$	9.77°	9.77°	-2.31°	-2.31°	9.99°	10.05°	-2.01°	-1.99°	0.22°	0.28°	0.31°	0.32°	0.28°

The next peak to be analyzed was located at approximately  $t = 108$  ms. The results are shown in Table 5.6 and are nearly identical to the actual arrivals.

**Table 5.6** Results for Peak 3,  $t = 108$  ms, for a numerically modeled, long, narrow room.

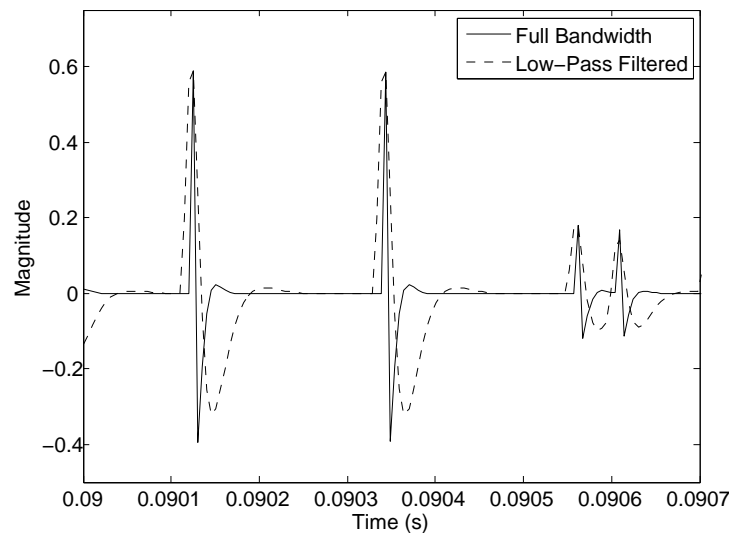
	Actual		Measured		Difference		Avg. Error
$\phi$	-35.63°	-144.37°	-35.65°	-144.35°	0.02°	0.02°	0.02°
$\theta$	-1.94°	-1.94°	-2.03°	-2.03°	0.09°	0.09°	0.09°

No erroneous image source solutions were returned. This shows how accurate the STCM can be in ideal circumstances.

For the final peak at  $t = 119$  ms, there were no useful results returned due to the length in time after the direct sound arrival (approximately 60 ms), showing that there clearly is a maximum usable time length as described in Sec. 4.10. Despite its amplitude, the peak was apparently not sufficiently stronger than the surrounding arrivals to produce meaningful results.

## 5.6 Low-Pass Filtered Room Models

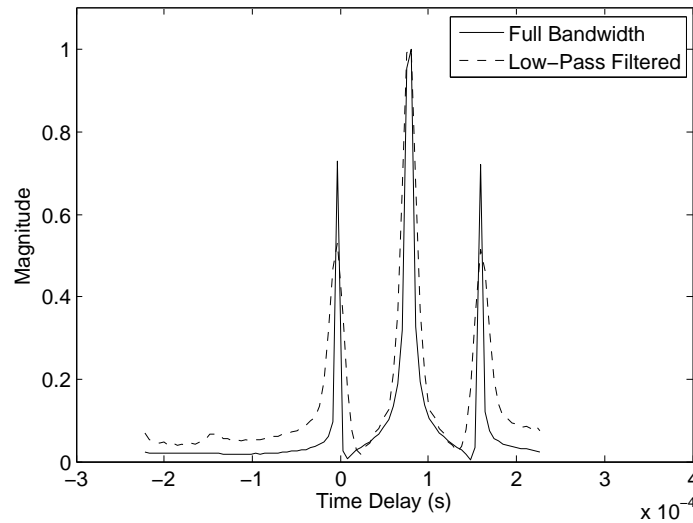
In all of the above numerical models, the IRs contained significant frequency content up to the Nyquist frequency (96 kHz). However, as discussed in Ch. 4, a physical system will typically exhibit a natural low-pass filter effect (see Sec. 4.3). This filtering effect is due to the frequency response characteristics of the loudspeaker and microphone used to measure the room IR. In order to more closely model a typical measured room IR, the responses generated for the two modeled rooms above were low-pass filtered with an  $N = 2$  Butterworth filter with cutoff frequency  $f_c = 22$  kHz. The cutoff frequency was chosen to model the high frequency roll off of the Mackie HR-824 loudspeaker used in the experiments (the microphone's high frequency roll off was negligible compared to the loudspeaker's). Figure 5.9 shows the effects of the high frequency roll off on the IR.



**Figure 5.9** Image source IR with a low-pass filtered version of it superposed.

With the responses filtered, the cross-correlation function is also altered. Figure 5.10 shows the original cross correlation function for a certain peak in the IR

along the  $x$ -axis with the filtered version superposed over it. As seen, the low-pass filter creates a smearing of the data in time. The results for the first three peaks analyzed in the modeled room with uniform absorption are shown in Tables 5.7 through 5.9. As seen, the results for the first two peaks are still accurate, though the accuracy is slightly worse than the full-bandwidth case. The results for the third peak, however, are incomplete, as the smearing of the information made it impossible to localize the third arrival composing the peak in the IR. This could be do to the fact that information from additional arrivals that did not compose the peak of interest was present and, when combined with the filtering, smeared the information such that at least one time-delay component of the third arrival was not distinguishable from other time delay values.



**Figure 5.10** Original cross correlation along the  $x$ -axis with a low-pass filtered version of it superposed.

The results for the first three peaks of the modeled room with nonuniform absorption are shown in Tables 5.10 through 5.12. As seen in the tables, all the arrivals were found, but the accuracy of the calculated results is less than in the full-bandwidth



**Table 5.7** Results for Peak 1,  $t = 25$  ms, for a numerically modeled room with uniform absorption.

	Actual			Measured			Difference			Avg. Error
$\phi$	45.0°	-61.7°	151.7°	43.32°	-61.62°	151.89°	1.68°	0.08°	0.19°	0.65°
$\theta$	-52.7°	25.36°	25.36°	-52.1°	26.86°	26.86°	0.6°	1.50°	1.50°	1.2°

**Table 5.8** Results for Peak 2,  $t = 29$  ms, for a numerically modeled room with uniform absorption.

	Actual			Measured			Difference			Avg. Error
$\phi$	-135.0°	-61.7°	151.7°	-133.37°	-59.66°	151.13°	1.63°	2.04°	0.57°	1.41°
$\theta$	20.8°	-41.36°	-41.36°	23.69°	-40.93°	-40.93°	2.89°	0.43°	0.43°	1.25°

**Table 5.9** Results for Peak 3,  $t = 64$  ms, for a numerically modeled room with uniform absorption.

	Actual			Measured			Difference			Avg. Error
$\phi$	-74.36°	-61.7°	151.7°	-75.53°	-65.48°	N/A	1.17°	3.78°	N/A	N/A
$\theta$	53.41°	-70.2°	-70.2°	47.03°	-65.17°	N/A	6.39	5.03°	N/A	N/A

case. Despite the decrease in accuracy, the overall average error of the arrivals here is  $1.41^\circ$ .

**Table 5.10** Results for Peak 1,  $t = 62$  ms, for a numerically modeled, long, narrow room with nonuniform absorption.

	Actual		Measured		Difference		Avg. Error
$\phi$	$-15.37^\circ$	$20.55^\circ$	$-17.09^\circ$	$17.76^\circ$	$1.72^\circ$	$2.79^\circ$	$2.25^\circ$
$\theta$	$14.21^\circ$	$-3.35^\circ$	$13.73^\circ$	$-1.59^\circ$	$0.48^\circ$	$1.76^\circ$	$1.12^\circ$

**Table 5.11** Results for Peak 2,  $t = 90$  ms, for a numerically modeled, long, narrow room.

	Actual				Measured				Difference				Avg. Error
$\phi$	$-169.61^\circ$	$-10.38^\circ$	$14.03^\circ$	$165.96^\circ$	$-167.49^\circ$	$-12.06^\circ$	$12.39^\circ$	$167.14^\circ$	$2.12^\circ$	$1.68^\circ$	$1.64^\circ$	$1.18^\circ$	$1.40^\circ$
$\theta$	$9.77^\circ$	$9.77^\circ$	$-2.31^\circ$	$-2.31^\circ$	$11.11^\circ$	$11.14^\circ$	$-1.61^\circ$	$-1.58^\circ$	$1.34^\circ$	$1.37^\circ$	$0.70^\circ$	$0.73^\circ$	$1.03^\circ$

**Table 5.12** Results for Peak 1,  $t = 108$  ms, for a numerically modeled, long, narrow room.

	Actual		Measured		Difference		Avg. Error
$\phi$	$-35.63^\circ$	$-144.37^\circ$	$-37.60^\circ$	$-141.75^\circ$	$1.97^\circ$	$2.62^\circ$	$2.29^\circ$
$\theta$	$-1.94^\circ$	$-1.94^\circ$	$-1.21^\circ$	$-1.23^\circ$	$0.73^\circ$	$0.70^\circ$	$0.72^\circ$

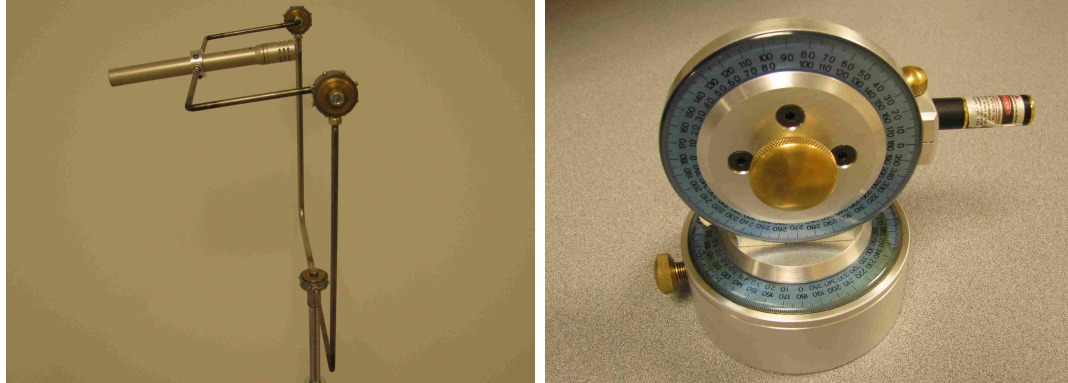
As the results above show, the low-pass filtering effect due to a room system can impact the results to varying degrees. This shows that if there is not sufficient high frequency content then there is a possibility of obtaining less accurate results or, in extreme cases, eliminating the possibility of extracting an arrival. The effect is

particularly noticeable in densely populated portions of the IR because there is not sufficient resolution to separate all of the information from separate arrivals due to the smearing effect of the filtering.

## 5.7 Experimental Results

To determine the usability the STCM when using real data, experiments were first set up and carried out in an anechoic chamber. The chamber used measures 8.7 x 5.7 x 3.4 m and is anechoic over the frequency range of 80 Hz to 20 kHz. The purpose of using the chamber was to allow for control in the placement of reflective surfaces and accurate measurements of the actual angles of arrival. This allowed a meaningful assessment of the average errors produced by the DIR. The EASERA room acoustics software package was used for measuring the IRs with a Mackie HR824 loudspeaker used as a source and a free-field, 1/2" type-1 precision microphone used for the receiver. For these experiments, seven measurements were taken sequentially using the custom-made microphone positioner shown in Fig. 5.11(a). As discussed in Ch. 4, the cross correlation of subsequent measurements in typical room acoustics scenarios is equivalent to the cross correlation of simultaneous measurements if the system is time invariant(see Sec. 4.9). The actual angles of arrival were found using a laser pointer mounted in an altazimuth positioner as shown in Fig. 5.11(b), along with mirrors to reflect the laser beams from the acoustically reflective surfaces. To find the arrival angles with the altazimuth-mounted laser, the center of rotation of the altazimuth was placed at the same point in space as the microphone array origin and the  $\phi = 0^\circ$  direction was oriented in the  $+x$  direction of the array. Direct sound arrival angles were measured by adjusting the elevation and azimuth until the laser beam was pointed at the loudspeaker. In the cases of reflections, the laser was adjusted

until the reflected beam hit the loudspeaker.



(a)

(b)

**Figure 5.11** Two pieces of custom equipment used for the experimental setup; (a) microphone positioner and (b) altazimuth-mounted laser pointer.

For these experiments, the array diameter was 7.62 cm and the sampling frequency was 192 kHz. In many cases, some erroneous solution sets were produced due to the number of peaks in the cross correlation functions. These solutions are ignored in the following results as fully-developed versions of proposed solutions are outside of the scope of this thesis (see Secs. 5.8.4 and 5.8.5).

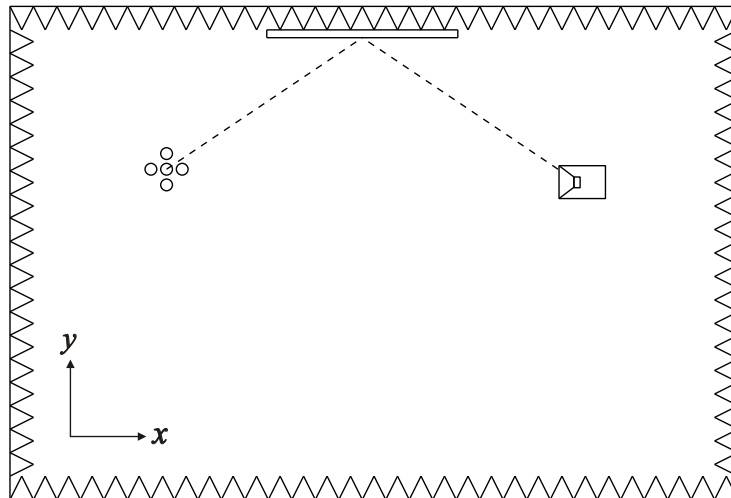
### 5.7.1 Margin of Error

Measuring arrival angles with a laser in an altazimuth mount can produce very accurate results. However, the reference markings on the altazimuth used were only accurate to about one half of a degree. In addition, the placement of the altazimuth in an anechoic chamber is often subject to errors due to deflection of the cable tension floor when a person stands near the altazimuth for adjustment. The accuracy of the elevation angles particularly is prone to error. Because of these inherent problems, a margin of error exists on the order of 1 to 4 degrees. The average error results given

for the experiments above are relative to the measured arrival angles as though they were absolutely correct. This is not always the case, as there is some uncertainty regarding the exact arrival angles.

### 5.7.2 Single Reflector Experiment

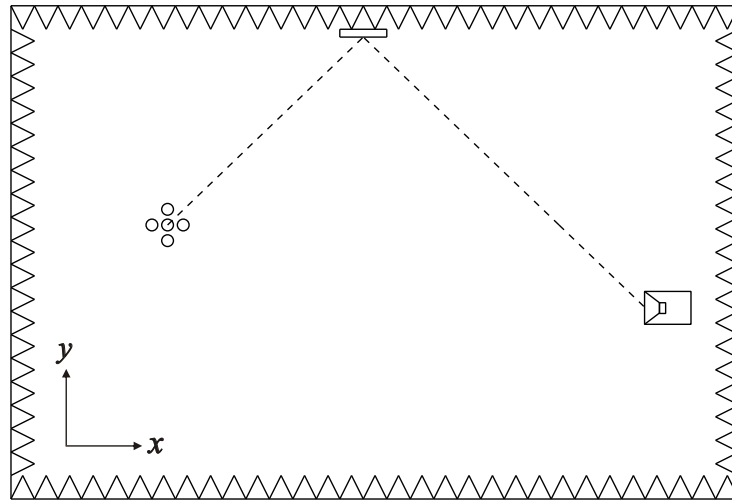
One of the simplest scenarios that can be conceived is a single source and a single reflector in a free-field environment. The reflecting panel was a standard-size (2.44 m x 1.22 m) sheet of MDF board and was placed along one wall of the anechoic chamber. The microphone array was positioned such that the direct sound arrival came from  $\phi = 0.25^\circ$ ,  $\theta = 0^\circ$ . The actual angle of arrival for the single reflection was determined to be  $\phi = 35.5^\circ$ ,  $\theta = -3^\circ$ . An illustration of the experimental setup is shown below in Fig. 5.12. For this experiment, the arrival angles for the direct sound calculated by the STCM algorithm were  $\phi = 1.2^\circ$ ,  $\theta = -1^\circ$  while those for the single reflection were  $\phi = 34^\circ$ ,  $\theta = -6^\circ$ , resulting in an average error of  $2.2^\circ$ .



**Figure 5.12** Experimental setup used for a single reflector case in an anechoic chamber where dotted lines represent rays traced by reflections.

### 5.7.3 Single Small Reflector Experiment

Another situation of interest relates to the size of the reflecting surface that creates an unwanted arrival. In order to investigate this, a single, small reflecting plywood panel was used in an anechoic chamber. The panel measured 60.5 cm x 61 cm and was placed along one wall of the chamber. A diagram of the setup is shown in Fig. 5.13.



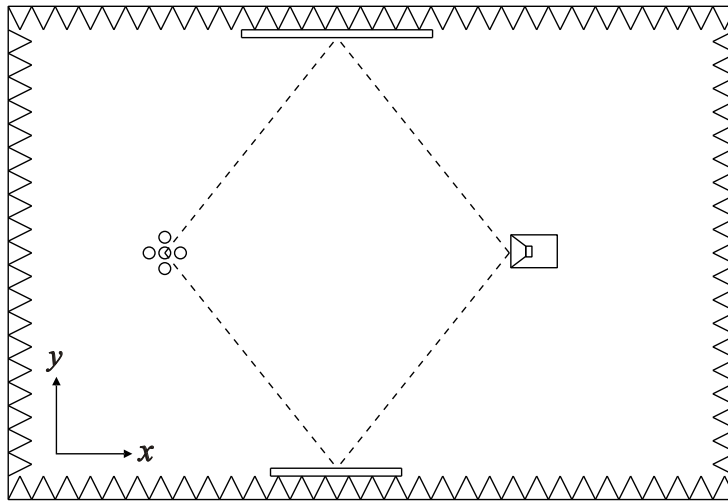
**Figure 5.13** Experimental setup used for a single, small reflector case in an anechoic chamber where dotted lines represent rays traced by reflections.

The actual direct sound arrival angles were  $\phi = -25^\circ$ ,  $\theta = 5.5^\circ$  and the arrival angles for the small panel were  $\phi = 44.5^\circ$ ,  $\theta = 1^\circ$ . The results from the STCM were  $\phi = -24^\circ$ ,  $\theta = 6.7^\circ$  for the direct sound and  $\phi = 43.05^\circ$ ,  $\theta = 2.02^\circ$  for the reflection, resulting in an average error of  $1.25^\circ$ .

### 5.7.4 Dual Symmetric Reflector Experiment

The next experiment was a dual reflector arrangement depicted in Fig. 5.14. The reflecting panels measured 2.44 m x 1.22 m and 1.75 m x 1.22 m and were placed on

opposite sides of the chamber with the microphone array and loudspeaker centered between the two. Because the peaks resulting from the reflecting panels had widths of more than one sample, the centering was performed by changing the loudspeaker position until the peak resulting from the two reflections was maximized. This indicated that the arrivals were overlapping closely, though there was likely some uncertainty in path lengths of approximately 5 - 10 mm. This was done because one of the most difficult scenarios to resolve is the case of simultaneous arrivals that are symmetric about an axis.

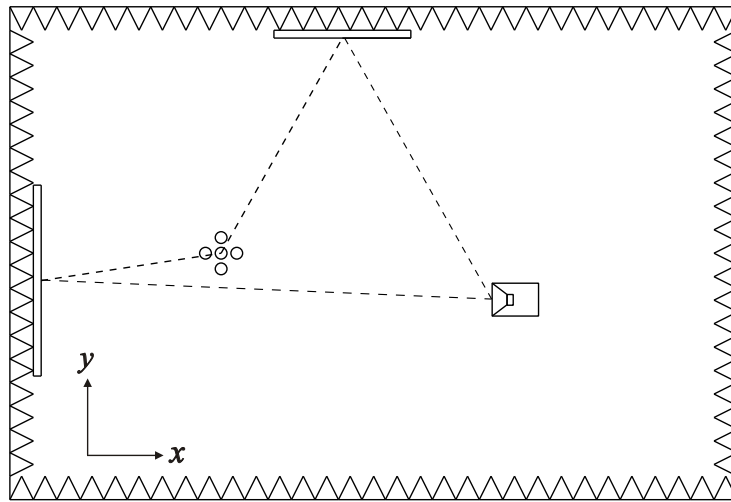


**Figure 5.14** Experimental setup used for a dual symmetric reflector case in an anechoic chamber where dotted lines represent rays traced by reflections.

The actual direct sound arrival was at  $\phi = 2^\circ$ ,  $\theta = 3^\circ$ , while those of the simultaneous reflections were  $\phi = 49^\circ$ ,  $\theta = -4.5^\circ$  and  $\phi = -45^\circ$ ,  $\theta = -0.5^\circ$ . The STCM calculated arrival angles for the direct sound of  $\phi = 2^\circ$ ,  $\theta = 7^\circ$  and reflection arrival angles of  $\phi = 42^\circ$ ,  $\theta = -3.5^\circ$  and  $\phi = -39^\circ$ ,  $\theta = -3.5^\circ$ , respectively. These numbers result in an average error of approximately  $4^\circ$ .

### 5.7.5 Dual Asymmetric Reflector Experiment

An additional experiment utilized two reflecting panels placed asymmetrically about the microphone array as shown in Fig. 5.15. This was implemented to investigate how well the STCM works with extreme azimuthal arrival angles. The reflectors had the same dimensions as in the previous experiment.



**Figure 5.15** Experimental setup used for a dual asymmetric reflector case in an anechoic chamber where dotted lines represent rays traced by reflections.

The actual angles of arrival for the direct sound in this arrangement were  $\phi = -15^\circ$ ,  $\theta = -1^\circ$  while those for the two simultaneous reflections were  $\phi = -176^\circ$ ,  $\theta = -6.5^\circ$  and  $\phi = 63^\circ$ ,  $\theta = -8^\circ$ . The calculated arrival angles were  $\phi = -17.5^\circ$ ,  $\theta = 0.5^\circ$  for the direct sound and  $\phi = -178^\circ$ ,  $\theta = -5^\circ$  and  $\phi = 63^\circ$ ,  $\theta = -6^\circ$  for the simultaneous arrivals, resulting in an average error of approximately  $1.2^\circ$ .

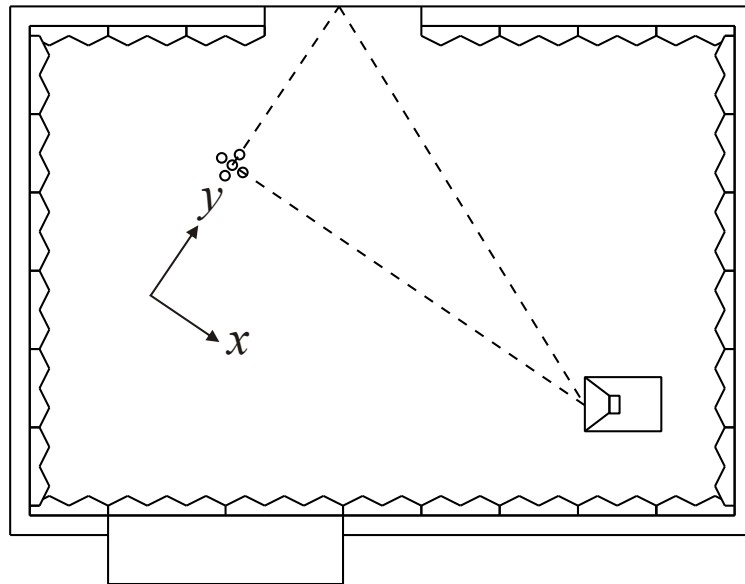


### 5.7.6 Variable Acoustics Chamber Experiment

As a final controlled experiment, the loudspeaker and microphone array were placed in a room designed for variable acoustics work. The chamber measures approximately 3.8 m x 2.7 m x 2.5 m. The ceiling and walls have removable absorptive panels. The panels do create some reflections at very low frequencies ( $< 100$  Hz) and high frequencies ( $> 5$  kHz), but the reflections were dominant from the exposed room surfaces. The floor is hard tile over concrete and the walls and ceiling behind the absorptive panels are concrete finished with hard plaster and paint. The room exhibits a very strong resonance near 43 Hz, so the IRs were bandpass-filtered between the frequencies of 100 Hz and 30 kHz.

For the experiment, two adjacent wall panels were removed and the microphone array and loudspeaker were placed in opposite corners of the room, as shown in Fig. 5.16. The microphone array was positioned such that the reflections from the wall and floor would arrive simultaneously and a second order reflection was created by a path from the loudspeaker, to the floor, to the wall and finally to the array.

The actual arrival angles for the direct sound were  $\phi = 1^\circ$ ,  $\theta = -20^\circ$ . The actual arrival angles for the simultaneous (see Sec. 5.7.4) first-order reflections were  $\phi = 97.5^\circ$ ,  $\theta = -15^\circ$  for the wall and  $\phi = 1^\circ$ ,  $\theta = -48^\circ$  for the floor. Finally, the actual angles of arrival for the second order reflection were  $\phi = 98^\circ$ ,  $\theta = -37.5^\circ$ . For the direct sound, the STCM algorithm produced arrival angles of  $\phi = 0.5^\circ$ ,  $\theta = -23^\circ$ . For the first order reflections, the STCM results were  $\phi = 102^\circ$ ,  $\theta = -16^\circ$  for the wall and  $\phi = 2^\circ$ ,  $\theta = -46^\circ$  for the floor reflection. The arrival angles for the second-order reflection were calculated to be  $\phi = 102^\circ$ ,  $\theta = -39^\circ$ . These results yield an average error of approximately  $3.25^\circ$ .



**Figure 5.16** Experimental setup used in a variable acoustics chamber where dotted lines represent rays traced by reflections.

### 5.7.7 Summary of Results

In summary, the results from the experimental setups discussed in Secs. ?? through 5.7.6 are presented in Table 5.13.

Table 5.13 Results for experimental setups in controlled acoustical environments.

	Single	Single Small	Dual Symmetric	Dual Asymmetric	Variable Chamber
Direct <sub>Actual</sub> ( $\phi/\theta$ )	0.25°/0.1°	-25.0°/5.5°	2.0°/3.0°	-15.0°/-1.0°	1.0°/-20.0°
Direct <sub>Measured</sub> ( $\phi/\theta$ )	1.2°/-1.1°	-24.0°/6.7°	2.0°/7.0°	-17.5°/0.5°	0.5°/-23.0°
First Order <sub>Actual</sub> ( $\phi/\theta$ )	35.5°/-3.0°	44.5°/1.05°	49.0°/-4.5° -45.0°/-0.5°	-176.0°/-6.5° 63.0°/-8.0°	97.5°/-15.0° 1.0°/-48.0°
First Order <sub>Measured</sub> ( $\phi/\theta$ )	34.1°/-6.0°	43.0°/2.04°	42.0°/-3.5° -39.0°/-3.5°	-178.0°/-5.0° 63.0°/-6.0°	102.0°/-16.0° 2.0°/-46.0°
Second Order <sub>Actual</sub> ( $\phi/\theta$ )	N/A	N/A	N/A	N/A	98.0°/-37.5°
Second Order <sub>Measured</sub> ( $\phi/\theta$ )	N/A	N/A	N/A	N/A	102.0°/-39.0°
Avg. Error	2.2°	1.25°	4.05°	1.2°	3.25°

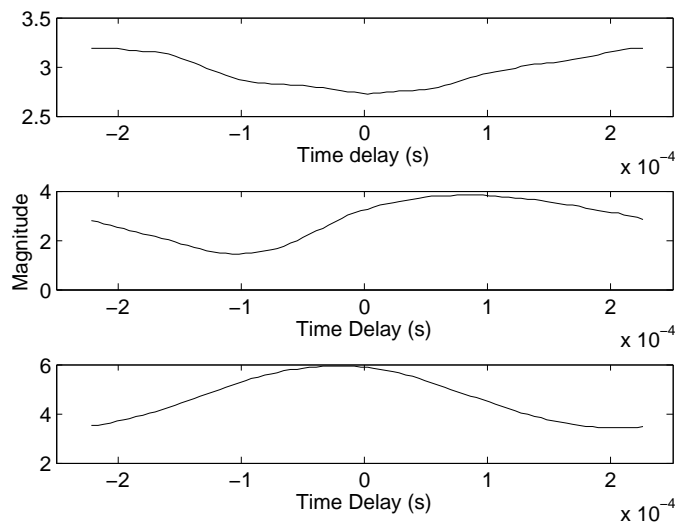
### 5.7.8 Off-Axis Experiments

Previously we discussed how loudspeakers and microphones exhibit a type of low-pass filtering effect on the system due to their limited frequency response characteristics. In addition to these effects that, until now, were assumed to be caused solely by the imperfect response characteristics of the sources and receivers, the enclosures of these also have a filtering effect. These effects are caused primarily by diffraction and the enclosure characteristics of the microphones or loudspeakers. Because we used precision omnidirectional microphones, these effects are less than those for the loudspeaker. The loudspeaker was very much *not* omnidirectional and its orientation can play a very large role in whether or not a reflection can be accurately localized. To determine the extent of this effect, measurements using the same experimental setups discussed previously were taken with the loudspeaker turned both  $90^\circ$  and  $180^\circ$  from their optimal orientation. Only the results for the dual symmetric reflector and the dual asymmetric reflector experiments are presented here.

For the dual symmetric reflector case (see Fig. 5.14) with the speaker facing directly away from both the microphone array and the reflectors, the frequency content was rolled off dramatically. Localization of the direct sound was possible, but the error was on the order of  $2.5^\circ$ . The two simultaneous reflections, however, were not able to be resolved at all due to the extreme smearing caused by the lack of high-frequency content. Upon further examination, the high frequencies began to roll off in the range of 1 - 2 kHz when the speaker was facing the opposite direction. This explains why there was no possible way to determine the arrival angles for the simultaneous arrivals. With the speaker rotated  $90^\circ$  from its original position, the direct sound and the simultaneous reflections were localized, but with an average error of over  $4^\circ$ .

For the dual asymmetric reflector case (see Fig. 5.15), the results were an even

better illustration of the effects that an off-axis measurement can have. With the source rotated  $90^\circ$  from its original position, the reflector at  $\phi = 63^\circ$ ,  $\theta = -8^\circ$  was located only slightly off-axis while the reflector located at  $\phi = -178^\circ$ ,  $\theta = -6.5^\circ$  was located strongly off-axis. When the STCM was run on these IRs, the panel slightly off-axis was localized with an average error of approximately  $6^\circ$  while the panel far off-axis was not able to be localized at all. The direct sound was also not able to be localized accurately. When the loudspeaker was turned to face  $180^\circ$  relative to the original orientation, both panels were located well off-axis. The direct sound was not able to be localized with this speaker orientation, either. The simultaneous arrivals were not able to be localized either. The smearing effect of the off-axis orientation was too great (see Fig. 5.17).



**Figure 5.17** Cross correlation function for simultaneous arrivals in the dual asymmetric reflector experiment with the loudspeaker oriented  $180^\circ$  from its original position.

Due to the sometimes dramatic reduction in usability of the STCM with signals containing little high frequency energy, it is important to ensure enough the source produces enough high frequency energy to provide enough detail in the measurements.

Use of a wide-bandwidth, low-directionality source would ensure proper energy at all directions, however, such a source is difficult to find. Even standard dodecahedrons exhibit limited high frequency response and spatial beaming at higher frequencies which increases directionality. For further discussion of the importance of high frequency content, see Sec. 5.8.3.

## 5.8 Discussion

During the course of the experiments, many important issues came to light regarding the STCM method and the experimental setup itself. This section explores these issues and other details.

### 5.8.1 EASERA

The signal processing used within EASERA also proved to be somewhat problematic. It was found that the use of band-limited excitation signals was not as effective as using full bandwidth sweeps and then filtering afterwards. The reason for this is unknown, but it appears that the method used to generate the bandlimited excitation signals produces some energy outside of the specified frequency band. This additional energy is primarily at lower frequencies, which caused excitation of the lowest-order room modes in the variable acoustics chamber (see Sec. 5.7.6). Additionally, the loudspeaker compensation function, which utilizes a deconvolution process to remove the effects of the loudspeaker from the measurement, creates some very strange artifacts, such as doubled peaks, when used. For this reason, all signal processing beyond derivation of the IR was done in Matlab. The IR could have been extracted in Matlab, but the IRs resulting from EASERA were calculated quickly and accurately, eliminating the need to do so in Matlab.

### 5.8.2 Erroneous Solutions and Cross Correlation Peak Selection

To help increase the effectiveness of the method, a subset of four microphones from the seven-microphone array was used to generate solution sets to be compared with the results from the seven-microphone array for both numerical simulations and experiments. This helped greatly in the numerical simulations, but proved to be problematic in some experimental cases. It was found that the four-microphone cross-correlation functions did not always contain sufficient data for the angles of arrival to be accurately determined. This resulted in the inability to compare the two sets of results, resulting in more erroneous solutions because the time-delay constraint was the only safeguard. It appears that the solution to this problem is to ensure sufficient high frequency content in order to provide the time resolution necessary to accurately determine the angles of arrival (see Sec. 5.8.3).

The selection of the peaks in the cross correlation functions proved to be more difficult than originally anticipated. At times, the peaks in the cross correlation necessary to produce the correct arrival angles were either smeared due to a lack of time resolution or they were of lower amplitude than other peaks, many of which did not relate to physical arrivals. This is due to the calculation of the cross correlation function. Further discussion of this issue and a proposed solution are presented in Sec. 5.8.5.

### 5.8.3 High-Frequency Content

High-frequency energy was found to be vital in maintaining a high level of detail in the cross-correlation function. If the function is considered in light of a Fourier transform, then it is readily apparent that a lack of high frequency information in the

function results in a decrease in the amount of detail that can be expressed within the cross correlation time window. When a wider bandwidth can be used, there is greater potential for better results when using the STCM. One must therefore be selective in both the source (loudspeaker) and receiver (microphone) used in the measurements. If one can compensate for their response deficiencies, the cross correlation will have greater detail and the measured response will more closely represent the unfiltered room response. An obstacle to completely removing the loudspeaker or microphone effects from a measurement is the fact that a deconvolution is only possible for a single angular axis at a time. This makes it unfeasible to remove all effects of the loudspeaker or microphone for every direction.

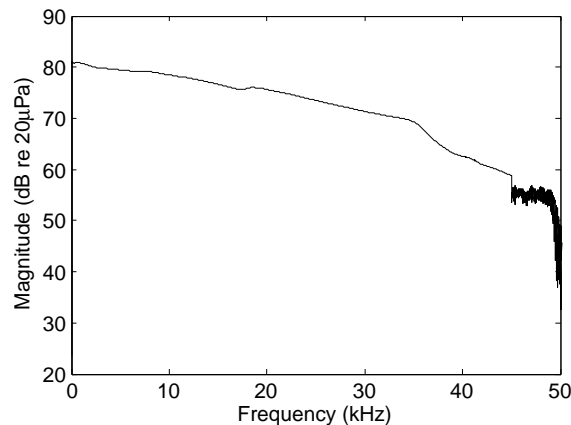
One possible advantage of the STCM when working with off-axis measurements is the fact that the sampling frequency can be reduced. As long as there is not significant frequency content above the Nyquist frequency, the method can be used successfully with virtually any sampling frequency. Higher sampling rates can continue to be used with the method despite a lack of some high frequency. The higher rates will essentially interpolate the data without providing any finer detail in time. Zero padding may also be used for interpolation to provide more closely spaced samples while not providing any additional information [41].

#### 5.8.4 Frequency Magnitude Compensation

Due to the importance of high-frequency content in the IRs and the cross correlation functions, at least some method for compensating for the loudspeaker or microphone high frequency rolloff is desired. Several methods could be implemented to remedy a lack of high frequency energy, but a typical loudspeaker will have a non-uniform radiation pattern in both the elevation and the azimuth. This makes some methods impractical. A general high-frequency magnitude boost might be acceptable, but



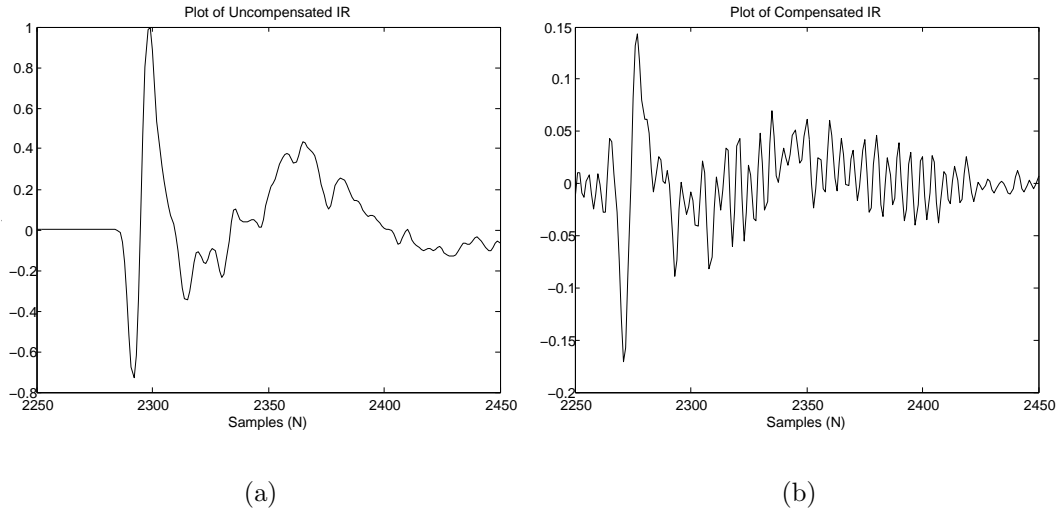
one would need to determine what the equalization curve should be. To eliminate some of the detrimental effects of using an on-axis response for off-axis measurements, the on-axis frequency-response magnitude was simply smoothed over octave bands. The smoothed frequency response magnitude was then inverted before multiplying the measurement frequency response magnitude. The smoothed frequency response curve of the powered two-way loudspeaker used in the experiments is shown in Fig. 5.18. One will note that above approximately 45 kHz, the signal turns to noise due to the limited amount of high-frequency signal present in the loudspeaker-microphone system.



**Figure 5.18** Smoothed frequency-response curve of a Mackie HR-824 loudspeaker used for frequency magnitude compensation.

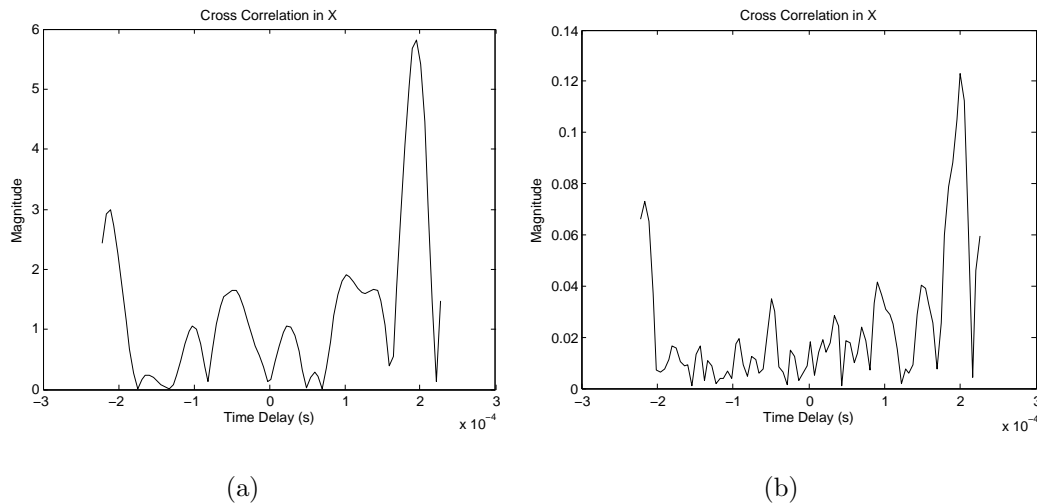
To illustrate the effects of the general frequency-resonse magnitude compensation, Fig. 5.19 shows a section of an IR before and after the process. The compensation increases detail and introduces a change in the amplitude scale. However, because the relative amplitude values are most important in both the IR and the cross correlation function, the latter change does not have a strong impact on the usefulness of the

DIR. Only the relative amplitudes are important for determining the arrival angles.



**Figure 5.19** Section of an IR (a) before and (b) after the frequency magnitude compensation process.

As seen in the figure, the magnitude compensation also introduces substantial enhancement of high frequency noise that was not as prominent previously and which has an effect on the cross-correlation function. Figure 5.20 shows an example cross-correlation function before and after the frequency magnitude compensation process. It is clear that the compensation increases the number of peaks in the cross-correlation function. This is particularly undesirable for an automated peak detection process because the greater the number of those peaks, the greater the chances are that erroneous solutions will be calculated that satisfy the constraint equations. To gain the advantages of the frequency magnitude compensation process while minimizing the negative effects, a solution is now presented.



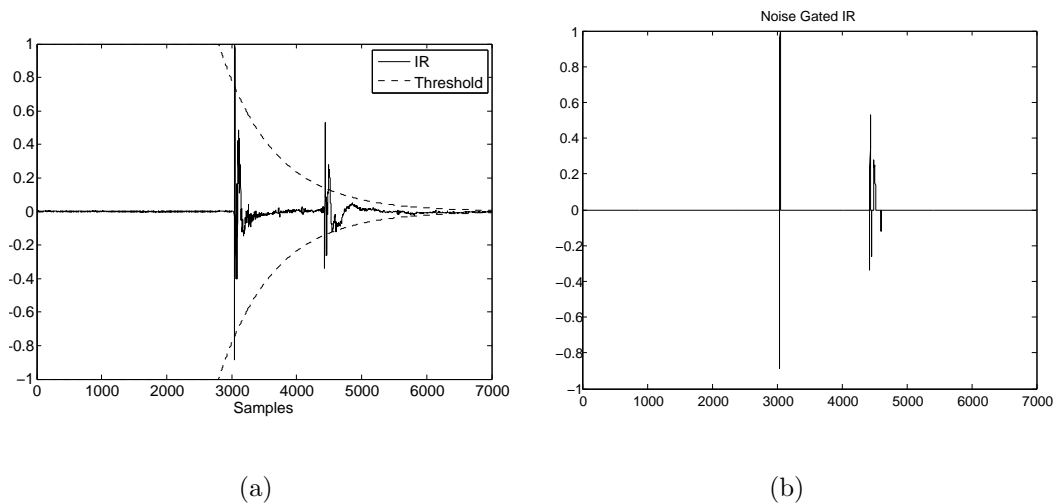
**Figure 5.20** Example cross correlation function (a) before and (b) after the magnitude compensation process.

### 5.8.5 Noise Gating

The frequency-response magnitude compensation thus has some benefits, but also some serious drawbacks. The greatly increased number of peaks in the cross correlation potentially poses a very serious threat to the utility of this method, especially if an automated peak detection feature is desired. In an effort to take advantage of the magnitude compensation's benefits while avoiding the drawbacks, one can begin by revisiting the cross-correlation function. It, like the convolution, can be considered graphically as well as mathematically. To graphically derive the time-windowed cross correlation function, the time windowed IRs from two different microphones are overlaid, multiplied, and summed. One of the windowed IRs remains stationary while the other is shifted by a sample in one direction. The process of multiplying and summing are repeated for each shift in both the positive and negative directions, where the shift represents the time delay between the signals and the value of the sum is the cross power. With this understanding, it is easy to see how the boosted

high-frequency noise in the magnitude-compensated IR will increase the noise in the cross-correlation function. If a method for eliminating the noise induced in the IR by the magnitude compensation process could be implemented without affecting the increase in high-frequency content in the IR peaks of interest, the quality of the STCM cross correlation calculation could be significantly enhanced.

One simple way to accomplish this is through the use of a noise gate. A noise gate was created with a decaying exponential threshold for this work. It was defined by the starting point (in samples) and the decay factor. All samples in the IR that did not exceed this threshold were set to zero and only the samples exceeding the threshold remained intact. Figure 5.21 shows an IR from the dual symmetric reflector experiment with the superposed threshold. It also shows the resulting noise-gated IR.



**Figure 5.21** An IR shown (a) with a superposed noise gate threshold and (b) the modified IR after a noise gate has been applied.

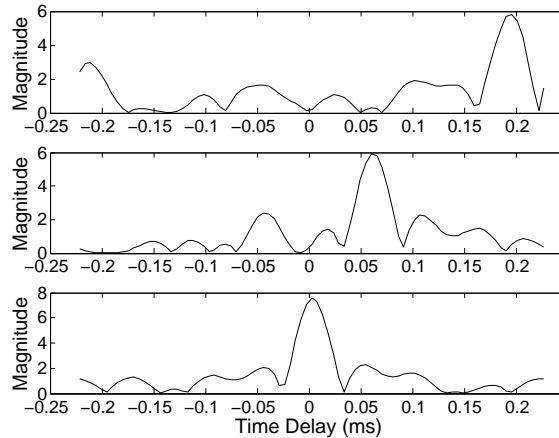
As seen in the figure, there are many places where the cutoff is very abrupt. This is sometimes referred to as “center clipping,” which can introduce high-frequency artifacts in the frequency response due to the sudden nature of the cutoff [42]. This

high frequency content was filtered out using a 15<sup>th</sup> order Chebyshev low-pass filter with 0.5 dB pass-band ripple and a cutoff frequency of 30 kHz. The Chebyshev filter was chosen because of its uniform phase shift, which helps keep the cross-correlation function intact after filtering.

The point of the noise gating was to reduce the noise associated with the magnitude compensation process. Figures 5.22 through 5.24 show the cross correlation along the  $x$ ,  $y$ , and  $z$  axes for the dual symmetric reflector case discussed in Sec. 5.7.4 using the original IRs, the gated IRs, and finally the compensated and gated IRs. The gating process combined with the compensation process provides clean cross correlation functions. While they can be somewhat coarse due to the simplistic nature of the noise gate, a more sophisticated gate would most likely produce better results. However, a more elaborate gating procedure would likely require complicated adaptive signal processing and is outside of the scope of this thesis. The preliminary results merely demonstrate that the concept shows promise and should be considered in pertinent future work.

## 5.9 Chapter Conclusions

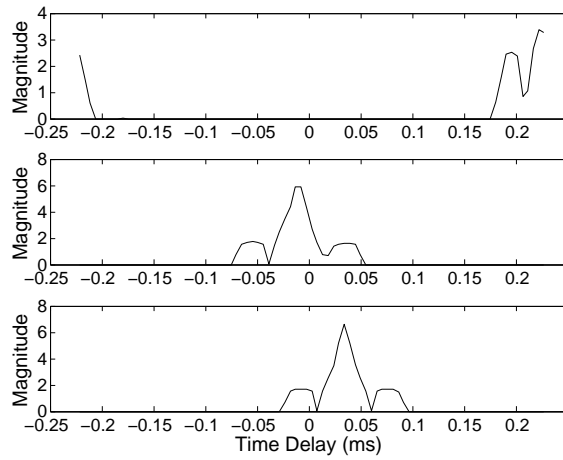
A variety of numerical simulations have been carried out to test the accuracy and usability of the STCM method described in Chs. 3 and 4. The STCM was found to be accurate, even in realistic cases where the impulse responses were densely populated and had many interesting features. Experiments were also developed and carried out to test the DIR's ability to work in real-world circumstances. It proved to be effective in these simple experiments, allowing one to conclude that the use of real data should not be problematic. One might argue that there should be very low errors (approaching zero) in the numerical simulations, but a few factors that may



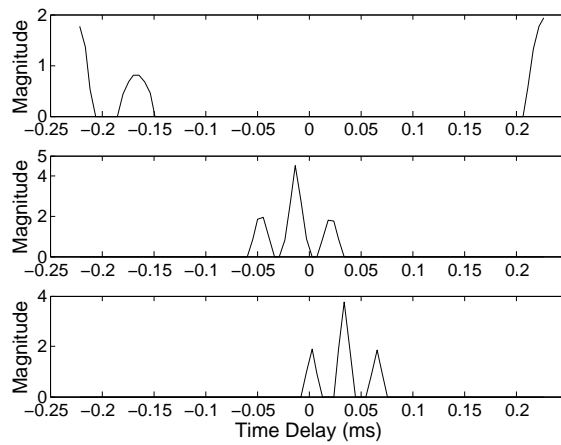
**Figure 5.22** Cross Correlation functions for the dual symmetric reflector experiment using original IRs.

contribute to them are the selection of peaks manually and the time quantization (sampling) of the signals, both of which introduce uncertainty into the measurement.

Certain drawbacks of the STCM have also been discussed. These drawbacks include its inability to be used for arrivals long after the direct sound arrival (except in cases of strong late arrivals) and the statistical likelihood that erroneous results will occasionally be produced, despite the constraints and other preventative measures that have been implemented. Additional considerations for the STCM that were discovered during the course of experimental work were discussed and plausible solutions to those were presented. Despite these limitations, the STCM has been shown to perform well in situations in which many of the previous methods would fail. A comparison between the STCM and the most common of the previous methods (the Polar ETC) will be presented in the next chapter.



**Figure 5.23** Cross Correlation functions for the dual symmetric reflector experiment using noise gated IRs.



**Figure 5.24** Cross Correlation functions for the dual symmetric reflector experiment using noise-gated and magnitude-compensated IRs.





# Chapter 6

## Comparison of Methods

The primary motivation for developing a new DIR method was to improve upon previous methods. While it was derived with very few assumptions and appears to give good results under specified conditions, a comparison with previous methods is needed to ascertain if significant improvements were really made. This chapter presents the results of a comparison between the STCM and the Polar ETC methods using both numerically simulated data and experimental data measured for the same setups discussed in Sec. 5.6.

### 6.1 Polar ETC

#### 6.1.1 Polar ETC Program

Before presenting the results of the comparison, a brief description of the Polar ETC program used is warranted. The commercial Polar ETC software is currently licensed and developed by GoldLine and implemented in the TEF analyzer. Consequently, the program can only be used with energy time curves (ETCs) measured by the TEF, which has a sampling frequency of 48 kHz. In order to ensure fair numerical compar-

isons, a sampling frequency of 192 kHz was needed. Interpolation can yield higher sampling frequencies, but it doesn't provide the additional information required. In addition, the ETC has been shown to not actually measure of the energy density in a room [31]. Finally, since we are using numerical simulations, there would be no way to use the actual numerical models used in the previous chapter with the commercial Polar ETC program. A Matlab program was accordingly written to perform the Polar ETC using squared IRs and the equations given in the original Polar ETC papers [5,6]. The code was benchmarked using numerically generated data to ensure proper operation (see Sec. 6.2.1). The code is included in Appendix A.

To properly implement the program, the IRs had to be generated as though they were detected with ideal cardioid microphones. The Berkley-Allen IR generation program was altered to produce the cardioid response pointing in each of the six required Cartesian directions. The code for the modification is also presented in Appendix A. It involved using the equation of a cardioid solid and applying to the the known direction of arrival from each image source. Specifically, the weighting equations used were

$$W_{+X} = 1 + \cos(\phi)\cos(\theta) \quad (6.1)$$

$$W_{-X} = 1 - \cos(\phi)\cos(\theta) \quad (6.2)$$

$$W_{+Y} = 1 + \sin(\phi)\cos(\theta) \quad (6.3)$$

$$W_{-Y} = 1 - \sin(\phi)\cos(\theta) \quad (6.4)$$

$$W_{+z} = 1 + \sin(\theta) \quad (6.5)$$

$$W_{-z} = 1 - \sin(\theta) \quad (6.6)$$

In these expressions,  $W$  indicates the weighting function and the subscript denotes the direction the virtual microphone points. The weightings were applied to every arrival in the IR.

For the experimental comparisons, the commercial Polar ETC package was implemented using a TEF-20 analyzer, a Shure SM-81 cardioid microphone, a Cartesian microphone positioner (see Fig. 5.11(a)), and TEF Soundlab Windows software. The excitation source was a Mackie HR-824 loudspeaker driven with a swept sine. One should note that the sampling frequency of the TEF analyzer and other parameters differ between the Polar ETC and STCM measurement setups. These differences are for the purpose of comparing results from the Polar ETC using a typical measurement setup with results from the STCM using a measurement setup that is anticipated to be typical of it. Erroneous results were generated by both the Polar ETC and the DIR, but they have been omitted to simplify the comparison process.

## 6.2 Comparison Results

The results for both the numerical and experimental scenarios are presented here. The results for the STCM method were presented in Ch. 5 so for the sake of brevity, only average errors are presented here for comparison.

## 6.2.1 Numerical Comparisons

### Small Rectangular Room

The first of the numerical comparisons is for the numerically modeled room in Sec. 5.4. In order to ensure proper working order of the Polar ETC code, the angles of arrival for the direct sound were also calculated. For the given source/receiver configuration, the actual direct sound angles of arrival are  $\phi = 45^\circ$ ,  $\theta = 35.26^\circ$ . The Polar ETC returned angles of arrival of  $\phi = 45^\circ$ ,  $\theta = 35.26^\circ$ , which are exactly the same. This indicated that the program was working correctly. The calculated angles of arrival for the STCM method were also  $\phi = 45^\circ$ ,  $\theta = 35.26^\circ$ .

To demonstrate the improved utility of the STCM over the Polar ETC, the strongest peak in the room IR was analyzed, which corresponds to three simultaneous arrivals (see Fig. 5.7). The Polar ETC returned arrival angles  $\phi = 45^\circ$ ,  $\theta = 35.26^\circ$  while the actual arrival angles were  $\phi = 45^\circ$ ,  $\theta = -52.70^\circ$ ,  $\phi = -61.7^\circ$ ,  $\theta = 25.36^\circ$ , and  $\phi = 151.7^\circ$ ,  $\theta = 25.36^\circ$ . Interestingly, those angles are identical to those of the direct sound arrival and result in an average error of  $69^\circ$ . By comparison, the STCM method produced an average error of  $0.5^\circ$ . This is a clear example of where the Polar ETC becomes ineffective due to the assumption of a single arrival within a given sample.

For the second peak analyzed in Sec. 5.4 (located at  $t = 25$  ms), the Polar ETC returned arrival angles of  $\phi = -135^\circ$ ,  $\theta = -35.26^\circ$ . Again, we see that when there are multiple arrivals within a given sample, the Polar ETC fails to provide meaningful results. The average error here was approximately  $31^\circ$  as opposed to the DIR's average error of about  $1^\circ$ . By comparison, the STCM method begins to break down with a very high arrival density within a given cross correlation time window, as discussed in Sec. 4.10.

### Long Narrow Room

To further compare the Polar ETC to the STCM method numerically, the long, narrow, rectangular room of Sec. 5.5 was used. The Polar ETC was used to analyze the first peak discussed there ( $t = 62$  ms) which was composed of two simultaneous arrivals. The Polar ETC calculated angles of arrival of  $\phi = 3.42^\circ$ ,  $\theta = 5.42^\circ$ . These angles are not very close to either of the actual values of  $\phi = -15.37^\circ$ ,  $\theta = 14.21^\circ$  or  $\phi = 20.55^\circ$ ,  $\theta = -3.35^\circ$ . This resulted in an average error of approximately  $14^\circ$ . By contrast, the STCM returned arrival angles of  $\phi = -14.62^\circ$ ,  $\theta = 14.18^\circ$  and  $\phi = 20.7^\circ$ ,  $\theta = -4.23^\circ$ , resulting in an average error of about  $0.5^\circ$  for the same peak.

The tallest peak in the IR was also analyzed. This peak is composed of four simultaneous arrivals. As expected, the Polar ETC only produced a single set of arrival angles,  $\phi = 90^\circ$ ,  $\theta = 57.8^\circ$ . This resulted in an average error of approximately  $72^\circ$  while the STCM method returned four solutions which resulted in an average error of approximately  $0.2^\circ$ .

### 6.2.2 Experimental Comparisons

Additional comparisons between the STCM and Polar ETC methods were conducted in an experimental setting. The setups used were the same as those presented in Sec. 5.7. The Polar ETC measurements were taken directly after the IRs for the STCM to reduce any chance of error due to non-identical setups. For each method, only the valid solution sets were considered, as there were erroneous solutions for both the Polar ETC and the STCM method.

### Single Reflector Experiment

For the single reflector case discussed in Sec. 5.7.2, the Polar ETC calculated angles of arrival of  $\phi = 1.9^\circ$ ,  $\theta = -2.4^\circ$  and  $\phi = 45.1^\circ$ ,  $\theta = -7.5^\circ$  for the direct and reflected arrivals, respectively. The actual arrival angles were  $\phi = 0.25^\circ$ ,  $\theta = 0^\circ$  and  $\phi = 35.5^\circ$ ,  $\theta = -3^\circ$  for the direct and reflected arrivals, respectively. This results in an average error of approximately  $4.5^\circ$ . This may be acceptable accuracy for some applications, though the azimuth of the reflection was off by about  $10^\circ$ . The STCM method performed better, producing an average error of approximately  $2.25^\circ$ .

### Single Small Reflector

In the case of a single, small reflecting surface, the Polar ETC again performed quite well. The actual arrival angles for the direct sound were  $\phi = -25^\circ$ ,  $\theta = 5.5^\circ$  while those calculated by the Polar ETC were  $\phi = -28.8^\circ$ ,  $\theta = 6.3^\circ$ , resulting in an average error of  $2.3^\circ$ . The STCM method had an average error of approximately  $1.1^\circ$ , which is slightly better, but when dealing with a difference of 1 degree, the performance of both could be considered good.

For the reflection, the Polar ETC produced angles of arrival of  $\phi = 47.9^\circ$ ,  $\theta = 2.9^\circ$ , resulting in an average error of  $2.4^\circ$  when compared to the actual arrival angles of  $\phi = 44.5^\circ$ ,  $\theta = 1^\circ$ . This still compares relatively well to the average error of  $1.25^\circ$  for the STCM method. All in all, both the STCM and Polar ETC methods performed quite well for this case.

### Dual Parallel Reflectors

The previous experiments were suitable for the Polar ETC because there were no simultaneous arrivals. The dual parallel reflector case discussed in Sec. 5.7.4, however, proved more of a challenge. The Polar ETC produced highly accurate angles of arrival

for the direct sound of  $\phi = 2.5^\circ$ ,  $\theta = 3.1^\circ$  resulting in an average error of  $0.3^\circ$ . This outperformed the DIR's average error of  $2^\circ$ . For the first-order reflections, the Polar ETC returned the angles  $\phi = -28.3^\circ$ ,  $\theta = -7.1^\circ$  resulting in an average error of about  $22^\circ$ . This was much worse than the DIR's average error of  $4^\circ$ .

### Dual Asymmetric Reflectors

The dual asymmetric reflector case discussed in Sec. 5.7.5 produced results that are not surprising, considering the simultaneous arrivals. The Polar ETC was able to localize the direct sound accurately with angles  $\phi = -15.1^\circ$ ,  $\theta = -0.7^\circ$ . However, it was unable to successfully localize either of the two simultaneous reflections, producing a single arrival at  $\phi = 54.3^\circ$ ,  $\theta = -16.7^\circ$ . The actual arrival angles were  $\phi = -15^\circ$ ,  $\theta = -1^\circ$  for the direct sound and  $\phi = -176^\circ$ ,  $\theta = 6.5^\circ$ ,  $\phi = 63^\circ$ ,  $\theta = -8^\circ$  for the reflections, resulting in an average error of  $0.4^\circ$  for the direct sound and  $50^\circ$  for the reflections. In contrast, the STCM identified the direct sound and both reflections, producing average errors of  $1.5^\circ$  and  $2^\circ$  for the direct and reflected arrivals, respectively.

### Variable Acoustics Chamber

The Polar ETC did not perform as well in the variable acoustics chamber. The direct sound was calculated at  $\phi = -5.5^\circ$ ,  $\theta = -15.1^\circ$ , resulting in an average error of about  $5.5^\circ$  while the STCM method averaged an error of approximately  $2^\circ$ . For the simultaneous first-order reflections, the Polar ETC returned the angles  $\phi = 89.6^\circ$ ,  $\theta = -18.1^\circ$ , which is actually quite close to the reflection off the wall, but it ignored the other reflection completely. The average error for just the wall reflection is  $6^\circ$  compared to  $3^\circ$  for the DIR. However, the Polar ETC's average error for the floor reflection is  $60^\circ$  compared to  $1.5^\circ$  for the DIR. Again, for the single second-order reflection, the Polar ETC produced very good results. The calculated angles were  $\phi$



$= 99.4^\circ$ ,  $\theta = -39.5^\circ$ , resulting in an average error of  $1.75^\circ$ , slightly better than the average error of  $2.5^\circ$  for the DIR.

### **Summary of Results**

In summary, Table 6.1 shows the average of the average errors for each scenario using both the Polar ETC and the STCM.

**Table 6.1** Average errors for the Polar ETC and STCM for the cases presented.

	Rectangular Room	Long Narrow Room	Single	Single Small	Dual Parallel	Dual Asymmetric	Variable Chamber
Polar ETC	33.4°	43.0°	4.5°	2.35°	11.05°	25.2°	34.8°
STCM	0.5°	0.35°	2.25°	1.18°	3.0°	1.75°	3.75°

Table 6.2 shows the average errors for both the Polar ETC and the STCM in the categories of single and multiple arrivals for both numerical and experimental cases.

**Table 6.2** Average errors for the Polar ETC and STCM classified by type.

	Numerical		Experimental	
	Single	Multiple	Single	Multiple
Polar ETC	0.0°	46.5°	2.1°	35.0°
STCM	0.5°	0.35°	1.6°	2.5°

### Off-Axis Results

In addition to the experiments just discussed, the effect of loudspeaker directivity on the measurements was assessed. This was done by rotating the speaker to point in a direction other than towards the microphone location. The results showed the reflecting surfaces closer to the radiation axis of the loudspeaker were weighted more strongly in the composite response. In some cases, the energy radiated toward one reflector dominated to such an extent that one reflector was correctly localized by the Polar ETC. For example, the reflector located at  $\phi = -176^\circ$ ,  $\theta = -8^\circ$  in the dual asymmetric reflector case was calculated to be  $\phi = -165.2^\circ$ ,  $\theta = -9.6^\circ$ . This is because the loudspeaker had its back almost directly facing the other reflecting panel, thus radiating very little energy in that direction. The average error in this case was approximately  $5.5^\circ$  for that one reflector, a great improvement over how the Polar ETC typically performs with simultaneous arrivals.

## 6.3 Discussion

The numerical and experimental comparisons brought to light some the issue of choosing a threshold using the commercial Polar ETC package. This proved to be an issue critical to extracting useful results from the Polar ETC in single arrival cases. The threshold is a user-defined level and only the arrivals in the composite ETC that exceed the threshold are localized. Several peaks in the ETC that were of interest spanned more than one sample, and therefore more than one angle was determined. In some cases, the discrepancy between these two calculated angles was as large as  $15^\circ$ . This has the potential to not only be confusing, but also to produce erroneous results. This is similar to having erroneous solution sets that meet the constraints in the DIR. The Polar ETC, therefore, also has problems with erroneous solution sets. This is an aspect of the measurement system that has, to the author's knowledge, not been previously disclosed or discussed. This is solely an artifact that is created when digital signals are used but, due to the need for processing of the signals, will always occur when a peak in the ETC spans more than one sample.

## 6.4 Chapter Conclusions

The STCM method has been compared with the Polar ETC for cases in which the Polar ETC results are valid and has been shown to perform favorably. The Polar ETC has been shown to be accurate for many single arrival cases, even outperforming the STCM at times. However, the fact that the Polar ETC is unable to distinguish between simultaneous arrivals is a demonstrated weakness. In multiple-arrival cases, the STCM has been shown to be far superior to the Polar ETC by identifying the separate arrivals with an acceptable degree of accuracy. The fact that the different arrivals were distinguishable is a large improvement over all methods that assume a

single arrival within a given sample.

# Chapter 7

## Conclusions

Over the years, many methods have been developed to determine the directions of acoustic arrivals in room acoustics problems. When one can accurately detect offending surfaces or room features that result in a nonideal room response, the cause is usually straightforward to fix. Most methods developed in the past have suffered from at least one of the following shortcomings: (1) simplified assumptions that are commonly invalidated when considering nonideal room responses, (2) a large amount of expensive equipment to adequately take the measurements, (3) long measurement times, and (4) insufficient accuracy.

This thesis has discussed the basics of some of these previous methods and has proposed a new short-time correlation method (STCM) for the measurement of directional information in rooms. Theoretical developments have been derived without the simplifying assumptions of past methods. The approach minimizes the amount of equipment and time necessary for practical implementation while still maintaining a high degree of accuracy. Many practical considerations necessary for its best implementation have also been presented.

## 7.1 Summary of Findings

The thesis has also provided numerical simulations to verify the STCM functionality. They have shown that the method is very accurate under typical conditions over a limited time range after the direct sound arrival. Beyond this time range, the results begin to be mixed with erroneous solution sets created by combinations of time delay components that meet the constraints.

Simple experiments have also been conducted in controlled acoustical environments to verify the utility of the method with measured data. When the proper steps were taken, the method shows great promise, though further work is required on signal-processing-related issues to make the method more robust and free from erroneous solutions. Two steps have been introduced preliminarily to help accomplish this goal: frequency-response magnitude compensation and noise gating. The combination helps increase the detail in the IRs (and therefore the cross correlation functions) while reducing noise. It thus allows for more accurate results while reducing the number of erroneous solution sets for given constraints.

Finally, the accuracy of the STCM has been compared to the results given by the Polar ETC method. It has been shown that for many of the most “interesting” features in modeled room IRs, the Polar ETC gives the directional result of a single arrival determined by the weighted sums given in Sec. 2.2, while for most cases, the new method is able to break down single peaks into constituent reflections with corresponding angles of arrival. Despite its known benefits, some characteristics of the STCM require further investigation and resolution. These include filtering effects and off-axis response characteristics of loudspeakers, and the comparison of results between the seven-microphone array and a four-microphone subset for enhanced sorting and elimination of erroneous solutions.

## 7.2 Contributions

This work has made several contributions to the field of Directional Impulse Response measurements:

1. An in-depth investigation and generalized analysis of the Polar ETC method that is not available in the literature.
2. Needed theoretical documentation of cross correlation techniques for DIRs.
3. Discovery and documentation of the errors produced by simultaneous arrivals for both the Polar ETC and cross correlation methods.
4. A new DIR measurement platform which includes:
  - (a) The ability to resolve simultaneous arrivals.
  - (b) The ability to conduct DIR measurements using either sequential or simultaneous measurements, allowing the user to choose between a minimum of equipment or a minimum of measurement time.
  - (c) A computationally efficient algorithm.
  - (d) A higher degree of accuracy than past methods.

It should be noted that the STCM is not the only solution to the DIR problem, nor is it a perfect solution. It has its own limitations. Nevertheless, the benefits of the method appear to outweigh its inconveniences or disadvantages when it is set up and implemented correctly. It is the author's intent that the STCM be developed into a more streamlined package that will allow both the professional acoustician and the dedicated hobbyist to better diagnose and correct room problems, and thus enhance listener enjoyment of program material.



### 7.3 Future Work

While this work has made significant progress in the development of an advanced DIR tool, there is still much work to be done to make it more user friendly and robust. The following tasks should be included in this effort:

1. Implement a more accurate frequency-response magnitude compensation to increase useful detail in the cross correlation function.
2. Implement an adaptive noise gating procedure to reduce noise in the cross correlation function, thereby reducing the number of erroneous solutions.
3. Investigate zero padding of the cross spectrum and other interpolation methods to further reduce errors of calculated solutions.
4. Develop automatic peak detection for the cross correlation function.
5. Create an intuitive graphical user interface (GUI).
6. Develop a method of accurate relative-amplitude determination for simultaneous arrivals.
7. Further reduce erroneous solution sets in the solution-finding algorithm by enhancing constraint equations.
8. Automate the determination of necessary constants for constraint equations.
9. Implement built-in filtering functions to eliminate the effects of nonideal microphone directivities.
10. Investigate ways to mitigate the detrimental effects produced by off-axis radiation from loudspeakers.

- 
11. Use every possible microphone pairing to further reduce the occurrence of erroneous solutions.
  12. Investigate the possibility of implementing the DIMUS or CLEAN-SC algorithms for deconvolving the spatial IR of the microphone array if every possible microphone pair is utilized.

The author encourages additional work in these and other related areas.



# Bibliography

- [1] S. Bech, “Timbral aspects of reproduced sound in small rooms. I,” *J. Acoust. Soc. Am.* **97**, 1717–1725 (1995).
- [2] S. Bech, “Timbral aspects of reproduced sound in small rooms. II,” *J. Acoust. Soc. Am.* **99**, 3539–3548 (1996).
- [3] S. Bech, “Timbral aspects of reproduced sound in small rooms. III,” *J. Acoust. Soc. Am.* **103**, 434–445 (1998).
- [4] R. Thiele, “Richtungsverteilung und Zeitfolge der Schallrückwürfe in Räumen,” *Acustica* **3**, 291–302 (1953).
- [5] P. D’Antonio, J. Konnert, F. Becker, and C. Bilello, “Sound Intensity and Interaural Cross-Correlation Measurements Using Time-Delay Spectrometry,” *J. Audio Eng. Soc.* **37**, 659–673 (1989).
- [6] F. Becker, “The Polar Energy Time Curve,” In *Proc. AES 6th Int. Conf. on Sound Reinforcement (Nashville, TN, May 1988)*,
- [7] Y. Yamasaki and T. Itow, “Measurement of spatial information in sound field by closely located four point microphone method,” *J. Acoust. Soc. of Jpn.* (E) **10**, 101–110 (1989).

- 
- [8] K. Sekiguchi, S. Kimura, and T. Hanyuu, "Analysis of Sound Field on Spatial Information Using a Four-Channel Microphone System Based on Regular Tetrahedron Point Peak Method," *Applied Acoustics* **37**, 305–323 (1992).
- [9] C. Choi, L. Kim, Y. Oh, S. Doo, and K. Sung, "Measurement of Early Reflections in a Room with Five Microphone System," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences* **E86**, 3283–3287 (2003).
- [10] B. Gover, J. Ryan, and M. Stinson, "Microphone Array Measurement System for Analysis of Directional and Spatial Variations of Sound Fields," *J. Acoust. Soc. Am.* **112**, 1980–1991 (2002).
- [11] B. Gover, J. Ryan, and M. Stinson, "Measurements of Directional Properties of Reverberant Sound Fields in Rooms Using a Spherical Microphone Array," *J. Acoust. Soc. Am.* **116**, 2138–2148 (2003).
- [12] B. Gover, "Directional Measurement of Airborne Sound Transmission Paths Using a Spherical Microphone Array," *J. Audio Eng. Soc.* **53**, 787–795 (2005).
- [13] R. Essert, "Measurement of spatial impulse responses with a Soundfield Microphone (A)," *J. Acoust. Soc. Am.* **100**, 2837 (1996).
- [14] R. Essert, "Progress in concert hall design - developing an awareness of spatial sound and learning how to control it," *EBU Technical Review* **Winter 1997**, 31–39 (1997).
- [15] A. Farina, A. Capra, L. Conti, P. Martignon, and F. Fazi, "Measuring spatial impulse responses in concert halls and opera houses employing a spherical microphone array," In *19th International Congress on Acoustics (ICA) (Madrid, Spain, September 2007)*,

- 
- [16] A. Farina, A. Capra, P. Martignon, and S. Fontana, “Measuring impulse responses containing complete spatial information,” In *22nd AES-UK Conference 2007 (Cambridge, UK, April 2007)*,
- [17] A. Farina and L. Tronchin, “3D Impulse Response measurements on S.Maria del Fiore Church, Florence, Italy,” In *Proc. of ICA98 - International Conference on Acoustics - (Seattle, WA, June 1998)*,
- [18] C. Noël, V. Planeau, and D. Habault, “A New temporal Method for the Identification of Source Directions in a Reverberant Hall,” *J. Sound Vib.* **296**, 518–538 (2006).
- [19] A. Abdou and R. Guy, “Directional Accuracy of Transient Sound Employing an Intensity Probe,” *Applied Acoustics* **50**, 65–77 (1997).
- [20] L. Beranek, *Music, Acoustics and Architecture* (Wiley, New York, 1962).
- [21] L. Cremer and H. Müller, *Principles and Applications of Room Acoustics* (Applied Science, Essex, UK, 1982), Vol. 1.
- [22] A. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications* (Acoustical Society of America, New York, 1989).
- [23] G. Elko, in *Acoustic Signal Processing for Telecommunication*, S. Gay and J. Benesty, eds., (Kluwer Academic, Norwell, MA, 2000).
- [24] S. Neeley and J. Allen, “Invertibility of a room impulse response,” *J. Acoust. Soc. Am.* **66**, 165–169 (1979).
- [25] A. Oppenheim and R. S. with J.R. Buck, *Discrete-Time Signal Processing* (Prentice Hall, New Jersey, 1990).

- 
- [26] R. Heyser, “Determination of Loudspeaker Signal Arrival Times Part 1,” *J. Audio Eng. Soc.* **19**, 734–743 (1971).
- [27] R. Heyser, “Determination of Loudspeaker Signal Arrival Times Part 2,” *J. Audio Eng. Soc.* **19**, 829–834 (1971).
- [28] R. Heyser, “Determination of Loudspeaker Signal Arrival Times Part 3,” *J. Audio Eng. Soc.* **19**, 902–905 (1971).
- [29] J. Allen and D. Berkley, “Image method for efficiently simulating small-room acoustics,” *J. Acoust. Soc. Am.* **65**, 943–950 (1979).
- [30] J. Eargle, *The Microphone Book* (Focal Press, Woburn, MA, 2001), pp. 82–85.
- [31] J. Vanderkooy and S. Lipshitz, “Uses and abuses of the energy time curve,” *J. Audio Eng. Soc.* **38**, 819–836 (1990).
- [32] J. D.B. Keele, “The Analytic Impulse and the Energy-Time Curve: The Debate Continues,” In *93rd Convention of the AES (October 1992)*, (1992).
- [33] P. Fellgett, “Ambisonics. Part One: General System Description,” *Studio Broadcast Engineering* **17**, 20–40 (1975).
- [34] A. Schuster, *An Introduction to the Theory of Optics* (Edward Arnold, London, 1904).
- [35] C. Knapp and G. Carter, “The generalized correlation method for estimation of time delay,” *IEEE Transactions on Acoustics, Speech, and Signal Processing* **24(4)**, 320–327 (1976).
- [36] V. Anderson, “Digital Array Phasing,” *J. Acoust. Soc. Am.* **32**, 867–870 (1960).

- [37] P. Sitsjma, “CLEAN based on spatial source coherence,” *Int. J. Aeroacoustics* **6**, 357–374 (2007).
- [38] R. T. Chester, *Error Sensor Strategies for Active Noise Control and Active Acoustic Equalization in a Free Field* (M.S. Thesis, Brigham Young University, Provo, UT, 2008).
- [39] D. Nutter, *Sound Absorption and Sound Power Measurements in Reverberations Chambers Using Energy Density Methods* (M.S. Thesis, Brigham Young University, Provo, UT, 2006).
- [40] L. Kinsler, A. Frey, A. Coppens, and J. Sanders, *Fundamentals of Acoustics*, fourth ed. (John Wiley and Sons, Inc., New York, NY, 2000), p. 428.
- [41] J. Proakis and D. Manolakis, *Digital Signal Processing*, fourth ed. (Pearson Prentice Hall, Upper Saddle River, NJ, 2007), pp. 456–458.
- [42] W. Strong and G. Plitnik, *Music Speech Audio* (BYU Academic Publishing, Provo, UT, 1977), pp. 284–285.
- [43] G. Arfken and H. Weber, *Mathematical Methods for Physicists*, sixth ed. (Elsevier Academic Press, New York, NY, 2005), p. 250.





# Appendix A

The Matlab code for all of the programs used in this thesis is given here.

## A.1 Sources in Free Space IR Generation

This code is an adaptation of a program written by Ryan Chester.

```
1 clear all,      close all tic
2
3 FS=192000; fs=FS/2; c=343; rho=1.21; d=1.5*.0254;
4
5 %% Setting up the relationships between the frequency and time domains
6 N=2^15; dt=1/fs; t=0:dt:dt*(N-1); df=1/(dt*N); f=(df:df:fs)-df;
7 k=2*pi*f/c;
8
9 DT=1/FS; T=0:DT:DT*(N-1); DF=1/(DT*N); F=DF:DF:FS; K=2*pi*F/c;
10
11 %% Setting up sensor locations, source locations, source strengths etc.
12 mic=[0,0,0];
13
14 chord=[3.972, -0.86, -.07;
15        3.8978, 7.6498, -1.2066;
```

```
16         -8.5933, 0.6009, -.9815];
17 %% Pressure calculations
18
19 q0=(2+.2*j)./k; A=j*rho*c*k.*q0/(4*pi); p=zeros(size(k));
20 sch=size(chord); for n=1:sch(1)
21     r=sqrt((mic(1)-chord(n,1))^2+(mic(2)-chord(n,2))^2+(mic(3)-chord(n,3))^2);
22     p=p+A.*exp(-j*k*r)/r;
23 end clear sch r dx dy dz
24
25 %% Frequency response
26 pfr=p./A; % pressure frequency response
27
28 %% Calculation of impulse response
29 pfrL=[0,(pfr(2:end)),0,fliplr(conj(pfr(2:end)))];%Extended pressure frequency response
30
31 pirL=ifft(pfrL); pir=pirL(1:N);
32
33 ipir=ifft(1./fft(pir)); ipir=fftshift(ipir);
34
35
36 %% Plotting the responses
37 figure plot(t(1:length(pir)/2),pir(1:length(pir)/2),'k')
38 title('Impulse Response') xlabel('Time (s)') ylabel('Normalized
39 Magnitude')
40
41 % break
42 %%
43 wavwrite(pir,FS,'F:\3DIRS\Simulations\DualAsymmetricModelOrigin.wav')
```

## A.2 Allen-Berkley IR Generation

This code is a Matlab translation of the image source IR generation program originally written by Allen and Berkley. The code for counting the actual arrival density as shown in Fig. 4.16 along with the code for writing the .Wav file has been appended to the end.

```

1  %Program that generates impulse responses based on the image source method
2  %
3  %
4  %
5  %
6  %
7  %
8  %
9  %
10 %
11
12 clear all; close all;
13
14 %Define constants
15 c=343; FS=192000; T=1/FS; NPTS=2^16; HT=zeros(1,NPTS); tmax=NPTS/FS;
16 dt=T; t=0:dt:tmax-dt; d=.0254*1.5;
17 num=zeros(1,NPTS); %This is to count the number of arrivals in a single IR peak
18 ampind=ones(1,NPTS); angnum=ones(1,NPTS);
19
20 %Room Dimensions
21 lx=30; ly=8; lz=4.5; RL=[lx/(c*T) ly/(c*T) lz/(c*T)];
22
23 %Source position

```



```
24 x0=25; y0=1.5; z0=1.25; R0=[x0/(c*T) y0/(c*T) z0/(c*T)];
25
26 %Receiver position
27 x=5; y=7; z=2.5; R=[x/(c*T) y/(c*T) z/(c*T)];
28
29 %Reflection coefficients
30 betafront=.85; betaback=0.85; betaleft=0.85; betaright=0.85;
31 betaup=0.8; betadown=0.7; Beta=[[betaback betaleft
32 betadown];[betafront betaright betaup]];
33
34 N1=floor(NPTS/(2*RL(1))+1); N2=floor(NPTS/(2*RL(2))+1);
35 N3=floor(NPTS/(2*RL(3))+1);
36
37 tic
38
39 for nx=-N1:N1
40     for ny=-N2:N2
41         for nz=-N3:N3
42
43             NR=[nx ny nz];
44
45             DR=R;
46             DR0=R0;
47             Delp=zeros(1,8);
48             R2L=zeros(1,3);
49             RP=zeros(3,8);
50             I1=1;
51             for L=-1:2:1
52                 for J=-1:2:1
53                     for K=-1:2:1
54
```

```

55             RP(1,I1)=DR(1)+L*DR0(1);
56             RP(2,I1)=DR(2)+J*DR0(2);
57             RP(3,I1)=DR(3)+K*DR0(3);
58
59             I1=I1+1;
60
61             end
62         end
63     end
64
65     R2L(1)=2*RL(1)*NR(1);
66     R2L(2)=2*RL(2)*NR(2);
67     R2L(3)=2*RL(3)*NR(3);
68     theta=zeros(1,8);
69     phi=zeros(1,8);
70
71     for I=1:8
72
73         Rang=zeros(1,3);
74
75         delsq=0;
76         for J=1:3
77
78             R1=R2L(J)-RP(J,I);
79             delsq=delsq+R1^2;
80             Rang(J)=R1;
81
82         end
83
84         Delp(I)=sqrt(delsq);
85         theta(I)=acosd(Rang(3)/sqrt(Rang(1)^2+Rang(2)^2+Rang(3)^2))-90;

```



```
117         if phinew(IDtemp,angnum(IDtemp))<-180
118
119             phinew(IDtemp,angnum(IDtemp))=...
120             phinew(IDtemp,angnum(IDtemp))+360;
121
122         end
123
124         angnum(IDtemp)=angnum(IDtemp)+1;
125
126     end
127
128     end
129 end
130 end
131
132 end
133 end
134 end
135
136 amp=amp/max(amp);
137
138 toc
139
140 f=1000;    %Low pass cutoff
141 w=8*atan(1)*f; T=1e-4; R1=exp(-w*T); R2=R1; B1=2*R1*cos(w*T);
142 B2=-R1*R1; A1=-(1+R2); A2=R2; Y1=0; Y2=0; Y0=0; for a=1:NPTS
143     X0=HT(a);
144     HT(a)=Y0+A1*Y1+A2*Y2;
145     Y2=Y1;
146     Y1=Y0;
147     Y0=B1*Y1+B2*Y2+X0;
```



```

148 end
149
150 HT=0.99*HT/max(abs(HT));
151
152 figure subplot(2,1,1) plot(t,HT,'k') title('Impulse Response')
153 ylabel('Amplitude') subplot(2,1,2) plot(t,num,'k') title('Number of
154 Arrivals') xlabel('Time (s)') ylabel('Number of Arrivals')
155
156 wavwrite(HT,FS,'F:\3DIRS\OriginIR.wav')
157 % break
158 win=21; arrden=zeros(1,length(num)); for a=win+1:length(num)-win
159     arrden(a)=sum(num(a-win:a+win));
160 end
161
162 for a=length(num)-win+1:length(num)
163     arrden(a)=sum(num(a-win:end));
164 end
165
166 wint=win/192000;
167 arrdentheory=4/3*pi*c^3/(lx*ly*lz)*((t+wint).^3-(t-wint).^3);
168
169 figure plot(t,arrden,'g',t,arrdentheory,'k—') title('Arrival
170 Density') xlabel('Time (s)') ylabel('Arrivals within 2\Deltat')
171 legend('Calculated','Predicted')

```

### A.3 Modified Modal Expansion

The code used for the modified modal expansion in Sec. 4.4.3 is based on code written by Dave Nutter.

```
1 clear all; close all; clc;
2
3 tic Lx = 30; Ly = 8; Lz = 4.5; c = 343; rho = 1.21; V=Lx*Ly*Lz;
4 S=2*Lx*Ly+2*Lx*Lz+2*Ly*Lz;
5
6 d=1.5*.0254;
7
8 %Source position
9 x0=25; y0=1.5; z0=1.25;
10
11 %Array Origin Location
12 x=5; y=7; z=2.5;
13
14 alphax=1-.85^2; alphay=1-.85^2; alphaz=((1-.8^2)+(1-.7^2))/2;
15
16 A=(2*alphax*Ly*Lz+2*alphay*Lx*Lz+2*alphaz*Lx*Ly);
17 alphas=A/S; %spatially averaged absorption coefficient
18
19 fcut=1000; %Specify max frequency of bandwidth
20
21 T60=.161*V/(-log(1-alphas)*S); %Calculate Norris-Eyring T60
22 Δf=2.2/T60; %modal bandwidth in Hz
23
24 df=2; %Specify frequency resolution
25
26 Nx=ceil(2/c*fcut*Lx); %Calculate number of modes needed along each axis
27 Ny=ceil(2/c*fcut*Ly); Nz=ceil(2/c*fcut*Lz);
28 % break
29 m = 1;
30
31 fndata = zeros((Nx+1)*(Ny+1)*(Nz+1),7); for nx=0:Nx
```

```
32     for ny=0:Ny
33         for nz=0:Nz
34             fndata(m,1) = c/2*sqrt((nx/Lx)^2+(ny/Ly)^2+(nz/Lz)^2);
35             fndata(m,2)=nx; fndata(m,3)=ny; fndata(m,4)=nz;
36             if nx ==0
37                 fndata(m,5)=1;
38             else
39                 fndata(m,5)=2;
40             end
41             if ny ==0
42                 fndata(m,6) = 1;
43             else
44                 fndata(m,6) = 2;
45             end
46             if nz ==0
47                 fndata(m,7) = 1;
48             else
49                 fndata(m,7) = 2;
50             end
51             m=m+1;
52         end
53     end
54 end
55
56 fndata = sortrows(fndata)';
57
58 fmax=ceil(fndata(1,end));    %Specify max frequency present
59 %% Prelims
60 Lxinv=1/Lx; Lyinv=1/Ly; Lzinv=1/Lz;
61
62 fw2=10*ceil( $\Delta f$ );
```



```

94     psiS = x0psi*y0psi*z0psi;
95     psiR = xpsi'*ypsi*zpsi;
96     Almn=-j*rho*c*k*Q*LinV/V*psiS/(k^2-kn(N)^2-j*kn(N)/(4*V)*A);
97     p(findx) = p(findx)+Almn*psiR;
98
99     end
100
101     end
102     farray(findx)=f;
103     findx=findx+1;
104 end
105
106 pfr1=[farray;p(1:length(farray))];
107 % save /fslhome/briguy/pfrOriginLN.mat pfr1
108 toc

```

## A.4 STCM Program

The actual program used for calculating the angles of arrival is here. This is the finished form, though several versions have existed.

```

1  %Program for taking the seven cartesian impulse responses, taking the cross
2  %correlation, determining whether there are multiple simultaneous arrivals
3  %and then computing the polar angles of arrival
4
5  clc; clear all; close all;
6
7  c=343;           %Speed of sound
8  d=1.5*.0254;    %Distance from origin to mic in meters

```

```
9
10 %Read in the seven impulse response files and the sampling frequency
11 [Front,FS]=wavread('\\kirchhoff\acoustics\students\Brian
12 Thornock\Experimental Data\6 August 2009\SingleReflectorFront.wav');
13 Back=wavread('\\kirchhoff\acoustics\students\Brian
14 Thornock\Experimental Data\6 August 2009\SingleReflectorBack.wav');
15 Right=wavread('\\kirchhoff\acoustics\students\Brian
16 Thornock\Experimental Data\6 August 2009\SingleReflectorRight.wav');
17 Left=wavread('\\kirchhoff\acoustics\students\Brian
18 Thornock\Experimental Data\6 August 2009\SingleReflectorLeft.wav');
19 Up=wavread('\\kirchhoff\acoustics\students\Brian
20 Thornock\Experimental Data\6 August 2009\SingleReflectorUp.wav');
21 Down=wavread('\\kirchhoff\acoustics\students\Brian
22 Thornock\Experimental Data\6 August 2009\SingleReflectorDown.wav');
23 Origin=wavread('\\kirchhoff\acoustics\students\Brian
24 Thornock\Experimental Data\6 August
25 2009\SingleReflectorOrigin.wav');
26 % break
27
28 %Truncate IRs to speed up calculation
29 Front=Front(1:9000); Back=Back(1:9000); Right=Right(1:9000);
30 Left=Left(1:9000); Up=Up(1:9000); Down=Down(1:9000);
31 Origin=Origin(1:9000);
32
33 %Define Some More physical constants
34 t=0:1/FS:length(Front)/FS-1/FS; %total time record of the impulse response
35  $\Delta t_{max} = \text{round}(FS*2*d/c)$ ; %Max time delay (in samples) for mic spacing
36  $\Delta t_{max4} = \text{round}(FS*d/c)$ ; %Max time delay for four-mic configuration
37 degfac=1; %Factor eliminating degenerate solution sets
38 timefac=7; %Constraint factor for 7 mic config
39 timefac4=5; %Constraint factor for 4 mic config
```

```
40  ovlpfac=7;                %Factor for choosing overlapping solutions
41
42  %Display the Origin IR
43  figure(1) plot(t,Origin) title('Plot of the Origin IR') xlabel('Time
44  (s)') ylabel('Magnitude') ylim([1.1*min(Origin) 1.1*max(Origin)])
45  pause (1)
46
47  %Zoom in on the peak of interest for greater accuracy
48  disp('Pick a Range of the Peak You are Interested in')
49  [time,ampt]=ginput(2); tpremod=time(1):1/FS:time(2);
50
51  Originpremod=Origin(round(time(1)*FS):round(time(1)*FS)+length(tpremod)-1);
52
53  close all;
54
55  figure plot(tpremod,Originpremod) title('Zoomed plot of the origin
56  IR') ylim([1.1*min(Originpremod) 1.1*max(Originpremod)])
57
58  pause(1)
59
60  choice=input('Would you like to zoom more?(y=1/n=0) - ');
61
62  close all;
63
64  while choice==1
65      figure
66      plot(tpremod,Originpremod)
67      title('Plot of the Origin IR')
68      xlabel('Time (s)')
69      ylabel('Magnitude')
70      ylim([1.1*min(Origin) 1.1*max(Origin)])
```

```
71     pause (1)
72
73     %Zoom in on the peak of interest for greater accuracy
74     disp('Pick a Range of the Peak You are Interested in')
75     [time,ampt]=ginput(2);
76     tpremod=time(1):1/FS:time(2);
77
78     Originpremod=Origin(round(time(1)*FS):round(time(1)*FS)+length(tpremod)-1);
79
80     close all;
81
82     figure
83     plot(tpremod,Originpremod)
84     title('Zoomed plot of the origin IR')
85     ylim([1.1*min(Originpremod) 1.1*max(Originpremod)])
86
87     pause(1)
88
89     choice=input('Would you like to zoom more?(y=1/n=0) - ');
90
91 end
92
93 %Make the two vectors the same length
94 Originpremod=Origin(round(time(1)*FS):round(time(1)*FS)+length(tpremod)-1);
95 close all; clear ampt; clear time;
96
97 figure(2) plot(tpremod,Originpremod) title('Zoomed Plot of the
98 Origin IR') xlabel('Time (s)') ylabel('Magnitude')
99 ylim([1.1*min(Originpremod) 1.1*max(Originpremod)]) clc;
100
101 %This gives the sample numbers corresponding to the time values
```



```
102 %chosen
103 disp('Pick the Peak You Want to Analyze') t0=ginput(1);
104
105 %Pick off just the time component, not amplitude
106 t0=t0(1); tpeak=round(FS*t0); clc; close all;
107
108 %Take the time window of the IRs to compute the cross correlation about the
109 %certain peak in the IR
110 Frontmod=Front(tpeak- $\Delta$ tmax:tpeak+ $\Delta$ tmax);
111 Backmod=Back(tpeak- $\Delta$ tmax:tpeak+ $\Delta$ tmax);
112 Rightmod=Right(tpeak- $\Delta$ tmax:tpeak+ $\Delta$ tmax);
113 Leftmod=Left(tpeak- $\Delta$ tmax:tpeak+ $\Delta$ tmax);
114 Upmod=Up(tpeak- $\Delta$ tmax:tpeak+ $\Delta$ tmax);
115 Downmod=Down(tpeak- $\Delta$ tmax:tpeak+ $\Delta$ tmax);
116 Originmod=Origin(tpeak- $\Delta$ tmax:tpeak+ $\Delta$ tmax);
117
118 Frontmod4=Front(tpeak- $\Delta$ tmax4:tpeak+ $\Delta$ tmax4);
119 Leftmod4=Left(tpeak- $\Delta$ tmax4:tpeak+ $\Delta$ tmax4);
120 Upmod4=Up(tpeak- $\Delta$ tmax4:tpeak+ $\Delta$ tmax4);
121 Originmod4=Origin(tpeak- $\Delta$ tmax4:tpeak+ $\Delta$ tmax4);
122
123 clear Front; clear Back; clear Right; clear Left; clear Up; clear
124 Down;
125
126 tmod=(tpeak- $\Delta$ tmax)/FS:1/FS:(tpeak+ $\Delta$ tmax)/FS;
127
128 figure(2) plot(tmod,Originmod) title('Plot of section of IR to be
129 analyzed') xlabel('Time (s)') ylabel('Magnitude')
130
131 pause(1)
132
```

```
133 close all;
134
135 Frontfmod=fft(Frontmod); Backfmod=fft(Backmod);
136 Rightfmod=fft(Rightmod); Leftfmod=fft(Leftmod); Upfmod=fft(Upmod);
137 Downfmod=fft(Downmod);
138
139 Frontfmod4=fft(Frontmod4); Leftfmod4=fft(Leftmod4);
140 Upfmod4=fft(Upmod4); Originfmod4=fft(Originmod4);
141
142 Sfb=Backfmod.*conj(Frontfmod); Srl=Leftfmod.*conj(Rightfmod);
143 Sud=Downfmod.*conj(Upfmod);
144
145 Sfo=Originfmod4.*conj(Frontfmod4); Slo=Leftfmod4.*conj(Originfmod4);
146 Suo=Originfmod4.*conj(Upfmod4);
147
148 clear Frontfmod; clear Backfmod; clear Rightfmod; clear Leftfmod;
149 clear Upfmod; clear Downfmod; clear Frontfmod4; clear Leftfmod4;
150 clear Upfmod4; clear Originfmod4;
151
152 Rfb=fftshift(ifft(Sfb)); Rrl=fftshift(ifft(Srl));
153 Rud=fftshift(ifft(Sud));
154
155 Rfo=fftshift(ifft(Sfo)); Rlo=fftshift(ifft(Slo));
156 Ruo=fftshift(ifft(Suo));
157
158 clear Sfb; clear Srl; clear Sud; clear Slo; clear Sfo; clear Suo;
159
160 tcorr=-Δtmax/FS+1/(2*FS)-.03/FS:1/FS:Δtmax/FS+1/(2*FS)-.03/FS;
161 figure(4) plot(tcorr,abs(Rfb)) title('Cross Correlation in X')
162 xlabel('Time Delay (s)') ylabel('Magnitude')
163
```

```
164 [tx,xamp]=ginput; clear xamp;
165
166 close;
167
168 figure(5) plot(tcorr,abs(Rr1)) title('Cross Correlation in Y')
169 xlabel('Time Delay (s)') ylabel('Magnitude')
170
171 [ty,yamp]=ginput; clear yamp; ty=-ty;
172
173 close;
174
175 figure(6) plot(tcorr,abs(Rud)) title('Cross Correlation in Z')
176 xlabel('Time Delay (s)') ylabel('Magnitude')
177
178 [tz,zamp]=ginput; clear zamp;
179
180 clc; close all;
181 %% Seven Mic Processing
182 angind=1; theta=0; phi=0;
183
184 for l=1:length(tx)
185     for m=1:length(ty)
186         for n=1:length(tz)
187             if abs(sqrt(tx(l)^2+ty(m)^2+tz(n)^2)-2*d/c)<(2*d/c)/timefac
188
189                 theta(angind)=asind(tz(n)/sqrt(tx(l)^2+ty(m)^2+tz(n)^2));
190                 phi(angind)=(180/pi*atan2(ty(m),tx(l)));
191
192                 if phi(angind)<-180
193
194                     phi(angind)=phi(angind)+360;
```

```
195
196         end
197
198         if phi(angind)>180
199             phi(angind)=phi(angind)-360;
200         end
201
202         times(angind,:)=[tx(l) ty(m) tz(n)];
203
204         angind=angind+1;
205
206     else
207         end
208     end
209 end
210 end
211
212 %Implement loops to eliminate multiple solution sets
213 if length(theta)>1
214     for aa=1:length(theta)
215         %As long as we are not at the last theta entry, compare the
216         %current entry to all higher ones for equality
217         if aa<length(theta)
218             for bb=1:length(theta)-aa
219                 %The choosiness of how close solution sets can be is
220                 %chosen here. In this case, it is for theta and phi to
221                 %within one degree.
222                 if abs(theta(aa)-theta(aa+bb))≤degfac...
223                     && abs(phi(aa)-phi(aa+bb))≤degfac
224
225                     %Set theta and phi to bogus values that can be easily
```

```
226             %found in a loop and eliminated
227             theta(aa+bb)=370;
228             phi(aa+bb)=370;
229
230         end
231     end
232 end
233
234
235 end
236 %Find the indices with the bogus value and eliminate them
237 indextheta=find(theta==370);
238 indexphi=find(phi==370);
239 theta(indextheta)=[];
240 phi(indexphi)=[];
241
242 else end
243
244 %% Four Mic Selection
245 %Now do the four mic stuff
246 tcorr4=-Δtmax4/FS+1/(2*FS)-.13/FS:1/FS:Δtmax4/FS+1/(2*FS)-.13/FS;
247
248 figure(7) plot(tcorr4,abs(Rfo)) title('Cross Correlation in X')
249 xlabel('Time Delay (s)') ylabel('Magnitude')
250
251 [tx4,xamp]=ginput; clear xamp;
252
253 close;
254
255 figure(8) plot(tcorr4,abs(Rlo)) title('Cross Correlation in Y')
256 xlabel('Time Delay (s)') ylabel('Magnitude')
```

```
257
258 [ty4,yamp]=ginput; clear yamp;
259
260 ty4=-ty4;
261
262 close;
263
264 figure(9) plot(tcorr4,abs(Ruo)) title('Cross Correlation in Z')
265 xlabel('Time Delay (s)') ylabel('Magnitude')
266
267 [tz4,zamp]=ginput; clear zamp;
268
269 close;
270
271 %% Four Mic Processing
272 angind4=1;
273
274 for l=1:length(tx4)
275
276     for m=1:length(ty4)
277         for n=1:length(tz4)
278             if abs(sqrt(tx4(l)^2+ty4(m)^2+tz4(n)^2)-d/c)<(d/c)/timefac4
279
280                 theta4(angind4)=asind(tz4(n)/sqrt(tx4(l)^2+ty4(m)^2+tz4(n)^2));
281                 phi4(angind4)=(180/pi*atan2(ty4(m),tx4(l)));
282
283                 if phi4(angind4)<-180
284
285                     phi4(angind4)=phi4(angind4)+360;
286
287                 end
```

```
288
289         if phi4(angind4)>180
290             phi4(angind4)=phi4(angind4)-360;
291         end
292
293         times4(angind4,:)=[tx4(1) ty4(m) tz4(n)];
294
295         angind4=angind4+1;
296
297         else
298             end
299     end
300 end
301 end
302
303
304 if length(theta4)>1
305     for aa=1:length(theta4)
306         %As long as we are not at the last theta entry, compare the
307         %current entry to all higher ones for equality
308         if aa<length(theta4)
309             for bb=1:length(theta4)-aa
310                 %The choosiness of how close solution sets can be is
311                 %chosen here. In this case, it is for theta and phi to
312                 %within one degree.
313                 if abs(theta4(aa)-theta4(aa+bb))≤degfac...
314                     && abs(phi4(aa)-phi4(aa+bb))≤degfac
315
316                 %Set theta and phi to bogus values that can be easily
317                 %found in a loop and eliminated
318                 theta4(aa+bb)=370;
```

```
319             phi4(aa+bb)=370;
320
321             end
322         end
323     end
324
325
326     end
327     %Find the indices with the bogus value and eliminate them
328     indextheta4=find(theta4==370);
329     indexphi4=find(phi4==370);
330     theta4(indextheta4)=[ ];
331     phi4(indexphi4)=[ ];
332
333 else end
334
335 %Now use a loop to cross compare between 4 and 7 mic
336 newind=1; for ind1=1:length(theta4)
337     for ind2=1:length(theta)
338
339         if abs(theta4(ind1)-theta(ind2))≤ovlpfac...
340             && abs(phi4(ind1)-phi(ind2))≤ovlpfac
341                 %The user can choose to use the 4 mic, 7 mic or a combination
342                 %for what the resulting angle should be. In some cases, one is
343                 %better than the others, but that can only be determined
344                 %afterwards
345                 thetanew(newind)=theta4(ind1);
346                 phinew(newind)=phi4(ind1);
347                 newind=newind+1;
348             else
349                 end
```



```
350
351     end
352 end
353
354 if length(thetanew)>1
355     for aa=1:length(thetanew)
356         %As long as we are not at the last theta entry, compare the
357         %current entry to all higher ones for equality
358         if aa<length(thetanew)
359             for bb=1:length(thetanew)-aa
360                 %The choosiness of how close solution sets can be is
361                 %chosen here. In this case, it is for theta and phi to
362                 %within one degree.
363                 if abs(thetanew(aa)-thetanew(aa+bb))≤degfac...
364                     && abs(phinev(aa)-phinev(aa+bb))≤degfac
365
366                     %Set theta and phi to bogus values that can be easily
367                     %found in a loop and eliminated
368                     thetanew(aa+bb)=370;
369                     phinev(aa+bb)=370;
370
371                 end
372             end
373         end
374
375
376     end
377     %Find the indices with the bogus value and eliminate them
378     indexthetanew=find(thetanew==370);
379     indexphinev=find(phinev==370);
380     thetanew(indexthetanew)=[];
```

```

381     phinew(indexphinew)=[];
382
383 else end
384
385 %Scatter plot of arrivals
386 figure(10) scatter(phinew,thetaneu) title('Scatter plot of spherical
387 angles of arrival') xlabel('\phi (deg)') ylabel('\theta (deg)')
388 xlim([-180 180]) ylim([-90 90])
389
390 %Display the numerical values in command window
391 disp('phi = ') disp(phinew) disp('theta = ') disp(thetaneu)

```

## A.5 Polar ETC Programs

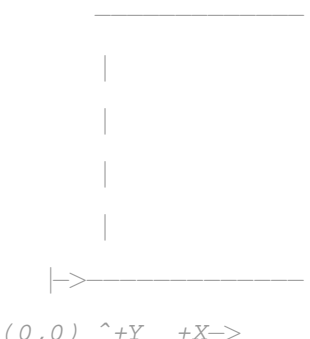
The different programs used for the Polar ETC (including the cardioid IR generation program) are included here.

### A.5.1 Cardioid IR Generation

```

1 %Cardioid impulse responses based on the image source method
2 %
3 %
4 %
5 %
6 %
7 %
8 %
9 %
10 %

```



```
11
12 clear all; close all;
13
14 %Define constants
15 c=343; FS=192000; T=1/FS; NPTS=2^16; HT=zeros(1,NPTS); tmax=NPTS/FS;
16 dt=T; t=0:dt:tmax-dt; d=.0254*.25;
17 num=zeros(1,NPTS); %This is to count the number of arrivals in a single IR peak
18 ampind=ones(1,NPTS); angnum=ones(1,NPTS);
19 % amp=zeros(NPTS,1);
20
21 %Room Dimensions
22 lx=30; ly=8; lz=4.5; RL=[lx/(c*T) ly/(c*T) lz/(c*T)];
23
24 %Source position
25 x0=25; y0=1.5; z0=1.25; R0=[x0/(c*T) y0/(c*T) z0/(c*T)];
26
27 %Receiver position
28 x=5; y=7; z=2.5; R=[x/(c*T) y/(c*T) z/(c*T)];
29
30 %Reflection coefficients
31 betafront=.85; betaback=0.85; betaleft=0.85; betaright=0.85;
32 betaup=0.8; betadown=0.7; Beta=[[betaback betaleft
33 betadown];[betafront betaright betaup]];
34
35 N1=floor(NPTS/(2*RL(1))+1); N2=floor(NPTS/(2*RL(2))+1);
36 N3=floor(NPTS/(2*RL(3))+1);
37
38 tic
39
40 for nx=-N1:N1
41     for ny=-N2:N2
```

```
42     for nz=-N3:N3
43
44         NR=[nx ny nz]; % modified by Buye. before was NR=[N1 N2 N3];
45
46         DR=R;
47         DR0=R0;
48         Delp=zeros(1,8);
49         R2L=zeros(1,3);
50         RP=zeros(3,8);
51         I1=1;
52         for L=-1:2:1
53             for J=-1:2:1
54                 for K=-1:2:1
55
56                     RP(1,I1)=DR(1)+L*DR0(1);
57                     RP(2,I1)=DR(2)+J*DR0(2);
58                     RP(3,I1)=DR(3)+K*DR0(3);
59
60                     I1=I1+1;
61
62                 end
63             end
64         end
65
66         R2L(1)=2*RL(1)*NR(1);
67         R2L(2)=2*RL(2)*NR(2);
68         R2L(3)=2*RL(3)*NR(3);
69         theta=zeros(1,8);
70         phi=zeros(1,8);
71
72         for I=1:8
```

```
73
74     Rang=zeros(1,3);
75
76     delsq=0;
77     for J=1:3
78
79         R1=R2L(J)-RP(J,I);
80         delsq=delsq+R1^2;
81         %I think R1 is the distance vector that I need to
82         %compare with R so that I can get the correct arrival
83         %angles
84         Rang(J)=R1;
85
86     end
87
88     Delp(I)=sqrt(delsq);
89     theta(I)=acosd(Rang(3)/sqrt(Rang(1)^2+Rang(2)^2+Rang(3)^2))-90;
90     phi(I)=180/pi*atan2(Rang(2),Rang(1));
91
92
93     end
94
95
96
97     I0=0;
98     angind=0;
99     for L=0:1
100         for J=0:1
101             for K=0:1
102
103                 angind=angind+1;
```

```

104         I0=I0+1;
105         ID=round(Delp(I0));
106         FDMI=ID;
107         ID=ID+1;
108         if ID>NPTS
109         else
110             %Origin IR
111             GID=Beta(1,1)^abs(nx-L)*Beta(2,1)^abs(nx)...
112                 *Beta(1,2)^abs(ny-J)*Beta(2,2)^abs(ny)...
113                 *Beta(1,3)^abs(nz-K)*Beta(2,3)^abs(nz)/FDMI;
114
115     %      %Front, Left, Back, Right facing IRs
116     %      GID=(1-cosd(phi(angind))*cosd(-theta(angind)))...
117     %      *Beta(1,1)^abs(nx-L)*Beta(2,1)^abs(nx)...
118     %      *Beta(1,2)^abs(ny-J)*Beta(2,2)^abs(ny)...
119     %      *Beta(1,3)^abs(nz-K)*Beta(2,3)^abs(nz)/FDMI;
120
121     %      %Up and Down facing IRs
122     %      GID=(1+sind(-theta(angind)))*Beta(1,1)^abs(nx-L)...
123     %      *Beta(2,1)^abs(nx)...
124     %      *Beta(1,2)^abs(ny-J)*Beta(2,2)^abs(ny)...
125     %      *Beta(1,3)^abs(nz-K)*Beta(2,3)^abs(nz)/FDMI;
126
127         HT(ID)=HT(ID)+GID;
128         num(ID)=num(ID)+1; %how many arrivals are in one sample
129
130         IDtemp=round(Delp(I0));
131         thetanew(IDtemp,angnum(IDtemp))=-theta(angind);
132         phinew(IDtemp,angnum(IDtemp))=phi(angind);
133
134         amp(IDtemp,angnum(IDtemp))=GID;

```

```
135
136         if phinew(IDtemp,angnum(IDtemp))<-180
137
138             phinew(IDtemp,angnum(IDtemp))=...
139             phinew(IDtemp,angnum(IDtemp))+360;
140
141         end
142
143         angnum(IDtemp)=angnum(IDtemp)+1;
144
145     end
146
147
148     end
149 end
150 end
151
152 end
153 end
154 end
155
156 amp=amp/max(amp);
157
158 toc
159
160 f=1000;    %Low pass cutoff
161 w=8*atan(1)*f; T=1e-4; R1=exp(-w*T); R2=R1; B1=2*R1*cos(w*T);
162 B2=-R1*R1; A1=-(1+R2); A2=R2; Y1=0; Y2=0; Y0=0; for a=1:NPTS
163     X0=HT(a);
164     HT(a)=Y0+A1*Y1+A2*Y2;
165     Y2=Y1;
```

```

166     Y1=Y0;
167     Y0=B1*Y1+B2*Y2+X0;
168 end
169
170 HT=2000*HT; %Don't normalize these IRs!
171
172 figure subplot(2,1,1) plot(t,HT) title('Impulse Response')
173 subplot(2,1,2) plot(t,num) title('Number of Arrivals') xlabel('Time
174 (s)')
175
176 wavwrite(HT,FS,'F:\3DIRS\Simulations\LongNarrowCardioidOrigin.wav')

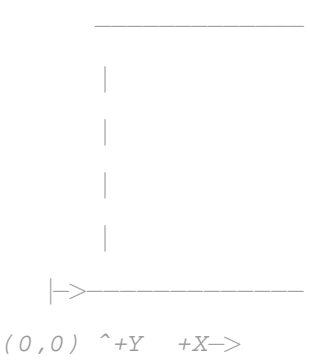
```

## A.5.2 Subcardioid IR Generation

```

1  %Subcardioid impulse responses based on the image source method
2  %
3  %
4  %
5  %
6  %
7  %
8  %
9  %
10 %
11
12 clear all; close all;
13
14 %Define constants
15 c=343; FS=192000; T=1/FS; NPTS=2^16; HT=zeros(1,NPTS); tmax=NPTS/FS;
16 dt=T; t=0:dt:tmax-dt; d=.0254*.25;

```





```
17 num=zeros(1,NPTS); %This is to count the number of arrivals in a single IR peak
18 ampind=ones(1,NPTS); angnum=ones(1,NPTS);
19 % amp=zeros(NPTS,1);
20
21 %Room Dimensions
22 lx=30; ly=8; lz=4.5; RL=[lx/(c*T) ly/(c*T) lz/(c*T)];
23
24 %Source position
25 x0=25; y0=1.5; z0=1.25; R0=[x0/(c*T) y0/(c*T) z0/(c*T)];
26
27 %Receiver position
28 x=5; y=7; z=2.5; R=[x/(c*T) y/(c*T) z/(c*T)];
29
30 %Reflection coefficients
31 betafont=.85; betaback=0.85; betaleft=0.85; betaright=0.85;
32 betaup=0.8; betadown=0.7; Beta=[[betaback betaleft
33 betadown];[betafont betaright betaup]];
34
35 N1=floor(NPTS/(2*RL(1))+1); N2=floor(NPTS/(2*RL(2))+1);
36 N3=floor(NPTS/(2*RL(3))+1);
37
38 tic
39
40 for nx=-N1:N1
41     for ny=-N2:N2
42         for nz=-N3:N3
43
44             NR=[nx ny nz]; % modified by Buye. before was NR=[N1 N2 N3];
45
46             DR=R;
47             DR0=R0;
```

```
48         Delp=zeros(1,8);
49         R2L=zeros(1,3);
50         RP=zeros(3,8);
51         I1=1;
52         for L=-1:2:1
53             for J=-1:2:1
54                 for K=-1:2:1
55
56                     RP(1,I1)=DR(1)+L*DR0(1);
57                     RP(2,I1)=DR(2)+J*DR0(2);
58                     RP(3,I1)=DR(3)+K*DR0(3);
59
60                     I1=I1+1;
61
62                 end
63             end
64         end
65
66         R2L(1)=2*RL(1)*NR(1);
67         R2L(2)=2*RL(2)*NR(2);
68         R2L(3)=2*RL(3)*NR(3);
69         theta=zeros(1,8);
70         phi=zeros(1,8);
71
72         for I=1:8
73
74             Rang=zeros(1,3);
75
76             delsq=0;
77             for J=1:3
78
```

```
79         R1=R2L(J)-RP(J,I);
80         delsq=delsq+R1^2;
81         %I think R1 is the distance vector that I need to
82         %compare with R so that I can get the correct arrival
83         %angles
84         Rang(J)=R1;
85
86     end
87
88     Delp(I)=sqrt(delsq);
89     theta(I)=acosd(Rang(3)/sqrt(Rang(1)^2+Rang(2)^2+Rang(3)^2))-90;
90     phi(I)=180/pi*atan2(Rang(2),Rang(1));
91
92
93 end
94
95
96
97 I0=0;
98 angind=0;
99 for L=0:1
100     for J=0:1
101         for K=0:1
102
103             angind=angind+1;
104             I0=I0+1;
105             ID=round(Delp(I0));
106             FDMI=ID;
107             ID=ID+1;
108             if ID>NPTS
109                 else
```

```

110      %Origin IR
111       $GID = \text{Beta}(1,1)^{\text{abs}(nx-L)} * \text{Beta}(2,1)^{\text{abs}(nx)} \dots$ 
112          *  $\text{Beta}(1,2)^{\text{abs}(ny-J)} * \text{Beta}(2,2)^{\text{abs}(ny)} \dots$ 
113          *  $\text{Beta}(1,3)^{\text{abs}(nz-K)} * \text{Beta}(2,3)^{\text{abs}(nz)} / \text{FDMI};$ 
114
115  %      %Front, Left, Back, Right facing IRs
116  %       $GID = (.7 - .3 * \cosd(\text{phi}(\text{angind})) * \cosd(-\text{theta}(\text{angind}))) \dots$ 
117  %          *  $\text{Beta}(1,1)^{\text{abs}(nx-L)} * \text{Beta}(2,1)^{\text{abs}(nx)} \dots$ 
118  %          *  $\text{Beta}(1,2)^{\text{abs}(ny-J)} * \text{Beta}(2,2)^{\text{abs}(ny)} \dots$ 
119  %          *  $\text{Beta}(1,3)^{\text{abs}(nz-K)} * \text{Beta}(2,3)^{\text{abs}(nz)} / \text{FDMI};$ 
120
121  %      %Up and Down facing IRs
122  %       $GID = (.7 + .3 * \text{sind}(-\text{theta}(\text{angind}))) * \text{Beta}(1,1)^{\text{abs}(nx-L)} \dots$ 
123  %          *  $\text{Beta}(2,1)^{\text{abs}(nx)} \dots$ 
124  %          *  $\text{Beta}(1,2)^{\text{abs}(ny-J)} * \text{Beta}(2,2)^{\text{abs}(ny)} \dots$ 
125  %          *  $\text{Beta}(1,3)^{\text{abs}(nz-K)} * \text{Beta}(2,3)^{\text{abs}(nz)} / \text{FDMI};$ 
126
127      HT(ID) = HT(ID) + GID;
128      num(ID) = num(ID) + 1; %how many arrivals are in one sample
129
130      IDtemp = round(Delp(I0));
131      thetanew(IDtemp, angnum(IDtemp)) = -theta(angind);
132      phinew(IDtemp, angnum(IDtemp)) = phi(angind);
133
134      amp(IDtemp, angnum(IDtemp)) = GID;
135
136      if phinew(IDtemp, angnum(IDtemp)) < -180
137
138          phinew(IDtemp, angnum(IDtemp)) = ...
139          phinew(IDtemp, angnum(IDtemp)) + 360;
140

```

```
141         end
142
143         angnum(IDtemp)=angnum(IDtemp)+1;
144
145     end
146
147
148     end
149 end
150 end
151
152 end
153 end
154 end
155
156 amp=amp/max(amp);
157
158 toc
159
160 f=1000;    %Low pass cutoff
161 w=8*atan(1)*f; T=1e-4; R1=exp(-w*T); R2=R1; B1=2*R1*cos(w*T);
162 B2=-R1*R1; A1=-(1+R2); A2=R2; Y1=0; Y2=0; Y0=0; for a=1:NPTS
163     X0=HT(a);
164     HT(a)=Y0+A1*Y1+A2*Y2;
165     Y2=Y1;
166     Y1=Y0;
167     Y0=B1*Y1+B2*Y2+X0;
168 end
169
170 HT=2000*HT; %Don't normalize these IRs!
171
```

```

172 figure subplot(2,1,1) plot(t,HT) title('Impulse Response')
173 subplot(2,1,2) plot(t,num) title('Number of Arrivals') xlabel('Time
174 (s)')
175
176 wavwrite(HT,FS,'F:\3DIRS\Simulations\LongNarrowCardioidOrigin.wav')

```

### A.5.3 Hypercardioid IR Generation

```

1 %Hypercardioid impulse responses based on the image source method
2 %
3 %
4 %
5 %
6 %
7 %
8 %
9 %
10 %
11
12 clear all; close all;
13
14 %Define constants
15 c=343; FS=192000; T=1/FS; NPTS=2^16; HT=zeros(1,NPTS); tmax=NPTS/FS;
16 dt=T; t=0:dt:tmax-dt; d=.0254*.25;
17 num=zeros(1,NPTS); %This is to count the number of arrivals in a single IR peak
18 ampind=ones(1,NPTS); angnum=ones(1,NPTS);
19 % amp=zeros(NPTS,1);
20
21 %Room Dimensions
22 lx=30; ly=8; lz=4.5; RL=[lx/(c*T) ly/(c*T) lz/(c*T)];

```

```
23
24 %Source position
25 x0=25; y0=1.5; z0=1.25; R0=[x0/(c*T) y0/(c*T) z0/(c*T)];
26
27 %Receiver position
28 x=5; y=7; z=2.5; R=[x/(c*T) y/(c*T) z/(c*T)];
29
30 %Reflection coefficients
31 betafront=.85; betaback=0.85; betaleft=0.85; betaright=0.85;
32 betaup=0.8; betadown=0.7; Beta=[[betaback betaleft
33 betadown];[betafront betaright betaup]];
34
35 N1=floor(NPTS/(2*RL(1))+1); N2=floor(NPTS/(2*RL(2))+1);
36 N3=floor(NPTS/(2*RL(3))+1);
37
38 tic
39
40 for nx=-N1:N1
41     for ny=-N2:N2
42         for nz=-N3:N3
43
44             NR=[nx ny nz]; % modified by Buye. before was NR=[N1 N2 N3];
45
46             DR=R;
47             DR0=R0;
48             Delp=zeros(1,8);
49             R2L=zeros(1,3);
50             RP=zeros(3,8);
51             I1=1;
52             for L=-1:2:1
53                 for J=-1:2:1
```

```
54         for K=-1:2:1
55
56             RP(1,I1)=DR(1)+L*DR0(1);
57             RP(2,I1)=DR(2)+J*DR0(2);
58             RP(3,I1)=DR(3)+K*DR0(3);
59
60             I1=I1+1;
61
62         end
63     end
64 end
65
66     R2L(1)=2*RL(1)*NR(1);
67     R2L(2)=2*RL(2)*NR(2);
68     R2L(3)=2*RL(3)*NR(3);
69     theta=zeros(1,8);
70     phi=zeros(1,8);
71
72     for I=1:8
73
74         Rang=zeros(1,3);
75
76         delsq=0;
77         for J=1:3
78
79             R1=R2L(J)-RP(J,I);
80             delsq=delsq+R1^2;
81             %I think R1 is the distance vector that I need to
82             %compare with R so that I can get the correct arrival
83             %angles
84             Rang(J)=R1;
```



```

85
86         end
87
88         Delp(I)=sqrt(delsq);
89         theta(I)=acosd(Rang(3)/sqrt(Rang(1)^2+Rang(2)^2+Rang(3)^2))-90;
90         phi(I)=180/pi*atan2(Rang(2),Rang(1));
91
92     end
93
94     I0=0;
95     angind=0;
96     for L=0:1
97         for J=0:1
98             for K=0:1
99
100                 angind=angind+1;
101                 I0=I0+1;
102                 ID=round(Delp(I0));
103                 FDMI=ID;
104                 ID=ID+1;
105                 if ID>NPTS
106                     else
107                         %Origin IR
108                         GID=Beta(1,1)^abs(nx-L)*Beta(2,1)^abs(nx)...
109                             *Beta(1,2)^abs(ny-J)*Beta(2,2)^abs(ny)...
110                             *Beta(1,3)^abs(nz-K)*Beta(2,3)^abs(nz)/FDMI;
111
112                 % Front, Left, Back, Right facing IRs
113                 GID=(.5-1.5*cosd(phi(angind))*cosd(-theta(angind)))...
114                     *Beta(1,1)^abs(nx-L)*Beta(2,1)^abs(nx)...
115                     *Beta(1,2)^abs(ny-J)*Beta(2,2)^abs(ny)...

```

```

116 %                               *Beta(1,3)^abs(nz-K)*Beta(2,3)^abs(nz)/FDMI;
117
118 %                               %Up and Down facing IRS
119 %                               GID=(.5+1.5*sind(-theta(angind)))*Beta(1,1)^abs(nx-L)...
120 %                               *Beta(2,1)^abs(nx)...
121 %                               *Beta(1,2)^abs(ny-J)*Beta(2,2)^abs(ny)...
122 %                               *Beta(1,3)^abs(nz-K)*Beta(2,3)^abs(nz)/FDMI;
123
124                               HT(ID)=HT(ID)+GID;
125                               num(ID)=num(ID)+1; %how many arrivals are in one sample
126
127                               IDtemp=round(Delp(I0));
128                               thetanew(IDtemp,angnum(IDtemp))=-theta(angind);
129                               phinew(IDtemp,angnum(IDtemp))=phi(angind);
130
131                               amp(IDtemp,angnum(IDtemp))=GID;
132
133                               if phinew(IDtemp,angnum(IDtemp))<-180
134
135                                   phinew(IDtemp,angnum(IDtemp))=...
136                                   phinew(IDtemp,angnum(IDtemp))+360;
137
138                               end
139
140                               angnum(IDtemp)=angnum(IDtemp)+1;
141
142                               end
143
144
145                               end
146                               end

```

```
147         end
148
149     end
150 end
151 end
152
153 amp=amp/max(amp);
154
155 toc
156
157 f=1000;    %Low pass cutoff
158 w=8*atan(1)*f; T=1e-4; R1=exp(-w*T); R2=R1; B1=2*R1*cos(w*T);
159 B2=-R1*R1; A1=-(1+R2); A2=R2; Y1=0; Y2=0; Y0=0; for a=1:NPTS
160     X0=HT(a);
161     HT(a)=Y0+A1*Y1+A2*Y2;
162     Y2=Y1;
163     Y1=Y0;
164     Y0=B1*Y1+B2*Y2+X0;
165 end
166
167 HT=2000*HT; %Don't normalize these IRs!
168
169 figure subplot(2,1,1) plot(t,HT) title('Impulse Response')
170 subplot(2,1,2) plot(t,num) title('Number of Arrivals') xlabel('Time
171 (s)')
172
173 wavwrite(HT,FS,'F:\3DIRS\Simulations\LongNarrowCardioidOrigin.wav')
```

#### A.5.4 Polar ETC Matlab Program

```
1 %Program to create the Polar SIR modification of the PolarETC Stripped
2 %down version uses seven impulse responses located on the Cartesian axes
3 %
4 %NOTE!!! The Polar ETC (and therefore the Polar SIR) does not account for
5 %multiple simultaneously arriving reflections. It can therefore give
6 %erroneous image source locations if two or more reflections arrive at the
7 %same instant in time
8
9 clear all; close all; clc;
10
11 c=343;
12
13 [Front,FS]=wavread('F:\3DIRS\Simulations\LongNarrowCardioidLeft.wav');
14 Back=wavread('F:\3DIRS\Simulations\LongNarrowCardioidRight.wav');
15 Left=wavread('F:\3DIRS\Simulations\LongNarrowCardioidBack.wav');
16 Right=wavread('F:\3DIRS\Simulations\LongNarrowCardioidFront.wav');
17 Up=wavread('F:\3DIRS\Simulations\LongNarrowCardioidUp.wav');
18 Down=wavread('F:\3DIRS\Simulations\LongNarrowCardioidDown.wav');
19 W=wavread('F:\3DIRS\Simulations\LongNarrowCardioidOrigin.wav');
20
21 t1=0; t2=length(W)/FS; dt=1/FS; t=t1:dt:t2-dt;
22
23 %Multiply the IR by it's complex conjugate to obtain the SIR
24
25 Fsq=(Front.*conj(Front)); Bsq=(Back.*conj(Back));
26 Lsq=(Left.*conj(Left)); Rsq=(Right.*conj(Right));
27 Usq=(Up.*conj(Up)); Dsq=(Down.*conj(Down)); Wsq=(W.*conj(W));
28
29 % break
30
31 clear FrIR; clear BaIR; clear LeIR; clear RiIR; clear UIR; clear
```

```
32 DoIR; clear Fr; clear Ba; clear Le; clear Ri; clear U; clear Do;
33 clear Mic1; clear Mic2; clear Mic3; clear Mic4; clear Mic5; clear
34 Mic6;
35
36 %Define the "energy density" E from the D'Antonio paper to be able to use
37 %the direction cosines u,v and w and therefore extract the directional
38 %information
39
40 E=1/2*sqrt((Fsq-Bsq).^2+(Rsq-Lsq).^2+(Usq-Dsq).^2);
41
42 disp('Select a range to examine and threshold')
43
44 figure plot(t,W) title('Pressure Impulse Response') xlabel('Time
45 (s)') ylabel('Magnitude') ylim([1.3*min(W) 1.3*max(W)]) xlim([t1
46 t2])
47 %      break
48 [time,amp]=ginput(2); tpremod=time(1):1/FS:time(2);
49
50 %Make the two vectors the same length
51 Originpremod=W(round(time(1)*FS):round(time(1)*FS)+length(tpremod)-1);
52 close all;
53 clear amp; %clear time;
54
55 % figure(1)
56 % plot(tpremod,Originpremod)
57 % title('Zoomed Plot of the Origin IR')
58 % xlabel('Time (s)')
59 % ylabel('Magnitude')
60 % ylim([1.1*min(Originpremod) 1.1*max(Originpremod)])
61 clc;
62 %
```

---

```

63 % disp('Press any key to continue and select a threshold')
64 % pause
65 %
66 % thresh=input('Set threshold value - ');
67 thresh=0.17; thresh=thresh*ones(1,length(W));
68
69 figure
70 plot(tpremod,Originpremod,'b',tpremod,thresh(1:length(tpremod)),'r—')
71 title('Composite IR with Threshold'); xlim([tpremod(1)
72 tpremod(length(tpremod))]); ylim([min(Originpremod)
73 max(Originpremod)]); xlabel('Time (s)'); ylabel('Magnitude');
74
75 % disp('Press any key to continue')
76 pause
77
78 %Find all peaks that are higher than the thresholds. This part may be
79 %obselete due to the same section that allows for user defineable time
80 %limits for narrowing the time window in which the user is interested
81
82 u=zeros(length(W),1); v=zeros(length(W),1); w=zeros(length(W),1);
83 theta=370*ones(length(W),1); phi=370*ones(length(W),1);
84 timeend=zeros(length(W),1);
85
86 for l=round(time(1)*FS):round(time(1)*FS)+length(tpremod)-1
87     if W(l)>thresh(l)
88
89         %Find the direction cosines using the squared impulse responses
90         %since energy is proportional to pressure squared (thus the squared
91         %impulse response)
92
93         u(l)=(Rsq(l)-Lsq(l))/(2*E(l));

```

```
94     v(l)=(Fsq(l)-Bsq(l))/(2*E(l));
95     w(l)=(Usq(l)-Dsq(l))/(2*E(l));
96
97     %     u(l)=(Right(l)-Left(l))/(2*E(l));
98     %     v(l)=(Front(l)-Back(l))/(2*E(l));
99     %     w(l)=(Up(l)-Down(l))/(2*E(l));
100
101
102     %Find the theta, phi and time values for the reflections.  Display
103     %them so that one can locate the peak on the graph using time and
104     %then find the offending surface using theta and phi where theta is
105     %the elevation from -90 to +90 degrees and phi ranges from zero
    to
106     %360 degrees, as is the physics convention
107
108     theta(l)=180/pi*atan2(w(l),sqrt(u(l)^2+v(l)^2));
109     phi(l)=- (180/pi*atan2(v(l),-u(l))+180);
110
111     if phi(l)<-180
112         phi(l)=phi(l)+360;
113     end
114
115     if phi(l)>180
116         phi(l)=phi(l)-360;
117     end
118
119     timeend(l)=1/FS;
120
121     sprintf('t(s) = %2.5g',timeend(l))
122     sprintf('theta(deg) = %2.5g', theta(l))
123     sprintf('phi(deg) = %2.5g',phi(l))
```

```
124
125     end
126 end
127
128 %Remove all the entries in phi and theta that correspond to samples that
129 %are not above the threshold
130 indextheta=find(theta==370); indexphi=find(phi==370);
131 theta(indextheta)=[]; phi(indexphi)=[];
132
133 close all;
134
135 scatter(phi,theta) title('Image Source Locations'); xlabel('\phi
136 (degrees)'); ylabel('\theta (degrees)'); xlim([-180 180]) ylim([-90
137 90])
```





# Appendix B

## B.1 Spherical Coordinate Rotations

To demonstrate that the weighting equations in Sec. 6.2.1 are correct for a cardioid solid with its main lobe facing along the Cartesian axes, let us consider the general case of rotating functions in three dimensions. This is a common practice in mechanics, but it is usually performed on a Cartesian function that one wishes to rotate about a fixed Cartesian axis. In practice, the coordinate system is rotated with respect to the function, but rotating the function while keeping the axes fixed yields the same result if the rotation angle is opposite in sign. In this work, the primary interest lies in rotating a function described in spherical coordinates as opposed to one defined in Cartesian coordinates.

To begin, let our function in spherical coordinates be defined as  $r(\theta, \phi)$ . The function can be described in Cartesian coordinates as

$$\begin{bmatrix} x(\theta, \phi) \\ y(\theta, \phi) \\ z(\theta, \phi) \end{bmatrix} = \begin{bmatrix} r(\theta, \phi)\cos(\theta)\cos(\phi) \\ r(\theta, \phi)\cos(\theta)\sin(\phi) \\ r(\theta, \phi)\sin(\theta) \end{bmatrix} \quad (\text{B.1})$$

where  $\theta$  is defined as the elevation angle relative to the horizontal ( $x$ - $y$ ) plane. The choice of rotation matrix depends on the axis to be fixed during a given portion of the

rotation process. An axis is described as fixed when the function is rotated in such a way that the components along the fixed axis remain unchanged. The Cartesian rotation matrices are [43]

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & \sin(\alpha) \\ 0 & -\sin(\alpha) & \cos(\alpha) \end{bmatrix} \quad (\text{B.2})$$

if the  $x$  axis is fixed,

$$\begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{bmatrix} \quad (\text{B.3})$$

if the  $y$  axis is fixed, and

$$\begin{bmatrix} \cos(\gamma) & \sin(\gamma) & 0 \\ -\sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (\text{B.4})$$

if the  $z$  axis is fixed. The angles  $\alpha$ ,  $\beta$ , and  $\gamma$  are arbitrary angles through which the function is rotated with the corresponding axis fixed.

By operating on the Cartesian definition of the spherical coordinate function with each of the rotation matrices defined in Eqs. (B.2) through (B.4), one obtains the following results:

$$\begin{aligned}
\begin{bmatrix} x(\theta', \phi') \\ y(\theta', \phi') \\ z(\theta', \phi') \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & \sin(\alpha) \\ 0 & -\sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} r \cos(\theta) \cos(\phi) \\ r \cos(\theta) \sin(\phi) \\ r \sin(\theta) \end{bmatrix} \\
&= \begin{bmatrix} r(\theta', \phi') \cos(\theta') \cos(\phi') \\ r(\theta', \phi') \cos(\theta') \sin(\phi') \cos(\alpha) + r' \sin(\theta') \sin(\alpha) \\ -r(\theta', \phi') \cos(\theta') \sin(\phi') \sin(\alpha) + r' \sin(\theta') \cos(\alpha) \end{bmatrix}, \quad (\text{B.5})
\end{aligned}$$

with rotation with the  $x$ -axis fixed,

$$\begin{aligned}
\begin{bmatrix} x(\theta', \phi') \\ y(\theta', \phi') \\ z(\theta', \phi') \end{bmatrix} &= \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{bmatrix} \begin{bmatrix} r \cos(\theta) \cos(\phi) \\ r \cos(\theta) \sin(\phi) \\ r \sin(\theta) \end{bmatrix} \\
&= \begin{bmatrix} r(\theta', \phi') \cos(\theta') \cos(\phi') \cos(\beta) + r' \sin(\theta') \sin(\beta) \\ r(\theta', \phi') \cos(\theta') \sin(\phi') \\ -r(\theta', \phi') \cos(\theta') \cos(\phi') \sin(\beta) + r' \sin(\theta') \cos(\beta) \end{bmatrix}, \quad (\text{B.6})
\end{aligned}$$

with the  $y$ -axis fixed, and

$$\begin{aligned}
\begin{bmatrix} x(\theta', \phi') \\ y(\theta', \phi') \\ z(\theta', \phi') \end{bmatrix} &= \begin{bmatrix} \cos(\gamma) & \sin(\gamma) & 0 \\ -\sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r \cos(\theta) \cos(\phi) \\ r \cos(\theta) \sin(\phi) \\ r \sin(\theta) \end{bmatrix} \\
&= \begin{bmatrix} r(\theta', \phi') \cos(\theta') \cos(\phi') \cos(\gamma) + r' \cos(\theta') \sin(\phi') \sin(\gamma) \\ -r(\theta', \phi') \cos(\theta') \cos(\phi') \sin(\gamma) + r' \cos(\theta') \sin(\phi') \cos(\gamma) \\ r(\theta', \phi') \sin(\theta') \end{bmatrix}, \quad (\text{B.7})
\end{aligned}$$

with the  $z$ -axis fixed. The  $'$  indicates the new coordinate system. In other words, after the rotation,  $r(\theta', \phi')$  is now defined in terms of the new coordinate system  $\phi'$

and  $\theta'$ . Depending on what one wants to achieve and what is known, the resulting equations can be solved for  $r(\theta', \phi')$  in terms of the original coordinate system (if the relationship between the old and new coordinate system is known), or they can be solved to find the relationship between the new coordinate system and the old [if  $r(\theta', \phi')$  is known].

## B.2 Rotation of a Cardioid Solid

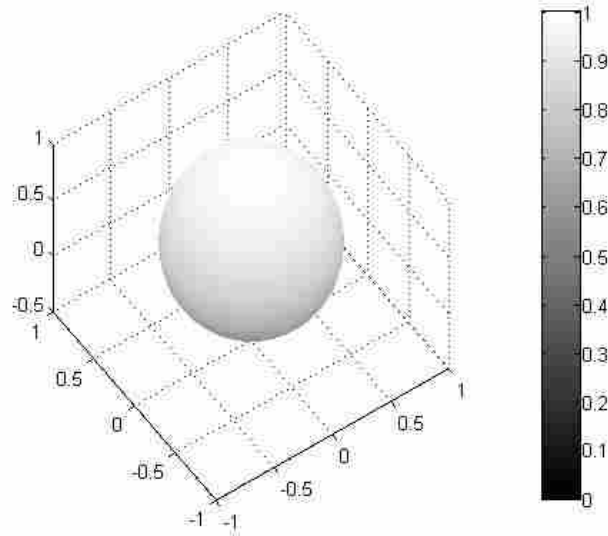
Let us now consider the case of a cardioid solid aligned along the  $+z$  axis. In two dimensions, the cardioid is described as

$$r(\theta) = 1 + \sin(\theta), \quad (\text{B.8})$$

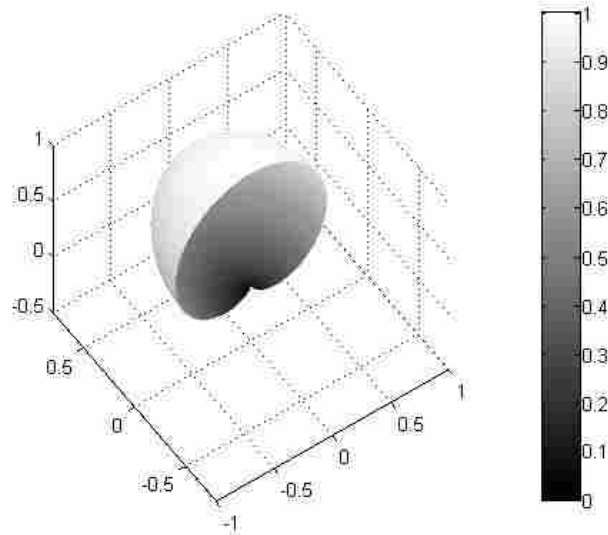
where  $\theta$  is the angle from the  $x$ - $y$  plane. If the two-dimensional cardioid is now rotated about the  $z$ -axis to create a solid, there is still no dependence on  $\phi$  (see Fig. ??), and the equation representing the solid in three dimensions remains exactly the same as the two-dimensional representation. If the solid is then rotated so that it is oriented along the  $+x$  axis, the  $y$  axis must be fixed during the rotation. The rotation angle  $\beta$  is  $90^\circ$ , so the resulting vector becomes

$$\begin{bmatrix} x(\theta', \phi') \\ y(\theta', \phi') \\ z(\theta', \phi') \end{bmatrix} = \begin{bmatrix} r(\theta', \phi') \sin(\theta') \\ r(\theta', \phi') \cos(\theta') \sin(\phi') \\ -r(\theta', \phi') \cos(\theta') \cos(\phi') \end{bmatrix} = \begin{bmatrix} [1 + \sin(\theta')] \sin(\theta') \\ [1 + \sin(\theta')] \cos(\theta') \sin(\phi') \\ -[1 + \sin(\theta')] \cos(\theta') \cos(\phi') \end{bmatrix} \quad (\text{B.9})$$

We desire to find  $r(\theta', \phi')$  as a function of the original, unprimed coordinate system which requires determining the new coordinates in terms of the old coordinates. Let us begin by considering the original Cartesian system. A plot of the axes and angular



(a)



(b)

**Figure B.1** A cardioid solid facing in the  $+z$  direction with views of (a) the entire solid and (b) the solid cut in half to show its cardioid profile.

definitions are shown in Fig. B.2(a). A plot of the axes after a  $90^\circ$  rotation about a fixed  $y$  axis with its angular definitions are shown in Fig. B.2(b).

It is also useful to define the angles in terms of the Cartesian components. These definitions can be found by inspection and are as follows:

$$\sin(\theta) = \frac{z}{\sqrt{x^2 + y^2 + z^2}} \quad (\text{B.10})$$

$$\cos(\theta) = \frac{\sqrt{x^2 + y^2}}{\sqrt{x^2 + y^2 + z^2}} \quad (\text{B.11})$$

$$\sin(\phi) = \frac{y}{\sqrt{x^2 + y^2}} \quad (\text{B.12})$$

$$\cos(\phi) = \frac{x}{\sqrt{x^2 + y^2}} \quad (\text{B.13})$$

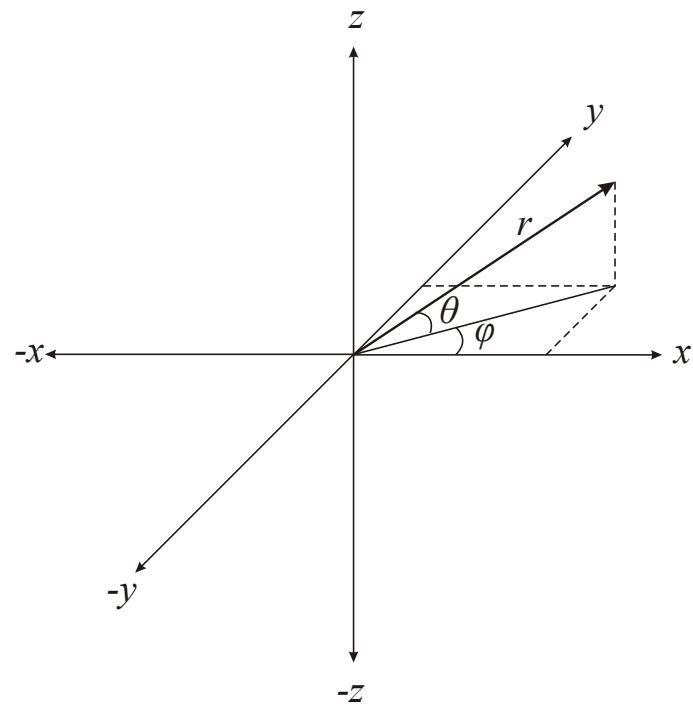
for the original coordinate system and

$$\sin(\theta') = \frac{z'}{\sqrt{x'^2 + y'^2 + z'^2}} \quad (\text{B.14})$$

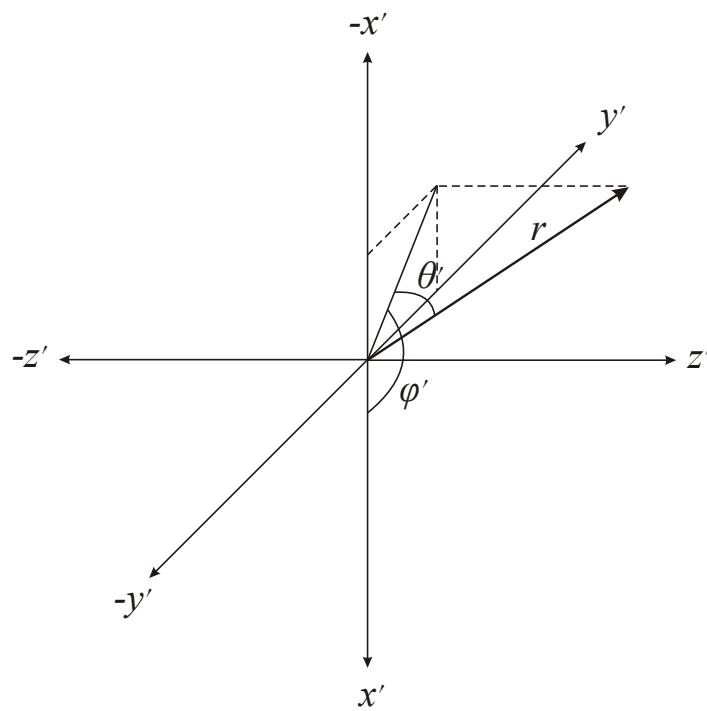
$$\cos(\theta') = \frac{\sqrt{x'^2 + y'^2}}{\sqrt{x'^2 + y'^2 + z'^2}} \quad (\text{B.15})$$

$$\sin(\phi') = \frac{y'}{\sqrt{x'^2 + y'^2}} \quad (\text{B.16})$$

$$\cos(\phi') = \frac{x'}{\sqrt{x'^2 + y'^2}} \quad (\text{B.17})$$



(a)



(b)

**Figure B.2** Cartesian coordinate system and angular definitions for (a) the original coordinate system and (b) the rotate coordinate system.



for the rotated coordinate system. As seen in Fig. B.2, the relationships between the Cartesian components in the two coordinate systems can be expressed as

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} -z \\ y \\ x \end{bmatrix} \quad (\text{B.18})$$

Substitution of the values in Eq. (B.18) into Eqs. (B.10) through (B.17) yields the following relations:

$$\sin(\theta') = \cos(\theta)\cos(\phi) \quad (\text{B.19})$$

$$\cos(\theta')\sin(\phi') = \cos(\theta)\sin(\phi) \quad (\text{B.20})$$

$$\cos(\theta')\cos(\phi') = -\sin(\theta) \quad (\text{B.21})$$

If these values are then substituted into Eq. B.9, obtain the values

$$\begin{bmatrix} x(\theta, \phi) \\ y(\theta, \phi) \\ z(\theta, \phi) \end{bmatrix} = \begin{bmatrix} r(\theta, \phi)\cos(\theta)\cos(\phi) \\ r(\theta, \phi)\cos(\theta)\sin(\phi)\cos(\theta) \\ r(\theta, \phi)\sin(\theta) \end{bmatrix} = \begin{bmatrix} [1 + \cos(\phi)\cos(\theta)]\cos(\theta)\cos(\phi) \\ [1 + \cos(\phi)\cos(\theta)]\cos(\theta)\sin(\phi) \\ [1 + \cos(\phi)\cos(\theta)]\sin(\theta) \end{bmatrix} \quad (\text{B.22})$$

From this equation it is obvious that the equation for  $r(\theta, \phi)$  becomes

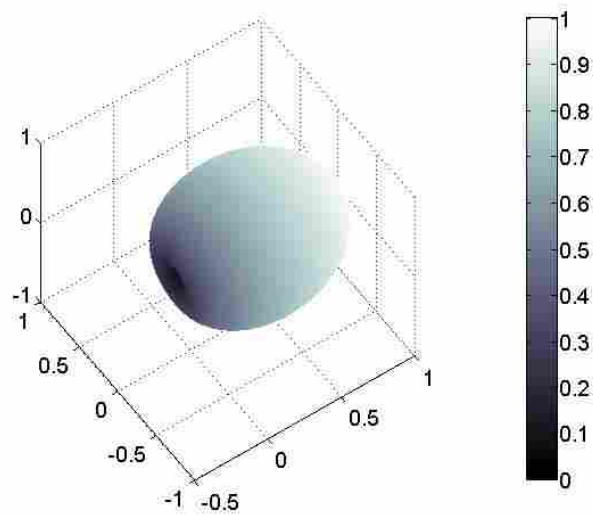
$$r(\theta, \phi) = 1 + \cos(\phi)\cos(\theta). \quad (\text{B.23})$$

A similar approach can be followed for rotating the solid from facing in the  $+z$  direc-

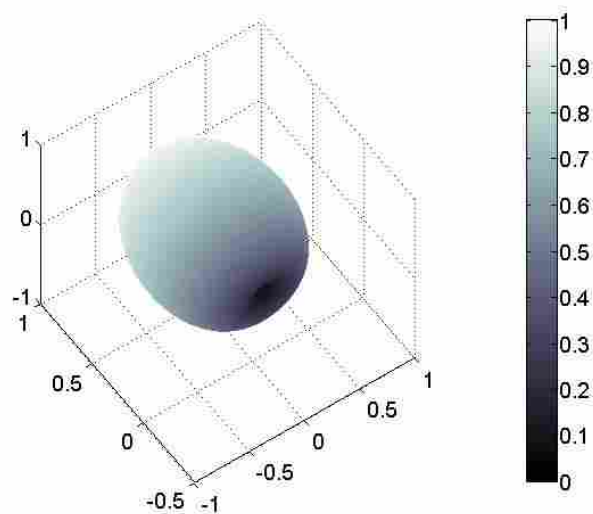
tion to the  $+y$  direction resulting in an equation for the solid of

$$r(\theta, \phi) = 1 + \sin(\phi)\cos(\theta). \quad (\text{B.24})$$

These and similar results show that the equations given in Eqs. (2.38) through (2.43) are valid. Three-dimensional plots of these two functions are shown in Fig. B.3.



(a)



(b)

**Figure B.3** Cardioid solids facing in the (a)  $+x$  direction and (b)  $+y$  direction.