

Sound From Ultrasound: The Parametric Array as an Audible Sound Source

F. Joseph Pompei

B.S. Electrical Engineering, Rensselaer Polytechnic Institute, 1995
M.M. Music Technology, Northwestern University, 1997

Submitted to the program in Media Arts & Sciences, School of Architecture and Planning
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the Massachusetts Institute of Technology
June 2002

©Massachusetts Institute of Technology, 2002
All Rights Reserved

11

Author

JF Pompei
Program in Media Arts & Sciences
May 22, 2002

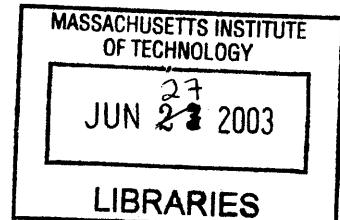
Certified by

M
Professor Barry Vercoe
Program in Media Arts and Sciences
Thesis Supervisor

KA

Accepted by

CD
Andrew B. Lippman
Chairperson
Departmental Committee on Graduate Students



Sound From Ultrasound: The Parametric Array as an Audible Sound Source

F. Joseph Pompei

Submitted to the program in Media Arts & Sciences, School of Architecture and Planning
on May 22, 2002 in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the Massachusetts Institute of Technology

Abstract

A parametric array exploits the nonlinearity of the propagation medium to emit or detect acoustic waves in a spatially versatile manner, permitting concise, narrow directivity patterns otherwise possible only with physically very large transducer geometries. This thesis explores the use of the parametric array as an audible sound source, permitting audible sound to be generated with very high directivity compared to traditional loudspeakers of comparable size.

The thesis begins with a review of basic underlying mathematics and relevant approximate solutions of nonlinear acoustic systems. Then, these solutions are used to construct suitable methods of ultrasonic synthesis for low-distortion audio reproduction. Geometrical modelling methods for predicting the acoustic distribution are presented and evaluated, and practical applications are explored experimentally. Issues of risk associated with ultrasonic exposure are presented, and the feasibility of a phased-array system for beam control is explored.

Thesis Supervisor: Barry L. Vercoe, D.M.A.
Professor of Media Arts and Sciences

This work was primarily supported by the Digital Life consortium of the MIT Media Laboratory, with additional support in the form of a fellowship by British Telecom.

Doctoral Dissertation Committee

Professor Barry Vercoe

Barry Vercoe is Professor of Music and Professor of Media Arts and Sciences at MIT, and Associate Academic Head of the Program in Media Arts & Sciences . He was born and educated in New Zealand in music and in mathematics, then completed a doctorate in Music Composition at the University of Michigan. In 1968 at Princeton University he did pioneering working in the field of Digital Audio Processing, then taught briefly at Yale before joining the MIT faculty in 1971. In 1973 he established the MIT computer facility for Experimental Music – an event now commemorated on a plaque in the Kendall Square subway station.

During the 1970's and early 1980's he pioneered the composition of works combining computers and live instruments. Then on a Guggenheim Fellowship in Paris in 1983 he developed a Synthetic Performer – a computer that could listen to other performers and play its own part in musical sync, even learning from rehearsals. In 1992 he won the Computer World / Smithsonian Award in Media Arts and Entertainment.

Professor Vercoe was a founding member of the MIT Media Laboratory in 1984, where he has pursued research in Music Cognition and Machine Understanding. His several Music Synthesis languages are used around the world, and a variant of his Csound and NetSound languages has recently been adopted as the core of MPEG-4 audio – an international standard that enables efficient transmission of audio over the Internet. At the Media Lab he currently directs research in Machine Listening and Digital Audio Synthesis, and is Associate Academic Head of its graduate program in Media Arts and Sciences.

<http://sound.media.mit.edu/%7Ebv/>

Professor Peter Cochrane

Peter Cochrane joined BT Laboratories in 1973 and has worked on a wide range of technologies and systems. In 1993 he was appointed as the Head of Research, and in 1999 he became BT's Chief Technologist. In November 2000 Peter retired from BT to join his own startup company - ConceptLabs - which he founded with a group out of Apple Computers in 1998 at Campbell CA, in Silicon Valley. A graduate of Trent Polytechnic (now Nottingham Trent University) and Essex University, he has been a visiting professor to UCL, Essex, Southampton, and Kent. He was the Collier Chair for The Public Understanding of Science & Technology at The University of Bristol and is a Member of the New York Academy of Sciences. He has published and lectured widely on technology and the implications of IT.

He led a team that received the Queen's Award for Innovation & Export in 1990; the Martlesham Medal for contributions to fibre optic technology in 1994; the IEE Electronics Division Premium in 1986; Computing and control Premium in 1994 and the IERE Bene-

factors Prize in 1994; the James Clerk Maxwell Memorial Medal in 1995; IBTE Best Paper Prize and Honorary Doctorates from Essex, Robert Gordon, Stafford, and Nottingham Trent Universities and was awarded an OBE in 1999 for his contribution to international communications. Peter was awarded an IEEE Millennium Medal in 2000 and The City & Guilds Prince Philip Medal in 2001.

<http://www.cochrane.org.uk>

Carl J. Rosenberg



Carl Rosenberg, a Fellow of the Acoustical Society of America and since 1970 a principal acoustical consultant with Acentech Inc., specializes in architectural acoustics, specifically performance spaces, sound isolation, and environmental noise.

Among the many projects Rosenberg has worked on are the ongoing residential soundproofing programs at Logan Airport in Boston and at San Jose International Airport; recently completed projects include the Performing Arts Center at Greenwich Academy, Center for Dramatic Arts at the University of North Carolina at Chapel Hill, Rogers Performing Arts Center at Merrimack College in North Andover, Massachusetts, and the Genko Uchida Building and performance hall for the Berklee College of Music; ongoing work includes the Sprint World Headquarters, Princeton Campus Center, University of the Arts in Philadelphia, and a new Humanities Building at Rice. He has recently edited the current acoustics documentation for the Supplemental Edition of Architectural Graphic Standards. A member of the American Institute of Architects and a registered architect in Massachusetts, Rosenberg came to MIT as a visiting lecturer in 1984. He earned a BA from Princeton University in 1965 and an MArch from MIT in 1971.

Mr. Rosenberg currently teaches a course in building technology and architectural acoustics at the MIT School of Architecture, as well as at Princeton.

<http://www.acentech.com>

<http://architecture.mit.edu/people/profiles/prrosenb.html>

Author Biography

Beginning his career in acoustics at 16 while in high school, starting as the first high school co-op and becoming the youngest engineer at Bose Corporation, Frank Joseph Pompei continued working part-time and summers for Bose while earning a degree in Electrical Engineering with an Electronic Arts Minor from Rensselaer Polytechnic Institute. Recognizing the importance and underutilization of spatialized sound, he decided to pursue research in psychoacoustics and application of auditory localization at Northwestern University, earning a Master's degree. Acutely aware of the limitations of traditional loudspeakers, he had the idea of using ultrasound as an acoustic projector, and successfully developed such a device while a student at the MIT Media Lab, earning his Ph.D.



Photo by Webb Chappell

Acknowledgements

First and foremost, I would like to thank Professors Barry Vercoe and Nicholas Negroponte for creating the best environment possible to cultivate this work - and, as importantly, for having enough interest in my unorthodox ideas to take a chance with them, even against the advice of those who did not believe this project possible. Also thanks to the many Media Lab Sponsors who have made this work possible.

I would also like to extend special thanks to my parents: my mother for a lifetime of support and constant encouragement to pursue everything that interests me, and my father, for his years of guidance and unwavering confidence in my abilities, as well as showing, through example, how rewarding the pursuit of innovation can truly be.

Also instrumental in this work were my friends and colleagues at the Media Lab, in particular J. C. Olsson, for his valuable technical contributions, in-lab entertainment, and supply of cold beverages and loud music. Thanks also goes to the present and former members of the Music, Mind & Machine research group, who not only offered solid support and advice over the years, but also put up with the endless flow of hardware, apparatus, and other clutter created in the course of this project.

Finally, I thank all of my colleagues and friends over the years who encouraged me along this path, and also providing the healthy distractions from my otherwise obsessed work-life. You know who you are.

Contents

1	Introduction	9
1.1	Contemporary methods of Sound Control	10
1.1.1	Loudspeakers and arrays	11
1.1.2	Speaker domes	11
1.1.3	Binaural audio	11
1.2	Origins of the Parametric Array	12
2	Mathematical Basis	14
2.1	Equation of State	14
2.2	Equation of Continuity	16
2.3	Euler's Equation	16
2.4	Combining the Equations	17
3	Low-Distortion Audio Reproduction	19
3.1	Basic Quasilinear Solution	19
3.2	Processing to reduce distortion	21
3.2.1	Effects of Bandwidth	21
3.3	Experimental Results	26
3.3.1	Ultrasonic Bandwidth	26
3.3.2	Distortion Versus Frequency	27
3.3.3	Distortion Versus Level	29
3.4	Modulation depth control	30
3.5	Ultrasonic Absorption	35
3.5.1	Performance effects	37
4	Geometric Characteristics	39
4.1	Ultrasonic Field	39
4.1.1	Far-field	40
4.1.2	Near field	41
4.2	Audible Field	45
4.2.1	Far field	45
4.2.2	Near field	48

CONTENTS	8
4.3 Experimental Results	52
4.3.1 Ultrasonic Field	52
4.3.2 Audible Field	52
4.3.3 Distortion Fields	53
5 Designing Applications	61
5.1 Dome-reflector loudspeakers	61
5.2 Projecting sound	66
5.2.1 Scattering	66
5.2.2 Scattering Experiment	66
5.3 Subwoofer augmentation	69
5.4 Precedence Effect	70
5.4.1 Psychoacoustics	70
5.4.2 Application to Parametric Loudspeakers	70
5.4.3 In-Place Performers	73
6 Beam Control	75
6.1 Basic Theory	75
6.1.1 Huygen's Principle	75
6.1.2 Fourier Analysis	76
6.1.3 Steering	77
6.1.4 Discrete Array	77
6.1.5 Dense Arrays	79
6.1.6 Steering Performance	80
6.2 Application to the parametric array	83
6.2.1 Simulation	84
7 Biological Effects	86
7.1 Effects on the body	87
7.2 Auditory effects	87
7.3 Subjective effects	88
8 Conclusions and Future Work	89
Bibliography	91
A Appendix A: Ultrasonic Exposure Study	95
A.1 Average Threshold Shifts	119
A.2 Stimulus Measurements	120
A.3 OSHA Standards for Ultrasound Exposure	126
A.4 Primary investigator C.V.	128
A.5 One Hour Exposure Supplement	131

Chapter 1

Introduction

Traditional sound systems are extremely good at filling rooms with sound and providing listeners with generally satisfying listening experiences. The art and science of loudspeaker design has been greatly developed over the last 75 years by many individual and groups of pioneers in the field.

While loudspeaker and sound systems have greatly evolved over the past several decades, there have been two areas which have received comparatively little attention, largely because available technology was unable to fully address them. By making a fundamental departure from traditional sound system design, as described in this thesis, these areas have enjoyed a renewed interest. The underlying theme is *control*, of which there are two aspects:

- Control of *position*

When listening to sound in a natural environment, the position of each sound source relative to the listener is a fundamental feature of each sound. A sound reproduction system with the ability to place each sound *exactly* where it should be could more faithfully mimic the real-world listening experience, and provide much greater realism.

- Control of *distribution*

In many cases, particularly in public areas, it is desirable to control *who hears what*. Traditional sound systems are extremely good at providing all listeners in a space with the same sounds, but without headphones, it is impossible to control which listeners receive the intended sounds.

This thesis presents a method of sound reproduction that was specifically designed to address these two points.

In order to control the distribution of sound waves through space, an analogy was drawn from our familiarity with lighting – while loudspeakers are like light bulbs, which fill rooms with light, until the development of the Audio Spotlight, no correlate to the spotlight or laser existed for sound.

The basic directivity of a sound source is related to the ratio of wavelength to size of source. For sources smaller than, or on the order of, a wavelength, sound propagates essentially omnidirectionally. For sound sources much larger than the wavelengths it is producing, the sound propagates with a high directivity.

Audible sound waves have wavelengths ranging from about an inch to several feet, so any loudspeaker of nominal size will be rather non-directional. However, if instead of transmitting audible sound waves, we transmit *only* ultrasound, with wavelengths of just a few millimeters, we can create a narrow beam of ultrasound, as the wavelengths are now much smaller than the sound source.

As ultrasound is completely inaudible, however, we need to rely on a nonlinearity in the propagation medium to cause the ultrasound to convert itself to audible sound. This nonlinearity exists as a perturbation in the speed of sound as a function of local air pressure or density.

As any weak nonlinearity can be expressed as a Taylor series, we can consider the first nonlinear term, which is proportional to the function squared.

If a collection of sine waves is squared, i.e.,

$$y = \left(\sum_i a_i \sin \omega_i \right)^2 \quad (1.1)$$

it can be shown that the nonlinearity creates a waveform containing frequencies at the sums and differences between each pairwise set of original frequencies:

$$y = \sum b_{i,j} (\sin(\omega_i + \omega_j) + \sin(\omega_i - \omega_j)) \quad (1.2)$$

If each $\omega_{i,j}$ exists within the ultrasonic frequency range, we can safely ignore the $\omega_i + \omega_j$ terms, as they result in additional ultrasonic frequencies. But more importantly, if frequencies are chosen correctly, $\omega_i - \omega_j$ is within the audible range.

This nonlinearity causes an explicit energy transfer from ultrasonic frequencies to audible frequencies. Therefore, the ultrasonic beam *becomes the loudspeaker* – the length of this beam, which is limited only by absorption of ultrasound into the air, can extend for several meters – which is much larger than most audible wavelengths. With a loudspeaker (even an invisible one) much larger than the wavelengths it is producing, a highly directional beam of audible sound results.

This dissertation is devoted to the investigation of this effect, termed a *parametric array* [1], and its application as an audible sound source.

1.1 Contemporary methods of Sound Control

Several methods have been adopted to provide control over either perceived sound location, or sound distribution, but rarely both. Common techniques make use of two basic techniques: physical, where the actual sound waves are controlled in some specific man-

ner, and psychoacoustical, which relies on perceptual processes of the listener to provide the intended effect.

1.1.1 Loudspeakers and arrays

The most familiar methods of sound control are the use of multiple loudspeakers distributed about the listener, as in a traditional stereo system, or the use of large arrays of loudspeakers, common to large musical venues. Of course, headphones, which are essentially small loudspeakers placed directly at the ears, are a very common way to provide individualized sound.

The use of multiple loudspeakers in a listening environment has evolved considerably since the advent of stereophonic sound in 1931 [2]. Extended from stereo, various other multichannel methods, such as quadraphonic, ‘surround sound’, 5.1 channel cinema sound, Logic 7, and AC-3 [3] have been introduced over the years. Each of these provide various embellishments of the basic idea of using multiple loudspeakers in a listening space, and each has their apparent niche and corresponding followers. But the dominant characteristic common to these methods is the use of simple panning between various loudspeakers placed about the listener. While largely effective for most environments, the spatial aspect of the positioned sound is often lacking, or at least relies on the presence of a loudspeaker at every desired location of the sound effect.

Loudspeaker arrays are very common when creating sound for large numbers of listeners, such as rock concerts. Because the physical size of these arrays is much larger than (most) audible wavelengths, the physical geometry and phase characteristics can be manipulated in many ways to provide a desired distribution of sound. However, because this type of geometrical manipulation requires a physically very large loudspeaker array, these methods are limited to very large venues.

1.1.2 Speaker domes

There are a variety of so-called “loudspeaker domes”, which typically rely on reflecting the output from a traditional (small) loudspeaker element against a curved reflector, which is generally parabolic or spherical. The intent is to ‘focus’ the sound to a person located just below the device.

While there is anecdotal evidence of their reasonable performance at short ranges, there have been very few formal objective analysis of their actual performance. In this dissertation, there will be a subsection devoted to reviewing the performance of these types of loudspeakers.

1.1.3 Binaural audio

The field of binaural or “3D” audio has enjoyed a substantial increase in attention, due in part to the prevalence of low-cost signal processing hardware and the strong marketing

of these systems. These systems work by taking advantage of the perceptual processes of directional hearing, thereby providing the listener with an illusion of sound existing at a particular point in space, or from a desired direction. Addressed in great detail in [4] and [5], this method of spatial audio presentation is generally limited to a stationary, solitary listener, preferably with a head-tracking apparatus.

1.2 Origins of the Parametric Array

The *parametric array* was developed in the late 1950s as a unique and enticing sonar technique [6], not only for improving the directivity (and directive consistency due to variation in wavelength) of the sonar beam, but also to increase the available bandwidth, resulting in shorter pulses and higher resolutions [7].

Throughout the next several decades, perhaps due to the generous investment from naval sources for its development, many researchers both in the US [8–12] and the USSR [13, 14] continued to develop theories and mathematical formalisms related to the nonlinear propagation of acoustic waves.

While there had been some level of discussion, speculation, and experimentation [8] regarding these nonlinear processes in air, it was not until 1975 [15] when a rigorous study was done of an airborne parametric array. These researchers were not intending to reproduce audible sound for listening applications (in fact, because the ‘ultrasound’ they were using was 18.6 kHz and 23.6 kHz, one would not want to be anywhere near their device), they nonetheless were able to show that the expected nonlinear effects did, indeed, exist in air.

In the early 1980s and later, several groups [16–19] had attempted to fabricate a loudspeaker that used these nonlinear effects to make audible sound. While they were able to create audible sound, significant problems with audible distortion, power requirements, ultrasonic exposure issues, and general device feasibility caused most of these researchers to abandon the technology.

More recently, other researchers such as [20] showed renewed interest in the technology, but these systems were essentially identical to those published in the early 1980s [21], used the very same transducers and signal processing techniques, and thus contained the very same shortcomings. To date, these groups have shown no published papers showing improvements over the earlier devices.

By recognizing the difficulties these earlier researchers had with the technology, as well as integrating much more of the early mathematical work by the sonar researchers, I was able to construct the very first audible, practical airborne parametric array with low distortion [22].

This thesis will draw on, and extend the earlier work in the field, particularly the mathematical formalisms developed for underwater acoustic beams, with their derivations adjusted for relevance to audio reproduction. The most important extensions to the earlier work provided by this thesis are:

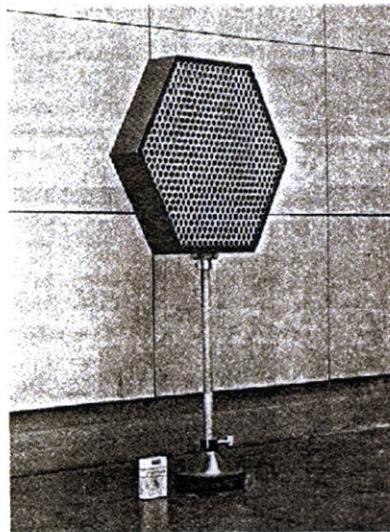


Figure 1.1: Early parametric array from 1983 [18]

- Distortion and fidelity

The early mathematical derivations were intended for sonar, not audio reproduction, so the mechanisms of demodulated distortion were generally not addressed. This thesis seeks to address these aspects of sound reproduction using ultrasound and to provide a rigorous set of experimental data detailing the results.

- Geometrical characteristics

The geometry of the beams has not been fully explored in the earlier literature. Earlier researchers generally used only axial measurements, or farfield polar plots, for describing the distribution of audible sound. By using novel measurement techniques, the thesis will present the full field of both ultrasound and audible sound with a two-dimensional colored graph.

- Ultrasonic exposure

As there is very little relevant literature describing the risks associated with exposure to airborne ultrasound, attention will be paid to this issue in the thesis. A thorough review of available literature will be presented, as well as the results of actual experiments done with an Audio Spotlight by Professor Martin Lenhardt of Virginia Commonwealth University.

- Phased arrays and steering

There appears to be no literature formally exploring the use of a phased array to steer a parametric audio source. Such a possibility has many applications, and deserves attention. The thesis will explore the possibility from a theoretical standpoint, but a physical device was not constructed due to time and cost constraints.

Chapter 2

Mathematical Basis

While there has been comparatively little mathematical derivation performed specifically for airborne parametric arrays, there is fortunately a substantial body of theoretical work devoted to the modeling of underwater parametric arrays. While the applications underwater are very different than those in air, which is reflected in the focus of the theoretical work, it does provide a solid theoretical foundation and invites straightforward extensions for airborne audio generation.

In this chapter, the basic derivation of the governing equations is presented. The derivations here are largely based upon those appearing in earlier work [7, 13, 23], but with guidance toward properties specific to airborne parametric arrays.

The derivation of equations describing the nonlinear propagation of acoustic waves is very similar to the familiar derivation of the standard linear wave equation. Except, of course, in this derivation, the ‘troublesome’ nonlinear terms usually discarded will be retained.

2.1 Equation of State

In the equation of state, the relationship between absolute pressure P , density ρ , and specific entropy s is established. For a general fluid, the pressure as a function of density and specific entropy, which is nonlinear, can be expanded with a Taylor series, with ambient pressure P_0 and while isentropic:

$$P = P(\rho, s) \tag{2.1}$$

$$= P_0 + \left(\frac{\partial P}{\partial \rho} \right)_{s,0} (\rho - \rho_0) + \frac{1}{2!} \left(\frac{\partial^2 P}{\partial \rho^2} \right)_{s,0} (\rho - \rho_0)^2 + \dots \tag{2.2}$$

The perturbed pressure and density are:

$$p = P - P_0 \quad (2.3)$$

$$\rho' = \rho - \rho_0 \quad (2.4)$$

making the substitution,

$$p = A \left(\frac{\rho'}{\rho_0} \right) + \frac{B}{2!} \left(\frac{\rho'}{\rho_0} \right)^2 + \frac{C}{3!} \left(\frac{\rho'}{\rho_0} \right)^3 + \dots, \quad (2.5)$$

where

$$A = \rho_0 \left(\frac{\partial P}{\partial \rho} \right)_{s,0} \quad (2.6)$$

$$B = \rho_0^2 \left(\frac{\partial^2 P}{\partial \rho^2} \right)_{s,0} \quad (2.7)$$

$$C = \rho_0^3 \left(\frac{\partial^3 P}{\partial \rho^3} \right)_{s,0}. \quad (2.8)$$

The speed of sound c in an isentropic fluid is defined and expanded as:

$$c^2 = \left(\frac{\partial P}{\partial \rho} \right)_s \quad (2.9)$$

$$= \left(\frac{\partial P}{\partial \rho} \right)_{s,0} + \left(\frac{\partial^2 P}{\partial \rho^2} \right)_{s,0} (\rho - \rho_0) + \frac{1}{2!} \left(\frac{\partial^3 P}{\partial \rho^3} \right)_{s,0} (\rho - \rho_0)^2 + \dots \quad (2.10)$$

$$= c_0^2 + B \frac{\rho'}{\rho_0^2} + \frac{C}{2} \frac{\rho'^2}{\rho_0^3} \quad (2.11)$$

After square-rooting, a binomial expansion allows a solution for c :

$$\frac{c}{c_0} = 1 + \frac{B}{2A} \left(\frac{\rho'}{\rho_0} \right) + \frac{1}{4} \left[\frac{C}{A} - \frac{1}{2} \left(\frac{B}{A} \right)^2 \right] \left(\frac{\rho'}{\rho_0} \right)^2 + \dots \quad (2.12)$$

where c_0 is the ambient (small signal) wave speed, and is equal to $\sqrt{A/\rho_0}$.

If the perturbations in density are assumed to be small compared to the density of air, the term containing quadratic and higher powers of (ρ'/ρ_0) can be omitted. For a plane wave, this approximation holds for $u \ll c_0$, where u is the particle velocity, or, equivalently, when $p \ll \rho_0 c_0^2 \approx 197$ dB SPL.

For an isentropic, diatomic gas (such as air), the equation of state can be written directly as:

$$p = P_0 \left(\frac{\rho}{\rho_0} \right)^\gamma \quad (2.13)$$

where $\gamma = \frac{c_p}{c_v}$ is the ratio of specific heats. In air, $\gamma = 1.4$, so

$$\frac{c}{c_0} = 1 + \frac{\gamma - 1}{2} \left(\frac{\rho'}{\rho_0} \right) . \quad (2.14)$$

2.2 Equation of Continuity

For a stationary cube of air with volume dV , face area dA and edge length dx , the rate of air mass flowing inward along the x axis minus the mass rate of air flowing out is equal to the rate of increase in air mass within the volume:

$$dV \frac{\partial \rho}{\partial t} = \rho u dA - \rho u|_{x+dx} dA \quad (2.15)$$

$$= \left[\rho u - \rho u - \frac{\partial(\rho u)}{\partial x} dx \right] dA \quad (2.16)$$

$$= \frac{\partial(\rho u)}{\partial x} dV \quad (2.17)$$

$$(2.18)$$

Combining the three spatial dimensions, and cancelling dV , the resulting equation is:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{u}) = 0 \quad (2.19)$$

2.3 Euler's Equation

Considering the another small volume dV of air, which is of mass dm , it is known from Newton's second law ($F = ma$) that:

$$d\vec{f} = \vec{a} dm = dP d\vec{A} . \quad (2.20)$$

The pressure on the left side of the volume is P , and on the right side is $P + \frac{\partial P}{\partial x} dx$, so the net force is:

$$d\vec{f} = -\frac{\partial p}{\partial x} dV , \quad (2.21)$$

and in three dimensions,

$$d\vec{f} = -\nabla P dV . \quad (2.22)$$

The acceleration of this volume of air can be written as:

$$\vec{a} = \frac{d\vec{u}}{dt} = \lim_{dt \rightarrow 0} \frac{\vec{u}_1 - \vec{u}_0}{dt} \quad (2.23)$$

where

$$\vec{u}_1 = \vec{u}_0 + \underbrace{\frac{\partial \vec{u}}{\partial x} u_x dt + \frac{\partial \vec{u}}{\partial y} u_y dt + \frac{\partial \vec{u}}{\partial z} u_z dt}_{(\vec{u} \cdot \nabla) \vec{u} dt} + \frac{\partial \vec{u}}{\partial t} dt \quad (2.24)$$

so the acceleration of the air mass is:

$$\vec{a} = \frac{\partial \vec{u}}{\partial t} + (\vec{u} \cdot \nabla) \vec{u} . \quad (2.25)$$

Substituting this into the force equation gives:

$$d\vec{f} = \vec{a} dm = \vec{a} \rho dV \quad (2.26)$$

so that Euler's equation becomes:

$$-\nabla p = \rho \left[\frac{\partial \vec{u}}{\partial t} + (\vec{u} \cdot \nabla) \vec{u} \right] . \quad (2.27)$$

2.4 Combining the Equations

At this point in the derivation of a linear wave equation, nonlinear terms such as $(\vec{u} \cdot \nabla) \vec{u}$ would be linearized, and the final wave equation would easily simplify to $\nabla^2 p = \frac{1}{c_0^2} \frac{\partial^2 p}{\partial t^2}$.

Because the nonlinear terms are required, and the nonlinearity is weak, a perturbation method is used to permit an approximate solution to be constructed.

If the perturbations in the sound field have a smallness μ ,

$$\frac{\rho'}{\rho_0} \sim \frac{p'}{p_0} \sim \frac{u}{c_0} \sim \mu \quad (2.28)$$

and waves along the x direction are considered, a common solution to these parameters can be written as an arbitrary function F propagating along the characteristic $t - \frac{x}{c_0}$:

$$\rho', u_x, p = F(t - \frac{x}{c_0}) . \quad (2.29)$$

The shape of the wave is altered by the nonlinearity as it travels, in both the transverse and axial directions. If it is assumed that the transverse changes (due to diffraction) are stronger than axial changes (due to nonlinearity), the solution to the wave equation is assumed to be of the form:

$$\rho', u_x, p' = \psi(t - \frac{x}{c_0}, \mu x, \sqrt{\mu} y) . \quad (2.30)$$

The arguments can be re-scaled, as:

$$\tau = t - x/c_0, x' = \mu x, y' = \sqrt{\mu}y . \quad (2.31)$$

Doing so makes the partial derivatives even smaller:

$$\frac{\partial \rho'}{\partial x'} \sim \frac{\partial u_x}{\partial x'} \sim \frac{\partial p'}{\partial x'} \sim \mu^2, \frac{\partial \rho'}{\partial y'} \sim \frac{\partial u_y}{\partial y'} \sim \frac{\partial p'}{\partial y'} \sim \mu^{3/2}, u_y \sim \frac{\partial u_x}{\partial y} \sim \mu^{3/2} . \quad (2.32)$$

When the equations found in the last three sections are re-scaled and combined, low-order smallness $O(\sqrt{\mu}), O(\mu), O(\mu^2)$ can be combined, and higher-order smallness can be dropped. After a several pages of equation manipulation¹, one (hopefully) arrives at the KZK equation, named for Khokhlov, Zabolotskaya, and Kuznetsov [24, 25]:

$$\frac{\partial^2 p}{\partial z \partial \tau} = \frac{c_0}{2} \nabla_r^2 p + \frac{\delta}{2c_0^3} \frac{\partial^3 p}{\partial \tau^3} + \frac{\beta}{2\rho_0 c_0^3} \frac{\partial^2 p^2}{\partial \tau^2} . \quad (2.33)$$

This equation is a good approximation for directional sound beams for points near the axis. Approximate solutions of this equation will be used to arrive at an algorithm for low-distortion audio reproduction in Chapter 3, and to predict audible beam geometry in Chapter 4.

¹Contact the author for a photocopy of the handwritten mess if you like; all the steps won't be reproduced here.

Chapter 3

Low-Distortion Audio Reproduction

The equations describing the physical process of nonlinear ultrasonic wave propagation presented in the previous chapter, most notably the KZK equation, provide the starting point for a time-domain solution. While the KZK equation is not exactly solvable for general geometries, there are various approximation methods that provide reasonably accurate results. Researchers have also used direct numerical simulation or iterative techniques [26–28] to solve the KZK equation, but, by themselves, these simulations do not provide sufficient information to construct an algorithm for real-time sound reproduction. This chapter focuses on approximate analytical solutions to the KZK equation, with appropriate extensions relevant to the goal of low-distortion sound reproduction.

3.1 Basic Quasilinear Solution

The most straightforward approximate solution to the KZK equation is known as the *quasilinear* solution, a method dating to the 1850s [29], and was used in early work in parametric arrays [1, 6]. Essentially, the governing equation (the KZK equation, in this case) is solved in two parts; first, the linear ultrasonic field is calculated by setting the nonlinear coefficient β to zero. Once the ultrasonic field is known, it is used as the source term for the nonlinear solution.

Beginning with the KZK equation:

$$\frac{\partial^2 p}{\partial z \partial \tau} = \frac{c_0}{2} \nabla_r^2 p + \frac{\delta}{2c_0^3} \frac{\partial^3 p}{\partial \tau^3} + \frac{\beta}{2\rho_0 c_0^3} \frac{\partial^2 p^2}{\partial \tau^2}, \quad (3.1)$$

the resulting field is assumed to consist of two components, $p = p_1 + p_2$, where p_1 is the solution to the (linear) ultrasonic field, and p_2 is the nonlinearly produced result.

The equation describing p_1 is a standard, linear, wave equation (here written in cylindrical coordinates, and including absorption) that can be solved using a variety of standard methods:

$$\frac{\partial^2 p_1}{\partial z \partial \tau} = \frac{c_0}{2} \nabla_r^2 p_1 + \frac{\delta}{2c_0^3} \frac{\partial^3 p}{\partial \tau^3}. \quad (3.2)$$

Once the ultrasonic field is known, it is used as the ‘source’ in the equation below to solve for p_2 :

$$\frac{\partial^2 p_2}{\partial z \partial \tau} = \frac{c_0}{2} \nabla_r^2 p_2 + \frac{\delta}{2c_0^3} \frac{\partial^3 p_2}{\partial \tau^3} + \frac{\beta}{2\rho_0 c_0^3} \frac{\partial^2 p_1^2}{\partial \tau^2}. \quad (3.3)$$

Solving the equation for p_2 on-axis gives [30]:

$$p_2(x, r, \tau) = \frac{\beta}{2\rho_0 c_0^4} \frac{\partial^2}{\partial \tau^2} \int_0^x \int_0^\infty p_1^2(x', r', \tau - \frac{r'^2}{2c_0(x-x')}) \frac{r' dr' dx'}{x-x'}. \quad (3.4)$$

If the ultrasonic field is generated by a piston vibrating with uniform velocity $\rho_0 c_0 p_1(0, r, t)$, and is perfectly collimated ($ka \gg 1$), and assumed planar, the solution for the ultrasonic field p_1 is [30]:

$$p_1(x, r, t) = p_1(0, r, t) e^{-\alpha x} H(a - r) \quad (3.5)$$

where $H(r)$ is the unit step function, and α is the absorption coefficient of the ultrasound used.

The ultrasonic signal is assumed to be an AM-modulated waveform of the form:

$$p_1(0, r, t) = P_0 E(t) \sin(\omega_0 t) \quad (3.6)$$

Inserting this into the equation for p_2 , making the far-field assumption $x \gg L$, and dropping the high-frequency terms, the equation for the demodulated signal becomes:

$$p_2 = \frac{\beta P_0^2 a^2}{16\rho_0 \alpha c_0^4 x} \frac{d^2}{d\tau^2} E^2(\tau) \quad (3.7)$$

This is essentially the same equation as developed by Berklay [12]. The equation illustrates several important points regarding the audible sound:

- The audible level is proportional to the *square* of the ultrasonic level. Because of the square term, the system will actually become much more efficient as the ultrasonic level used is increased. For every doubling of ultrasonic level, audible sound is quadrupled.
- The audible level is proportional to the *area* of the transducer. In the farfield limit, louder systems can be easily made simply by increasing the area of the transducer.

- Low frequencies require more ultrasound to generate. The second derivative in time creates an inherent equalization curve of +12 dB/octave. Of course, this can be corrected with equalization, but only at the expense of overall output level. For each octave one descends in frequency, the required level of ultrasound doubles.
- The audible result is proportional to the *square* of the modulation envelope. As shown in the next section, simply taking the square-root of the audio signal reduces distortion dramatically, provided the system has sufficient bandwidth to reproduce this signal accurately.

3.2 Processing to reduce distortion

Under the farfield assumption, it was shown that:

$$p_2(x, 0, \tau) \propto \frac{d^2}{d\tau^2} E^2(t). \quad (3.8)$$

Therefore, to reproduce an audio signal $g(t)$, a straightforward solution would be to simply double-integrate, and take the square root before modulating:

$$E(t) = \left(1 + m \iint g(t) dt^2 \right)^{\frac{1}{2}}. \quad (3.9)$$

The factor m corresponds to modulation depth. If the modulated signal $E(t) \sin \omega_c t$ can be transmitted precisely, the demodulated signal should be a distortion-free audio signal.

As shown in the next section, reproducing a modulated, suitably-processed audio signal accurately is no simple task for traditional ultrasonic systems. In particular, the limited bandwidth of available ultrasonic transducers, and their reproduction apparatus (i.e. amplifiers, etc.), can lead to unacceptably high levels of distortion.

3.2.1 Effects of Bandwidth

The main challenge is that the square-rooting operation introduces harmonics into the signal, and necessarily increases its bandwidth significantly, as shown in Figure 3.1. When the audio signal is modulated to ultrasound, the entire reproduction system must reliably reproduce the entire signal bandwidth as accurately as possible. Any limitations in bandwidth, or even nonuniformities in the ultrasonic frequency response, will lead to an increase in audible distortion.

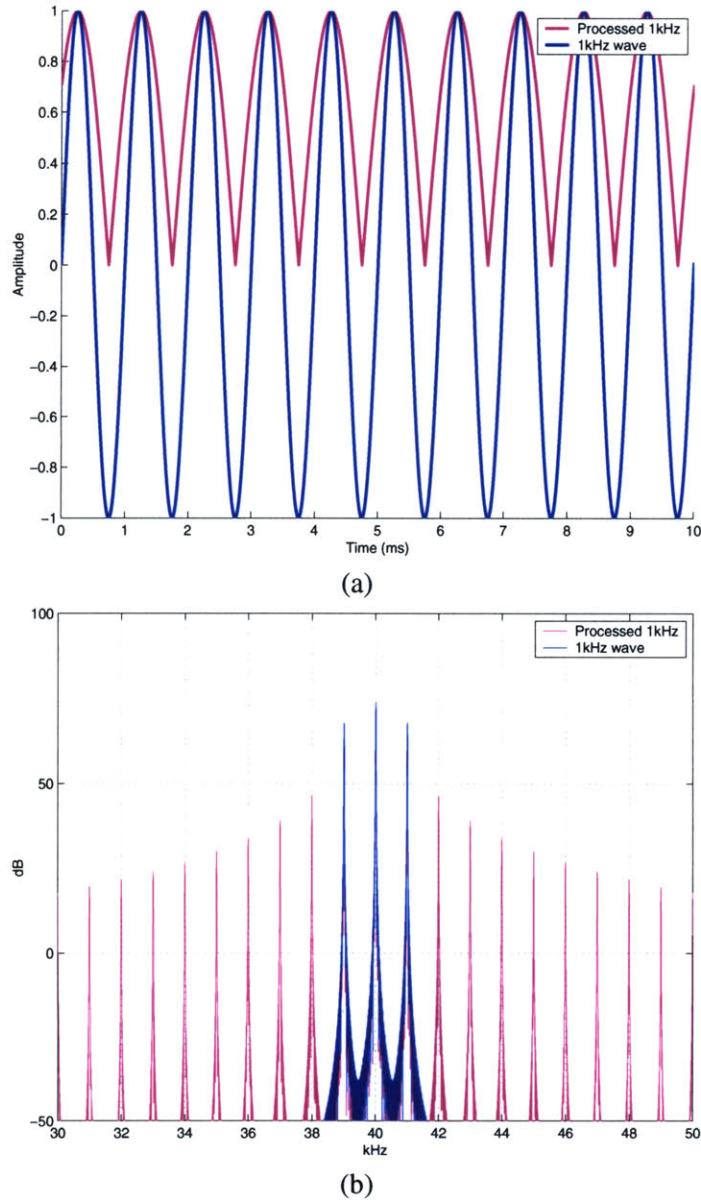


Figure 3.1: Example unprocessed and processed 1 kHz waveforms are shown. In (a), the unprocessed and processed 1 kHz audible signal are shown in the time domain, and their corresponding spectra are in (b). Note that the bandwidth of the processed spectra is much larger than before processing.

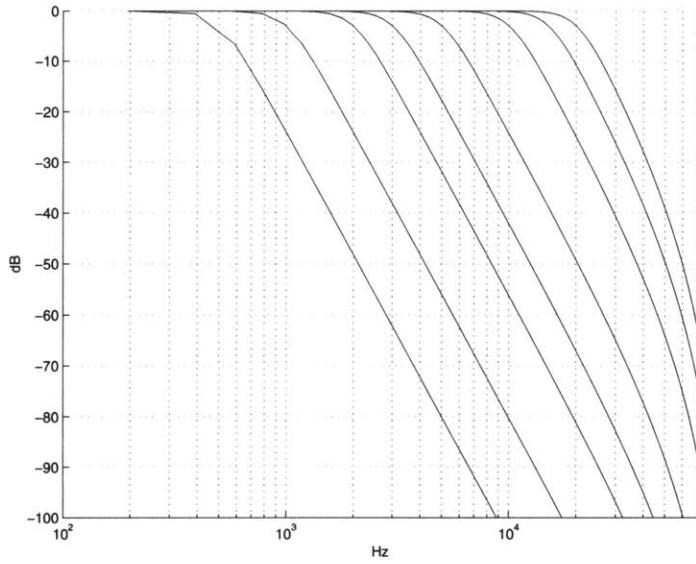


Figure 3.2: Various filters used to simulate the effects of a restricted bandwidth. Each is a fourth-order Butterworth filter, with -3 dB cutoffs set for several frequencies.

Distortion as a function of bandwidth

It has been shown theoretically [31] that when using the processing algorithm above the resulting audible distortion is correlated with the bandwidth of the ultrasonic system. The simulations given by [31] are expanded to include all audible frequencies (which had not been considered), and, since to the current approximation, the carrier frequency has no impact on the demodulated audio, that particular parameter will be omitted.

Filters simulating the various bandwidths of the ultrasonic system were created, and are shown in Figure 3.2. Each was generated with a fourth-order Butterworth, and simulates reproduction bandwidths of 1 kHz, 2 kHz, 4 kHz, 6 kHz, 10 kHz, 20 kHz, 30 kHz, and 40 kHz.

To carry out the simulation, the processed (offset and square-rooted) audio signal is first fed through one of the filters, and are then squared. The resulting audible harmonic distortion is computed, recorded, and plotted as a function of frequency and available bandwidth. The simulation was repeated with modulation depths of 1.0, 0.95, 0.9, 0.75, and 0.5.

From the graphs, several interesting results can be observed. First, the level of distortion as a function of frequency, regardless of bandwidth, is not uniform, but rather has a distinct peak whose location is essentially independent of the modulation depth. The peak location seems to correlate with the sharp transition between the flat portion of the filter and the strongly sloped section.

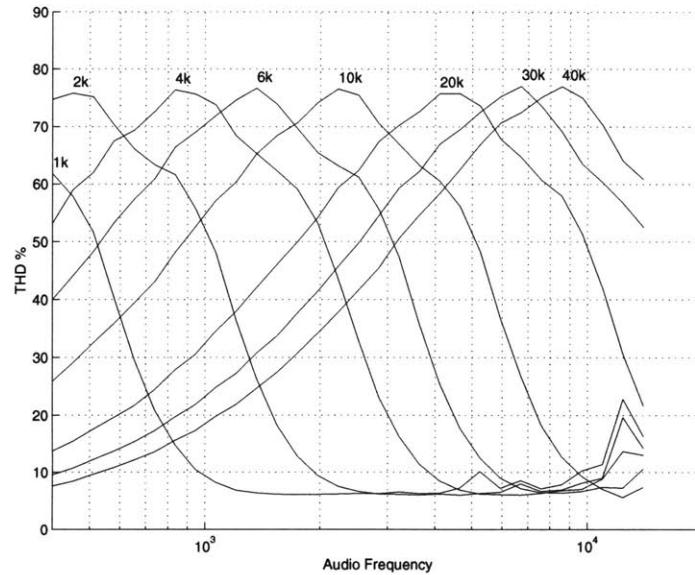


Figure 3.3: Audible distortion as a function of frequency and transducer bandwidth, for 100% modulation depth. Each curve corresponds to a specific transducer bandwidth as labelled. Note that a wide bandwidth is necessary for ensuring low distortion at low frequencies, but at higher frequencies, bandwidth is of less importance.

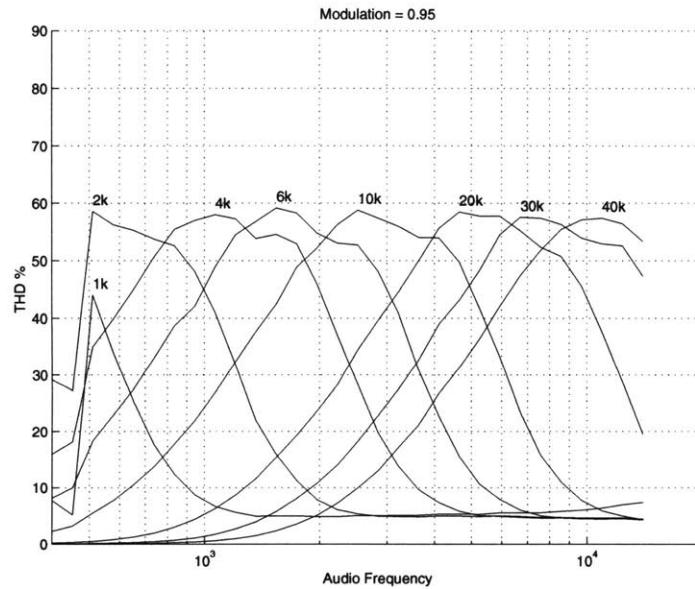


Figure 3.4: Audible distortion as a function of frequency and transducer bandwidth, at 95% modulation depth.

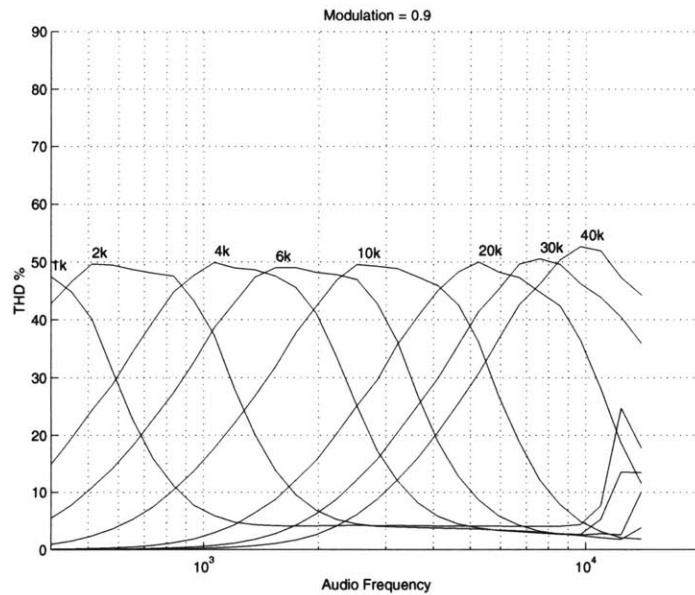


Figure 3.5: Audible distortion as a function of frequency and transducer bandwidth, at 90% modulation depth.

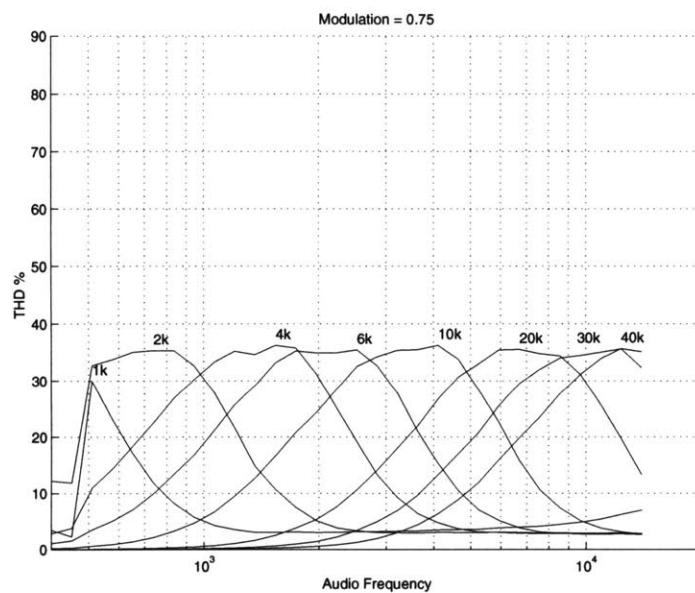


Figure 3.6: Audible distortion as a function of frequency and transducer bandwidth, at 75% modulation depth.

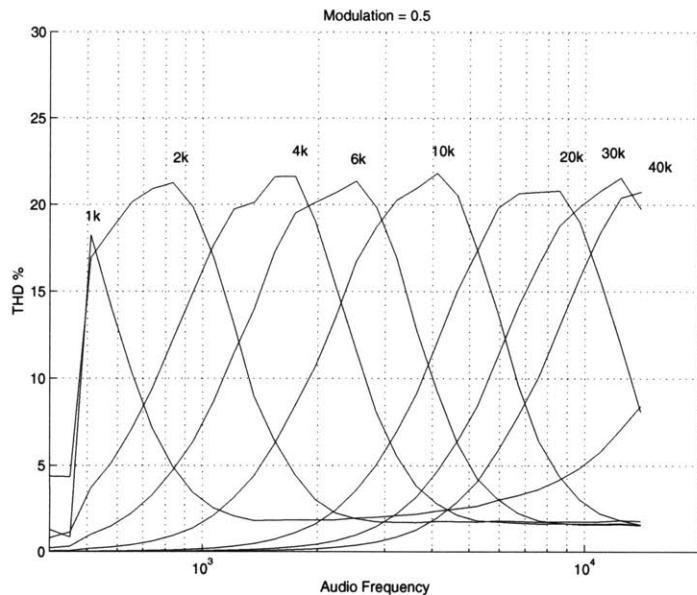


Figure 3.7: Audible distortion as a function of frequency and transducer bandwidth, at 50% modulation depth.

It is very clear that even a modest reduction in modulation depth can lead to a large reduction in distortion, most notably between 100% and 90%. The corresponding reduction in audible output by reducing modulation depth to 90% is less than 1 dB.

3.3 Experimental Results

To test the efficacy of this algorithm, a broadband ultrasonic source and associated electronics was constructed. The transducer used was approximately 40cm diameter, and measurements were made on-axis at a distance of 2m. The modulation depth was approximately 80%.

3.3.1 Ultrasonic Bandwidth

The ultrasonic frequency response is shown in Figure 3.8. Bandwidth of the ultrasonic system was measured by applying a slow sweep to the transducer while simultaneously measuring its fundamental frequency output. The results were smoothed before plotting.

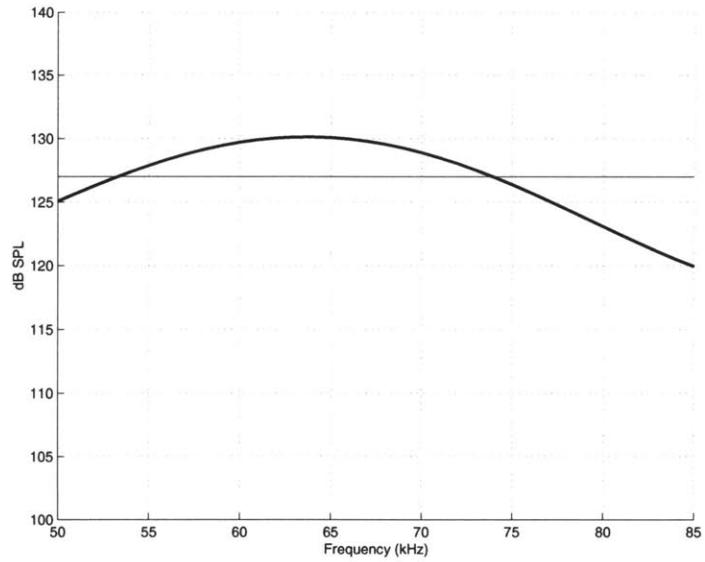


Figure 3.8: The transducer and amplifier frequency response is shown. The practical bandwidth of the system, measured by the -3 dB points (horizontal line), extends from approximately 53 kHz to 74 kHz, allowing 21 kHz total bandwidth.

For this amount of bandwidth, and with 80% modulation depth, we expect distortion to be well under 5% for frequencies lower than 2 kHz, and somewhat more thereafter.

Transducers (and amplifiers) used in earlier attempts contained a much narrower bandwidth – typically on the order of 4 kHz [18]. As predicted by these simulations, this restriction in bandwidth would raise distortion levels to approximately 60%, unless modulation depth was significantly reduced. Even at a modulation depth of 50%, distortion will be near 20% for frequencies near 1 kHz.

3.3.2 Distortion Versus Frequency

To measure distortion versus frequency, the input frequency to the system was swept, and the output was continuously analyzed for audible level and distortion. Two volume settings were used. The results for each are in Figure 3.9 and 3.10.

In the moderate volume setting (Figure 3.9), distortion level is rather uniform as a function of frequency, beginning at 5% for low frequencies and declining to 1% for higher frequencies. It is likely that the ambient room noise, which was on the order of 20 dB, has exaggerated low-frequency distortion somewhat.

The expected peak from the simulations of the last section was not present in this set of measurements. It is hypothesized that this is due to the absence of a sharp transition between the passband and stopband in the ultrasonic frequency response of the actual system.

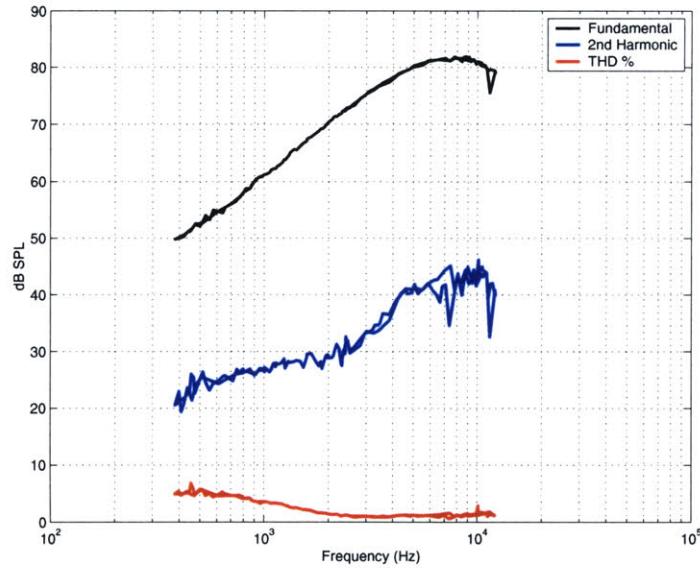


Figure 3.9: Output level, second harmonic level and %THD are plotted as a function of frequency for a moderate volume setting. Note that distortion level is uniformly very low, well under 5%, and as low as 1%. The nonflat frequency response is due to the absence of an input equalizer.

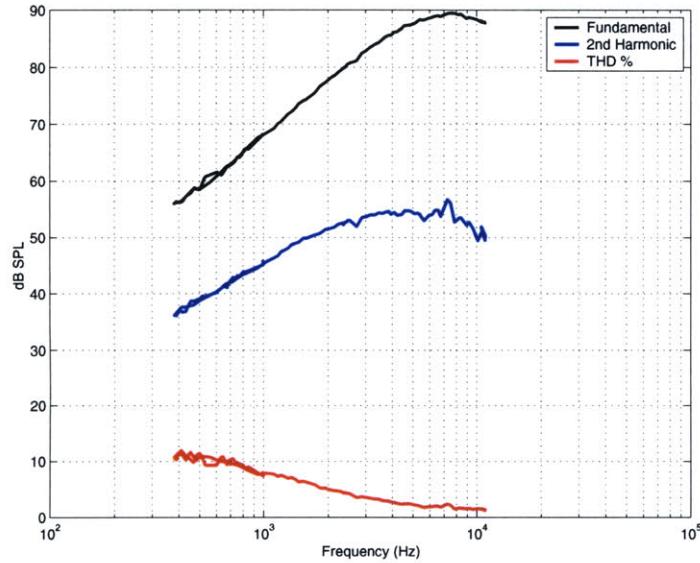


Figure 3.10: Output level, second harmonic level and %THD are plotted as a function of frequency for a higher volume setting. Note that distortion level is still quite low, although slightly higher than the previous case.

3.3.3 Distortion Versus Level

As the system has internal limitations on output level, it would be instructive to examine the relationship between the output level and level of distortion. In this experiment, a processed 1 kHz signal was slowly ramped up in amplitude while audible output level and distortion were computed and recorded. The results are shown in Figure 3.11.

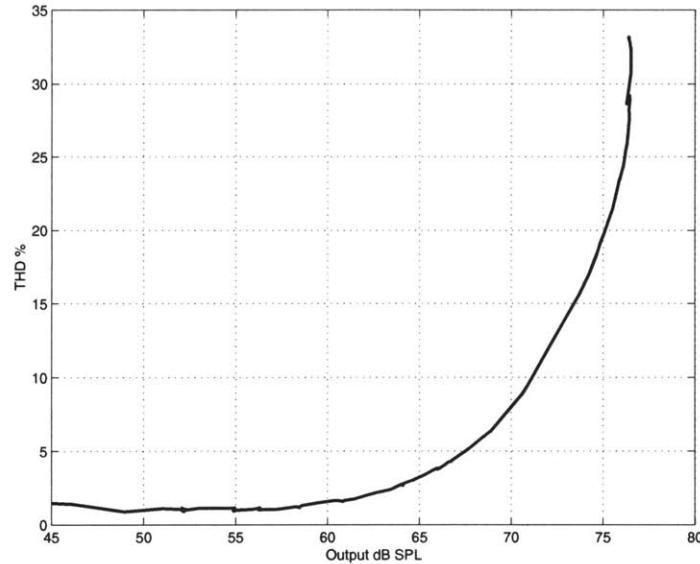


Figure 3.11: The distortion level of a 1 kHz signal is plotted against the output sound pressure level, measured at 2m distance. Note that distortion is extremely low for moderate sound pressure levels.

At moderate signal output levels, the distortion is very low – reaching 1% for most amplitudes. The actual distortion of the generated signal may, in fact, be even lower, as the detected distortion for low levels may simply be ambient background noise.

As output level increases, distortion increases exponentially, which is likely due to the approaching limitations of the amplifier (clipping) and/or transducer (excursion), as well as the limitations of the quasilinear approximation to predict the resulting audible field.

It is clear that the increased bandwidth of this system compared to those constructed in the past [18, 19, 32] has permitted a substantial reduction in audible distortion.

3.4 Modulation depth control

The general processing scheme, including offsetting, integrating, and square rooting, has been shown as an effective way to reduce distortion in an airborne parametric array. But note that when the audible signal $g(t)$ to be reproduced is zero, the system is still generating a significant amount of ultrasound (1/2 peak energy). Clearly this is undesirable, as it leads to significant wasted power and stress on the system.

If the unitary offset in Equation 3.2 is replaced with a time-varying function $L(t)$, the modulation envelope becomes:

$$E(t) = \left(L(t) + \iint g(t) \right)^{\frac{1}{2}}. \quad (3.10)$$

The function $L(t)$ effectively controls the amount of modulation depth, given a varying amplitude signal $g(t)$.

The demodulated waveform containing $L(t)$ is:

$$p_2(t) \propto \frac{d^2}{dt^2} E^2(t) \quad (3.11)$$

$$\propto \frac{d^2}{dt^2} \left(L(t) + \iint g(t) \right) \quad (3.12)$$

$$\propto \frac{d^2}{dt^2} L(t) + g(t) \quad (3.13)$$

The function $L(t)$ should be designed to follow the envelope $g(t)$, such that the quantity $L(t) + g(t)$ is minimized, but always positive, and that the second time derivative of $L(t)$ is inaudible.

Other researchers [32] have proposed the use of a traditional envelope follower for $L(t)$, but did not fully address the problem of overmodulation, which occurs when $L(t) + g(t) < 0$. Traditional envelope detectors are typically implemented by first taking the absolute value of the waveform, and applying a lowpass filter (LPF), so that:

$$L(t) = LPF[|g(t)|] \quad (3.14)$$

The results of this envelope detection for a voice signal is shown in Figure 3.12. If the lowpass filter is sufficiently low in frequency, it will be relatively inaudible compared to $g(t)$ due to the second derivative. However, since $L(t)$ is slowly moving compared to $g(t)$, any rapid increase in $g(t)$ will cause the sum $L(t) + g(t)$ to become negative, as seen in Figure 3.13, resulting in overmodulation and substantial transient distortion.

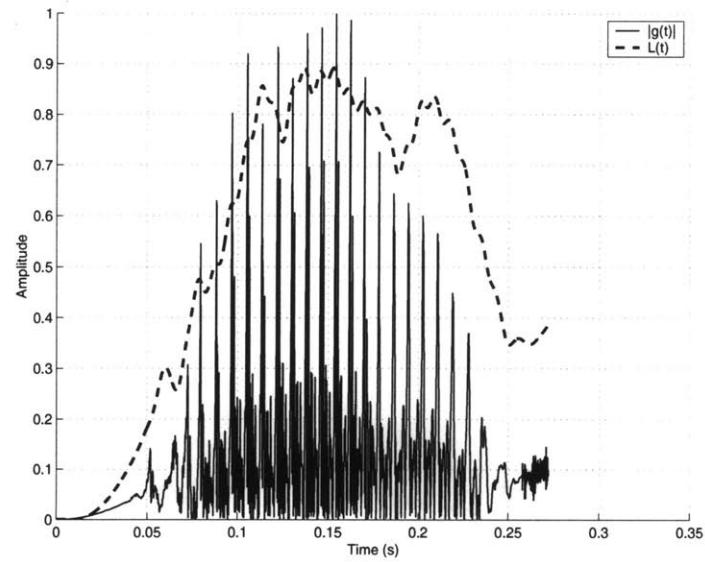


Figure 3.12: The instantaneous (voice) signal amplitude is plotted, along with the detected envelope. The detected envelope generally follows the contour of the input signal, but falls behind during rapid transitions.

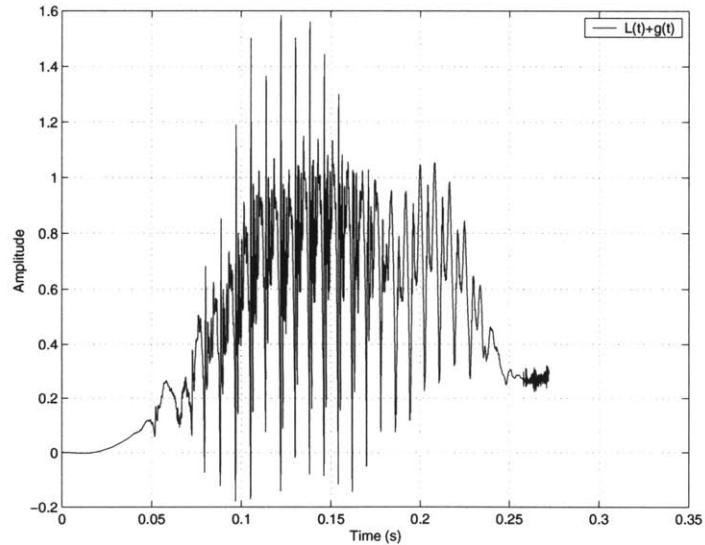


Figure 3.13: The resulting offset signal $E(t)$, used prior to ultrasonic modulation, is shown. The negative portion of the signal corresponds to overmodulation.

To remedy these problems, a temporally asymmetric offset function $L(t)$ is used, such as a peak detector with a slow decay (long time constant). An instantaneous attack allows $L(t)$ to react to abrupt changes in the incoming signal, while the slow decay removes the residual term in the expression for audible sound.

Assuming the target modulation depth is $m = 1$, the level detector $L(t)$ can be written as:

$$L(t) = U(t)e^{-rt} * \iint g(t)dt^2 \quad (3.15)$$

with $U(t)$ being a unit step, $*$ denoting convolution. The decay rate is set with r , typically with $r \lesssim 1$. Small values of r are allowed, as long as the distortion introduced by this term is of very low frequency (becoming inaudible).

The $L(t)$ generated due to an impulse is:

$$L(t) = U(t)e^{-rt} \quad (3.16)$$

Differentiating twice, the demodulated result $L''(t)$ is:

$$L'(t) = \delta(t)e^{-rt} - rU(t)e^{-rt} \quad (3.17)$$

$$L''(t) = \delta(t)e^{-rt} - \delta(t)re^{-rt} - r(\delta(t)e^{-rt} - rU(t)e^{-rt}) \quad (3.18)$$

$$= \delta(t)[1 - 2r] - r^2U(t)e^{-rt} \quad (3.19)$$

since $\delta(t) \approx 2r\delta(t)$ and r^2 is small, $L''(t)$ vanishes, and the resulting demodulated audio is then:

$$p_2(t) \propto g(t) \quad (3.20)$$

which is exactly the target audio output. For additional inaudibility, if necessary, the asymmetric envelope $L(t)$ can also be low-passed without significantly altering the results.

The new envelope using this algorithm is shown in Figure 3.14, and the resulting offset (summed) signal is shown in Figure 3.15. Note that the offset signal is uniformly positive, and thus overmodulation, and its coincident noise, is eliminated.

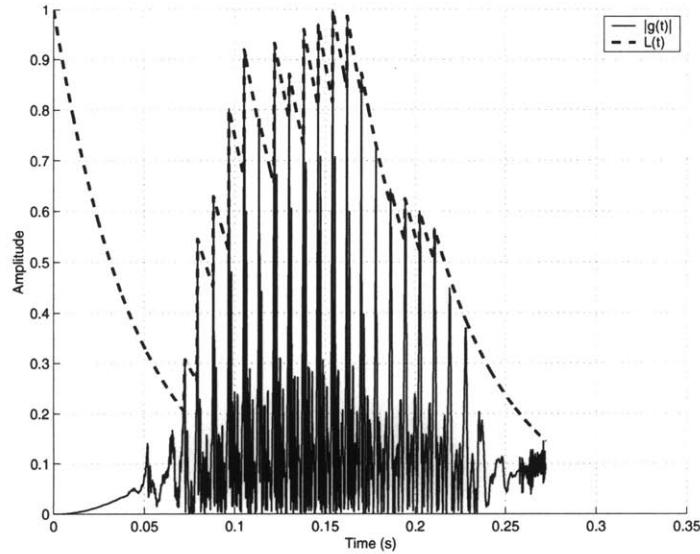


Figure 3.14: The instantaneous (voice) signal amplitude $|\int \int g(t) dt^2|$ is plotted, along with the detected envelope $L(t)$ using the proposed peak detect algorithm. Notice how this envelope successfully follows the contours of the peaks during fast transitions.

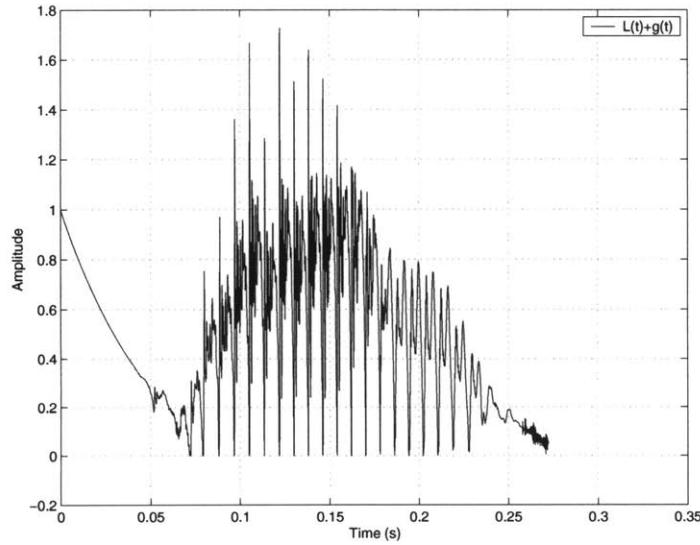


Figure 3.15: The offset signal, used prior to modulation, and with the peak detect method envelope is shown. There is no negative portion of this wave, so overmodulation is prevented.

The energy W required for an envelope algorithm can be estimated by integrating the output signal:

$$W = \int E^2(t)dt . \quad (3.21)$$

The relative amounts of energy, compared to a constant offset signal ($L(t) = 1$), is shown in Figure 3.16. The source signal was a short voice segment, approximately one second long. Clearly, energy savings through these methods are dramatic – straightforward envelope detection has reduced average energy requirements by approximately 70%, and for the peak-detect algorithm, energy use has been decreased by approximately 75%.

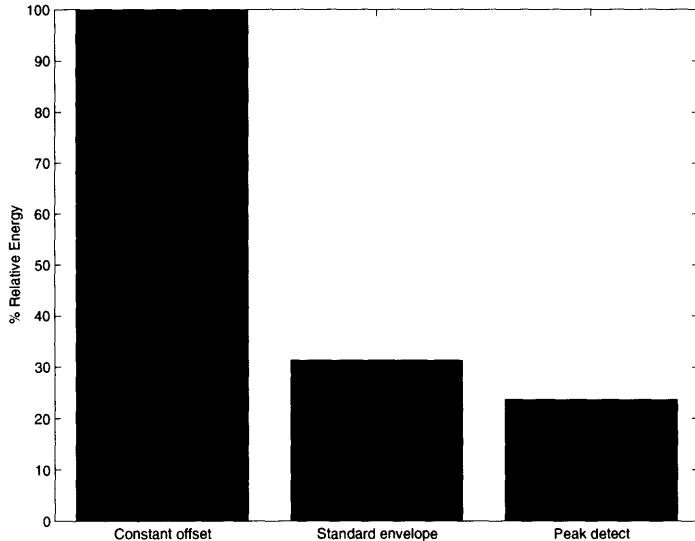


Figure 3.16: The relative energy use between various modulation level control methods for a typical voice signal is shown. The leftmost plot corresponds to the uncontrolled offset ($L(t) = 1$), the middle is the result of the use of a standard envelope detector, the the rightmost plot shows the novel peak-detect method. Energy savings are dramatic when either envelope detection algorithm is used.

This modulation control algorithm has several important benefits, in that it continually adjusts both the *modulation depth* and *ultrasonic output* so that *there is never any more ultrasound used than necessary to recreate the audio*. Overmodulation distortion is prevented, no matter the input source, and device efficiency is optimized. In addition, this algorithm is fully causal, and can be implemented readily into a working system.

3.5 Ultrasonic Absorption

Ultrasonic absorption is known to directly affect the audible sound output (Eq. 3.1), as well as the directivity (next chapter), and possibly amount of distortion in the audible sound.

The airborne absorption model used is one originally presented by [33], as updated by [34–36]. It offers a convenient, closed form solution for airborne sound absorption by the following formula:

$$\alpha = \frac{f^2}{p_s} \left[1.84 \times 10^{-11} \left(\frac{T}{T_0} \right)^{\frac{1}{2}} + \left(\frac{T}{T_0} \right)^{-\frac{5}{2}} \left[0.01275 \frac{e^{-2239.1/T}}{F_{r,O} + f^2/p_s^2 F_{r,O}} + 0.1068 \frac{e^{-3352/T}}{F_{r,N} + f^2/p_s^2 F_{r,N}} \right] \right] \quad (3.22)$$

where p_s is atmospheric pressure (atm), T is the air temperature in Kelvin, $T_0 = 293.15$ is a reference temperature, and $F_{r,O}$ and $F_{r,N}$ are the relaxation frequencies for oxygen and nitrogen, respectively, given by:

$$F_{r,O} = \left(24 + 4.04 \times 10^4 h \frac{0.02 + h}{0.391 + h} \right) \quad (3.23)$$

$$F_{r,N} = \left(\frac{T_0}{T} \right)^{1/2} \left(9 + 280 h e^{4.17[(T_0/T)^{1/3} - 1]} \right). \quad (3.24)$$

The absolute humidity h is related to relative humidity h_r through the saturation vapor pressure p_{sat} :

$$h = h_r \frac{p_{\text{sat}}}{p_s} \quad (3.25)$$

The units of α is nepers, which can be converted to dB/m by multiplying by $20 \log(e)$. At $T = 20^\circ\text{C}$ and $p_s = 1\text{ atm}$, values for acoustic absorption are plotted as a function of frequency in Figure 3.17. While the rate of absorption is very low for audible frequencies, ultrasonic absorption can be quite pronounced, particularly at higher frequencies. Further, there is a variation in absorption rates with respect to humidity, which, if not corrected, could lead to additional distortion in the audible result.

If air temperature is included as an additional variable, ultrasonic absorption is shown in Figure 3.18 to be slightly nonuniform over most ultrasonic frequencies.

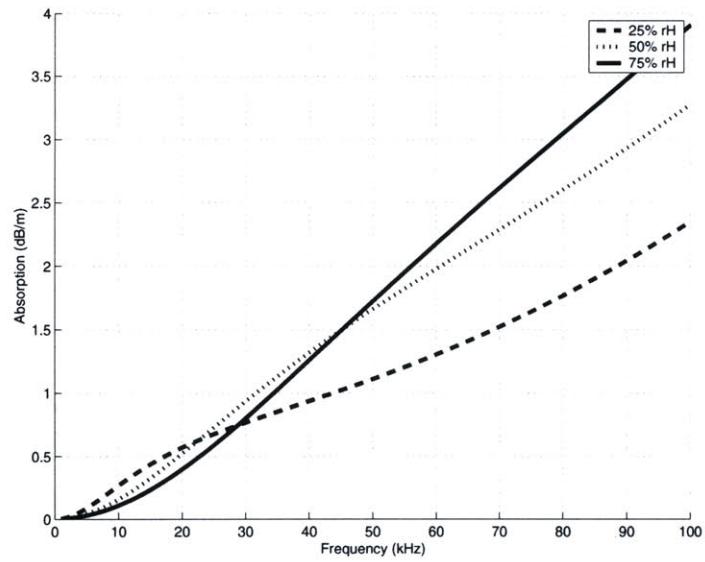


Figure 3.17: Acoustic absorption rates at 1 atm and 20C are shown as a function of frequency. Note that absorption becomes much more significant at higher frequencies.

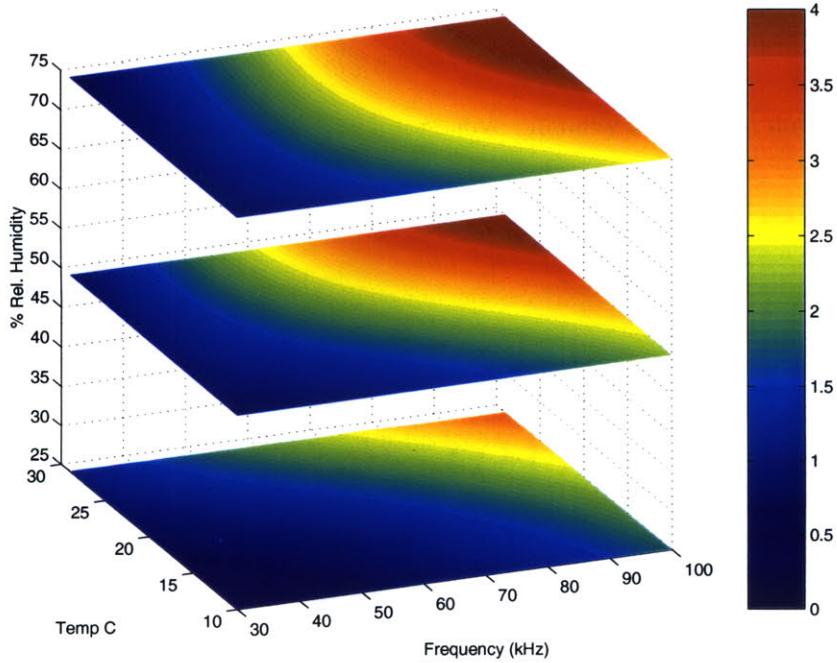


Figure 3.18: Acoustic absorption rates (dB/m) at 1 atm, for various temperatures and humidities.

3.5.1 Performance effects

Absorption can be expected to affect the performance of the parametric array in three primary ways. First, the rate of ultrasonic absorption directly impacts the audible level of the produced sound. To a first approximation for the farfield, audible level is inversely proportional to α . Second, because the ultrasonic signal usually contains significant bandwidth for the reproduction of audible sounds, any significant alteration in this set of harmonics could lead to increased distortion of the audio signal. This effect is most pronounced when ultrasonic absorption changes rapidly with frequency in the vicinity of the carrier. Third, the virtual array length, and therefore the farfield directivity, is directly related to the rate of ultrasonic absorption. This will be described in more detail in the next chapter.

Audible level

From the demodulation equation (Eq 3.1), the farfield audible level is inversely proportional to α . Therefore, in order to improve the efficiency of a parametric array, the rate of ultrasonic absorption should be minimized, which indicates that, in general, a lower-frequency ultrasonic carrier is preferable.

However, the effects of absorption are not necessarily significant; the change in audible level when using a carrier of 40 kHz versus 60 kHz is only about 3 dB – this translates to an additional requirement of only 1.5 dB of ultrasound to reproduce the same amount of audio. At much higher carrier frequencies, or in extremes of temperatures and/or humidities, the effect can be more pronounced. This generally limits the usefulness of airborne parametric arrays to ultrasonic frequencies less than 100 kHz.

Distortion

The effect on audible distortion can be evaluated by relating absorption effects to an effective change in the ultrasonic bandwidth of the system. Assuming that these effects will be most pronounced within 1m of the system source, the change in effective bandwidth can be estimated by through the coefficient α in dB/m.

The estimated change in bandwidth for three different carrier frequencies is plotted in 3.19. From this graph, it is clear the resulting change in absorption is very small, on the order of 1 dB. Since this error is far smaller than those imposed by limited transducer bandwidths, absorption's impact on audible distortion is minimal compared to other possible sources of distortion.

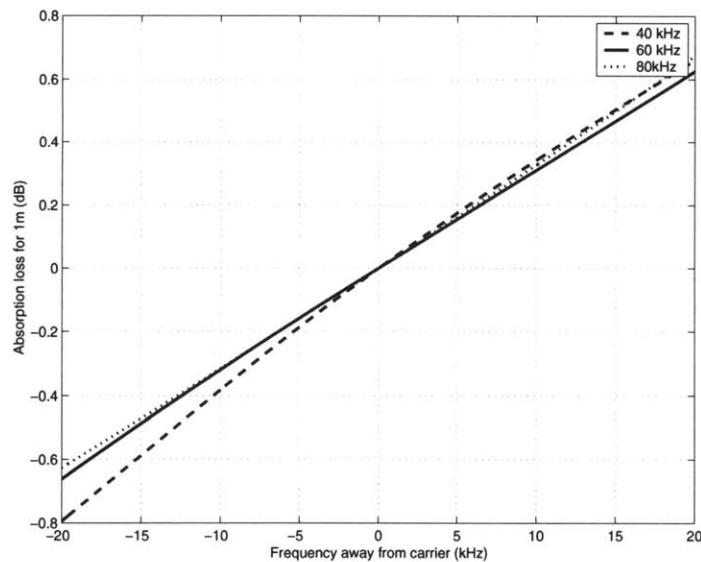


Figure 3.19: The estimated total absorption, leading to a change in ultrasonic bandwidth, is shown for three different carrier frequencies. In all cases the accumulated error is less than 1 dB, which indicates that this effect will not contribute significantly to audible distortion.

Chapter 4

Geometric Characteristics

The controlled, beam-like distribution of audible sound is the most striking feature of the airborne parametric array, but the precise geometry has not historically been well understood. Earlier researchers in the field typically limited themselves to one-dimensional plots, such as axial measurements [15] or polar plots [18]. While these methods do provide useful information about particular regions of the sound field, a complete understanding can be gained by examining the full acoustic field in two (or three) dimensions.

In this chapter, the methods for modeling the geometric characteristics of both the full ultrasonic and audible fields are presented, and compared with experimental results.

4.1 Ultrasonic Field

Under the quasilinear approximation, the ultrasonic field can be modelled by solving the linear wave propagation equation presented in Chapter 3 (Equation 3.1):

$$\frac{\partial^2 p_1}{\partial z \partial \tau} = \frac{c_0}{2} \nabla_r^2 p_1 + \frac{\delta}{2c_0^3} \frac{\partial^3 p}{\partial \tau^3}. \quad (4.1)$$

The boundary condition (source) in this case is a planar piston source of radius a . Rather than use the absorption model from the KZK equation, an exponential decay is used, as in the previous chapter.

There are two main regions of interest of the ultrasonic field. The near-field of the ultrasonic source exhibits strong self-interference effects from interfering emissions from different locations on the source. The far-field behavior, as it is sufficiently distant from the source, is a more well-behaved and uniform field.

The point of transition between the near field and far field is generally identified as occurring just after the last axial maximum of self-interference [37]:

$$z_{\text{far}} > \frac{a^2}{\lambda} - \frac{\lambda}{4}. \quad (4.2)$$

where a is the radius of the source, and λ is the wavelength of the transmitted ultrasound. A graph showing the relative lengths of the near and far-fields is shown in Figure 4.1.

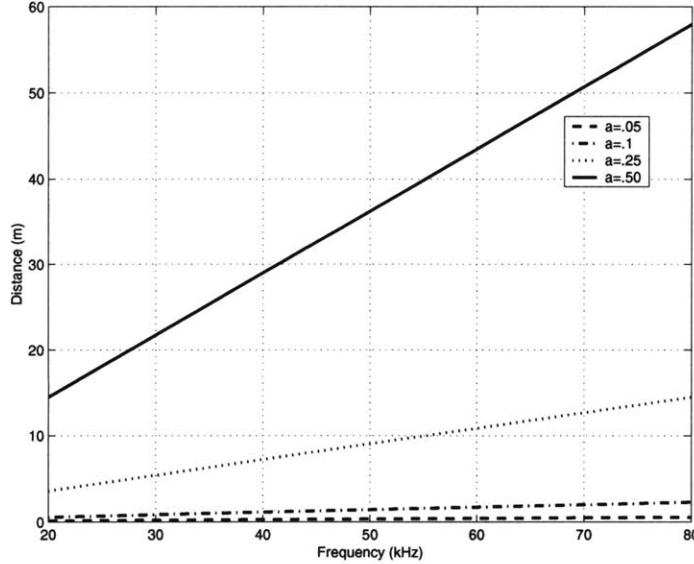


Figure 4.1: Transition point between the near and far fields as a function of frequency for various transducer radii a . Note that the near field becomes quite large for high frequencies and large transducers.

The graph shows quite clearly that for larger transducer sizes, near field effects can be present even over very long distances.

4.1.1 Far-field

Because the far field (by definition) does not contain interference features, a directivity function $D(\theta)$ can be defined as the ratio of sound pressure amplitude along direction θ to the on-axis pressure (at $\theta = 0$). A straightforward integration [37] provides a closed-form solution:

$$D(\theta) = \frac{2J_1(ka \sin \theta)}{ka \sin \theta} \quad (4.3)$$

where J_1 is the Bessel function of the first kind, and $k = \frac{2\pi}{\lambda}$ is the wave number.

The beam directivity angle can be defined as the angle at which half of the sound pressure exists, i.e., $D(\theta) = 0.5$. By solving this equation numerically, the ultrasonic beam angles for various transducer sizes as a function of frequency can be predicted. These are shown in Figure 4.2.

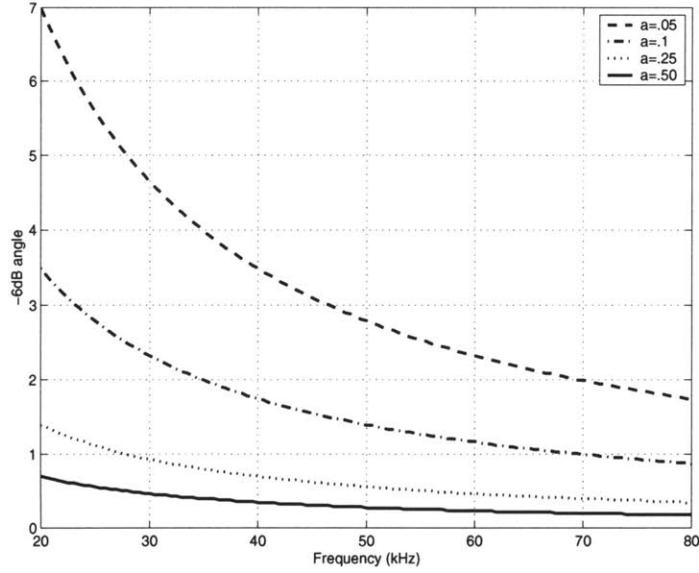


Figure 4.2: Farfield beam angles for ultrasonic sources of various radii a are shown as a function of frequency. The beam angle is defined as the angle at which the main lobe is at half amplitude (-6 dB).

4.1.2 Near field

Because of the presence of complex interference patterns, the nearfield of an acoustic source must be calculated through direct numerical integration of Eq. 4.1. Because the wavelengths of the ultrasound are much smaller than the geometry of the system, a very fine calculation grid is required, which would ordinarily require substantial calculation time. By taking advantage of the circular symmetry, and applying the transform described in [38], direct field response can be obtained much more efficiently.

Once the linear, lossless field $\tilde{p}(z, r)$ is calculated, the effects of absorption can be added with the following approximation:

$$p(z, r) = \tilde{p}(z, r)e^{-\alpha\bar{z}} \quad (4.4)$$

where \bar{z} is the average of the axial distance z and the distance from the observation point z to the edge of the transducer. This approximation holds as long as the transducer is smaller than the distance over which absorption significantly affects the ultrasonic level. The absorption factor α is calculated from equation 3.5.

The ultrasonic fields at 40 kHz, 60 kHz, and 80 kHz, including the effects of absorption (20C @ 50% rH), are generated by a 0.5m radius transducer are shown in Figure 4.3.

The 40 kHz beam appears to diverge somewhat more than the 60 kHz or 80 kHz beam, but absorbs into the air at a slower rate. There are also complex interference patterns present in the near field of the source, which are expected to change rapidly with wavelength. Because these features are much smaller than the audible wavelengths they generate, the fine features are not expected to substantially impact the audible field.

Repeating this analysis at a relative humidity of 30% (Figure 4.4) shows a diminished absorption effect at higher frequencies, particularly at 80 kHz. Other conditions of temperature and humidity will affect the absorption rate of the ultrasound, as shown in the previous chapter.

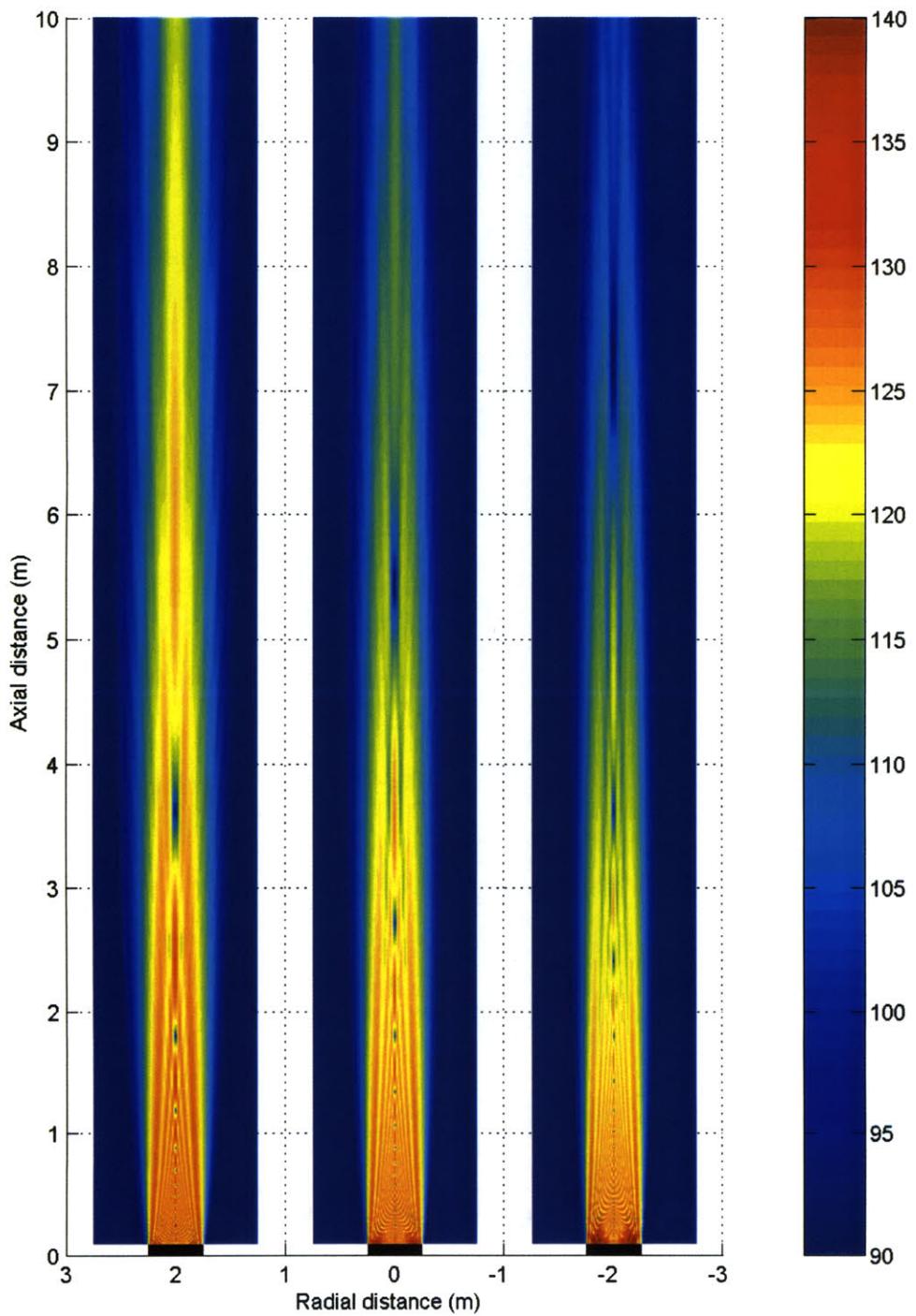


Figure 4.3: The ultrasonic fields, including absorption effects, from a 0.5m diameter transducer transmitting 40 kHz, 60 kHz, and 80 kHz, is shown. The transducer in each case is represented by the black rectangle. Absorption was calculated for 20C, 1 atm, and 50% rh.

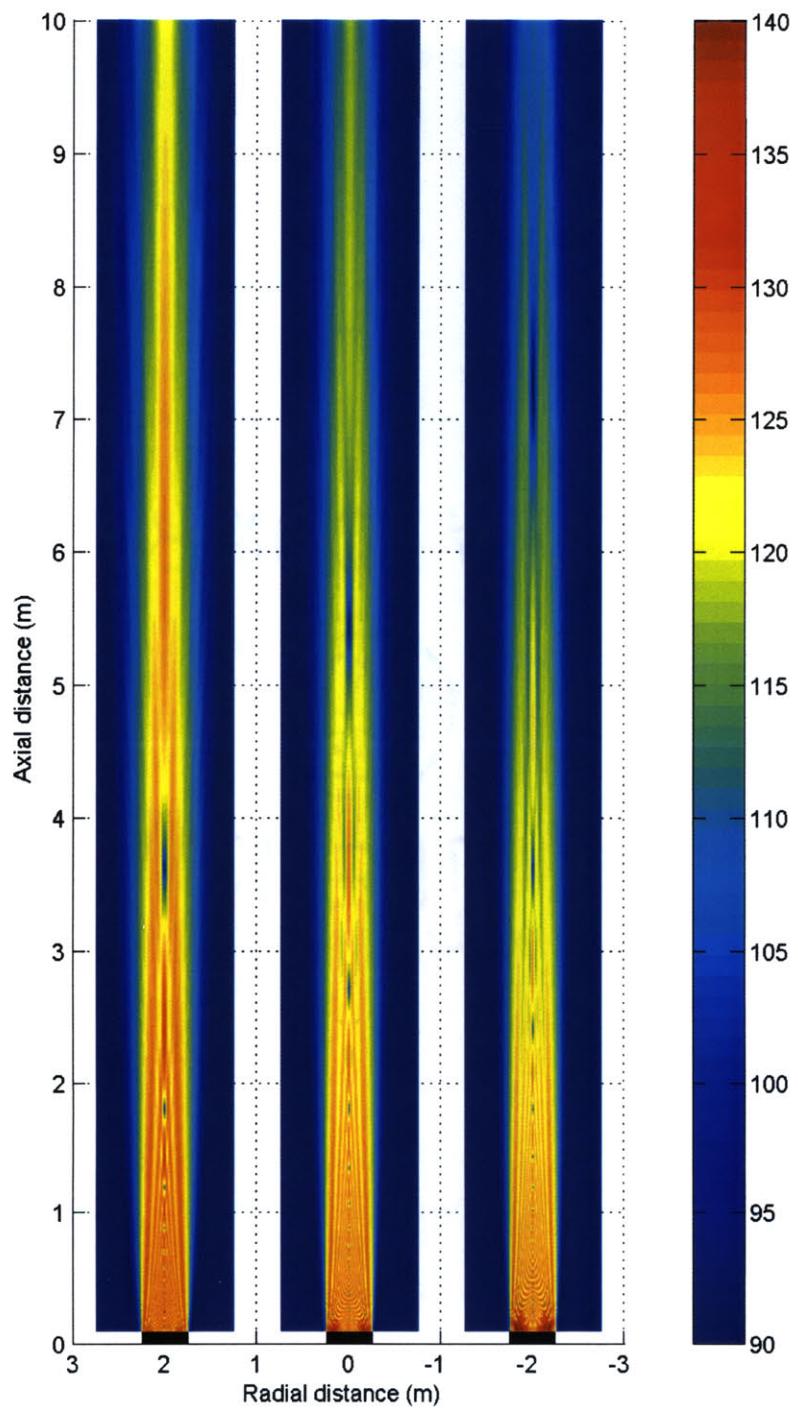


Figure 4.4: The ultrasonic fields, including reduced absorption effects, from a 0.5m diameter transducer transmitting 40 kHz, 60 kHz, and 80 kHz, is shown. The transducer in each case is represented by the black rectangle. Absorption was calculated for 20C, 1 atm, and 30% rh.

4.2 Audible Field

Using the quasilinear approximation presented earlier, the audible sound field can be predicted with reasonable accuracy, and without resorting to full-field numerical calculations. Recall that the ultrasonic field is now considered the source of audible sound, so that a solution for the audible field can be found by integrating the contributions of each point of ultrasound.

The geometry used in this analysis is shown in Figure 4.5.

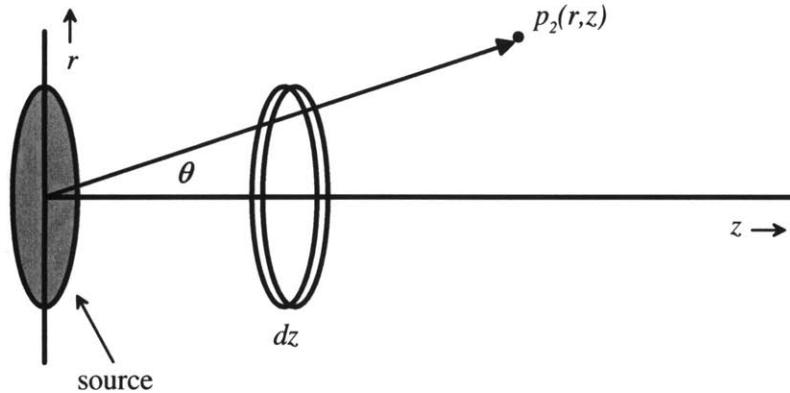


Figure 4.5: The geometry of the system used for directivity modeling.

4.2.1 Far field

The most common, and straightforward methods for predicting the far-field directivity of the demodulated beam, with their derivations, are described in detail in the literature [1, 7, 39]. Generally, the results are taken in the far-field of the virtual source, i.e., when the ultrasound has been substantially absorbed into the air ($z \gg 1/\alpha$), and for small angles θ .

From Westervelt's original analysis [1], the directivity function is:

$$D_W(\theta) = \frac{1}{1 + i(2k/\alpha) \sin^2(\theta/2)} \quad (4.5)$$

where k is the wavenumber of the demodulated signal, and α is the absorption rate of the ultrasound. A more precise solution includes an aperture factor [39] equal to the farfield directivity of a piston source (Eq. 4.1.1).

The farfield directivity of a 43cm diameter parametric source was calculated using a carrier frequency of 60 kHz ($\alpha = 0.23$), and is plotted in Figure 4.6. The solid lines show the directivity of the parametric source, while the dashed lines show the directivity of a traditional piston source of the same dimensions.

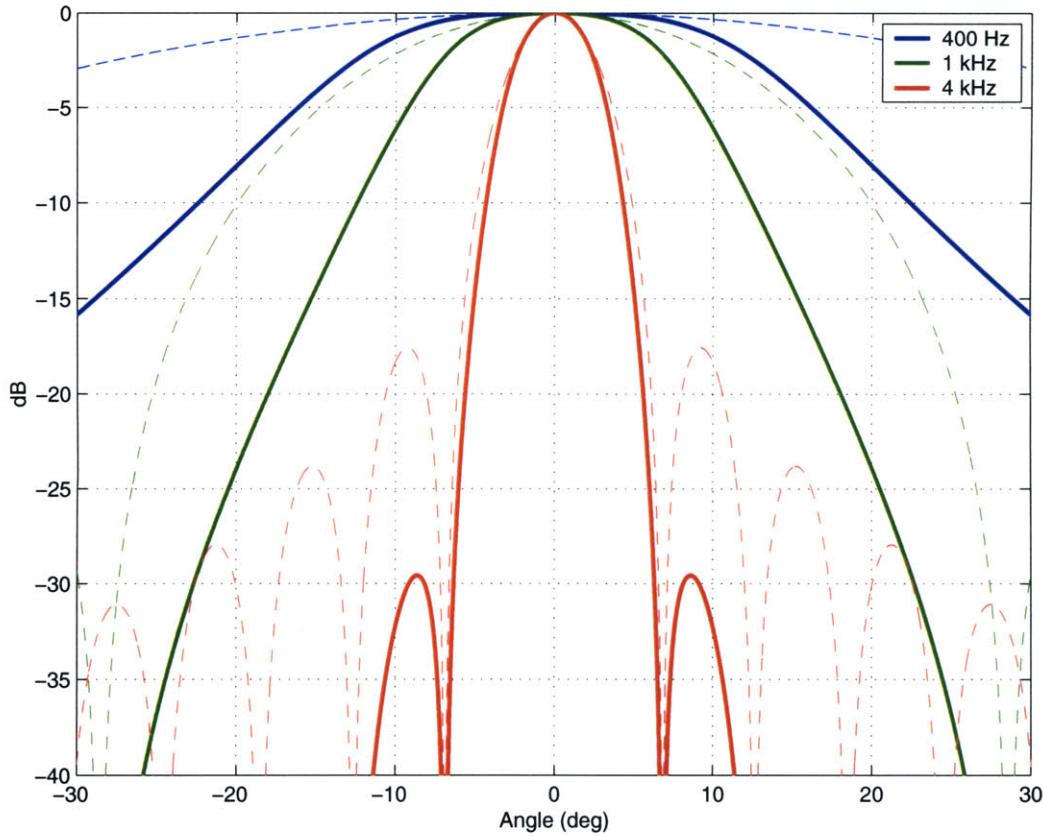


Figure 4.6: The far directivity of a parametric source (solid lines) versus a traditional source (dashed lines) is shown. Note that the change in directivity is much more pronounced at low frequencies. Sidelobes are also effectively suppressed.

For the analyzed frequencies, the change in directivity is substantial at lower frequencies (400 Hz and 1 kHz), but less so at a higher frequency (4 kHz), due to the reduced size of the audible wavelength. However, in all cases, sidelobes are greatly suppressed, a result of the parametric source behaving as an end-fire array.

The half-power angle can be calculated by solving $D_W(\theta) = 1/2$, or

$$\theta_{\frac{1}{2}} = 2 \sin^{-1} \left(\sqrt{\frac{\alpha}{2k}} \right) . \quad (4.6)$$

The half-power angle as a function of carrier frequency is shown in Figure 4.7.

Over the range of carrier frequencies considered, the farfield directivity appears to change very little. This is primarily due to the limited change of ultrasonic absorption over the octave of ultrasound considered.

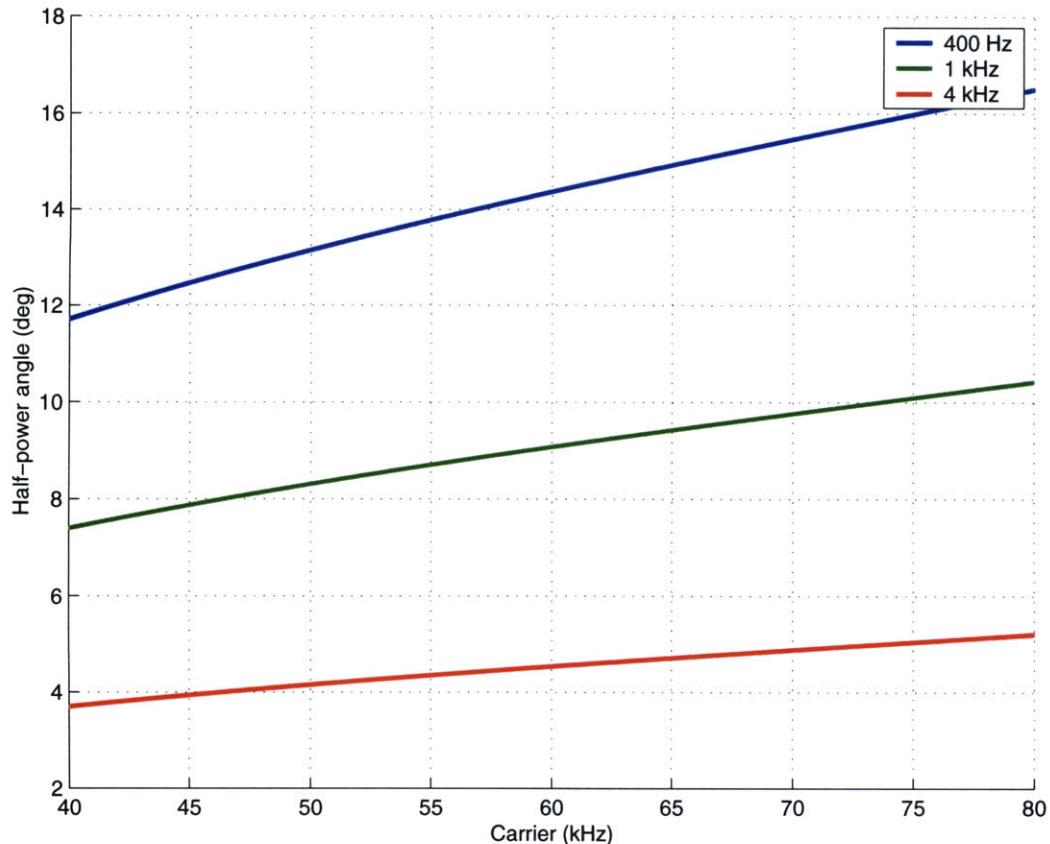


Figure 4.7: The far field half-power angle of a parametric source at various audible frequencies, as a function of carrier frequency. The directivity is reasonably consistent over practical ranges of carrier frequencies.

4.2.2 Near field

From the calculations in Figure 4.1, the near field of the ultrasonic source can be quite large – on the order of 10-20m. The calculated absorption rates of ultrasound show that the effective virtual array length (α^{-1}) is on the order of 4-7m. Thus, most practical ranges of interest fall well within the near field region of the device, where the far-field approximation in the previous section does not apply.

Because of the complexity of the governing equations, nearfield solutions have generally been performed through numerical simulations [27,40], particularly finite-difference techniques. While full-field analysis generally requires substantial computational effort, a reasonable approximation may be obtained by using the quasilinear, virtual-source model as was performed for the time domain analysis.

To calculate the distribution of audible sound, the ultrasonic field is first calculated from the standard linear wave equation, and the audible field then results by using the ultrasonic field as the ‘source’ for the audible frequency field.

Since the previous section showed that the ultrasonic beam is well collimated, the ultrasonic field p_1 can be written as:

$$p_1(z, r, \tau) = p_0 e^{-\alpha z} E(\tau) H(a - r) \quad (4.7)$$

where $H(r)$ is the Heaviside function, and $\tau = t - z/c_0$ is lag-time.

If the ultrasonic field is envisioned as being the ‘source’ of audible sound, a reasonable approximation can be calculated by numerically integrating the field due to a virtual source, and convolving this result over two-dimensional space of ultrasound. The analysis begins with an adaptation of the modeling equations provided in [7] for the field strength q :

$$q(r, z) \propto \int_0^z \int_0^\infty q_0^2(r', z') G(r, z|r', z') r' dr' dz' \quad (4.8)$$

where $q_0(r, z)$ is the strength of the ultrasonic field, and G is the Green’s function for the Helmholtz equation solution shown below:

$$G(r, z|r', z') = \frac{ik}{2\pi(z - z')} J_0\left(\frac{krr'}{z - z'}\right) e^{\left(-\alpha_L(z - z') - \frac{ik(r^2 - r'^2)}{2(z - z')}\right)}. \quad (4.9)$$

In this equation, k and α_L are the wavenumber and absorption coefficient, respectively, if the audible sound. For these calculations the absorption of audible sound can be neglected ($\alpha_L \approx 0$).

Substituting this into the integral, the audible field strength is:

$$q(r, z) \propto \int_0^z \int_0^\infty q_0^2(r', z') \frac{ikr'}{2\pi(z - z')} J_0\left(\frac{krr'}{z - z'}\right) e^{\frac{-ik(r^2 - r'^2)}{2(z - z')}} dr' dz'. \quad (4.10)$$

Since all terms containing z appear as $z - z'$, the equation can be solved by convolving the ultrasonic source strength $q(r, z)$ with the Green’s function G . The Green’s

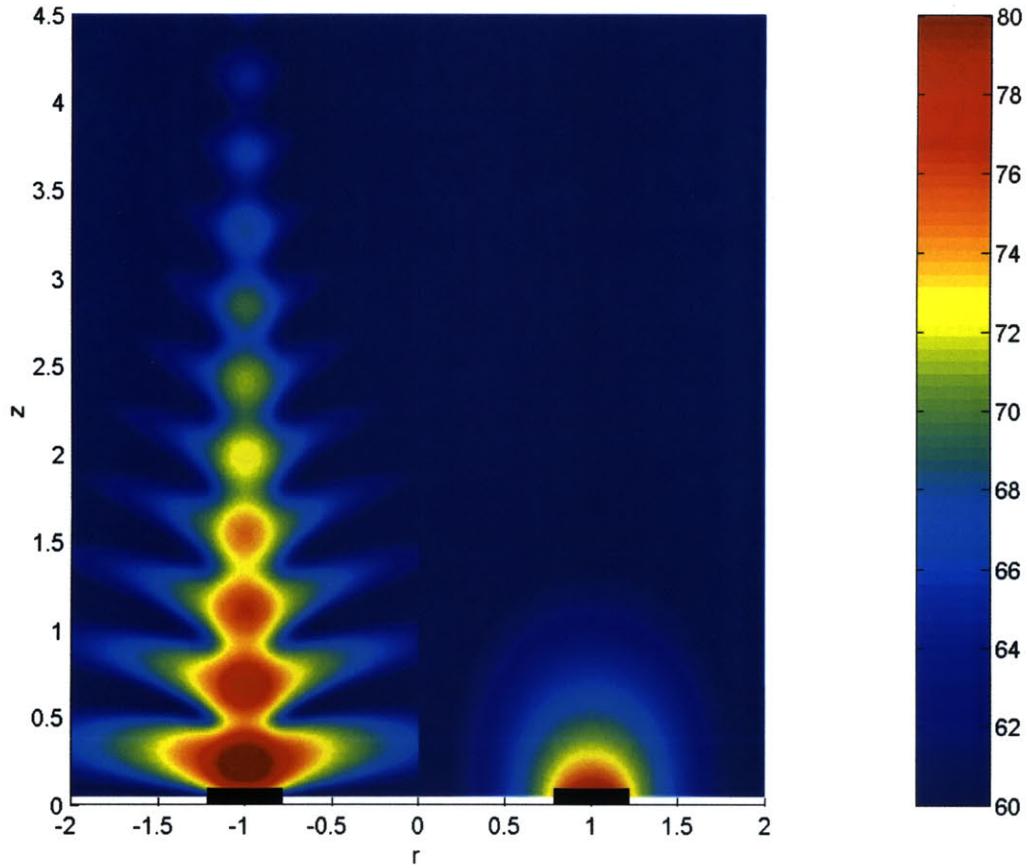


Figure 4.8: Simulated parametrically generated audible field of 400 Hz, compared to a traditional source of equal size.

function was calculated through straightforward numerical integration with a step size of .02m, and over a distance of 6m.

To make the convolution straightforward, the ultrasonic source is assumed to be perfectly collimated, decaying exponentially due to absorption, and is expressed by:

$$q_0(r, z) = p_0 e^{-\alpha z - ikz} H(a - r) . \quad (4.11)$$

The resulting simulated fields are shown in the following figures.

The predicted 400 Hz field has a peculiar energy pattern, and appears to naturally narrow in directivity as distance increases. This narrowing is due to the diminished contributions of ultrasound with greater axial distance. The fringes are caused by constructive and destructive interference between the ‘virtual’ sources.

The beam-like characteristics of the 1 kHz source are very apparent in this plot. Again, interference fringes are observed in the nearfield, but they are spatially quite small.

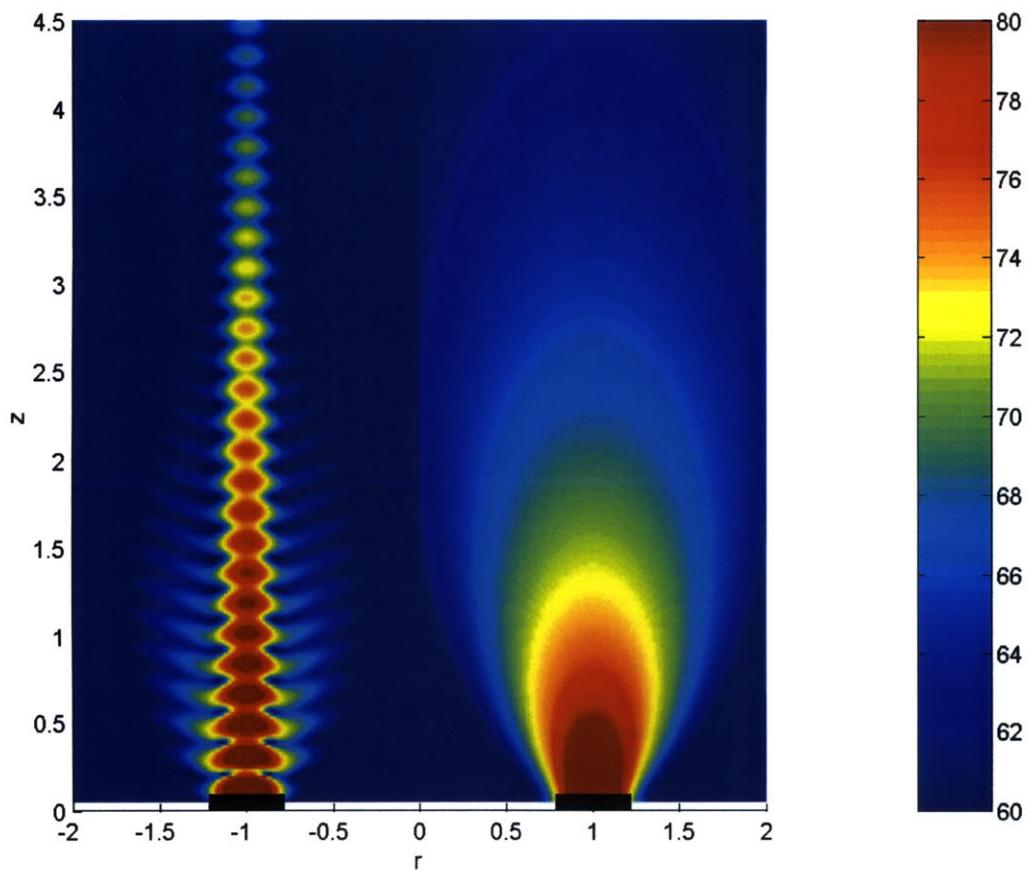


Figure 4.9: Simulated parametrically generated audible field of 1 kHz, compared to a traditional source of equal size.

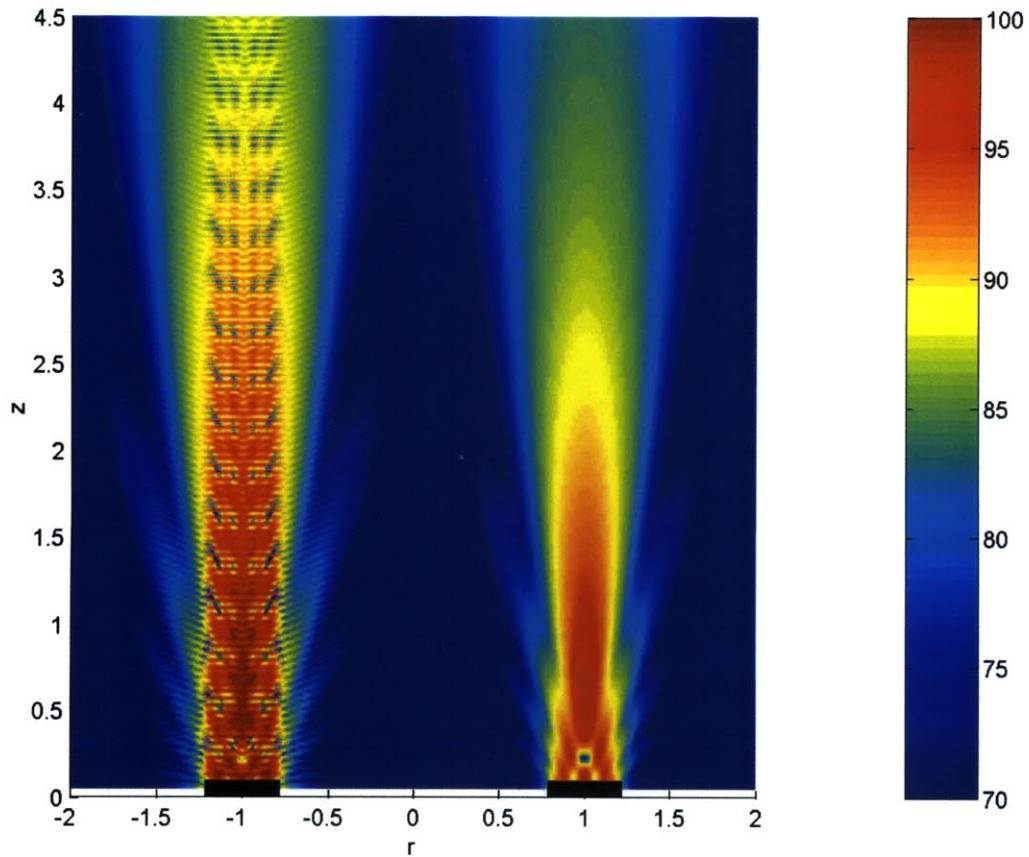


Figure 4.10: Simulated parametrically generated audible field of 4 kHz, compared to a traditional source of equal size.

The beam width narrows with distance, and appears to asymptotically approach a constant width of about half the transducer diameter.

At 4 kHz, the audible wavelengths (8.5cm) are much smaller than the diameter of the transducer, so the difference between a traditional source and a parametric source is less distinct. However, notice that the parametrically generated field is far longer, and more evenly distributed than the traditionally generated field. The sidelobes are also suppressed, compared to the main lobe.

4.3 Experimental Results

To verify the accuracy of these models, a test system was constructed using a 40cm circular transducer (to ensure symmetry) and appropriate amplification electronics. Processed ultrasonic signals were generated using the algorithms shown in Chapter 3, using a carrier of 60 kHz and a modulation depth of approximately 80%.

The system was measured by audible-range (B&K 4190) and ultrasonic (B&K 4138) microphones along the center plane of the transducer in a substantially-anechoic room. The microphone signal was continuously monitored and analyzed by a nearby PC while the microphone itself was manually moved about the room. At the commencement of each measurement cycle, the microphone position was logged using an automated tracking apparatus. The spatially correlated results were plotted in a two-dimensional grid, spatially interpolated, and colored according to sound pressure level.

4.3.1 Ultrasonic Field

The measured ultrasonic field for a nominal output level is shown in Figure 4.11. The carrier frequency used was 60 kHz. This particular ultrasonic field was used to create a pure 4 kHz tone, but it is expected the ultrasonic field will not change significantly for other stimuli.

As expected, the measured distribution of ultrasound is substantially well collimated, with only small deviations and fine features due to self-interference. The fine structure seen in the simulation could not be accurately measured, due to spatial undersampling.

4.3.2 Audible Field

Pure tone stimuli processed using the methods given in Chapter 3 were modulated and sent to the test apparatus, and the full field was measured at audible tones 400 Hz, 1 kHz, and 4 kHz. Each resulting field of audible sound is plotted next to the field of an equivalent traditional source of same physical dimensions.

The audible field at 400 Hz is shown in Figure 4.12. On the left is the actual 400 Hz measurements, and on the right is a traditional source of same physical dimensions. There is an apparent strong collimation of sound in front of the transducer, with sound levels nominally around 65 dB, and divergence almost nonexistent. In the very near field, an interesting interference field of much higher levels (> 70 dB) is observed. It is possible that some of these features could be caused by microphone nonlinearities, but these small, localized ‘pockets’ of audible sound can be readily heard by a human listener positioned at precise locations. It is likely more of this structure exists, but is not visible due to the limited spatial resolution of the measurements.

Further from the source, a secondary audible field can be observed in light blue and green (50-60 dB), which shows the diffraction of the main beam into off-axis regions. Further away, minor interference fringes are observed (3-4m), which are the result of re-

flections from the (imperfect) sound absorbing material on the wall of the measurement chamber.

The strong nearfield features predicted by the model in Figure 4.8 are not readily apparent, although there is some hint of them in the 1-3m range. This is likely due to spatial undersampling of the measured field, or, more likely, the departure of the nearfield of the ultrasonic source from a perfectly collimated planar wave. Any divergence from the planar wave will likely diminish features caused by off-axis constructive interference.

The audible field at 1 kHz is shown in Figure 4.13. A similar nearfield structure is observed to the 400 Hz field, but it extends much farther. A main, teardrop-shaped lobe extends for approximately 3.5m, after which diffraction appears to widen the field. Note that within a few meters from the source the beam is again almost perfectly collimated. A lateral movement of 0.5m can cause a 15-20 dB drop in level. Close to the transducer, regular fringes corresponding to those appearing in the simulation are noted, although they are obscured by measurement limitations. Again, minor interference fringes (appearing obliquely) are observed at further distances due to the reflection from the rear wall.

The directivity of a traditional source of this size at 4 kHz is fairly strong, but the sound field is much more evenly distributed in the parametric source. The audible field is very uniform for most distances shown, and divergence of the beam is very slight. In addition, sidelobes appear to be nonexistent.

4.3.3 Distortion Fields

Distortion levels, as indicated by the second harmonic of the audio signal, were measured simultaneously during audible field measurements. The distribution of audible distortion is shown in the following figures. In each case, distortion components appear relatively evenly distributed, and are as directional as the fundamental source.

For an audible frequency of 400 Hz (Figure 4.15, the second harmonic is distributed relatively evenly in space, at a level approximately 25 dB below the fundamental, corresponding to a harmonic distortion of less than 6% for most positions near the axis.

At 1 kHz (Figure 4.16, the field of distortion is also very well collimated, and evenly distributed with respect to distance, although it appears to decrease for further distances.

At 4 kHz (Figure 4.17, the distortion component is likewise highly directional and very well collimated. The amplitude of the harmonic is slightly higher than in the other cases, but it still reasonably low relative to the primary for most positions.

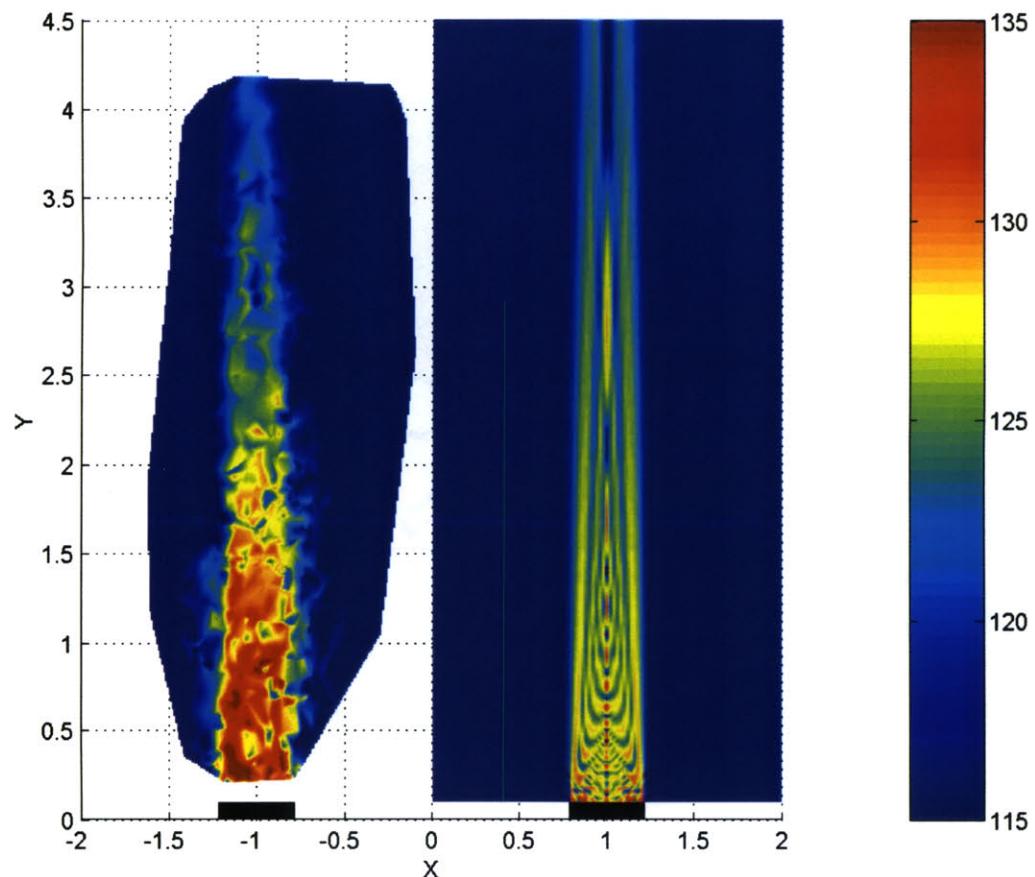


Figure 4.11: Ultrasonic field to generate 4 kHz tone (left), compared to the simulated ultrasonic field of the carrier alone (right). While it appears well collimated, there is a more irregular structure due to the increased complexity of the ultrasonic signal, particularly its bandwidth, as well as undersampling in space.

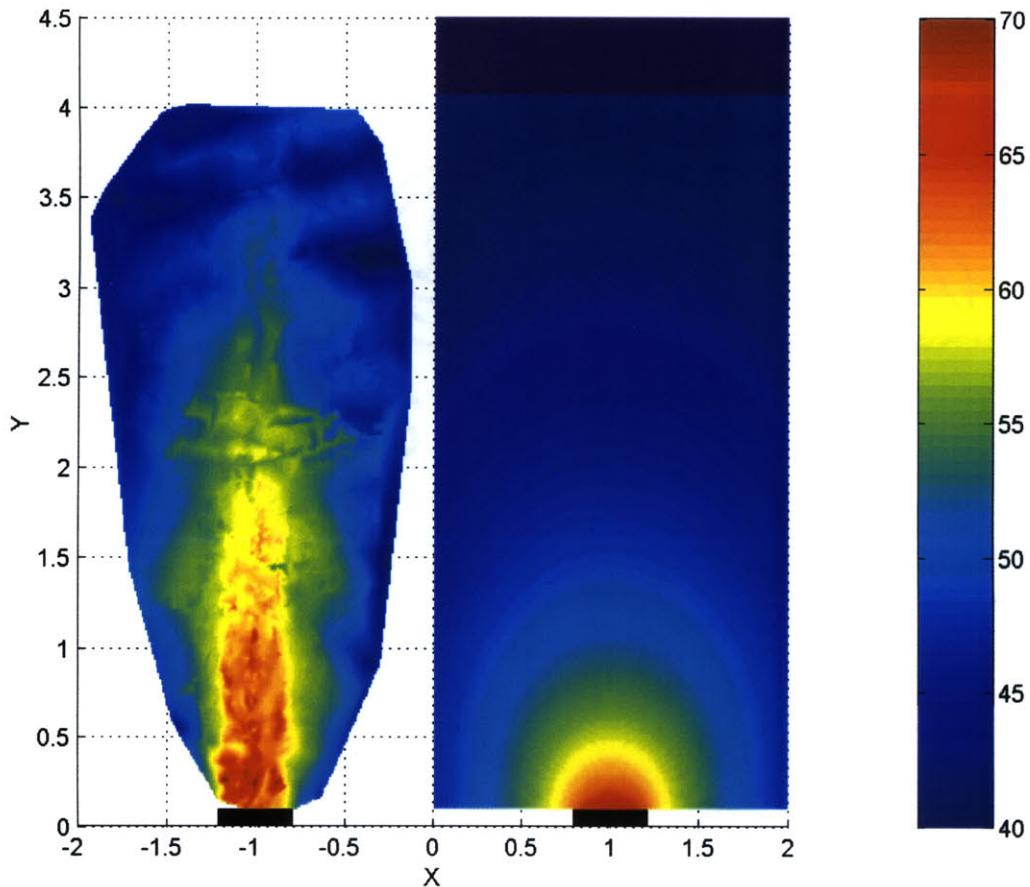


Figure 4.12: The measured audible sound field of a parametric source for a 400 Hz signal is shown on the left, along with a linear source of equal size on the right. The distribution of audible sound with axial distance is even over a much longer distance, and there is a >12 dB drop for locations 0.5m off-axis. The transducer is indicated by the black rectangle. Spatial dimensions are meters, and the colors represent dB SPL.

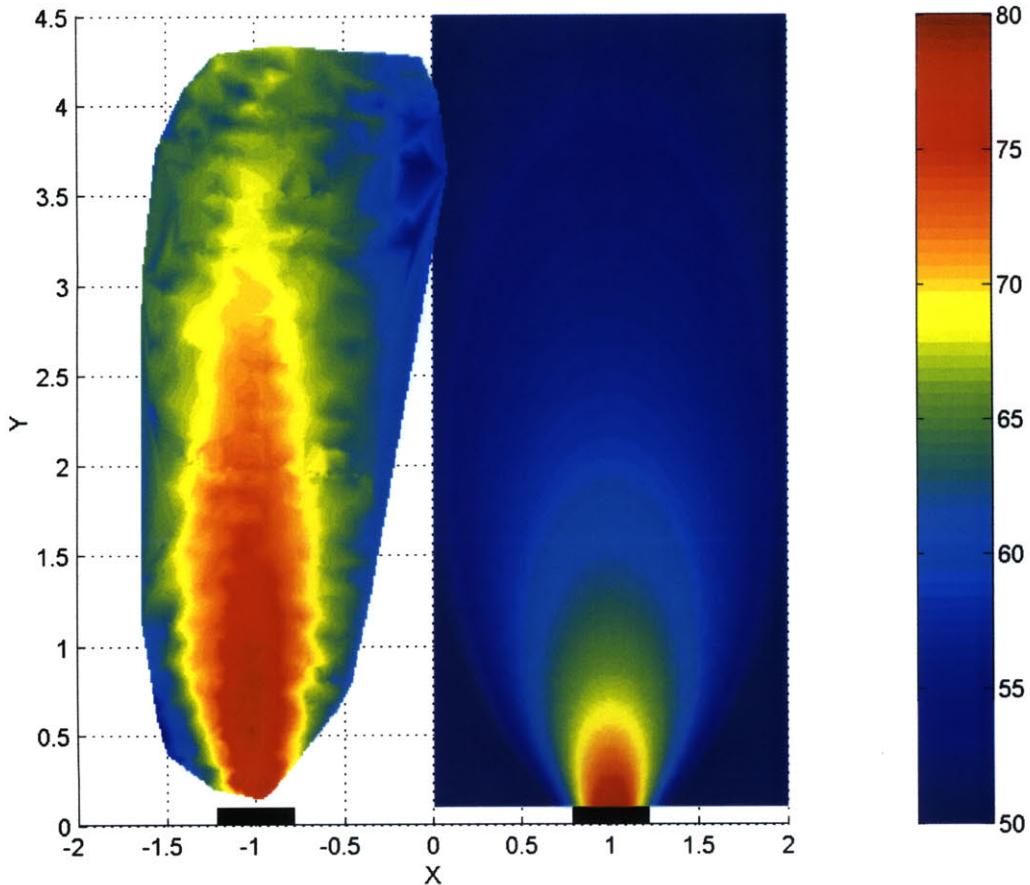


Figure 4.13: The measured audible sound field of a parametric source for a 1 kHz signal is shown on the left, along with a linear source of equal size on the right. The distribution of audible sound with axial distance is uniform over a long distance, and there is a >12 dB drop for locations 0.5m off-axis. The transducer is indicated by the black rectangle. Spatial dimensions are meters, and the colors represent dB SPL.

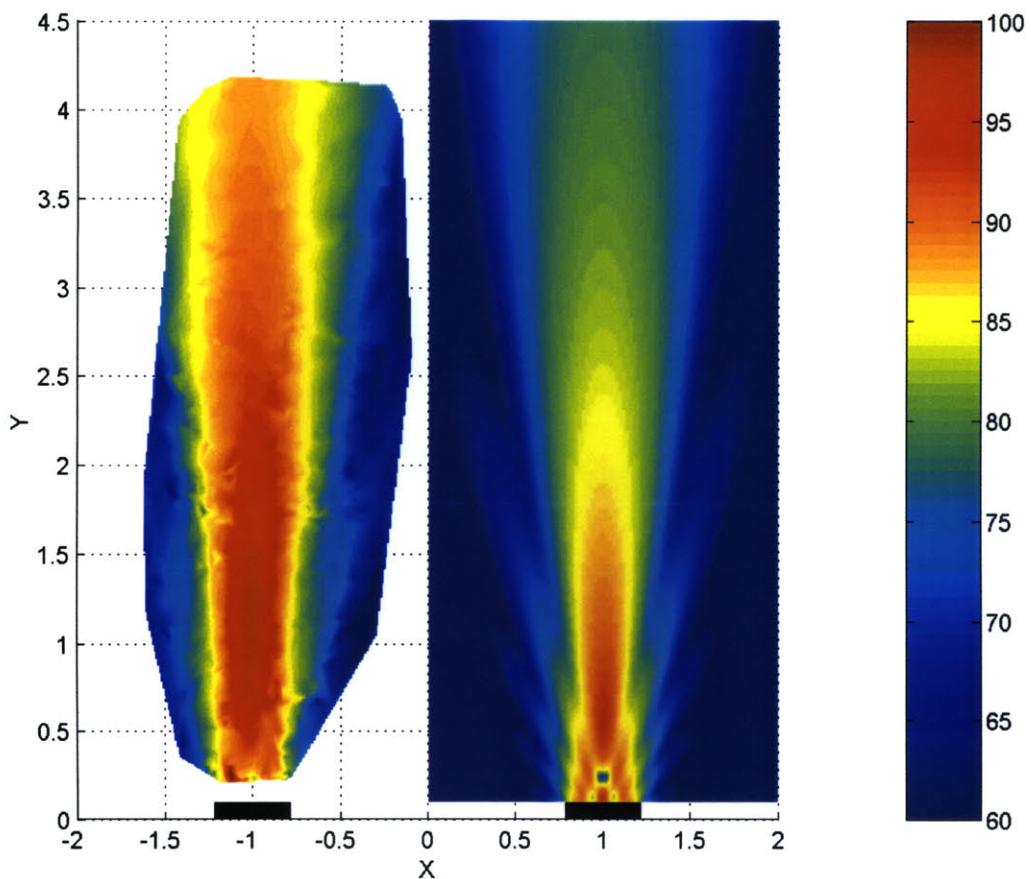


Figure 4.14: The measured audible sound field of a parametric source for a 4 kHz signal is shown on the left, along with a linear source of equal size on the right. The transducer is indicated by the black rectangle.

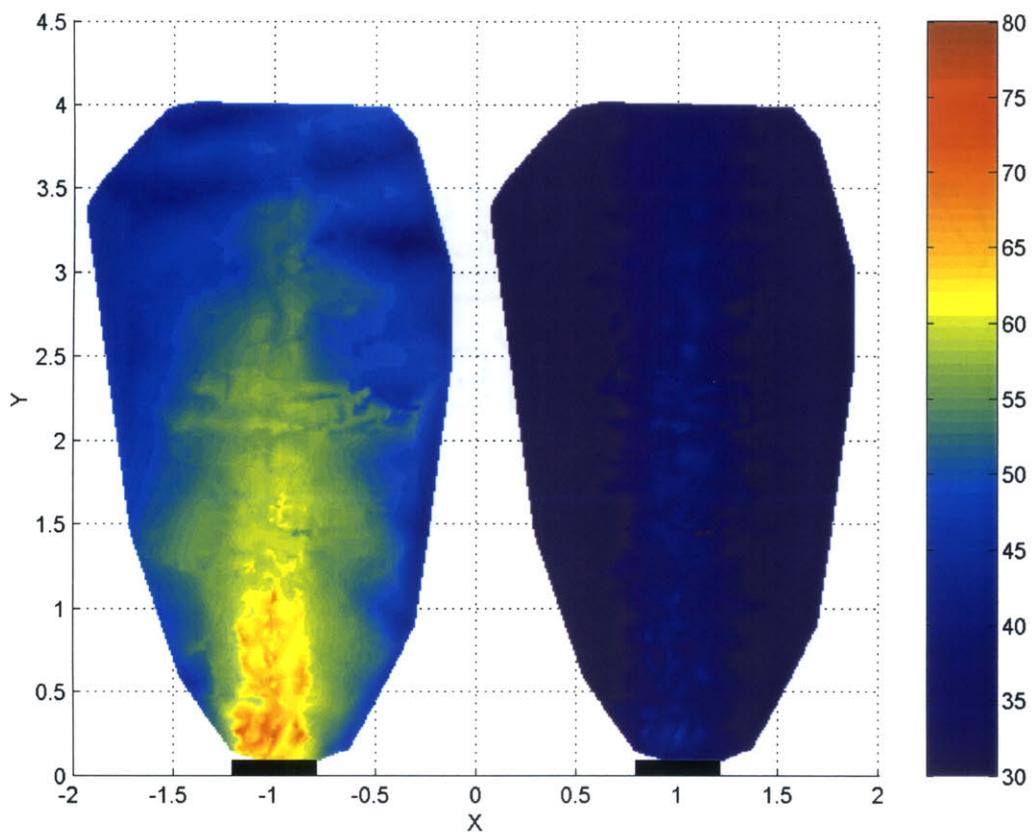


Figure 4.15: The measured audible sound field of a parametric source for a 400 Hz signal is shown on the left, along with the second harmonic (800 Hz) on the right. Distortion, in this case, is relatively evenly distributed about the audible field, and has slightly higher directivity.

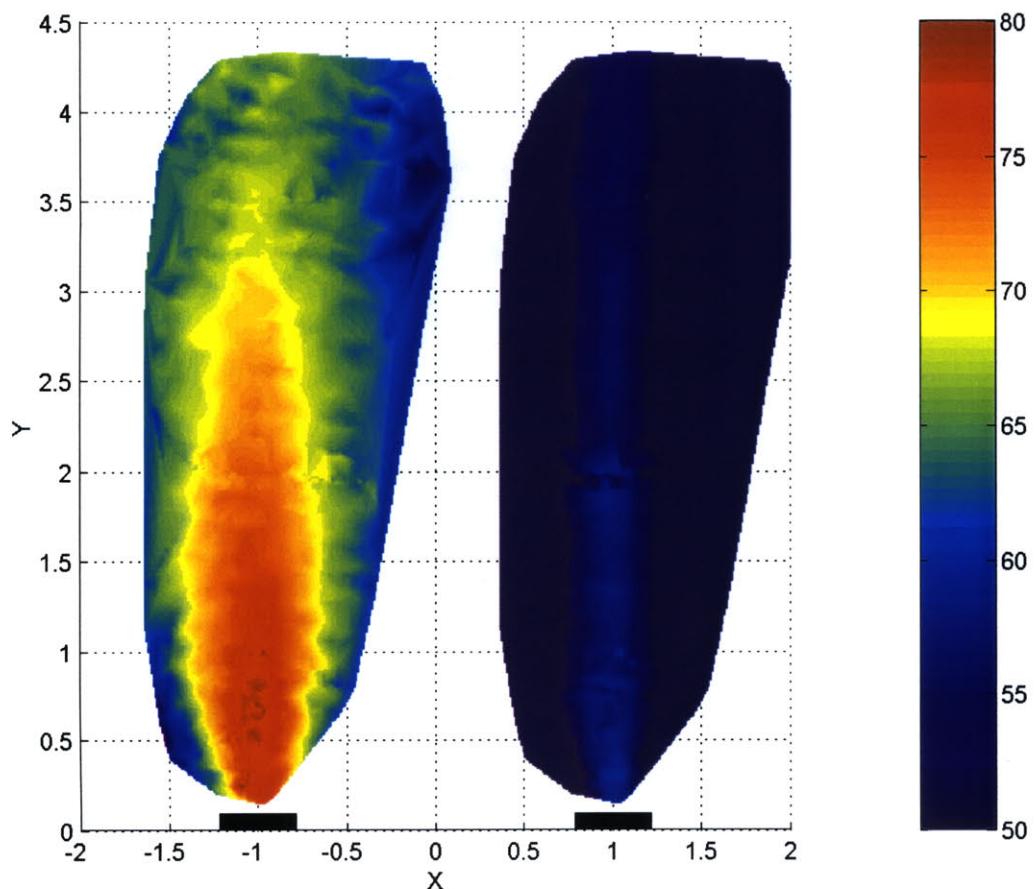


Figure 4.16: The measured audible sound field of a parametric source for a 1 kHz signal is shown on the left, along with the second harmonic (2 kHz) on the right. Distortion is also relatively evenly distributed about the audible field, and has slightly higher directivity.

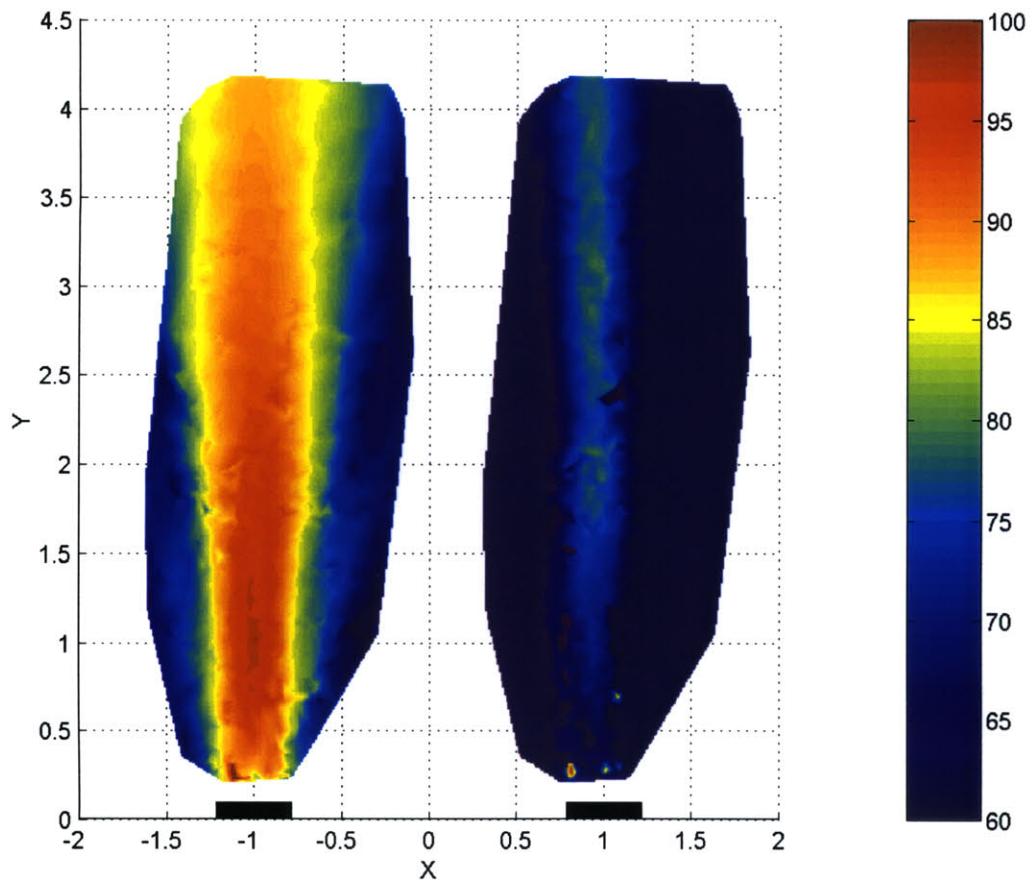


Figure 4.17: The measured audible sound field of a parametric source for a 4 kHz signal is shown on the left, along with the second harmonic (8 kHz) on the right.

Chapter 5

Designing Applications

Basic application of a highly directional acoustic source is reasonably straightforward. The lighting analogy (“Audio Spotlight™”) is very effective in providing the designer with an intuitive understanding of the behavior of the acoustic beam in general installations. Words like ‘shine’, ‘project’, and ‘shadow’, generally applied only to light, should be applied to these acoustic beams as well.

There are a few aspects of installation design which would benefit from a quantitative assessment. In particular, applications *acoustic isolation* and *sound projection* are addressed and carefully quantified through physical experimentation. Other system additions, such as subwoofer augmentation and the use of the precedence effect for auxiliary reinforcement are also considered in this chapter.

5.1 Dome-reflector loudspeakers

One popular contemporary method for creating localized sound fields is the dome-reflector loudspeaker. Because it has been widely applied (although not always successfully), the system merits attention.

This type of loudspeaker consists of a very large curved reflector (typically spherical or parabolic), and a small traditional loudspeaker mounted near its focal point. Commercial units are available from several manufacturers, such as Brown Innovations, SoundTube, and Museum Tools (Figure 5.1).

The general goal of these types of loudspeakers is to use a large curved reflector to focus sound waves directly to the vicinity of the listener’s ears, or, simply to increase the effective aperture of the loudspeaker. If the reflector is large enough, and the listener is positioned properly, the reflector/focuser should be effective.

To assess their efficacy, beam field plots were performed on a Brown Innovations’ “Localizer” speaker (far right of Figure 5.1), which consists of a transparent plastic hemispherical dome (76cm diameter), and a pair of small loudspeakers (presumably for stereo effects) placed within. The field measurements were taken identically to those in Chapter 4, but because the loudspeaker is not radially symmetric, measurements were performed



Figure 5.1: Various commercial ‘dome-style’ loudspeakers intended for localized audio reproduction. From left, loudspeakers manufactured by Museum Tools, SoundTube, and Brown Innovations.

approximately 25cm from the central axis. This was the manufacturer’s recommended position for a standing listener. While this is a stereo system, only one channel is used for measurements, as it is expected that both channels will be have similarly.

At 400 Hz (Figure 5.2), we see that the field created be the loudspeaker dome is reasonably well concentrated in the very near field (for distances less than 1m), but for further distances, audible sound is not well confined to the intended region. The complex pattern may be due to interference from different parts of the dome itself, or possibly small room reflections.

For 1 kHz sound waves, the isolation pattern is more distinct. The sound is not nearly as directive as the parametric source, but does have a distinct main lobe of a size consistent with a source of similar effective aperture. For distances greater than about 1.5m, however, the sound field is not well controlled.

At higher frequencies, such a large aperture is expected to cause natural beaming and sidelobes, as clearly seen in Figure 5.4. It is likely that at these frequencies, because the wavelengths are approaching the size of the originating loudspeaker (8.5cm), the speaker elements themselves are beginning to beam, resulting in only a portion of the reflector dish to be actively used.

In summary, it appears that the dome-reflector loudspeaker is reasonably effective in focusing sound to a small area directly in front of the dish opening, but its usable listening range is only about 1-2m. But as these are typically mounted overhead, the total distance to the floor may be up to 4m. By this point the sound has a tendency to spread significantly, which may lead to increased ‘spill’.

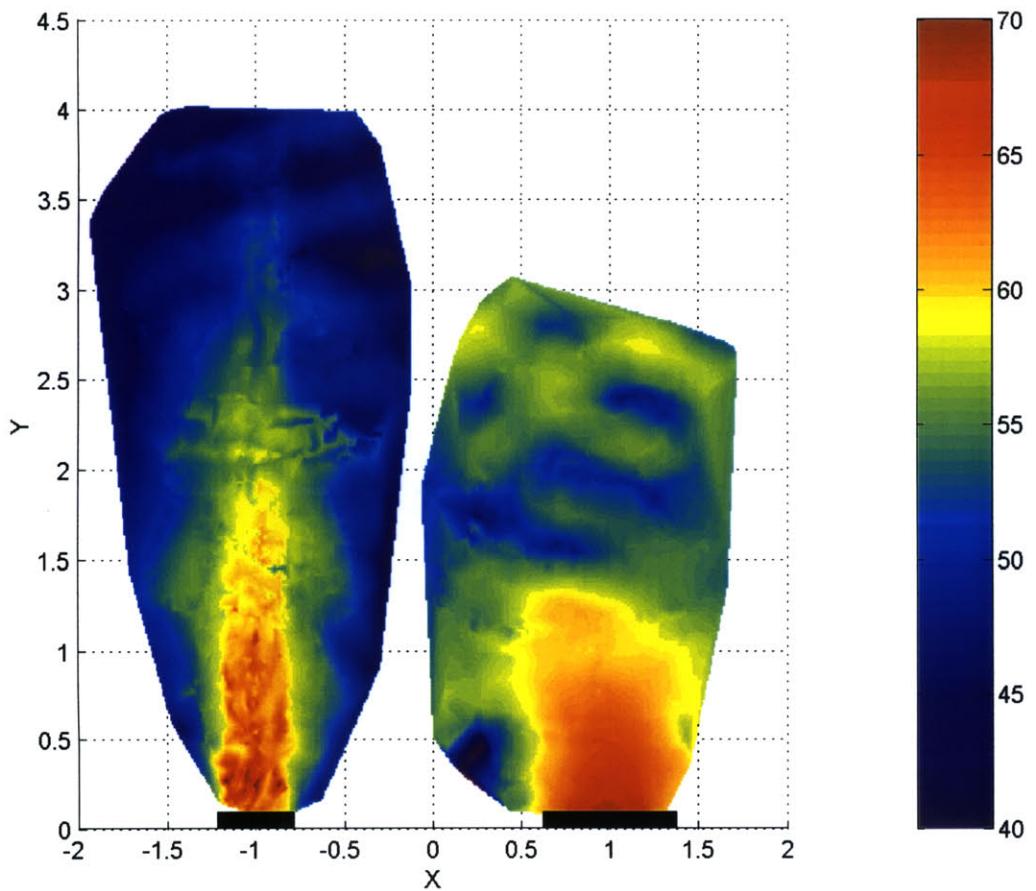


Figure 5.2: The sound fields from the parametric source (left) and dome-reflector loudspeaker (right) at 400 Hz.

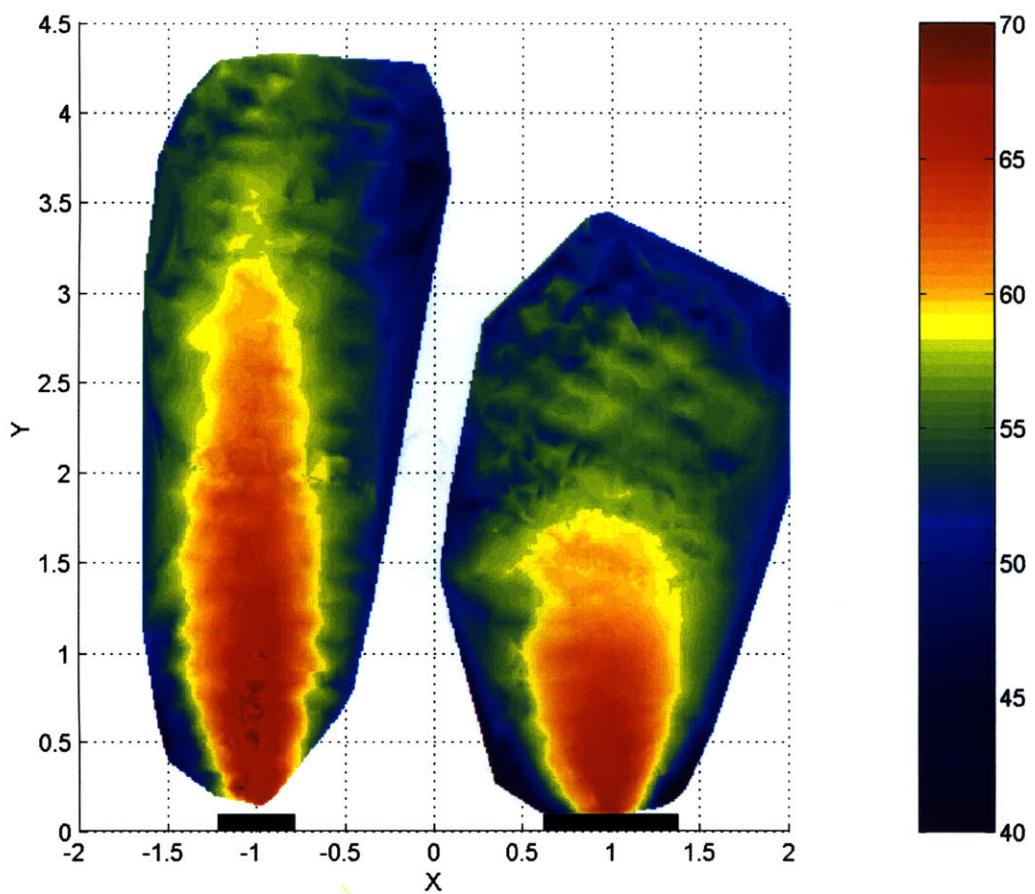


Figure 5.3: The sound fields from the parametric source (left) and dome-reflector loudspeaker (right) at 1 kHz.

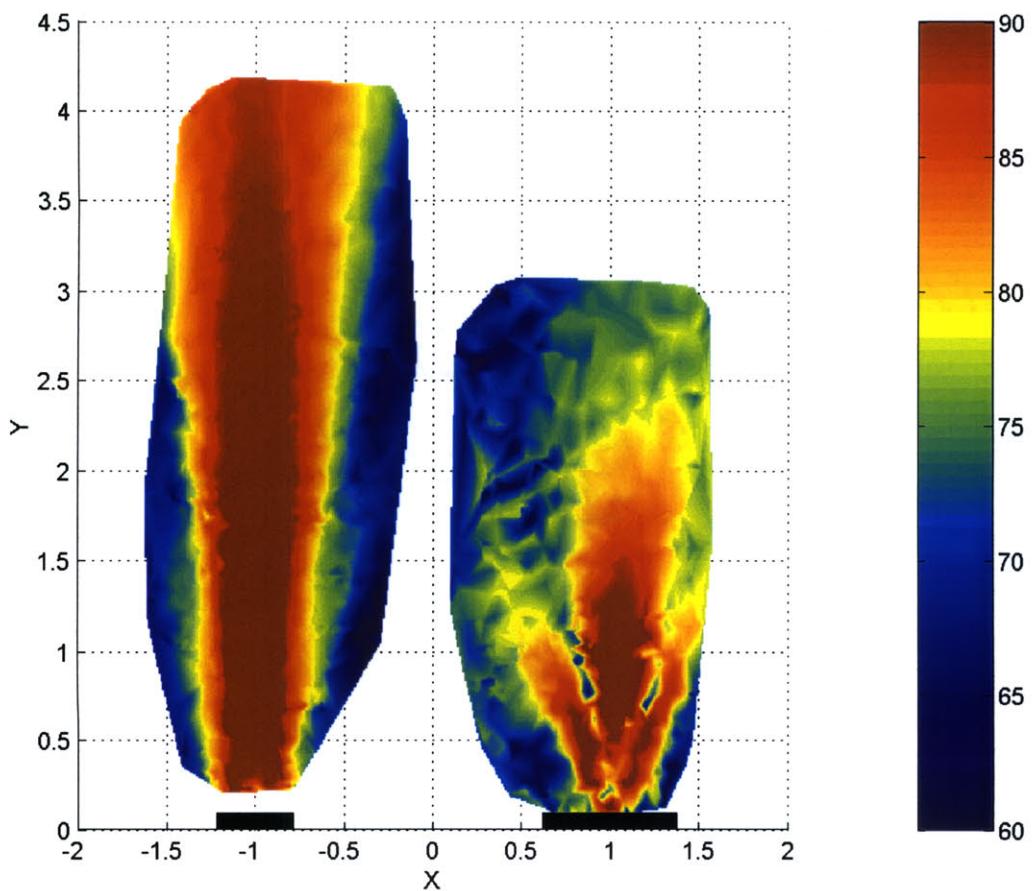


Figure 5.4: The sound fields from the parametric source (left) and dome-reflector loudspeaker (right) at 4 kHz.

5.2 Projecting sound

A novel and interesting application of an ultra-directive sound source is to ‘project’ the sound against an object or surface, much as one would project light from a movie projector or flashlight. Because directivity of the audible beam is maintained over long distances, a clear projection can take place.

5.2.1 Scattering

Just as with light, the absorptive and scattering properties of a projection surface will strongly influence the field of the reflected sound waves. A black wall will reflect very little light, a very smooth surface (mirror) will reflect light specularly, and an irregular surface (painted surface, or a movie screen) will reflect light in a diffuse manner.

Considerations for absorption of sound waves is straightforward – ‘shining’ sound against an acoustically absorbing target (foam, heavy curtains, etc.) will result in a relatively weak reflection, and a poor projected image. Shining sound against a highly reflecting surface (most solid objects, walls, glass) will result in a strong reflection.

Mechanisms of scattering from surfaces are somewhat more complex, but as a general rule of thumb, it is well known [41] that the degree of scattering depends on the wavelengths relative to the size of the scattering objects. In the case of light, irregularities on the scale of hundreds of nanometers will cause a diffuse reflection, but a diffuse (omnidirectional) reflection for sound waves requires irregularities on the scale of (at least) several inches.

5.2.2 Scattering Experiment

To test the effects of surface scattering, an airborne parametric array was aimed at a solid (painted drywall) surface in an otherwise anechoic room and the field was mapped using the methods in the previous chapter. The wall was then covered with a two-dimensional acoustic scattering material (‘Skyline’ from RPG diffusors, Figure 5.5) and the experiment was repeated.

The source signal was filtered white noise, containing frequencies from 400 Hz to 4 kHz. Pure tones were not used as they would cause strong interference fringes between the incident and reflected sound fields that would obscure the resulting fields.

The fields corresponding to a specular and diffuse reflection is shown in Figure 5.6. The transducer is on the upper right side in each case, and the reflection surface is shown with the black bar.

In the upper graph, the beam was simply reflected from a flat, plain wall. It is apparent that directivity is unaltered; a listener in the reflected field will perceive the sound as coming from the point of reflection (or ‘projection’), but it is only clearly heard when the listener is directly inside the reflected beam.

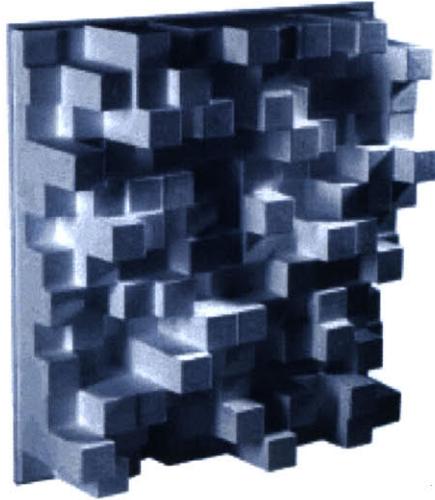


Figure 5.5: The acoustic scattering panel used in the experiment. A 2x2 array of these panels (a total of 4 ft x 4 ft) was used.

When a scattering surface is added, there is a clear loss of directivity, combined with a loss of intensity. The loss of intensity can be attributed to the redistribution of sound energy in all directions; a one-dimensional diffusor would have limited scattering to the horizontal plane, and a greater SPL for more listening positions would be observed. Regardless, a listener essentially anywhere in the room perceives the sound as originating at the projection point, although with less loudness than a direct specular reflection.

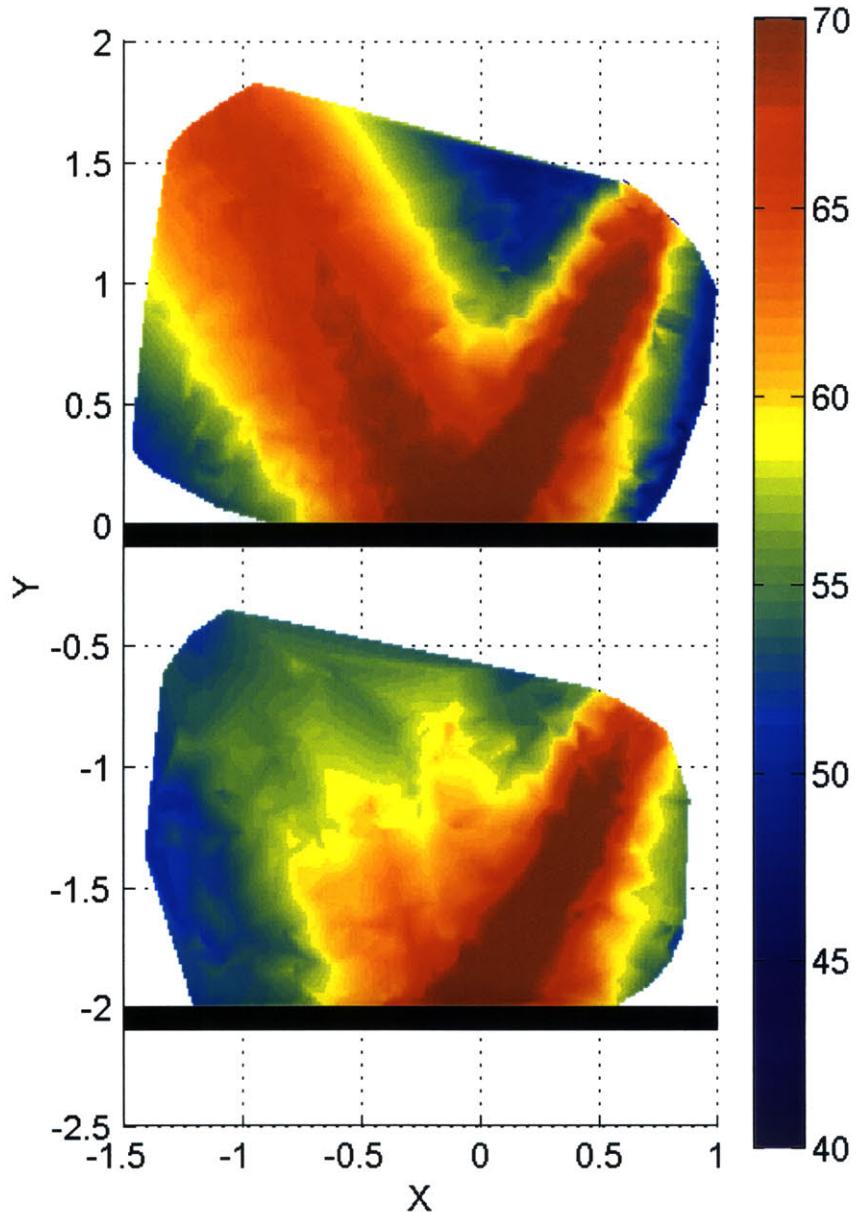


Figure 5.6: In the upper plot, an acoustic beam of broadband noise is reflected from a featureless wall. Clearly, the reflection is specular, and the reflected field maintains the same directivity as the incident beam. In the lower plot, an irregularly-shaped material was placed on the wall, resulting in a diffuse reflection, and a complete loss of directivity. Spatial units are in meters, and the colors indicate audible SPL.



Figure 5.7: Two small parametric arrays were mounted overhead in this automobile experiment, and subwoofers were embedded directly within the seats. It was found that the low-frequencies were effectively heard by the intended listener, but were not noticeable by the other passenger.

5.3 Subwoofer augmentation

Because of the limited low-frequency output of the airborne parametric array, an augmentation with traditional subwoofer systems can be beneficial. It has been observed that as long as the subwoofer output is of reasonably low level, it does not cause much distraction or annoyance to those outside of the audible sound beam. However, when listening to sounds well-correlated between the audible beam and subwoofer, the low-frequency reinforcement is very effective.

This was most dramatic when used in the close confines of automobiles. In one installation, two overhead parametric arrays were directed to a driver and passenger separately, and each also had their own subwoofer mounted directly into the seat (Figure 5.7). While objective performance experiments were not performed, the listeners reported that the ‘leaked’ subwoofer channel was not noticeable.

5.4 Precedence Effect

The parametric loudspeaker is extremely effective in projecting sound to a distant surface, creating a ‘virtual’ sound at an arbitrary point distant from the transducer. One compelling application is for use within theaters to allow precise positioning and motion of auditory objects.

A drawback of the parametric speaker is that it cannot reliably reproduce low frequencies, and the amplitude of the device is somewhat limited compared to traditional loudspeakers. Ordinary loudspeaker systems, on the other hand, can clearly produce high amplitude, and at low frequencies, but obviously cannot steer sound cleanly.

By taking advantage of the psychoacoustic effect known as the *precedence effect*, we can cause a listener to localize the sound from the parametric loudspeaker (which has been positioned), while they still hear (but do not substantially localize) the traditional reinforcing loudspeakers.

5.4.1 Psychoacoustics

The basic tenet of the precedence effect states that a sound is localized (its position is identified) by the *first* sound a listener hears [42, 43]. In a reflective room, where one is surrounded by strong echoes, this allows a listener to localize to the ‘true’ source (sound directly from the source), which has not been reflected. The reflections are certainly heard, but do not substantially contribute to the listener’s perception of whence the sound is originally coming.

In a simpler case, consider two loudspeakers in front of a listener. If the same sound is played through both speakers, as long as the speakers are not too far apart, the listener will tend to think all of the sound is coming from the *closer* speaker. This separation in space can be substituted with a delay, so that for speakers the same distance away, the listener will identify the loudspeaker they hear first as the overall source of the sound. The amount of delay generally has to be under 25ms [43].

5.4.2 Application to Parametric Loudspeakers

This effect can be used to reinforce the apparent loudness of the parametric loudspeaker with traditional loudspeakers which are *delayed* by some small amount, ensuring that the listener will always hear the parametric speaker first.

When this is accomplished, the full loudspeaker system contributes to the received sound, but the listener perceives the sound as substantially coming from the point of reflection. Thus, the advantages of both the parametric loudspeaker and traditional loudspeakers are used together.

The actual delay amount, which would be applied to only the traditional loudspeakers, is very dependent on the geometry of the theater, and also somewhat on the source

material. In addition, filtering and level changes (between loudspeakers) can also be used to emphasize certain frequency bands or create a stronger spatial impression.

As an example, a possible scenario is shown in Figure 5.8. The beam from the parametric source is projected against a surface, and is intended to be received by a listener located at the X. Two loudspeakers are on either side of the screen.

Based on the precedence effect, it is desirable to receive the sound from the projected image first, before those of the loudspeakers. An advance time of around 5ms is sufficient, so the time delay t applied to each loudspeaker would be:

$$t_L = 5 + (d_1 + d_2 - d_L)/c \quad (5.1)$$

$$t_R = 5 + (d_1 + d_2 - d_R)/c \quad (5.2)$$

There are good opportunities for future research in effective application of this technique, particularly for more complex geometries and with additional psychoacoustic studies. It should be quite feasible for a sound system to be constructed that controls the aiming of the beam and the delay of traditional loudspeakers in real time.

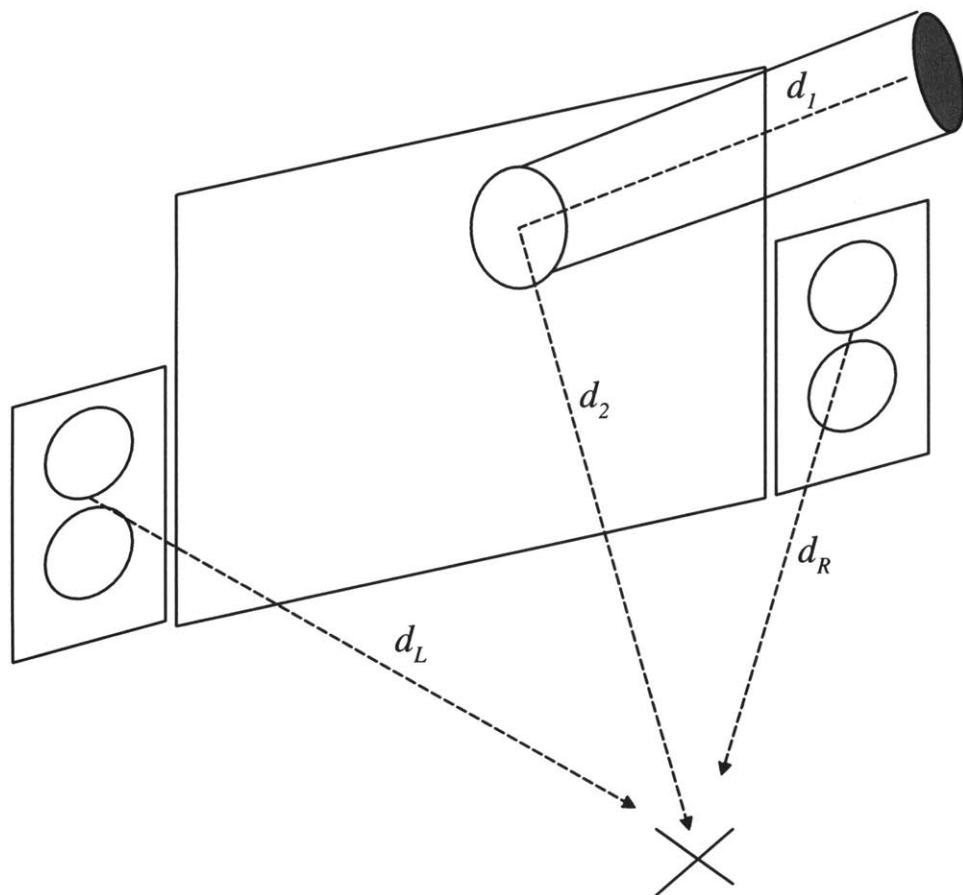


Figure 5.8: The geometry for a projected audio image against a surface, while being reinforced with traditional loudspeakers.

5.4.3 In-Place Performers

A particularly compelling and creative installation was designed during the spring of 2000 to celebrate the opening of Media Lab Europe in Dublin. “In-Place Performers”, designed by Professor Barry Vercoe, uses three parametric array systems (“Audio Spotlights™”) to create a unique musical composition that changes with the position of the listener.

Shown installed in Adelaide, Australia in Figure 5.9, three Audio Spotlights (yellow circle outlines) are mounted on the wall well above the listeners, and each is aimed at a slightly different angle (arrows). Each of the three listeners thus hears a different soloist – either a trumpet player, vocalist, or violin player. Traditional loudspeakers were used to reproduce the rhythm section, so that it may be heard everywhere. The result is a coherent jazz quartet, with the soloist chosen by the listener simply by moving to another place in the room.

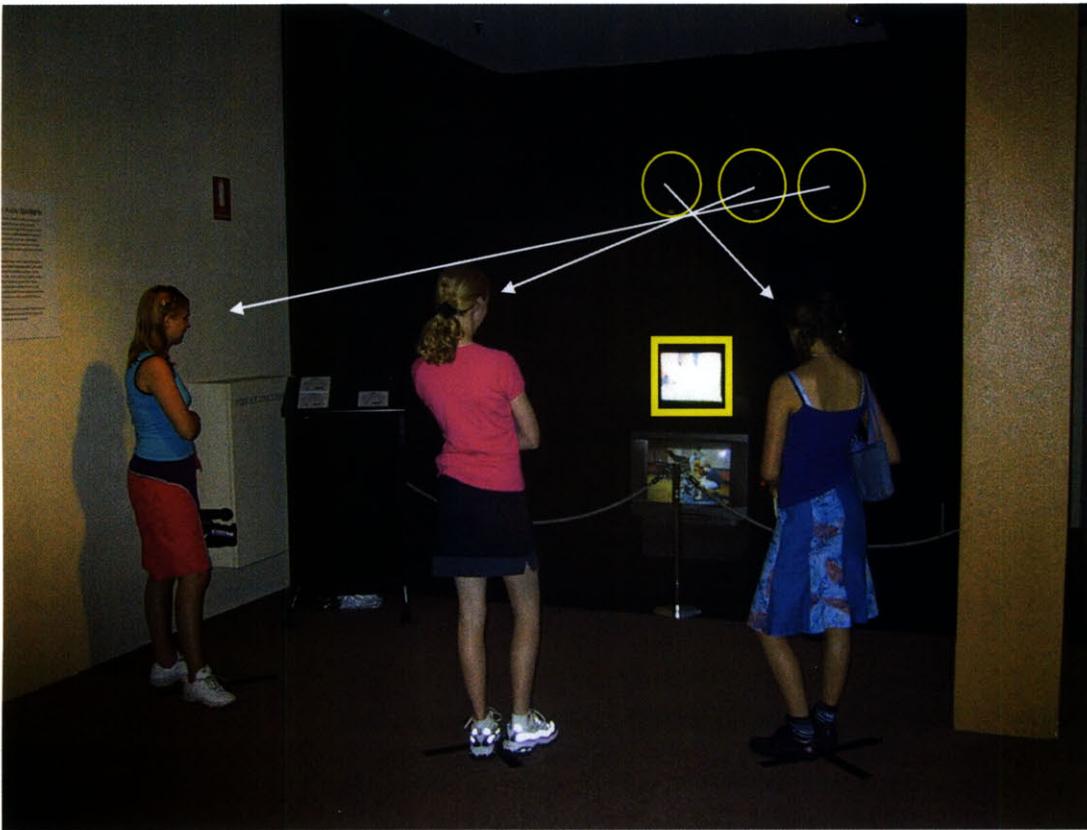


Figure 5.9: The musical installation “In-Place Performers”, installed at the South Australian Museum in Adelaide. Each Audio Spotlight (yellow circles) directs a different soloist over a jazz quartet to each listener, and is coordinated with a directional video display (yellow rectangle) which provides video of each performer.

To provide a complete experience, a special video system was used, designed by Professor Steven Benton and the Spatial Imaging Group of the MIT Media Lab. A traditional television display (seen from anywhere) is used to display the rhythm section, while a special directional video display (yellow square outline), consisting of projectors, Fresnel and lenticular lenses, and a holographic diffuser, displays the appropriate soloist for each listener. Thus, as the listener moves to another position, the sound *and* video transitions from soloist to soloist.

While a compelling musical experience, the system can also be developed further to permit, for example, multiple family members to share a single television set, but allow each person to *see and hear* the show of their choice, or to provide audio in multiple languages to public areas.

Chapter 6

Beam Control

While the airborne parametric array offers tremendous control over the distribution of audible sound in an environment, control of the beam direction itself has historically been rather inelegant, typically relying on motorized aiming or reflecting plates. It would be highly preferable to control the direction of the beam through a purely non-mechanical means.

One such method is the use of a phased array, commonly used in radar and sonar systems. The basic principle is that, by adding a spatially depending time delay (or phase shift) along an axis of the transducer, one can steer the beam about that axis.

In this chapter, we will present the theoretical approach to designing a phased array system for steering the audible sound beam, and comment on its feasibility. Through simulation, we will show that such a system is possible, but due to time and cost constraints a physical device was not constructed for this dissertation.

6.1 Basic Theory

The far-field response of an arbitrary source condition can be derived using Huygen's principle, which may then be tailored to describe the specific case of a traditional linear array with uniformly spaced elements. This is largely based on the analysis given in [44, 45] and [46], with excerpts used from a recent publication [47].

6.1.1 Huygen's Principle

Huygen's principle states that any wave producing source can be modelled as an infinite number of individual sources distributed identically to the original source. A pulsating simple source radiates energy in spherical waves described by the equation:

$$p_{\text{point}}(R) = \frac{dp_0}{R} e^{j(\omega t - kR)} . \quad (6.1)$$

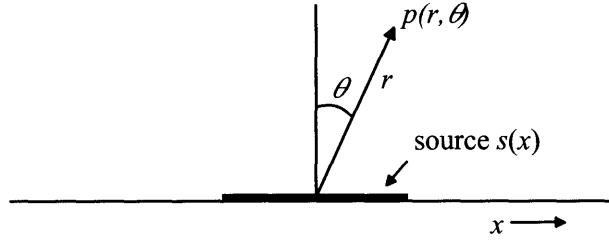


Figure 6.1: The geometry of the one-dimensional source under consideration.

where dp_0 is the source amplitude, R is the distance from the source to the point of interest, ω is the driving frequency, k is the wavenumber, and j is a unit imaginary number.

With an infinite baffle assumption, modeling a one-dimensional source with arbitrary source amplitude distribution $s(x)$, as shown in Figure 6.1, results in the pressure distribution in an integral form:

$$p(r, \theta) = e^{j\omega t} \int \frac{s(x)}{R} e^{-jkR} dx . \quad (6.2)$$

The source amplitude $s(x)$ is the contribution of source element dx to the pressure at point p , and relates to the source's normal vibration velocity amplitude $u(x)$ as $s(x) = \frac{\rho_0 \omega}{2\pi} u(x)$.

The $e^{j\omega t}$ term may be omitted for a linear and nondispersive system, and R can be approximated as

$$\begin{aligned} R &= (r^2 + x^2 - 2rx \sin \theta)^{1/2} \\ &\approx r - x \sin \theta + \frac{x^2}{2r} . \end{aligned} \quad (6.3)$$

Substituting this approximation into the integral gives:

$$p(r, \theta) = \frac{1}{r} e^{-jkr} \int s(x) e^{jkx \sin \theta} e^{-jk \frac{x^2}{2r}} dx . \quad (6.4)$$

The last factor in the integral is significant only for small r , when $r \ll \frac{1}{2} kx^2$, the Fresnel distance. The other factors are not dependent on r , and therefore it describes the pressure in the *far field* of the source. For simplicity, this analysis will address only far-field effects, limiting the extent of the array to those dimensions for which this approximation is satisfied.

6.1.2 Fourier Analysis

If the change in variables $x' = x/\lambda$ and $\beta = \sin \theta$ is made while disregarding the amplitude scaling with r , the integral describing the far field directivity becomes

$$H(\beta) = \int s(x') e^{j2\pi x' \beta} dx' . \quad (6.5)$$

This is the Fourier transform of the source distribution function scaled by $1/\lambda$. This important result makes the analysis of the far field straightforward and intuitive.

As an alternative, which is more suitable for Discrete Fourier Transform (DFT) simulation, one can substitute $\Omega = \beta/\lambda$, and arrive at:

$$H(\beta/\lambda) = f(\Omega) = \int s(x)e^{j2\pi\Omega x} dx . \quad (6.6)$$

In this case, the scaling with λ is taken *after* the Fourier transform, so that downsampling prior to taking the DFT is not necessary.

Simulation Steps:

1. Given the source distribution $s[n]$, where $n = xT$ is a sufficiently fine spatial sampling of the source function
2. Zero pad $s[n]$
3. Compute the DFT; $H(\Omega) = \text{DFT}\{s[n]\}$
4. Rescale the angle axis $\beta = \lambda\Omega$
5. Normalize (if desired)

6.1.3 Steering

Beam steering is simply shifting the overall response with respect to β , such that:

$$H'(\beta) = H(\beta + \beta_s) \quad (6.7)$$

where β_s is the steering angle. This constant shift in the ‘transform’ domain is equivalent to multiplying by $e^{jk\beta_s x}$ in the ‘source’ domain, i.e.,

$$s'(x) = s(x)e^{jk\beta_s x} . \quad (6.8)$$

This factor is simply a space-dependent phase shift, or time delay, distributed linearly across the array.

6.1.4 Discrete Array

A source $s(x)$ comprised of a discrete array of identical elements, shown in Fig. 6.2(a), can be treated as an infinite number of copies of the element source function $s_0(x)$ multiplied by an overall aperture function $w(x)$, which is generally rectangular:

$$s(x) = s_0(x) * \left(w(x) \cdot \sum_{n=-\infty}^{\infty} \delta(x - x_n) \right) \quad (6.9)$$

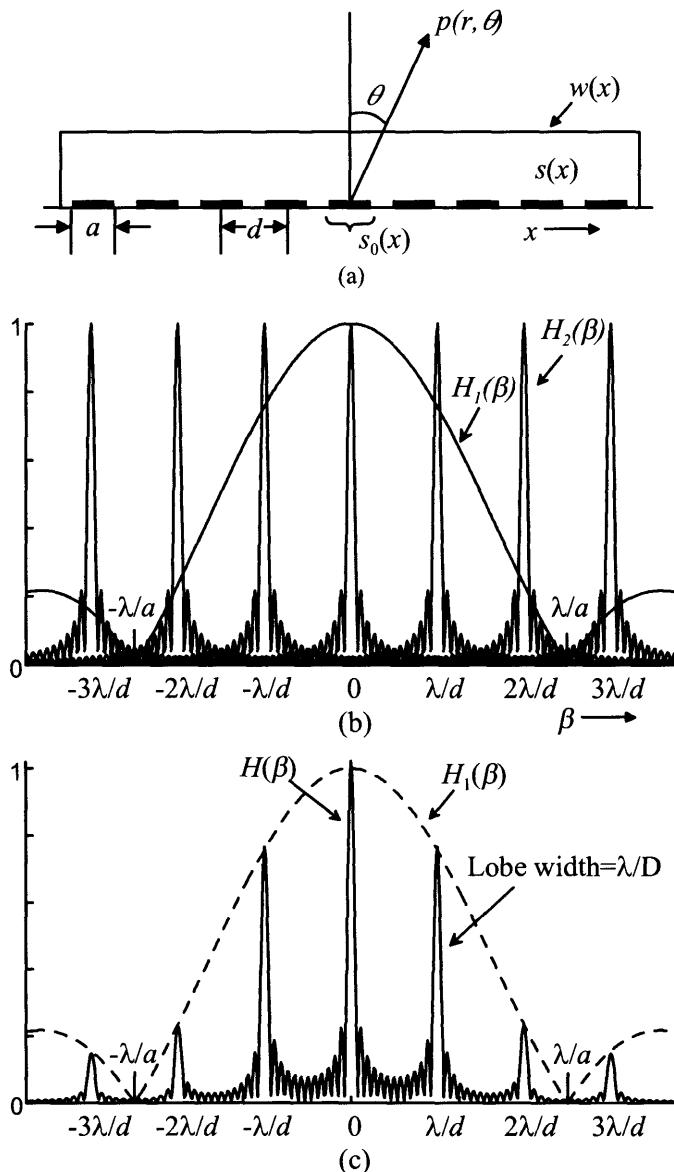


Figure 6.2: A typical linear array, with rectangular elements. The physical components of the array, including the elements of width a , interelement spacing d , and overall array size D , are shown in (a). The contributions of the directivity function due to the element size and overall array size are shown in (b). The final directivity function $H(\beta)$ is shown in (c). Recall that $\beta = \sin \theta$.

where x_n is the location of the n^{th} element, and the asterisk denotes convolution. For a periodic array, where $x_n = nd$, the far field response $H(\beta)$, given by the Fourier transform of the source function, is:

$$H(\beta) = S_0(\beta) \cdot \left(W(\beta) * \sum_{m=-\infty}^{\infty} \delta \left(\beta - \frac{m\lambda}{d} \right) \right). \quad (6.10)$$

This can be conveniently written as the product of two directivity functions, $H_1(\beta)$ and $H_2(\beta)$:

$$H(\beta) = H_1(\beta)H_2(\beta) \quad (6.11)$$

with $H_1(\beta) = S_0(\beta)$ and $H_2(\beta)$ equal to the terms in parentheses. In this study, $H_1(\beta)$ corresponds to the performance of the beam steering itself, particularly grating lobes, and $H_2(\beta)$ will predict sidelobe characteristics. Because a practical parametric array will require a large surface area (to ensure sufficient energy for demodulation), it can be assumed that ultrasonic sidelobes will be minimal.

An example array with rectangular elements is shown in Fig. 6.2(a). Here, the rectangular elements have width a and the overall array has a width of D . The normalized responses from these components, plotted in Fig. 6.2(b) are:

$$s_0(x) = \text{rect}(a) \Leftrightarrow H_1(\beta) = \text{sinc}(\pi a \beta / \lambda) \quad (6.12)$$

$$w(x) = \text{rect}(D) \Leftrightarrow W(\beta) = \text{sinc}(\pi D \beta / \lambda). \quad (6.13)$$

The rectangular function $\text{rect}(a)$ is unity for $-a/2 \leq x \leq a/2$ and zero otherwise, and $\text{sinc}(x) = \sin(x)/x$. The constant interelement spacing of d creates periodic lobes at intervals d/λ in $H_2(\beta)$. The final response $H(\beta)$ is shown in Fig. 6.2(c). Only the interval $-1 \leq \beta \leq 1$ maps to physical space, so this interval is usually termed the *visible region*. Any additional lobes present in this region are termed *grating lobes*.

This result illustrates two important intuitive points:

- The far-field response consists of an infinite set of copies of the overall aperture response $W(\beta)$, spaced by λ/d .
- This total response is then modulated by an *element* response, or the directivity envelope, $H_1(\beta)$.

Thus, in the ‘source’ domain, the individual element response influences the overall response of the array in the transform domain, while the overall window in the source domain affects the individual elements in the transform domain.

6.1.5 Dense Arrays

In many applications, it is desirable to maximize sensitivity per unit length of the array, which necessitates the maximization of the density of the array. In a traditional linear

array, this is accomplished by making the interelement spacing d equal as close as possible to the element size a . This type of array may be called a *dense array*.

For an array with element size a equal to its interelement spacing d , the grating lobes in an unsteered beam are coincident with zeros in the directivity envelope $H_1(\beta)$, as shown in the upper plot of Fig. 6.3.

For clarity, the magnitudes are plotted with respect to a normalized angle $\beta' = \beta d/\lambda$, so that variations in wavelength will simply re-scale the horizontal axis, and affect the visible region, not the shape of the response.

As the beam is steered, the grating lobe increases in magnitude, while the main lobe decreases. The main lobe and grating lobe have equal amplitudes when $\beta'_s = \frac{1}{2}$, or equivalently when $\beta_s = \lambda/2a$. This sets a limit on useful steering angle to a small range of angles where $\beta_s < \lambda/2a$.

6.1.6 Steering Performance

The most important indicators of phased array performance which depend on array geometry are (1) the main lobe magnitude M , (2) the grating lobe magnitude G , and (3) the ratio between the amplitude of the main lobe versus that of the grating lobe $\zeta = M/G$. Recall that the beam width is independent of the element shape, and is simply associated with the overall extent of the array D . As the characteristics of interest are due only to elements themselves, an infinitely large array is assumed, so that the lobes are impulses.

For a dense rectangular array, where $d = a$, the main lobe magnitude M and grating lobe magnitude G as a function of steering angle are:

$$M_{\text{rect}}(\beta_s) = \text{sinc}(\pi\beta_s a/\lambda) \quad (6.14)$$

$$G_{\text{rect}}(\beta_s) = \text{sinc}(\pi(\beta_s a/\lambda \pm 1)) \quad (6.15)$$

and the ratio ζ of the main lobe to grating lobe is:

$$\zeta_{\text{rect}}(\beta_s) = \frac{\text{sinc}(\pi\beta_s a/\lambda)}{\text{sinc}(\pi(\beta_s a/\lambda \pm 1))}. \quad (6.16)$$

In normalized angular coordinates $\beta'_s = \beta_s d/\lambda$,

$$M_{\text{rect}}(\beta'_s) = \text{sinc}(\pi\beta'_s) \quad (6.17)$$

$$G_{\text{rect}}(\beta'_s) = \text{sinc}(\pi(\beta'_s \pm 1)) \quad (6.18)$$

$$\zeta_{\text{rect}}(\beta'_s) = \frac{\text{sinc}(\pi\beta'_s)}{\text{sinc}(\pi(\beta'_s \pm 1))}. \quad (6.19)$$

These results (in decibels) are plotted in Fig. 6.4. Note that, while the main lobe amplitude is strong for modest angles, the ratio ζ diminishes very quickly as the beam is steered. An improved phased array system would have a more uniform main lobe response and lobe ratio ζ for a desired steering sector. As these functions are wholly dependent on the element source function $s_0(x)$, this source function can be tailored to improve the lobe ratio ζ .

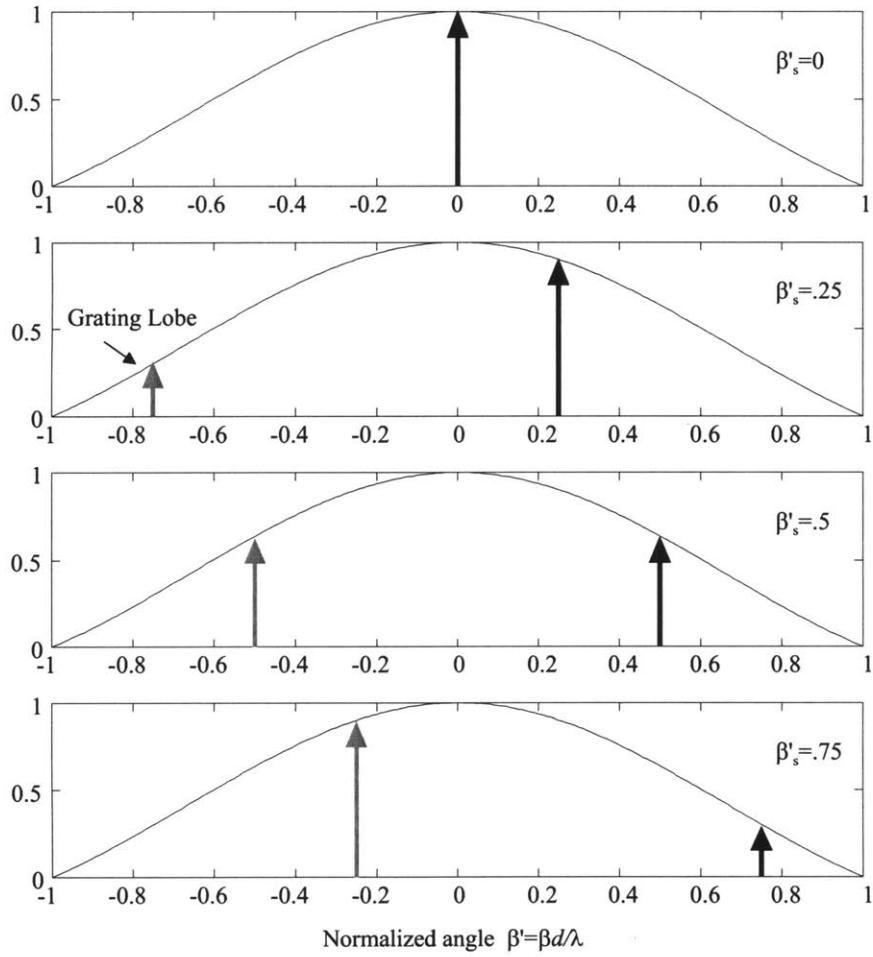


Figure 6.3: The directivity of a dense array with element size a equal to the element spacing d . The plots show the changing position and amplitude of the main lobe and grating lobe as the beam is steered. Notice that the magnitude of the grating lobe is equal to that of the main lobe when $\beta_s = \lambda/2a$.

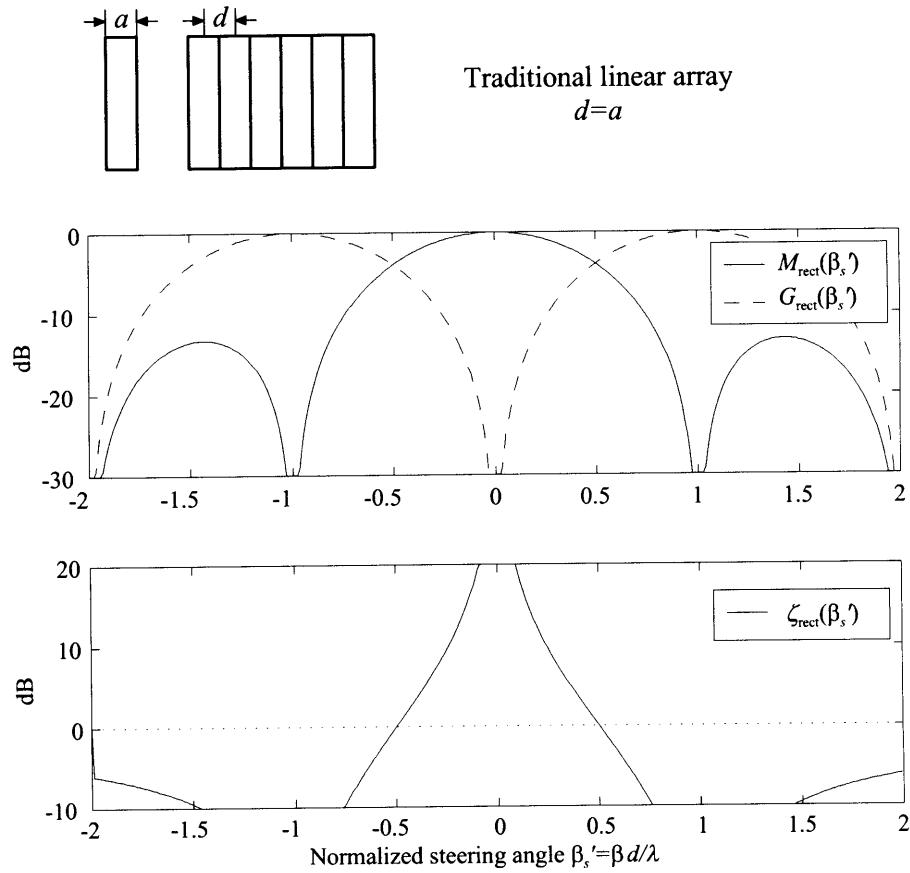


Figure 6.4: A rectangular element dense linear array, where $d = a$. The upper plot shows the magnitude of the main lobe M and grating lobe G as a function of normalized steering angle β_s' , and the bottom plot shows the ratio ζ of main lobe to grating lobe amplitude. Note that, while the magnitude of the main lobe does not diminish substantially for small angles, the main lobe to grating lobe ratio vanishes quickly for even small steering angles.

6.2 Application to the parametric array

By steering the beam of ultrasound, the effective source of the audible sound is steered to a desired angle, and the audible byproducts should be likewise steered. This direct application of a phased array to the parametric array is straightforward, although it apparently has not been reviewed in the literature.

As discovered in the previous section, the primary concern with designing a phased array system for this application is the sufficient suppression of grating lobes, which disappear completely only when the interelement spacing d is less than half of the wavelength of ultrasound used. This implies that, for a 50cm array running at 60 kHz, one would require 174 individually-driven transducers, each of dimension less than 3mm across. If each channel requires data at a sample rate of 200 kHz, the total data throughput is approximately 35×10^6 samples per second. This data rate, combined with the expected necessity of a separate amplifier for each transducer, shows that the construction of a phased-array beamsteered system will be a substantial undertaking. For a two-dimensional array, the data rate would be on the order of 6×10^9 samples per second, and would require over 30,000 channels!

There are, thankfully, several ways of reducing the cost and complexity of such a system, which may make at least a one-dimensional steering array practical in the near future:

- Relaxing the interelement spacing

In Chapter 4, it was shown that the audible result is approximately proportional to the square of the ultrasound generated at a particular region of space. Because of this square relationship, the effect of the grating lobe(s) on the audible sound will be greatly diminished, and the interelement spacing requirement can be somewhat relaxed. For all steering angles, the audible amplitude difference in decibels between the main and grating lobe will be doubled, so that the requirement that $d < \lambda/2$ may be relaxed.

- Element shapes

If the shapes of the elements are chosen carefully, interelement spacing can be reduced *without* creating new grating lobes. This technique is described in detail by this author elsewhere [47], and allows interelement spacing to be relaxed in exchange for a limitation in maximum steer angle. For example, if the steering angle is restricted to 30 degrees off-axis, the interelement spacing can be relaxed by a factor of two. This halves the required number of ultrasonic channels with no corresponding decrease in performance.

- Tapped delay line

The data rate and amplifier count are some of the most significant sources of system complexity and cost for a phased array system. Because each of the array element channels are nothing more than delayed copies of a single signal, the use of a tapped

delay line would reduce the data rate to that of a single channel (200 kHz). If the delay line is employed mechanically through creative transducer design, this may also eliminate the need for a single amplifier for each ultrasonic transducer.

6.2.1 Simulation

To demonstrate the phased array, a proposed design was simulated. To ensure reasonable suppression of grating lobes, an element size and interelement spacing of 3mm was used with 60 kHz ultrasound ($\lambda = 5.75\text{mm}$, 133 channels). The effects of absorption were included. The simulation was then repeated for an interelement spacing of 5mm (80 channels). In both cases, the array is 40cm across. The results are shown in Figure 6.5.

For an interelement spacing approximately equal to half of a wavelength (left column), beam steering at all angles is very effective, and no grating lobes are noted. However, when the interelement spacing is slightly increased (right column), the grating lobe is apparent, and becomes as prominent as the main lobe. If the maximum steer angle is restricted, it may be possible to use a physical barrier (absorber) to block the grating lobe, but it is far preferable to use an array with the proper interelement spacing.

While the simulation was performed for 60 kHz ultrasound, it is clear that similar results would hold for nearby frequencies, although the interelement spacing would have to be reduced to satisfy the highest frequency of interest (i.e. 80 kHz).

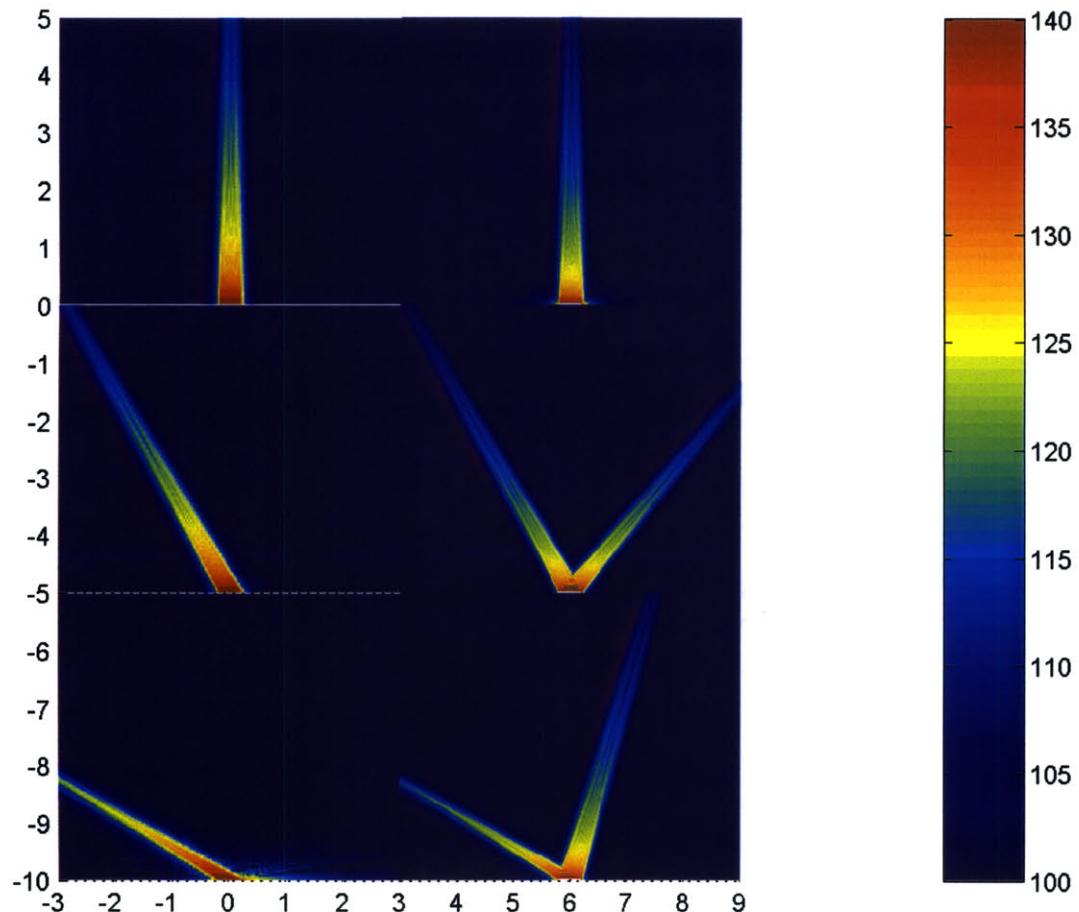


Figure 6.5: The 60 kHz ultrasonic field generated by steered arrays is shown. The left column reflects an interelement spacing of 3mm, and on the right, 5mm. Each row is a different steer angle (0, 30, and 60 degrees). For the smaller interelement spacing, grating lobes are clearly absent, but they become very prominent with only a small increase in interelement spacing.

Chapter 7

Biological Effects

Because this system, by its nature, can expose listeners to unusually high levels of ultrasound, a thorough review and exploration of the related risks is merited.

While there have been no reports of adverse affects when listening to the parametric array described in this dissertation (which operates at a 60 kHz carrier), several listeners have reported experiencing fatigue, dizziness, and even substantial auditory pain while listening to an earlier device running at a much lower ultrasonic frequency (40 kHz, [20]). At the time our modern device was being designed, it was hypothesized that the use of intense ultrasound near the audible range is the cause of the ill effects, and that higher frequency ultrasound should instead be used.

There currently exist no specific governmental restrictions governing ultrasonic exposure, but the Occupational Safety and Hygiene Association (OSHA) has published recommendations for exposure limits, based on those suggested by the American Congress of Governmental Industrial Hygiene (ACGIH). They currently recommend an exposure level not exceeding 145 dB SPL [48], but offer no specific reasons for this choice. Adding to the confusion, just a few years prior, the recommended maximum level was only 115 dB SPL – an increase of a factor of over thirty!

As it is unusual for safety guidelines to be relaxed to such a degree, a member of the ACGIH was asked informally about this decision. His response was, simply, that they felt the limitation was overly restrictive, and they knew of “no reason” not to increase it.

The lack of concrete data from the ACGIH, combined with the somewhat sparse literature on the subject of the exposure to airborne ultrasound, indicated a clear need for specific study to address these effects. A group of experts, led by Professor Martin Lenhardt, a bioacoustician and specialist in health effects due to airborne ultrasound exposure, supervised by Professor Richard Wilson, of the Center for Risk Analysis of Harvard University, was commissioned to research and analyze the risk associated with use of the specific parametric loudspeaker described in this dissertation. The study, included in the appendix, clearly showed that the listening risk of this specific parametric acoustic source was comparable to that of a traditional loudspeaker, indicating that the ultrasound itself was harmless. Again, this is likely due to the prudent selection of an ultrasonic frequency band

(50 kHz–70 kHz) well outside of the audible band, in contrast to earlier devices [18, 20, 32] which used ultrasound of much lower frequency (30 kHz–40 kHz).

The risks associated with exposure to airborne ultrasound are grouped into three areas – physiological effects on the body, those affecting the auditory system directly, and subjective effects.

7.1 Effects on the body

Airborne ultrasound is a form of radiated energy; any absorption of this radiated energy will create heat, which must be within safe limits. Our upper allowable ultrasonic limit, then, depends on both the absorptive nature of the human body (which is frequency dependent) and the maximum allowable heating.

Skin is a very poor absorber of airborne ultrasound, typically reflecting 99.90% [49] of the ultrasonic energy. Because of this, only under extremely high intensities would ultrasound cause heating of the skin. Hair, clothing, and fur, on the other hand, are better absorbers of ultrasound, promoting heating at lower energy levels.

Studies show that measurable (but slight) heating occurs near 145 dB [50] for small furred animals, and near 159 dB [51] for the human skin. These measurements were done at relatively low ultrasonic frequencies (20–30 kHz), and because of the impedance mismatch between skin and air, it is probable that these effects diminish at higher frequencies.

This heating effect can be related to a more familiar radiation exposure – sunlight. If heat alone is considered (ignoring UV-exposure effects and similar risks), the intensity of sunlight is about 1000W/m^2 at noon on a clear day, which corresponds to approximately 150 dB of ultrasound. Therefore, assuming perfect absorption (which is almost certainly not approached for ultrasound, even for ‘furry’ surfaces), when standing in a 150 dB beam one can expect to be warmed as if standing in a ray of sunlight.

Although we are reasonably sure that an upper ultrasonic exposure limit of around 145 dB will be safe, the absorptive nature of skin and hair (and perhaps fur) should be understood more completely. The question of whether or not absorption by hair or clothing increases or decreases with frequency is of particular future interest.

7.2 Auditory effects

Research in the auditory effects of airborne ultrasound exposure has primarily addressed industrial exposure, where the frequencies are generally near 20–40 kHz, and are usually fairly narrowband.

Several researchers (summarized in [51]) reported temporary hearing threshold shifts (temporary reduction in sensitivity), but no permanent damage, from exposure to frequencies of 17–37 kHz at 148–154 dB. This investigation, however, seemed to consider brief, not continuous, exposure. Similar studies failed to find any effects from the exposure to relatively high levels (130 dB) of ultrasound.

Some studies [52] speculate on the generation of subharmonics via the nonlinearity of the middle and inner ear, which could explain the influence of ultrasonic noise on the audible band. The degree to which this may damage the auditory system is not known. Then again, these subharmonics may be generated by the nonlinearity of the air itself, as in the operation of the device presented in this dissertation, a phenomenon which was not well understood by the prior investigators. If this is the case, any risks associated with the ultrasound may be simply those which are associated with exposure to standard audible-range sound that has been produced by the ultrasound.

In the absence of audible subharmonics, there is general consensus among early researchers [51–53] that, for airborne ultrasound of sufficiently high frequency, temporary hearing loss should not occur for levels below 140 dB.

7.3 Subjective effects

Other effects, such as dizziness, tinnitus, a sense of ear fullness, and even pain have been reported by those being exposed to high levels of ultrasound, but only at frequencies much lower than those being used in our system (40 kHz and less) [50]. However, as these effects are difficult to quantify, formal studies of them are sparse.

Chapter 8

Conclusions and Future Work

Through these experiments, it can be concluded that:

1. High quality audio can be made from ultrasound alone.

For the first time, high quality audio has been created in mid-air through nonlinear interactions of ultrasound waves. By properly modeling the nonlinear demodulation process, as well as building a high-accuracy reproduction system, distortion effects were nearly eliminated.

2. Audio can be generated with very high directivity, even from a physically small source.

Traditional loudspeakers must be physically extremely large to produce narrow beams of audible sound. Using the techniques described in this dissertation, very narrow (and small-diameter) beams of audible sound can be created from a much smaller source.

3. Phased arrays can be used to steer the ultrasound as well as the audible beam.

A reasonably straightforward implementation of a phased array for ultrasound reproduction was designed, and is expected to reliably steer the audible result while maintaining directivity. However, construction of such a system will be an expensive undertaking.

4. Properly designed, an ultrasonic system of this type does not present a health risk.

Through the study of earlier work in auditory physiology and the effects of exposure to high levels of ultrasound, the system described in this dissertation was specifically designed to prevent health risks due to ultrasonic exposure. The elimination of exposure risk was confirmed experimentally.

However, if the system was designed differently, particularly if lower-frequency ultrasound is used, it is likely there will be risks to the auditory health of the listener.

While this research effort resulted in a simple, practical, and safe parametric array system for audio reproduction, there are several opportunities for future research and developments.

While 1% distortion is by most standards very low, the system would benefit from further reduction in distortion, particularly at high levels of output. Accomplishing this requires additional accuracy in the approximations for nonlinear demodulation, additional system bandwidth, and more accurate reproduction apparatus.

According to the research in biological effects of ultrasound exposure, the frequencies used in this device may permit an additional 5–10 dB of ultrasound output, resulting in 10–20 dB of additional audible output for a transducer of the same size. These improvements can be obtained by further engineering development of the physical apparatus.

The phased array has been shown to be a compelling method of controlling the direction of the resulting audible beam, but the construction and evaluation of a physical device was outside the scope of this dissertation. With additional engineering and development, a phased array system can be constructed.

Additional field tests and actual application use of this device promises to uncover new techniques of employing this technology, including its integration with traditional loudspeaker systems, applying the system to larger audiences, and utilizing improved scattering surfaces for acoustic projection.

By curving the transducer, the ultrasonic beam can be focused or de-focused, resulting in interesting changes in the audible field according to informal calculations. A full, formal analysis and experiment of curved and shaped transducers would be a worthwhile subject of further investigation.

Bibliography

- [1] Peter J. Westervelt. Parametric acoustic array. *Journal of the Acoustical Society of America*, 35(4):535–537, April 1963.
- [2] Alan Dower Blumlein. Improvements in and relating to sound-transmission, sound-recording and sound-reproducing systems. British Patent Specification 394325, 1931.
- [3] Dr. Floyd E. Toole. Direction and space: The final frontiers. White paper, Harman International, 1999.
- [4] William G. Gardner. *3-D audio using loudspeakers*. PhD thesis, Massachusetts Institute of Technology, 1997.
- [5] D. Begault. Challenges to the successful implementation of 3-d sound. *Journal of the Audio Engineering Society*, 39(11):864–870, 1991.
- [6] Peter J. Westervelt. Scattering of sound by sound. *Journal of the Acoustical Society of America*, 29(2):199–203, 1957.
- [7] Mark F. Hamilton and David T. Blackstock, editors. *Nonlinear Acoustics*. Academic Press, Boston, 1998.
- [8] J. L. S. Bellin and R. T. Beyer. Experimental investigation of an end-fire array. *Journal of the Acoustical Society of America*, 34(8):1051–1054, August 1962.
- [9] Mark F. Hamilton. Effects of noncollinear interaction on parametric acoustic arrays in dispersive fluids. *Journal of the Acoustical Society of America*, 76(5):1493–1504, Nov 1984.
- [10] Mark F. Hamilton et al. Nonlinear effects in the farfield of a directive sound source. *Journal of the Acoustical Society of America*, 78(1):202–216, July 1985.
- [11] David T. Blackstock. Generalized Burgers' equation for plane waves. *Journal of the Acoustical Society of America*, 77(6):2050–2053, June 1985.
- [12] H. O. Berkay. Possible exploitation of nonlinear acoustics in underwater transmitting applications. *Journal of Sound and Vibration*, 2(4):435–461, 1965.

- [13] N. S. Bakhvalov et al. *Nonlinear Theory of Sound Beams*. American Institute of Physics, 1987. ISBN 0-88318-520-2.
- [14] N. S. Bakhvalov et al. Focused high-amplitude sound beams. *Soviet Physical Acoustics*, 24(1):10–15, Jan-Feb 1978.
- [15] M. B. Bennett and David T. Blackstock. Parametric array in air. *Journal of the Acoustical Society of America*, 57:562–568, 1975.
- [16] Toshiaki Nakamura. Waveform variation of pulsed parametric source. In *Proceedings of the 11th International Congress on Acoustics*, Toulouse, France, 1983.
- [17] Kenichi Aoki et al. A parametric loudspeaker - applied examples. *Electronics and Communications in Japan*, 77(1):64–73, 1994.
- [18] Masahide Yoneyama et al. The audio spotlight: An application of nonlinear interaction of sound waves to a new type of loudspeaker design. *Journal of the Acoustical Society of America*, 73(5):1532–1536, May 1983.
- [19] Yoneyama M. Kamakura, T. and K. Ikegaya. Developments of parametric loudspeaker for practical use. In *10th International Symposium on Nonlinear Acoustics*, Kobe, 1984.
- [20] Elwood G. Norris. The creation of audible sound from ultrasonic energy: A fundamental paradigm shift. In *Proceedings from the ASA 133rd meeting*, 2pEA5, 1997.
- [21] David T. Blackstock. Audio application of the parametric array. In *Proceedings of the ASA 134th Meeting*, 2pEA5, 1997.
- [22] F. Joseph Pompei. The use of airborne ultrasonics for generating audible sound beams. In *Proceedings of the Audio Engineering Society, 105th Convention*, San Francisco, 1998.
- [23] O. V. Rudenko B. K. Novikov and V. I. Timoshenko. *Nonlinear Underwater Acoustics*. Acoustical Society of America, New York, 1987.
- [24] E. A. Zabolotskaya and R. V. Khokhlov. Quasi-plane waves in the nonlinear acoustics of confined beams. *Soviet Physical Acoustics*, 15:35–40, 1969.
- [25] V. P. Kuznetsov. Equations of nonlinear acoustics. *Soviet Physical Acoustics*, 16:467–470, 1971.
- [26] Pierre Cervenka and Pierre Alais. Fourier formalism for describing nonlinear self-demodulation of a primary narrow ultrasonic beam. *Journal of the Acoustical Society of America*, 88(1):470–481, July 1989.

- [27] Sigurd Ivar Aanonsen et al. Distortion and harmonic generation in the nearfield of a finite amplitude sound beam. *Journal of the Acoustical Society of America*, 75(3):749–768, March 1984.
- [28] Sharad Singhal and John G. Zornig. Synthesis of arbitrary broadband signals for a parametric array. *Journal of the Acoustical Society of America*, 72(1):238–244, July 1982.
- [29] G. B. Airy. The astronomer royal on a difficulty in the problem of sound. *Phil. Mag.*, 34(3):401–405, 1849.
- [30] Michalakis A. Averkiou et al. Self-modulation of amplitude and frequency-modulated pulses in a thermoviscous fluid. *Journal of the Acoustical Society of America*, 94(5):2876–2883, Nov 1993.
- [31] Thomas Kite, John T. Post, and Mark F. Hamilton. Parametric array in air: Distortion reduction by preprocessing. In P. K. Kuhl and L. A. Crum, editors, *Proc. 16th Int. Cong. Acoust.*, volume 2, pages 1091–1092, New York, 1998. Acoustical Society of America.
- [32] Kenichi Aoki et al. Parametric loudspeaker - characteristics of acoustic field and suitable modulation of carrier ultrasound. *Journal of the Acoustical Society of America*, 74(3):76–82, 1991.
- [33] H. E. Bass et al. Absorption of sound by the atmosphere. *Physical Acoustics*, 17:145–232, 1984.
- [34] H. E. Bass et al. Atmospheric absorption of sound - update. *Journal of the Acoustical Society of America*, 88(4):2019–2021, 1996.
- [35] H.E. Bass et al. Atmospheric absorption of sound: Further developments. *Journal of the Acoustical Society of America*, 97(1):680–683, January 1995.
- [36] H. E. Bass et al. Atmospheric absorption of sound: Further developments. *Journal of the Acoustical Society of America*, 99(2):1259–1259, 1996.
- [37] Lawrence E. Kinsler, Austin R. Frey, Alan B. Coppens, and James V. Sanders. *Fundamentals of Acoustics*. John Wiley & Sons, 4th edition, 2000.
- [38] Chihng-Tsung Liauh and Win-Li Lin. Fast numerical scheme of computing acoustic pressure fields for planar circular ultrasound transducers. *Journal of the Acoustical Society of America*, 105(4):2243–2247, April 1999.
- [39] J. Naze and S. Tjøtta. Nonlinear interaction of two sound beams. *Journal of the Acoustical Society of America*, 37:174–175, 1965.

- [40] Tomoo Kamakura et al. Nonlinearly generated spectral components in the nearfield of a directive sound source. *Journal of the Acoustical Society of America*, 85(6):2331–2337, June 1989.
- [41] Phillip M. Morse. *Vibration and Sound*, chapter 7, pages 346–347. Acoustical Society of America, 2 edition, 1981.
- [42] William A. Yost and George Gourevitch. *Directional Hearing*. Springer-Verlag, 1987.
- [43] Jens Blauert. *Spatial Hearing*. MIT Press, 1983.
- [44] Bernard D. Steinberg. *Principles of Aperture and Array System Design*. John Wiley & Sons, 1976.
- [45] Shi-Chang Wooh and Yijun Shi. Optimum beam steering of linear phased arrays. *Wave Motion*, 29:245–265, 1999.
- [46] Shi-Chang Wooh and Yijun Shi. Influence of phased array element size on beam steering behavior. *Ultrasonics*, 36:737–749, 1998.
- [47] F. J. Pompei and S. C. Wooh. Phased array element shapes for suppressing grating lobes. *Journal of the Acoustical Society of America*, 111(5):2040–2048, May 2002.
- [48] Occupational Safety and Health Administration. OSHA technical manual. Technical Report Section III, Chapter 5, 2002.
- [49] C. Wiernicki and W. J. Karoly. Ultrasound: Biological effects and industrial hygiene concerns. *American Industrial Hygiene Association Journal*, 46(9):488–496, 1985.
- [50] C. H. Allen et al. Some biological effects of intense high frequency airborne ultrasound. *Journal of the Acoustical Society of America*, 20(1):62–65, Jan 1947.
- [51] W. I. Acton. The effects of industrial airborne ultrasound on humans. *Ultrasonics*, 12:124–128, 1974.
- [52] H. O. Parrack. Effects of airborne ultrasound on humans. *International Audiology*, 5:294–308, 1966.
- [53] W. I. Acton. Criterion for the prediction of auditory and subjective effects due to airborne noise from ultrasonic sources. *Annals of Occupational Hygiene*, 11:227–234, 1968.

Appendix A

An Assessment of Airborne Ultrasound Risk from the Audio Spotlight™

Martin L. Lenhardt, Ph.D., Au.D.

Professor, Biomedical Engineering, and Director, Bioacoustics Laboratory
Virginia Commonwealth University

Douglas G. Richards, Ph.D.

Alan G. Madsen

BiosonX, Inc.
Richmond, VA

An Assessment of Airborne Ultrasound Risk from the Audio Spotlight™

Executive Summary

The Audio Spotlight™ produces a beam of airborne sound with components in both the ultrasonic and audible ranges. This report examines the risks of exposure to this sound beam, and to low frequency (20-100 kHz) airborne ultrasound in general. A review of the literature suggests that exposure to sound in this range should pose little risk for threshold changes or hearing damage at or below the current OSHA limit for exposure to low frequency airborne ultrasound, 145 dB. Pressure effects on the tympanic membrane may produce a sense of fullness in the ear at lower levels, but this effect should only be temporary. There currently exist no assessment data on this subjective effect. Other physiological effects such as heating of the skin should be negligible.

For the assessment of the Audio Spotlight™, twenty (20) subjects listened to five minutes of stimulation by the system at one (1) meter distance, using 2 kHz as the audio input. The output of the Audio Spotlight™ consisted of wide-band ultrasound with a peak intensity of 121 dB SPL at 58 kHz, and a demodulated audio signal with intensity of 80 dB SPL at 2 kHz. Hearing thresholds and tympanograms were measured before and after exposure to the Audio Spotlight™ output and to the audio signal alone.

1. Under the conditions of the study (five minutes of 58 kHz ultrasonic stimulation at 121 dB SPL and 2 kHz at 80 dB SPL), no statistically significant auditory effects, either temporary or permanent, were found. The lack of any hearing effects suggests that listening to audio frequencies via a loudspeaker or the Audio Spotlight™ at equivalent intensities (e.g., 80 dB SPL) is comparable.
2. This conclusion is not to be interpreted as that there was no ultrasonic absorption by the body, but the absorption was not sufficient to be detected in the testing methods employed.
3. No subjective effects were reported by the subjects under the present conditions.
4. Under the conditions tested, the Audio Spotlight™ is in full compliance with the ACGIH and OSHA regulations in regard to hearing safety for airborne ultrasound.

Acknowledgments

The authors wish to thank Professor Richard Wilson of the Harvard University Center for Risk Analysis, School of Public Health, for his guidance and assistance with designing this research effort, and to Holosonic Research Labs, Inc., for sponsoring and funding this study.

Airborne Ultrasound and the Audio Spotlight

The Audio Spotlight™ produces a beam of airborne sound with components in both the ultrasonic and audible ranges. This report examines the risks of exposure to this sound beam, and to low frequency (20-100 kHz) airborne ultrasound in general. Although the inner ear can detect ultrasound when directly coupled through bone conduction, airborne ultrasound is not easily detectable due to the large impedance mismatch in the middle ear. It should pose little risk for threshold changes or hearing damage at or below the current OSHA limit for exposure to low frequency airborne ultrasound, 145 dB. Pressure effects on the tympanic membrane may produce a sense of fullness in the ear at lower levels, but this effect should only be temporary. There currently exist no assessment data on this subjective effect. Other physiological effects such as heating of the skin should be negligible. Levels of the demodulated audio frequency sound from the Audio Spotlight™ should conform to current OSHA exposure standards (e.g., 85 dB for eight hours).

Definitions:

Airborne ultrasound: energy >20 kHz propagated in air with no fluid or solid coupling;

Audible ultrasound: energy >20 kHz that results in an audible sensation characterized by some degree of pitch, loudness, duration and timbre;

Audiometric frequencies: energy <10 kHz;

High audio frequencies: energy from 10-20 kHz;

High frequency ultrasound: energy >100 kHz;

Low frequency ultrasound: energy from 20-100 kHz.

Characterizing the Sound Emitted by the Audio Spotlight

The output of the Audio Spotlight™ is a beam of airborne sound about 35 cm wide at the source. It consists of a carrier at 58 kHz with a maximum intensity of about 121 dB at 1 meter from the source, and sidebands resulting from amplitude modulation, of varying bandwidth and strength, depending on the spectrum of the modulation source. The modulation source can be any audio signal below 20 kHz. In addition, the demodulation of the modulated ultrasound in air generates sound in the audio range below 20 kHz, typically 30 dB or lower in intensity than the ultrasound carrier.

Figure 1 is an example of the output of the Audio Spotlight™ with an input of a 2 kHz pure tone audio signal. These measurements were made at 1 meter from the transducer using a B&K 4939 microphone and a B&K real time analyzer with Pulse software. Note that the sound is quite complex and wideband in the ultrasonic range, even with a pure tone input, but that the demodulated signal in the audio range is a relatively pure tone at 2 kHz. The most intense of the ultrasonic sidebands are below the carrier at 121 dB.

The form of coupling of the ultrasound to the human subject is important. Ultrasound can be airborne, or it can be fluid or solid-coupled. Because a substantial impedance mismatch exists between the air and most solids or fluids (including the human body), with airborne ultrasound very little energy is actually transferred into the object. For the case of airborne ultrasound impacting human skin, less than 0.1% of the energy is absorbed (Wiernicki & Karoly, 1985). On the other hand, solid or liquid coupling is far more efficient, and is typically used in industrial applications. One source of risk in industrial applications is the unintended transfer of acoustic energy into the human operator through solid or liquid coupling. The Audio Spotlight uses only air-coupled ultrasound. There is never an opportunity to come into contact with the transducer.

Nevertheless, the transducer mounting could conceivably provide a source of coupling. We have used an accelerometer to measure the ultrasound conducted through our system of mounting. With the transducer array suspended from a microphone stand by a wire attached by screws at two points on its periphery, we measured acceleration at the base of the stand at the carrier frequency of 58 kHz with an

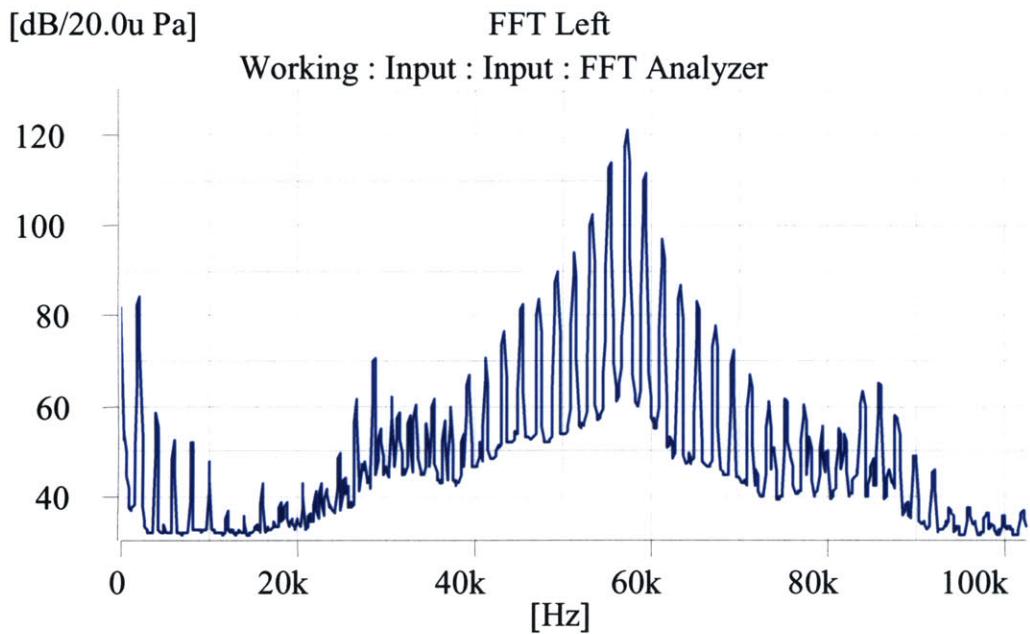


Figure 1. Spectrum of the maximum output of the Audio Spotlight at a 1 meter distance with a 2 kHz audio input. The sound at 1 meter from the transducer was input to a B&K real time analyzer with Pulse software, using a B&K 4939 microphone. The audio input to the Audio Spotlight was adjusted to produce the maximum possible audio and carrier output with distortion of the audio approximately 30 dB below the 2 kHz fundamental (see Figure 5). Note the absence of significant energy in the high audio range (10-20 kHz).

output of 121 dB in air at 1 meter. There is a negligible amount of coupling to the support system, therefore the stimulus is air only. The level of the measured acceleration at 58 kHz, with the accelerometer coupled to the back of the transducer assembly, was -20 dB re 1 m/s². This is well below the OSHA threshold level value of +15 dB re 1 g rms (1 g = 9.80665 m/s²) measured at the mastoid of the head and not on the source.

Risks of Exposure to Airborne Ultrasound

Potential risks of exposure to ultrasound include specific effects on the auditory system, and more general physiological effects on the body. Specific effects on the auditory system include temporary and permanent threshold shifts and tinnitus caused by damage to the inner ear, and the effects of pressure on the tympanic membrane and middle ear. The more general physiological effects include heating due to ultrasound absorption, and other physiological effects such as nausea that have been reported in some studies.

Most of the research on the adverse effects of ultrasound deals with directly coupled ultrasound. Little research has been performed to directly assess the risks of low frequency airborne ultrasound. Current exposure standards are based on the concept that detectability and the potential for damage to hearing are related. The current American Conference of Governmental Industrial Hygienists (ACGIH) TLVs accepted by OSHA offer guidelines based on two lines of argument: the bottom up approach addressing detectability of directly coupled ultrasound and the top down approach based on evidence of damage from exposure. Below we discuss the reasoning and data from these two approaches, but emphasize the absence of data relevant to purely airborne ultrasound, and the need to collect data specifically for the Audio Spotlight™.

Detectability: The Bottom-Up Approach

The bottom up approach addresses detectability of directly coupled ultrasound referenced to either acceleration (re: 1 g rms) (Lenhardt et al., 1991), or sound pressure in water (re: 20 µPa [not 1 µPa]) (Corso, 1963). If ultrasonic thresholds are plotted in reference to thresholds at 8 kHz, the data of Lenhardt et al. and Corso are in good agreement. In the range from 60 to 100 kHz, the typical threshold for detection of directly coupled ultrasound is above the present level. These data provide only the sensitivity of the inner ear to ultrasound when there is direct and efficient coupling, and therefore cannot be used directly to determine thresholds for airborne ultrasound. The medium is not considered to play any role other than its impedance matching to skin. In the case of airborne ultrasound the medium may have a more substantial influence on not only the intensity but on which frequencies are detected. In particular, nonlinearities in the sound path may result in the generation of subharmonics in the audible range.

Subharmonics

The reports in the 1960s (e.g., Parrack, 1966) often made reference to the effects of ultrasound subharmonics in the audible range that could be a partial or even primary cause of adverse reactions. Ultrasonic sources near 20 kHz would be expected to generate energy in the high audio frequencies. It seemed that it was ultrasound subharmonics that represented the risks for hearing and related effects. However, while subharmonics must be considered as high audio frequency sound exposure and should be treated appropriately, one cannot rely on subjective reports that the energy is only in the high audio frequencies. There is considerable evidence that the pitch of audible ultrasound (directly coupled) is identical to that of the audio frequencies from about 7 – 16 kHz. Both would sound the same, and the presence of both may increase any damage risk associated with either.

The source of subharmonics, assuming that the ultrasound source is totally above 20 kHz, can be due to vibrational nonlinearities in the (1) air, (2) canal/tympanum, (3) middle ear, and (4) cochlea. The Audio Spotlight™ is based on the principle of demodulation by nonlinearities in the air, and the intensity of the audio frequency sound will be addressed later in this report. The remaining three sources are in the ear.

Spectral measurements in the ear are straightforward, but identifying the site of subharmonic generation, if any, will be difficult. A probe microphone inserted in the ear can record pressure and spectra at the tympanum. Unfortunately since the tympanum and middle ear ossicles are coupled, there will not be site specificity of the subharmonic source. Measuring cochlear subharmonics will be even more elusive. It is well established that the basilar membrane is non-linear and that there is an active motor system (outer hair cells) in the cochlea that increases non-linearity. Intense ultrasounds, like intense high audio frequencies, would be expected to generate lower frequencies not present in the stimulating source.

The potential role of cochlear subharmonics is not trivial, in that auditory scientists will expect their presence and assume that they are part of what is perceived, even in the absence of measured data. Providing convincing evidence for lack of cochlear subharmonics resulting from ultrasonic exposure will be difficult, if that is indeed the case. Von Gierke (1950) proposed that the middle ear is activated at subharmonics of intense ultrasonic exposure, resulting in subharmonics recordable at the tympanum (cited as personal communication in Karl Kryter, 1960, p. 538). Henning von Gierke repeated his concept to me in November of 1996, further suggesting that the middle ear then stimulates the inner ear at these subharmonics in the same fashion as conventional air conduction hearing. The ossicles are suspended from the middle ear cavity and lag behind (inertia) during bone conduction stimulation. The resonance of the middle ear is about 2.5 kHz, with little transmission over 3 kHz imparted with impact stimulation. If there is overpressure in the canal, the middle ear ossicles could be displaced, with induced oscillation at their natural frequency. This would be observed as a 2-3 kHz sound in the canal with the tympanum acting as a reverse driver.

If audio subharmonics are eliminated, there is a general consensus by early researchers that airborne pure ultrasound (>25 kHz) would not cause any temporary hearing loss even at levels up to 140 dB SPL (Parrack, 1966; Acton, 1968; 1974; cited by Wiernicki and Karoly, 1985). This observation is consistent with the current OSHA standard for uncoupled ultrasound.

The thresholds for airborne ultrasound itself will be higher than the thresholds for directly coupled ultrasound due to the filter characteristics of the ear canal, tympanum and middle ear and the lack of absorption of ultrasound by the body (<0.1%). Threshold changes due to inner ear damage can be either temporary (temporary threshold shift or TTS) or permanent. For airborne ultrasound between 20 and 35 kHz, Parrack (1966) found a TTS with exposure to about 150 dB, but this result may be due to ear-generated subharmonics (Acton, 1974). From the bottom-up approach, airborne low frequency ultrasound, in the absence of any subharmonics in the audible range, should not be detectable at levels less than 145 dB SPL. Therefore, exposure at levels less than 145 dB should not result in any inner ear damage.

Evidence From Exposure: The Top-Down Approach

Equally difficult is the quantifying of the so-called subjective effects of ultrasound: dizziness, tinnitus, and a sense of ear fullness. This leads to the second approach to assessing risk, the top-down strategy, in which reports of these adverse effects at particular sound levels are used to predict risk. Again, there exist no quantitative data for airborne ultrasound at the intensity emitted by the Audio Spotlight™, although there are laboratory observations.

The two studies that established the human ultrasonic audiogram are Corso (1963) and Lenhardt et al. (1991), but Deatherage et al. (1953) was the first to urge caution with ultrasonic power. Bruce Deatherage (personal communication, 1992) indicated that he developed permanent high frequency hearing loss and life long tinnitus from exposure to a 50 kHz beam in which his head was coupled (submerged). Converting his threshold to the current convention (dB SPL re: 1 uPa) the level would be 160 dB @ 50 kHz. He presumably received damage by making suprathreshold loudness judgments. He did not specify level, but reported a rapid growth in loudness with no sense of pain. The dynamic range from threshold to unpleasantly loud in our experience is about 30 dB. My best guess at the damage level at 50 kHz is ~180 dB SPL (water).

Tinnitus may be induced at lower intensities for ultrasonic frequencies (< 30 kHz) because it has been reported in industrial exposures. The spectral purity of industrial exposure is always an open question.

Clearly the presence of tinnitus creates the concern of cochlear damage. Tinnitus is difficult to ascertain objectively, except under some circumstances with imaging procedures (Lockwood et al., 1998). Subjective questionnaires and loudness and frequency matching evaluations can be used to document the extent of tinnitus.

Persistent dizziness or vertigo has not been reported at the levels of ultrasound emitted by the Audio Spotlight™. Allen et al. (1948) noted that loss of equilibrium and dizziness do not occur until about 160-165 dB at 20 kHz. However, temporary unsteadiness has been a common report in our lab when we used ultrasound at < 15 dB sensation level (SL). This sensation is short lasting, but is not induced at levels just above threshold. We did not explore the relation between ultrasonic intensity, frequency and unsteadiness systematically. My experience is that unsteadiness occurs in the first one and two steps after cessation of high level stimulation when listeners rise from seated positions. My interpretation is that some saccular stimulation by ultrasound reduced the sensitivity of this otolith organ and it is the rising from the chair that miscued some vestibular ocular-motor reaction. While recovery is very rapid, individuals with vestibular dysfunction, as well as some elderly, could be at risk of falling. As with tinnitus, a questionnaire approach can be a good start at assessing dizziness and unsteadiness.

Ear fullness is a feeling that can be induced by overpressure in the ear canal referenced to ambient middle ear pressure or to changes in middle ear pressure in reference to ambient atmospheric pressure. The most likely cause of ear fullness is due to canal overpressure pushing the tympanum, with the middle ear inadequately equalizing the pressure. This is a common symptom at infrasonic frequencies at about 130 dB or with whole body vibration in excess of 0.25 g RMS. Negative pressures observed in my laboratory for infrasound are on the order of about -100 mm H₂O, a level associated only with fullness, not pain.

Other Physiological Effects: Ultrasonic Heating

Airborne ultrasound is simply radiated energy. Any absorption of this radiated energy will create heat. The upper allowable limit of ultrasound related to heat effects depends on both the absorptive nature of the human body (which is frequency dependent) and the maximum allowable heating.

Skin is a very poor absorber of airborne ultrasound, typically reflecting 99.90% of the ultrasonic energy (Wiernicki & Karoly, 1985). Because of this, only under extremely high intensities would ultrasound cause heating of the skin. Acton (1974) and Allen et al. (1948) have shown that measurable (but slight) heating occurs near 159 dB for human skin and 145 dB for small, furred animals. These measurements were done at relatively low ultrasonic frequencies (20-30 kHz), and because this an impedance mismatch phenomenon, it is probable that these effects diminish at higher frequencies. Thus heating is unlikely to pose a significant risk for humans.

Audio Frequency Sound

The principle of the Audio Spotlight™ is that the beam of modulated ultrasound demodulates in air to yield an audio frequency component. With a 121 dB ultrasound carrier, the level of demodulated audio at 1 m from the transducer at 2 kHz has been measured at 83 dB, and at 4 kHz at 91 dB. The 91dB level could be damaging to hearing with long term exposure. Audio was not the subject of this study, as its management is the same as for conventional loudspeakers.

Because the slope of the demodulated sound goes up by about 6 dB per octave, a wideband modulating sound (e.g., equal levels at 1 kHz and 16 kHz), could, when demodulated to provide an 83 dB level at 1 kHz (not unusual for listening to music for entertainment), provide a 107 dB level at 16 kHz, which might damage hearing with extended exposure. Music and speech do not normally have equal energy at 1 and 16 kHz, but, especially in synthesized music, such sounds are conceivable.

Current regulations and limits: application to Audio Spotlight™

Wiernicki and Karoly (1985) review the evolution of exposure standards for airborne ultrasound since the 1960s. These have been generally in the region of 110-120 dB for maximum exposure limits for ultrasound greater than 20 kHz, but reflect a concern for subharmonics as well as ultrasound, and are not supported by much actual data on the risks of pure ultrasound exposure. A paper by the International Radiation Committee published in 1984 (Jammet et al., 1984) recommended that the limits of continuous occupational exposure be set at 110 dB, and 100 dB for exposure to the general public. The recommended limits for exposure to airborne ultrasound have increased over the years, as more has been understood about associated risks.

Until recently the Occupational Safety and Health Administration (OSHA) recommended limit was 115 dB for continuous exposure. OSHA, whose findings are based on the American Conference of Governmental Industrial Hygienists (ACGIH), explicitly stated that this was not based on risks associated with the ultrasound itself (which they deemed to be harmless), but on the risks associated with subharmonics usually present when using industrial ultrasonic equipment.

Based on recent findings from the ACGIH, OSHA has increased its published maximum exposure limit by 30 dB to 145 dB (Appendix C). This level is in line with the power output from the Audio Spotlight, and the following studies with the Audio Spotlight itself were performed to justify the safety of this level.

Assessments for Audio Spotlight™

To fully assess the safety of the output of the Audio Spotlight, the following measurements were deemed most important:

1. Detectability of ultrasonic carrier. Can listeners detect something other than the demodulated audio? Assessment in the form of questioning.
2. Temporary threshold shift (TTS). Compare the TTS (if any) caused by the audio component alone (e.g., an 80 dB tone at 2 kHz) compared with the full modulated ultrasound waveform (including demodulated audio at the same level).
3. Impedance audiometry to detect effects of pressure on the tympanic membrane, e.g., retraction. Combine with questionnaire on feeling of ear fullness.

Methods

Subjects: The subjects were 20 individuals, between the ages of 11 and 63 years of age. Seventeen had normal audiograms. Three had some degree of high frequency loss. All were without symptoms of otological or neurological disease.

Set Up: The experimental set up is depicted in Figure 2. The subject is seated in the beam at one meter. Pictures of the recording and analysis set up are in Figures 3 and 4.

Playback Levels: Playback levels were chosen to maximize the ultrasound output of the Audio Spotlight™, with the demodulated audio at low distortion. Figure 5 shows the ultrasound and audio levels at 2 kHz for ultrasound intensities ranging from 110 dB to 125 dB, and the second harmonic distortion. The second harmonic distortion rises abruptly at about 121 dB of ultrasound. Therefore, this study used 121 dB as the maximum level of ultrasound. This resulted in a demodulated audio level at 2 kHz of about 80 dB.

Practically, the levels were adjusted before each playback to just below the distortion level, and this gave a typical 121 dB ultrasound and 80 dB audio. Appendix B gives graphs of the ultrasound playback spectra

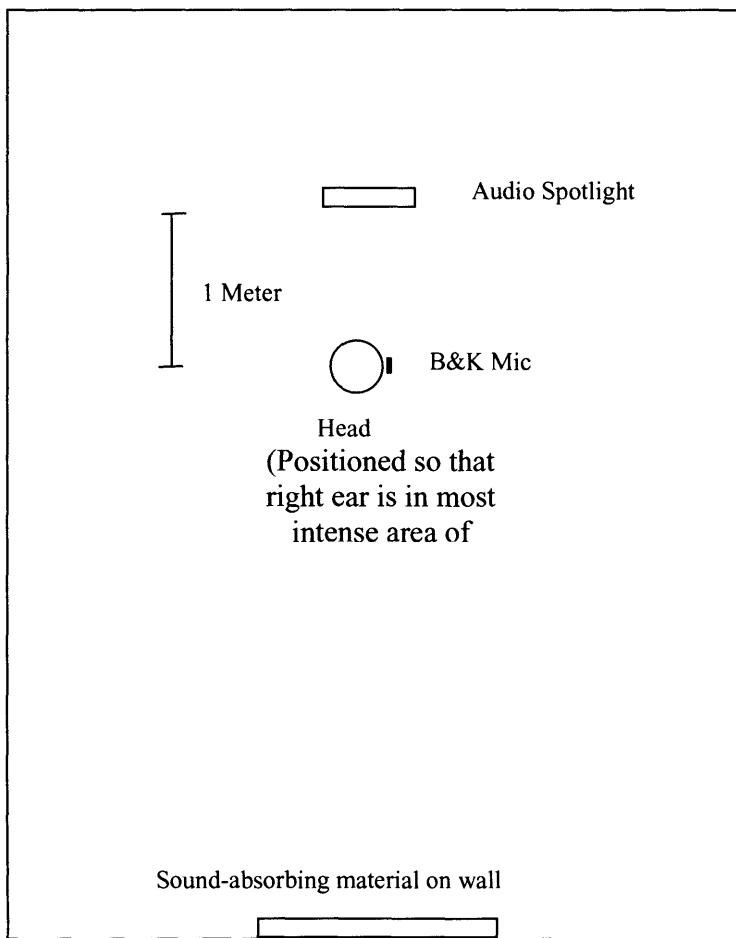
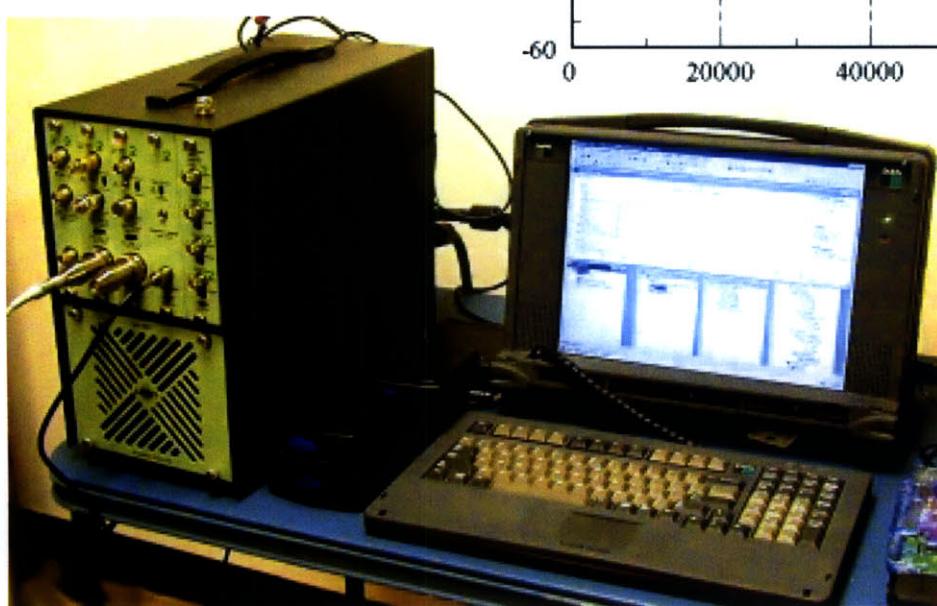
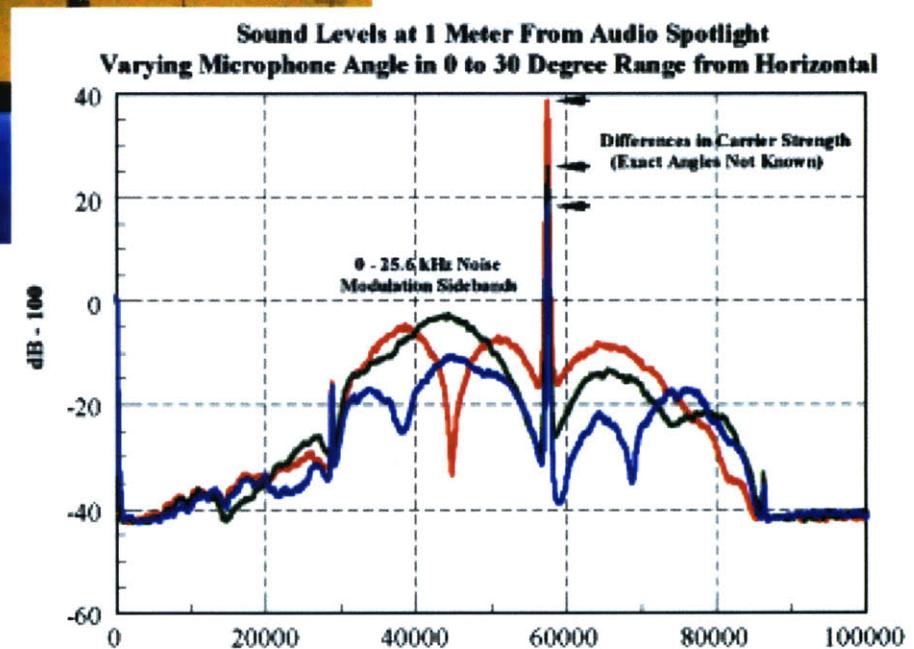
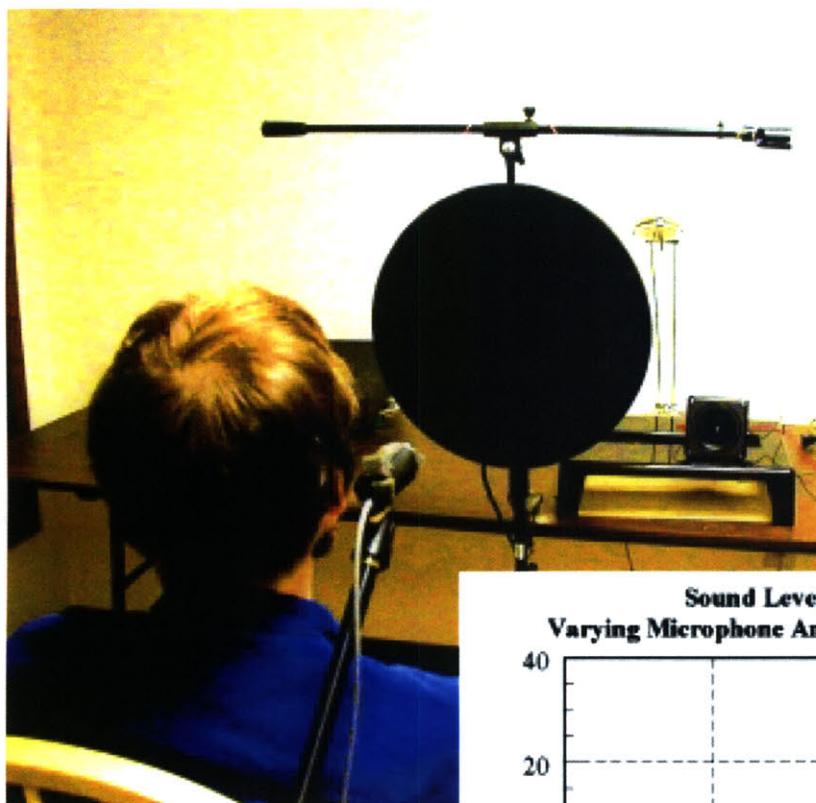
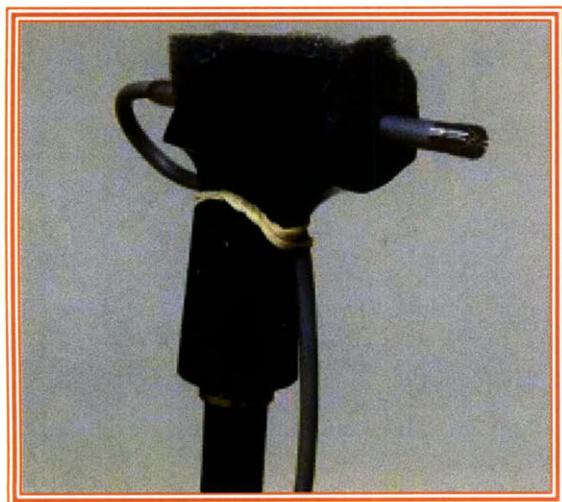


Figure 2. Experimental setup for exposure to the beam from the Audio Spotlight. Note the position of the right ear in the most intense level of the beam, with the B&K 4939 microphone next to the ear.





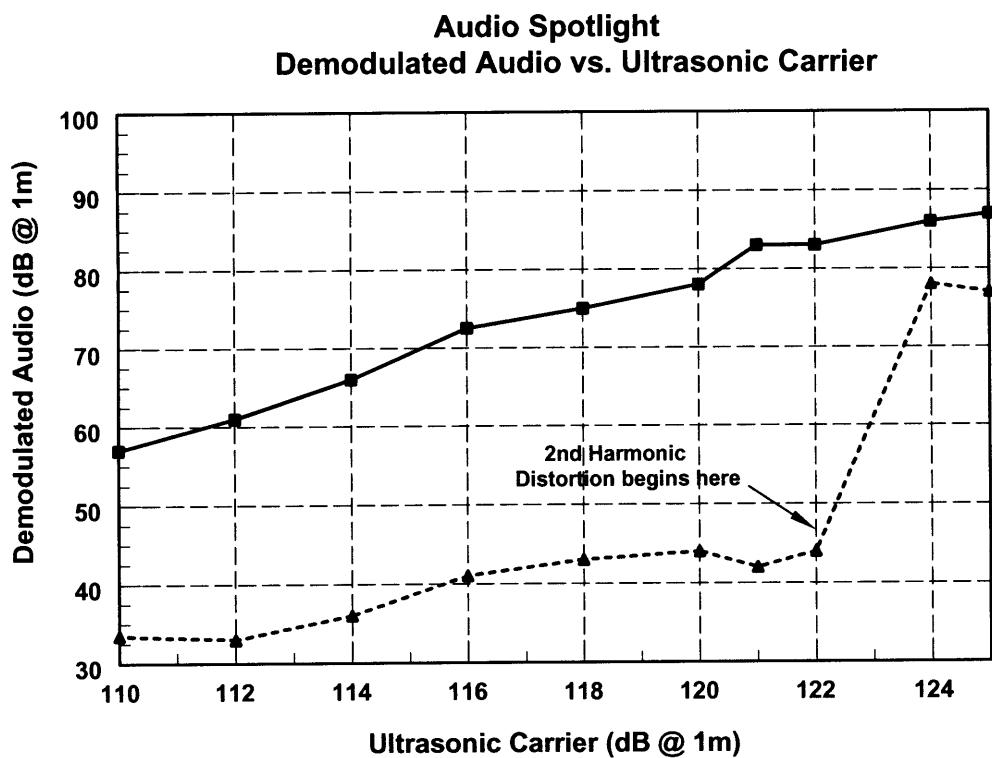


Figure 5. Comparison of demodulated audio and 2nd harmonic distortion with the level of the ultrasonic carrier at 1 meter from the transducer. There is an abrupt rise in distortion after 122 dB; 121 dB was chosen as the maximum usable carrier level.

for each subject. For the audio alone comparison, the audio level was adjusted to 80 dB at the location of the subject.

Audio Playback: The stimulus for the audio playback was the demodulated 2 kHz tone audio output from the Audio Spotlight™, recorded on DAT tape and played back through a Crown amplifier and a Radio Shack tweeter speaker. The FFT analysis of the playback stimulus (Figure 6) shows a pure tone at 2 kHz with harmonic distortion at least 30 dB below the level of the tone, and no ultrasound or high frequency audio present.

Playback Duration: The duration of the exposure to the audio and the Audio Spotlight™ was set at five (5) minutes, although two minutes of audio stimulation is sufficient to elicit a temporary threshold shift (TTS). TTS is a reversible shift in hearing as a result of sound exposure. The recovery of a TTS depends upon the stimulating sound properties, including duration. Five minute exposure durations are common in auditory research (Ward, 1962; Melnick, 1998). The five minute duration was also chosen to minimize any effects of the 80 dB demodulated tone, since sounds between 80 and 105 dB, with durations less than eight hours, carry a risk of TTS that increases linearly with sound pressure (Melnick, 1998). Parrack (1966) reported TTSSs for frequencies between 17- 34 kHz at intensities of 148-154 dB SPL for durations of 5 minutes. Contact ultrasound, applied to the round window for 20 minutes, resulted in broad hair cell damage to the base and second turn of the cochlea (Barnett, 1980). Based on the above studies, exposure duration of five minutes would be sufficiently long to elicit any auditory threshold effect. Documentation of TTS is important in that repeated occurrences may gradually cause permanent hearing loss.

Positioning of Subject: The goal was to expose the subject to the maximum ultrasound SPL. The Audio Spotlight™ beam is narrow (Figure 7), and the ultrasound intensity varies considerably within the beam. Figures 9 and 10 show the ultrasound intensity at 58 kHz across the width of the beam at 1 and 2 meters distance, respectively. The most intense sound is not in the center of the beam, but rather about 3-4 inches off center. The ultrasound is more intense and uniform at 1 meter distance. Therefore, the subject was positioned with the right ear near the most intense part of the beam at a distance of one meter. SPL was measured with a B&K microphone (4939) next to the right ear. Ambient noise in the testing room is depicted in Figure 8. Appendix B gives the FFT analysis of the Audio Spotlight™ stimulus for each subject. The levels were set for 121 dB ultrasound, and about 80 dB audio; however, due to the inherent spatial variability of the Audio Spotlight™ output (e.g., Figures 9 and 10), it is likely that the actual sound reaching the right ear varied between about 119 and 123 dB. The stimulus spectra for the individual subjects reflect this inherent variability (Appendix B). The ultrasound at the left ear, positioned outside the most intense part of the beam, was probably about 10 dB lower.

Carrier stability: Ultrasonic carrier stability was good with mild drifting from 57-58 kHz.

Tympanometry: Tympanometry was performed using a Virtual Model 310 tympanometer. Both ears were measured (1) prior to sound exposure, (2) immediately after (within 2 minutes) exposure to 2 kHz audio alone at 80 dB, and (3) immediately after (within 2 minutes) exposure to the Audio Spotlight (58 kHz carrier at 121 dB, with 2 kHz demodulated audio at 80 dB).

Hearing Thresholds: Hearing thresholds were measured using a Virtual Model 320 audiometer. Testing was performed in a quiet room and the ambient noise levels were lower than the ANSI S1.1977 standard (see Figure 8 for the noise spectrum of the room). Frequencies of 1, 2, 3, 4, 6, 8, 9, 10, 11.2, 12.5, 14, 16, 18 and 20 kHz were used. Hearing thresholds were measured immediately after tympanometry for the three tests. Hearing testing occurred approximately two minutes post exposure, which is widely accepted as a uniform reference in assessing TTS (specified as TTS₂; Melnick, 1998).

Exit Interview: Comments were solicited about the listening experience in an unstructured manner.

Sequence of Events: The sequence of the session for each subject was as follows:

1. Baseline tympanometry and audiometry (15 minutes).
2. Five minutes exposure to 2 kHz audio at 80 dB.

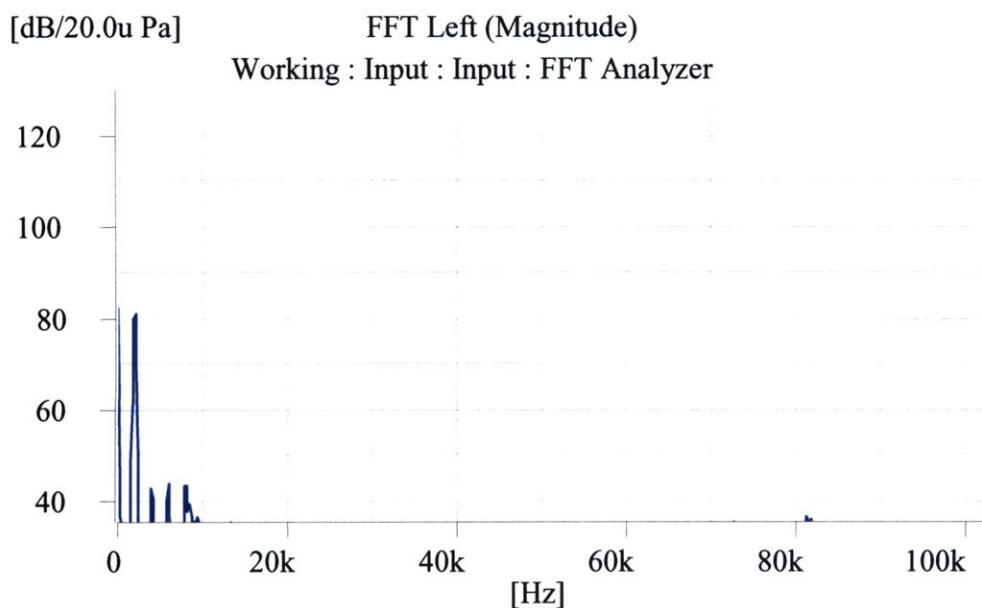


Figure 6. Spectrum of the 2 kHz audio stimulus used in the playback experiment (recorded from the Audio Spotlight at 1 meter). The ultrasonic carrier and modulation were filtered out using a combination of an audio frequency microphone and a digital low pass filter. The spectrum was produced on a B&K real time analyzer using Pulse software.

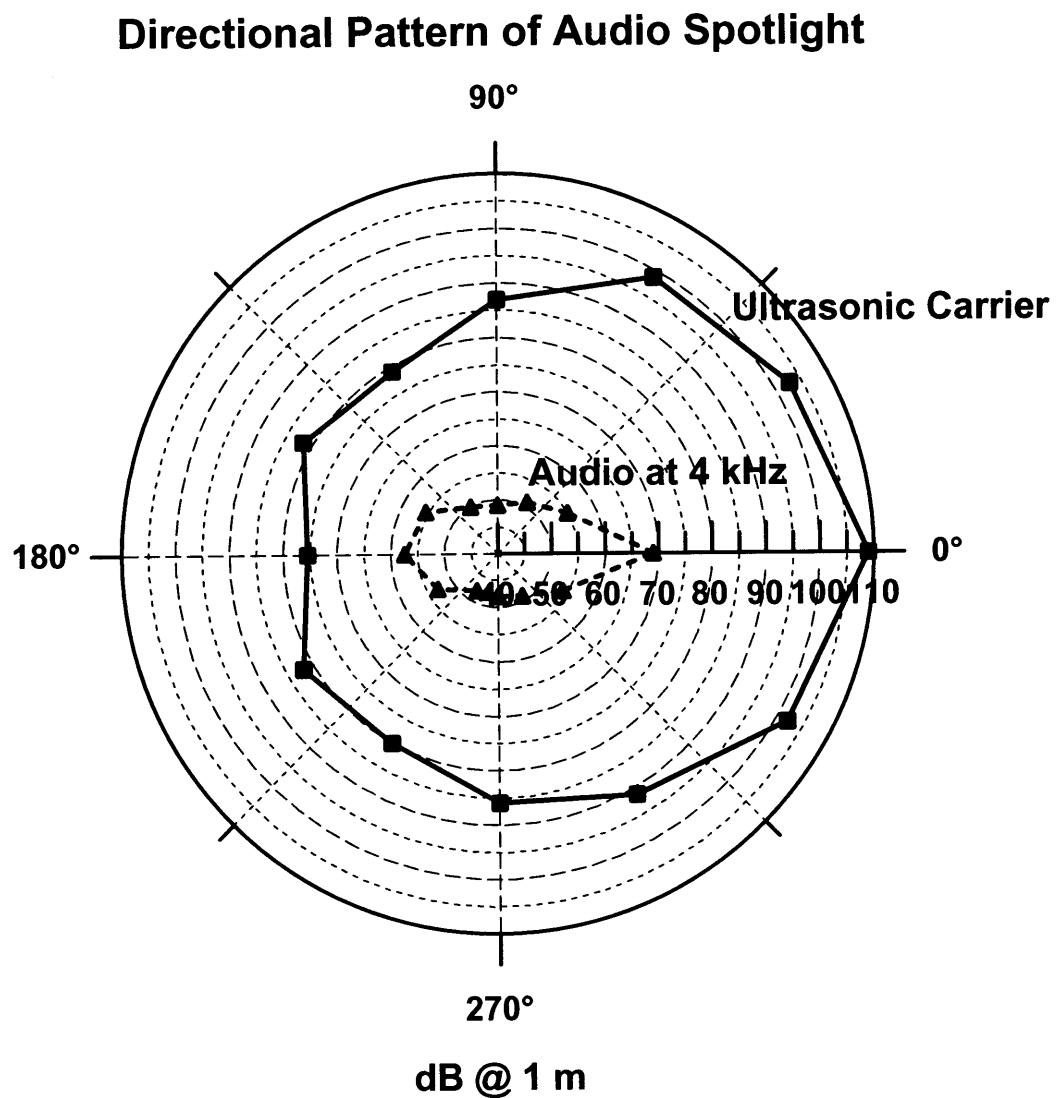


Figure 7. The directional pattern of the Audio Spotlight, measured at 1 meter in the testing room using a B&K 4939 microphone.

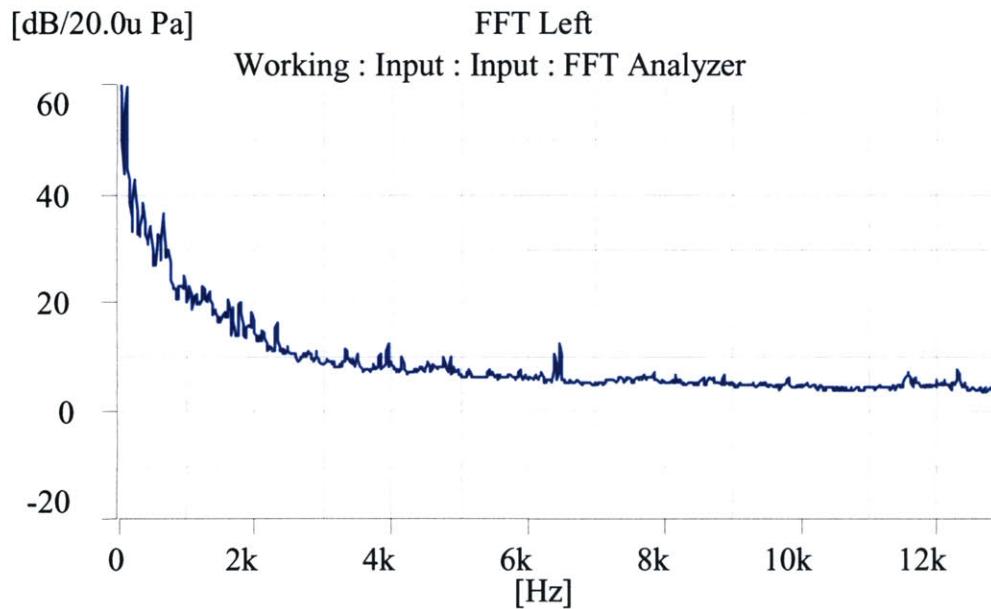


Figure 8. Ambient room noise for hearing tests, measured using a B&K 4939 microphone. Note that this background conforms to ANSI S1.1977 for audiometric testing at frequencies from 1000 to 20,000 Hz.

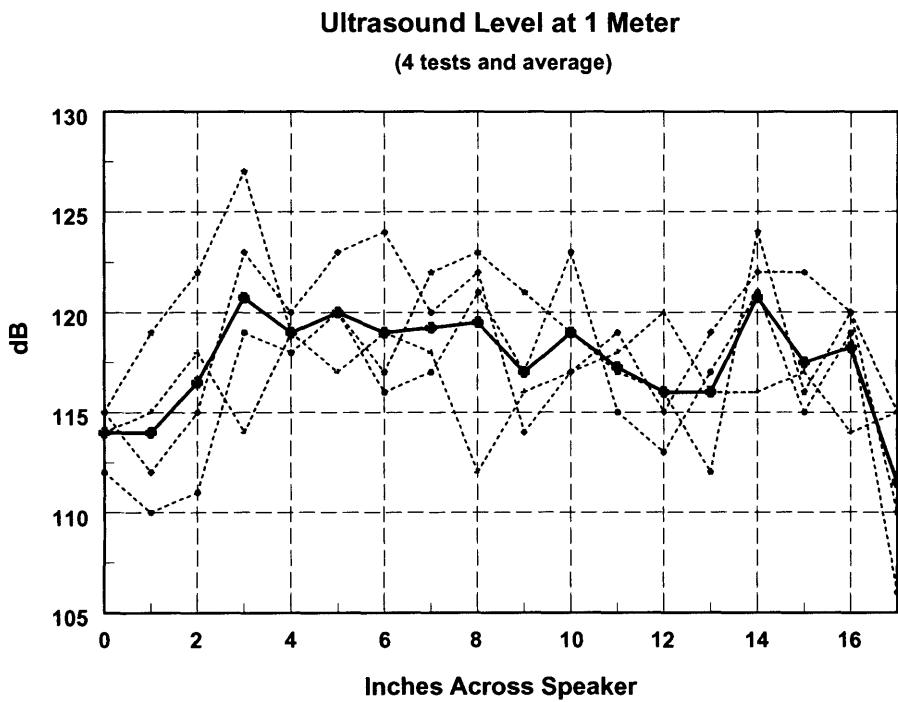


Figure 9. Ultrasound level measured across beam of Audio Spotlight at 1 meter distance with B&K 4939 microphone; 4 tests and average.

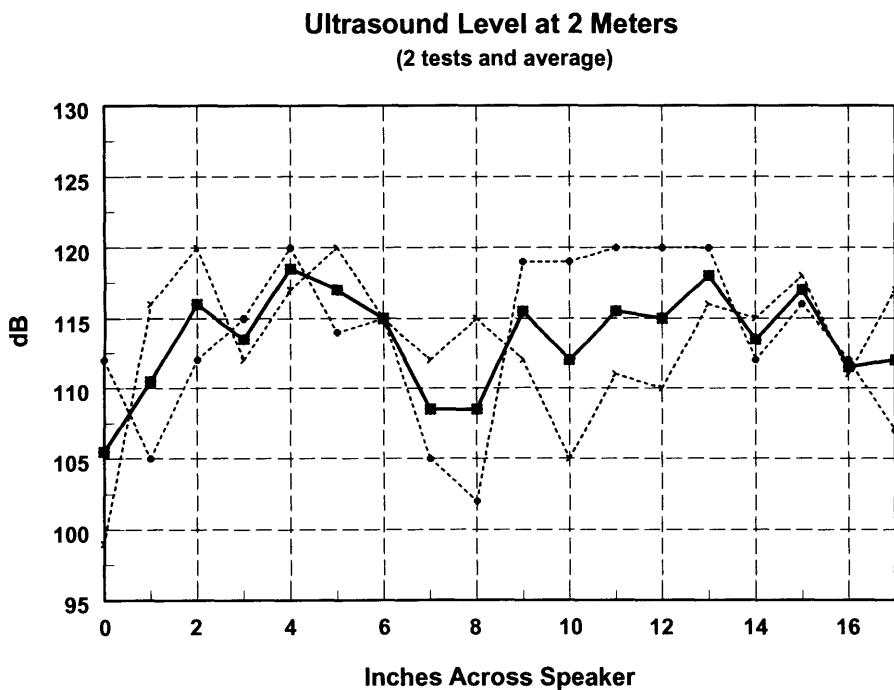


Figure 10. Ultrasound level measured across beam of Audio Spotlight at 2 meter distance with B&K 4939 microphone; 2 tests and average.

3. Tympanometry and audiometry (15 minutes).
4. Five minutes exposure to Audio Spotlight™ (ultrasound and demodulated audio).
5. Tympanometry and audiometry (15 minutes).
6. Exit interview.

Ultrasonic Threshold Determination: Separately from the playback experiments, a bone conduction threshold was determined at the ultrasonic carrier frequency (58 kHz), and equivalent SPL in air was calculated. To determine the threshold, the intensity of a just detectable 58 kHz tone coupled to the head through an ultrasonic transducer was measured using a B&K Model 4374 accelerometer. This value was compared to that of Lenhardt et al. (1991) and Corso (1963), and is used as an estimate of ear detectability with complete bodily coupling.

Results

Hearing Thresholds: To determine the effect of the ultrasound on hearing thresholds, the hearing thresholds obtained after exposure to the audio alone were subtracted from the thresholds obtained after exposure to the ultrasound. Figure 11 gives the means and standard deviations at each frequency. The mean shift at all frequencies was less than 1 dB. There was little difference between the right and left ears. For each ear, for each subject, an average of the threshold change after exposure to ultrasound at 2, 4 and 6 kHz was calculated. None of the subjects had a shift of 10 dB or greater, the current OSHA standard, and the differences among the means (< 1 dB) were not significant (*t* test; $p = 0.59$ right; 0.42 left). The average hearing threshold data, converted to dB SPL, is presented for the right and left ears in Figures 12 and 13 respectively. Appendix A is a table of threshold changes for each subject. Appendix B is the individual stimulus for each subject.

Tympanometry: To determine the effect of the ultrasound on the tympanic membrane and middle ear, the tympanograms obtained after exposure to the audio alone were subtracted from the tympanograms obtained after exposure to the ultrasound. Tables 1 and 2 give the means and standard deviations at each frequency. The mean shift and left/right difference was small, subclinical and not significant. The eardrums were not retracted from exposure to the ultrasonic beam for five continuous minutes. The normal range of middle ear pressure is from +50 to -100 daPa. There were no indications of any even transient middle ear effects from the five minute exposure to the Audio Spotlight™.

Table 1. Right ear peak pressure comparisons (daPa).

	Audio-Baseline	Ultrasound-Baseline	Ultrasound-Audio
Mean	5.6	6.9	1.3
Standard Deviation	24.3	25.6	20.3

Table 2. Left ear peak pressure comparisons (daPa).

	Audio-Baseline	Ultrasound-Baseline	Ultrasound-Audio
Mean	0.4	-2.4	-2.8
Standard Deviation	21.9	10.7	20.9

Exit Interview: None of the subjects volunteered any negative comments about their experience with the Audio Spotlight. No listener was able to detect the ultrasonic carrier; that is both the audio alone and the audio modulated by ultrasound has the same percept.

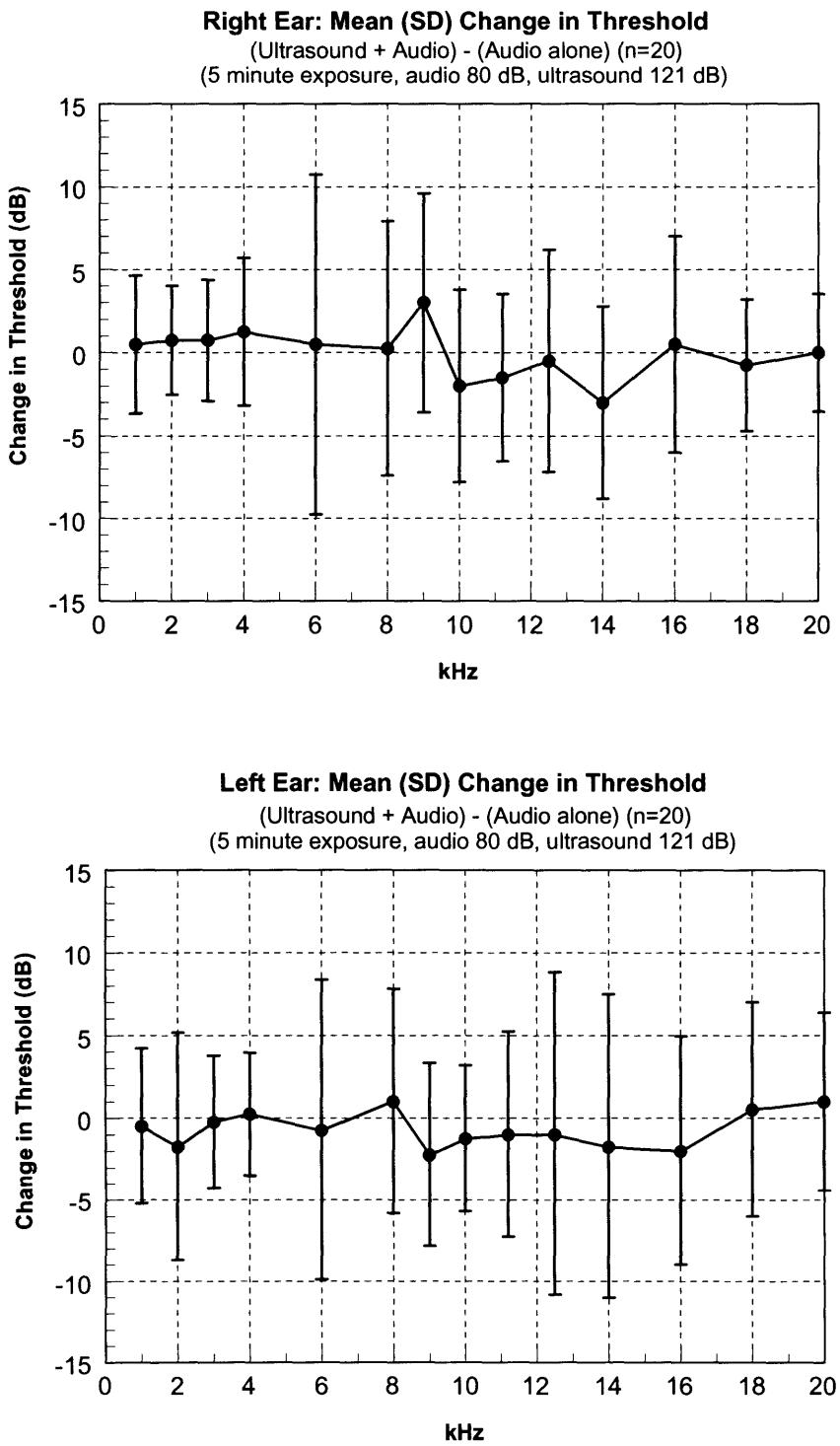


Figure 11. Changes in hearing thresholds: ultrasound with demodulated audio compared to audio alone (5 minute exposure, audio 80 dB, ultrasound 121 dB) (means and standard deviations).

Right Ear: Average Hearing Thresholds (SPL) for Baseline, Audio, and Ultrasound

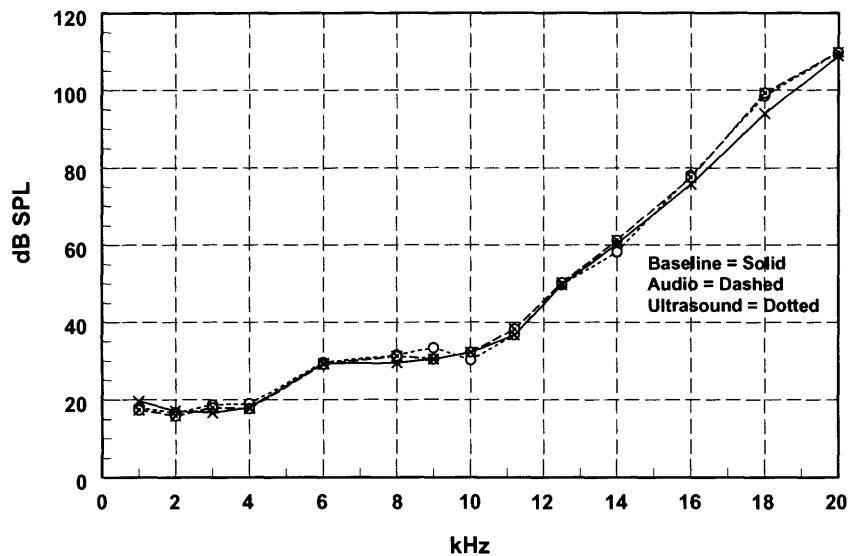


Figure 12. Right ear: average hearing thresholds for baseline, after exposure to audio, and after exposure to ultrasound plus audio (n = 20 subjects).

Left Ear: Average Hearing Thresholds (SPL) for Baseline, Audio, and Ultrasound

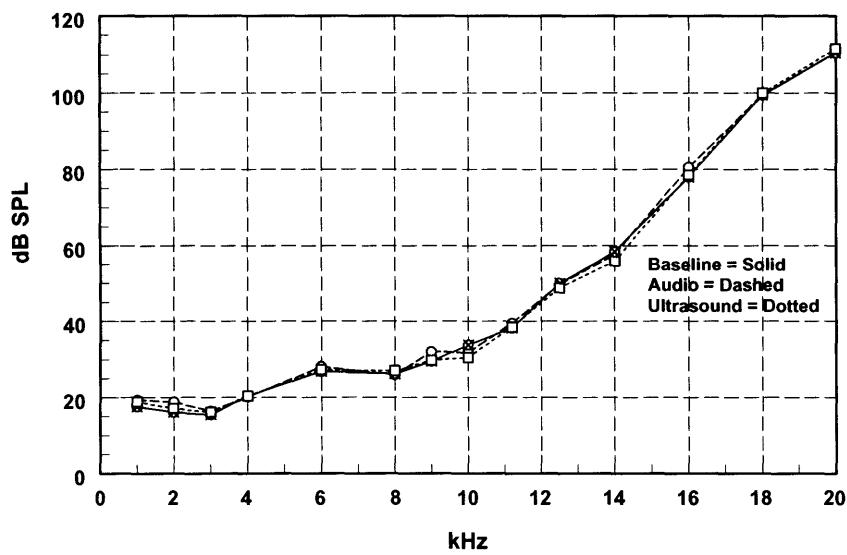


Figure 13. Left ear: average hearing thresholds for baseline, after exposure to audio, and after exposure to ultrasound plus audio (n = 20 subjects).

Ultrasonic Threshold Determination: The ultrasonic threshold allows the calculation of the extrapolated lowest possible ultrasonic detection in the case of direct bodily coupling. For the carrier frequency of interest (~60 kHz) the extrapolated threshold of audibility is 140 dB SPL. Deatherage reported a similar threshold (150 dB @ 60 kHz) with damage occurring at about 180 dB (see graph for other ultrasonic frequencies in Figure 14).

Discussion

The concept of ultrasound being above the range of hearing is a misnomer. Under proper coupling conditions frequencies thought to be inaudible are audible. Further, exposure to intense ultrasound, in a coupled medium, has resulted in hearing loss and tinnitus. Under the current ACGIH and OSHA individuals thought to be at low risk are those that are stimulated with airborne ultrasound as contrasted to those having water or substrate coupling. The large impedance difference between sound in air and the body affords some degree of protection. The device under evaluation uses airborne ultrasound as a carrier to send audiofrequencies in a beam to listeners. The basic question asked is the beam safe in regard to hearing risk. The data from the five minute exposure at 121 dB SPL are quite clear, the thresholds of hearing in the audible range revealed no systematic effect of the ultrasound, i.e. threshold shift, nor was there any indication of negative middle ear pressure. No subject spontaneously reported any negative impressions of the ultrasonic exposure. The exit interview was unstructured and was not designed to elicit any specific responses in regard to pleasantness, safety or sound quality.

Although this study is in agreement with the ACGIH's position that the older TLVs of 115 dB SPL were too stringent when applied to airborne ultrasound, the study is limited in terms of carrier frequency (~60 kHz) and exposure time. There are no available assessment data on other airborne ultrasonic frequencies or duration of exposure. Human sensitivity (threshold of audibility) does decrease with frequencies lower than 60 kHz, and there is no empirical evidence to support hearing safety in this region. Duration of exposure has not been systematically studied; hence only five minutes of exposure can be said to be safe in regard to hearing risk.

The output of the Audio Spotlight™ is broadband ultrasound, with the highest spectral peak being the carrier at 58 kHz. This peak presents the greatest risk to hearing if the body absorption is equal across frequencies. High frequency sensitivity, even in the ultrasonic range, is poorer with increasing frequency up to approximately 60 kHz, the point where asymptote is reached (see Figure 14). The extension of energy into the high audio range (10-20 kHz) may also pose a threat in that the ear is relatively more sensitive. With the 58 kHz carrier, there is no appreciable energy in the high audio range. Any increase of the bandwidth of the ultrasound down into the high audio range could present an additional hearing risk. In addition, high audio energy might appear if the carrier were to be moved down in frequency (e.g., to 40 kHz).

With the present carrier frequency and bandwidth of the Audio Spotlight™, it should be possible to raise the level by 10-20 dB (e.g., 132-142 dB SPL) and still not cause hearing threshold shifts for short periods of exposure (e.g., five minutes in the present study). However, this is only speculation at this time without specific data. Alternatively, if the carrier is shifted upward in frequency say to 80-90 kHz, the bandwidth would also shift upward possibly affording an increased safety margin.

How ultrasonic energy is summed in the ear is not known. A conservative approach is to assume that the equal energy hypothesis holds. That is, equal amounts of sound energy will produce equal amounts of hearing impairment regardless of how the sound energy is distributed in time. Accordingly a very high frequency carrier with a narrow bandwidth should be preferred over energy spread into the high audio frequencies.

Hearing sensitivity is another variable not adequately explored in this study. The selection of seventeen subjects ranging from sub-teenagers to mature adults was a reasonable sample for the initial study. Further, three young –older adults with some high frequency sensorineural loss were also included. Although generalizations are not possible on such a small group, those with hearing loss did reveal more

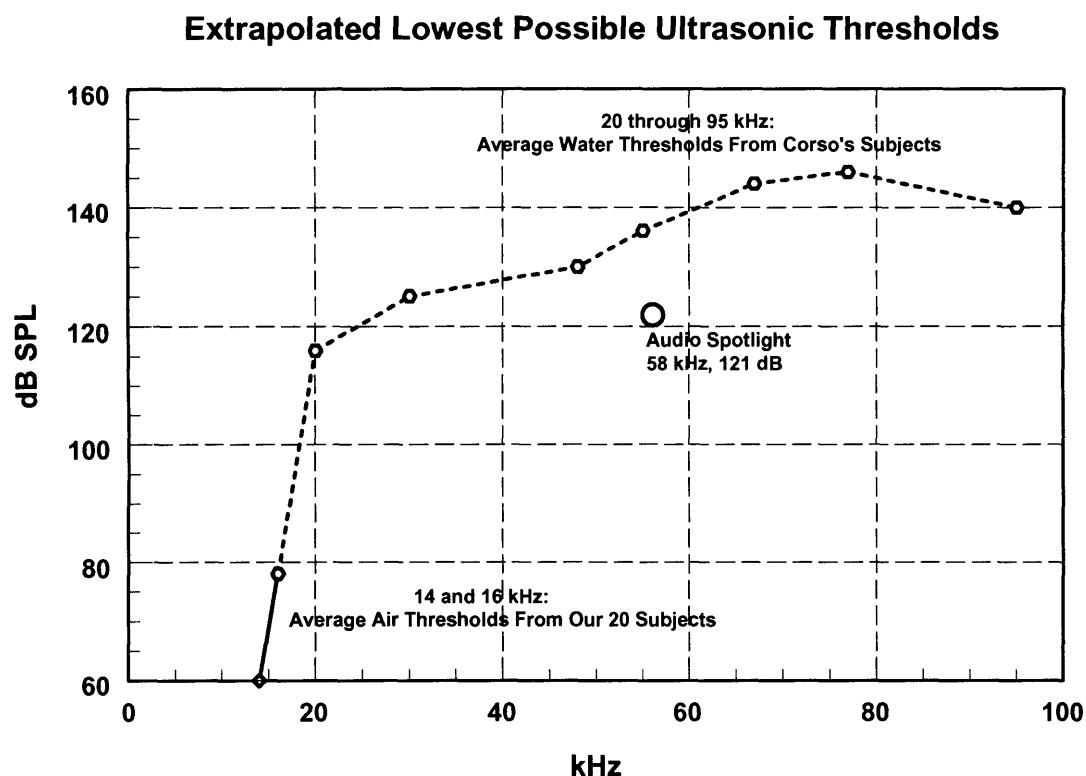
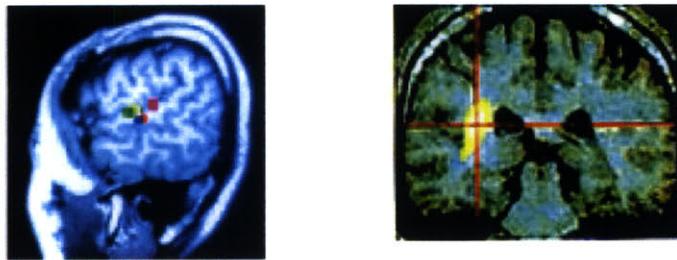


Figure 14. Estimate of lowest possible ultrasonic detection thresholds in air, extrapolated from the average thresholds in this study from 20 subjects (14 and 16 kHz), combined with data from Corso (1963) from 14 to 95 kHz.

variability in threshold pre and post testing. There may exist some sensitivity to ultrasound in hearing impaired listeners, although intuitively just the opposite might be assumed. That is to say some hearing loss may not decrease the risk of additional hearing loss, but might itself be a risk factor. Young children with very sensitive hearing, or children and adults with hypersensitivity may also respond differently than the subjects in this study.

It is important to recognize that ultrasound is processed in the auditory brain much like audio frequencies. Below are the imaging results of Hosoi et al., 1998 (left) and Imaizumi et al., 2001 (right). The primary auditory cortex is tonotopically organized and ultrasonics activate near the region activated by audio frequencies. There should be no doubt that the human nervous system can be stimulated by ultrasound if the body is coupled to the source. There is very poor airborne coupling, hence the lack of detection of the 121 dB carrier by listeners in this study seated at one meter away from the source.



Just as the brain is tonotopically organized so is the cochlea. Ultrasound is assumed to be coded at the basal tip with audio frequencies code progressively apically as a function of frequency. In this study using 2 kHz as the modulating frequency, stimulation would occur some 15 mm from the base. Thus there would be little likelihood of any spatial overlap between the carrier and modulator. If, however, a complex high frequency stimulus such as music was modulated with an ultrasonic carrier, the potential for intracochlear interaction exists, but only if there is some degree of ultrasonic coupling. Also, because of the nature of the beam, higher modulating frequencies are more intense.

Report Conclusions

1. Under the conditions of the study (five minutes of 58 kHz ultrasonic stimulation at 121 dB SPL and 2 kHz at 80 dB), no statistically significant auditory effects, either temporary or permanent were found. The lack of any hearing effects suggests that listening to audio frequencies via a loudspeaker or the Audio Spotlight™ at equivalent intensities (e.g. 80 dB SPL) is comparable.
2. This conclusion is not to be interpreted as that there was no ultrasonic absorption by the body, but the absorption was not sufficient to be detected in the testing methods employed.
3. No subjective effects were reported by the subjects under the present conditions.
4. Under the conditions tested, the Audio Spotlight is in full compliance with the ACGIH and OSHA regulations in regard to hearing safety for airborne ultrasound.

References

- Acton WI Criterion for the prediction of auditory and subjective effects due to airborne noise from ultrasonic sources. *Ann. Occup. Hyg.* 1974;11:227-234.
- Acton WI. The effects of industrial airborne ultrasound on humans. *Ultrasonics* 1974;12:124-128.
- Allen CH, Frings H, Rudnick I. Some biological effects of intense high frequency airborne sound. *J. Acoust. Soc. Am.* 1948;20(1):62-65.
- American Conference of Governmental Industrial Hygienists (ACGIH): <http://www.acgih.org>
- Barnett, SB. The effects of ultrasonic irradiation on the structure integrity of the inner ear labyrinth. *Acta Otolaryngol.* 1980;89:424-432.
- Corso, J.F. B one conducted threshold for sonic and ultrasonic frequencies. *J. Acoust. Soc. Am.* 1963;35:1738-1743.
- Deatherage, B.H., Jeffers, L.A. and Blodgett, H.C. J. A note on the audibility of intense ultrasound. *J. Acoust. Soc. Am.* 1953;26:583
- Hosoi, H et al. Activation of the auditory cortex by ultrasound. *Lancet* 1998;351:496-497.
- Imaizumi, S. et al. Ultrasound activates the auditory cortex of profoundly deaf subjects. *NeuroReport* 2001;12:583-586.
- Jammet HP, et al. Interim guidelines on limits of human exposure to airborne ultrasound. *Health Physics* 1984;46(4):969-974.
- Kryter K. The effects of noise on man. New York, 1960
- Lenhardt ML, Skellett R, Wang P, Clarke AM. Human ultrasonic speech perception. *Science* 1991;253:82-85
- Lenhardt ML. Audible ultrasound revisited. *The Hearing Review* 1998;5:50-52.
- Lockwood AH, Salvi RJ, Coad ML, Towsley ML, Wack DS, Murphy BW. The functional anatomy of tinnitus: Evidence for limbic system links and neural plasticity. *Neurology* 1998;50:114-120.
- Melnick W. Hearing loss from noise exposure. In Harris CM (ed.). *Handbook of Acoustical Measurements and Noise Control*. Woodbury, NY: Acoustical Society of America, 1998:18.1-18.19.
- Occupational Safety and Health Administration (OSHA): <http://www.osha.gov>
- OSHA Guidelines for Ultrasound Exposure: http://www.osha-slc.gov/dts/osta/otm/otm_iii/otm_iii5.html#table_iii:5_4
- Parrack HO. Effects of airborne ultrasound on humans. *International Audiology* 1966;5:294-308.
- Von Gierke, H.E., Subharmonics generated in the ears of humans and animals at intense sound levels. *Federation Proceedings*, 1950, 9, 130-145.
- Ward WD. *J. Acoust. Soc. Am.* 1962;34:1610.
- Wiernicki C, Karoly WJ. Ultrasound: biological effects and industrial hygiene concerns. *American Industrial Hygiene Association Journal* 1985;46(9):488-496.

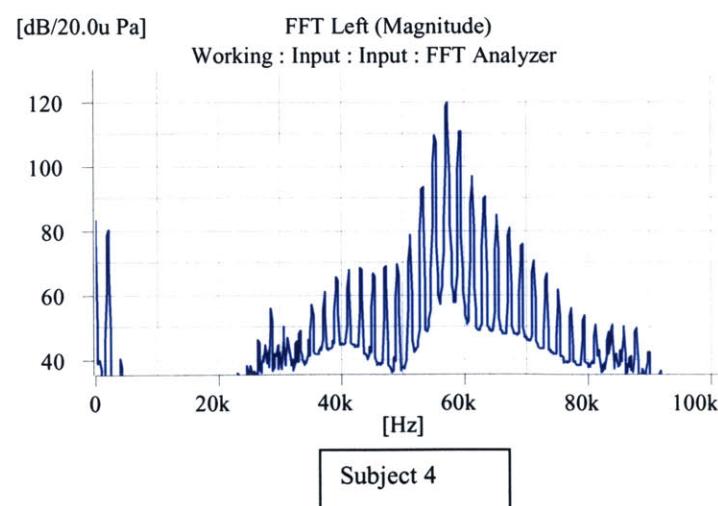
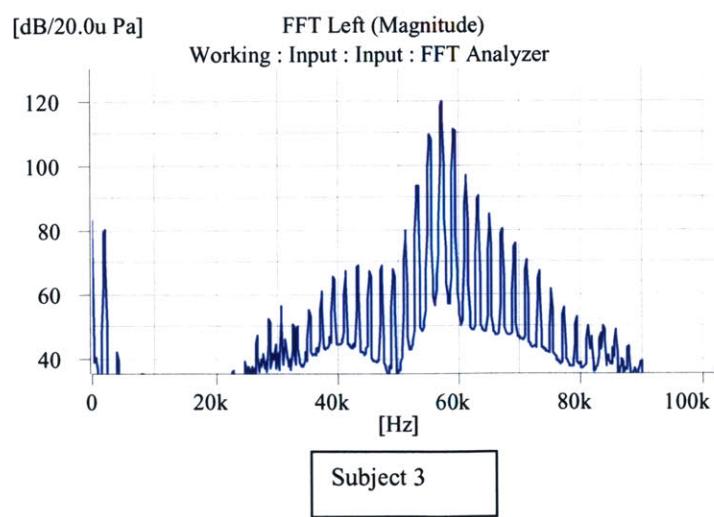
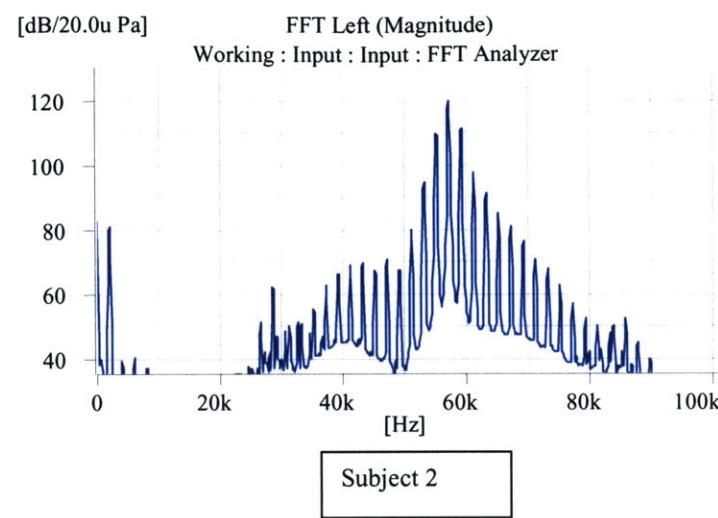
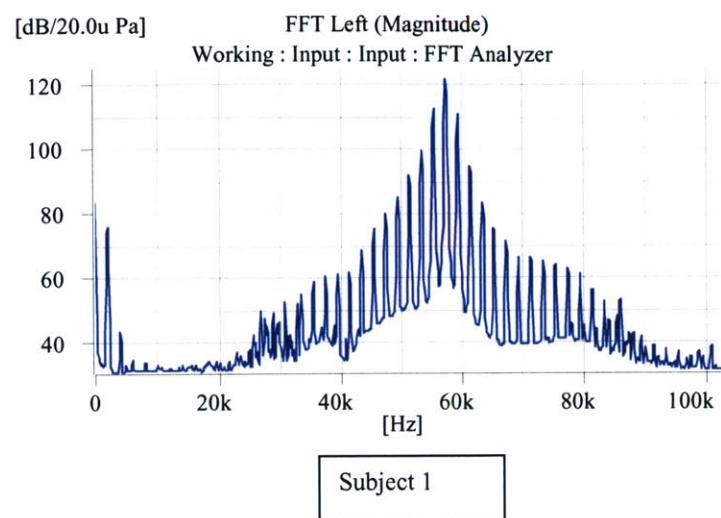
Appendix A.1

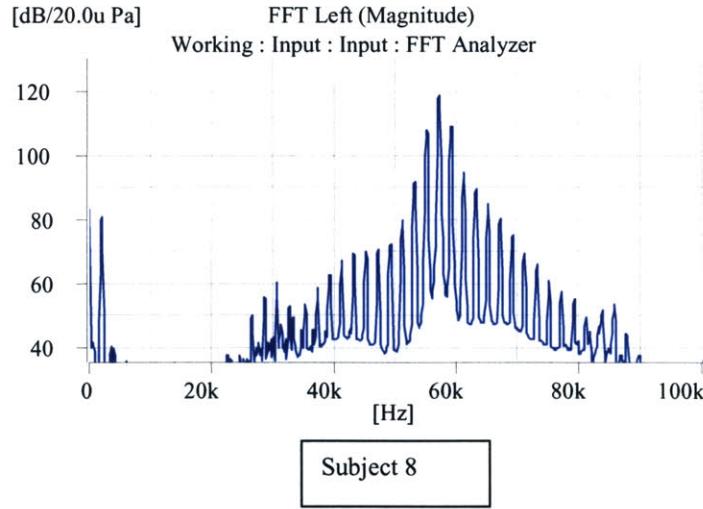
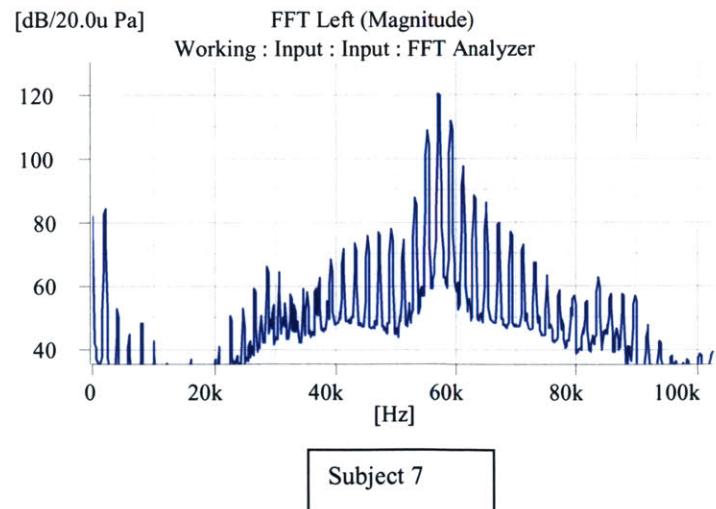
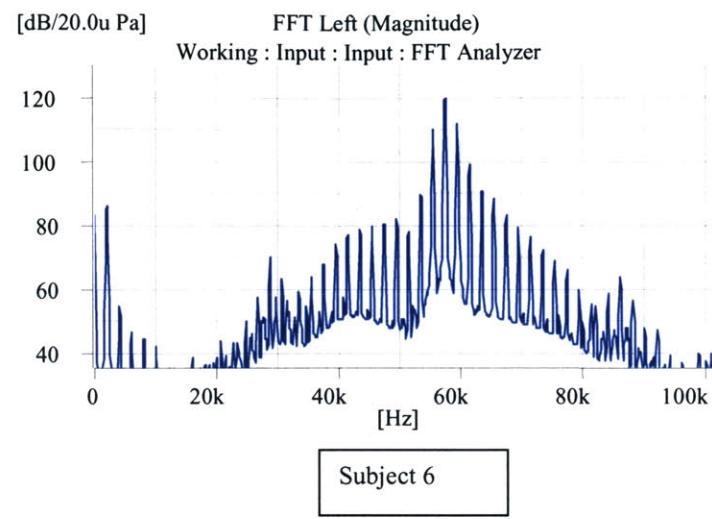
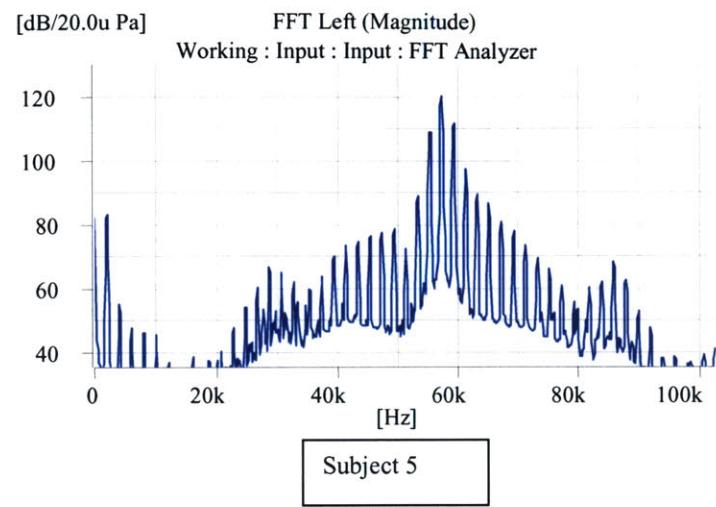
Average of threshold shifts (dB) at 2, 4 and 6 kHz for Audio-Baseline, Ultrasound-Baseline, and Ultrasound-Audio, and paired t-test of the mean difference. Numbers in bold type compare the exposure to ultrasound and audio, the primary measure in this study.

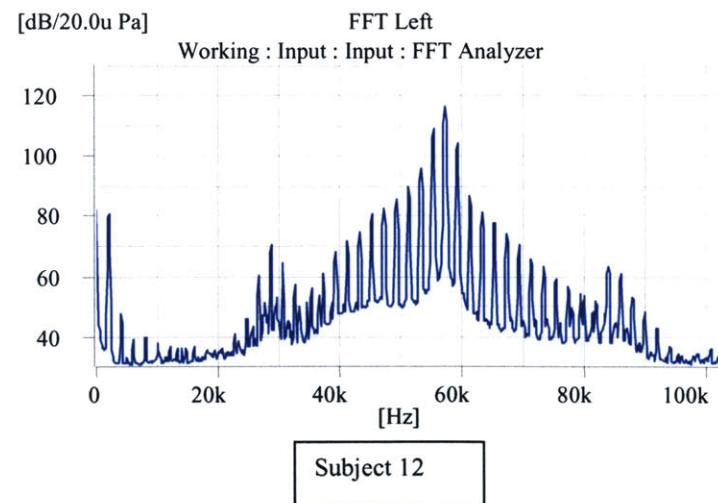
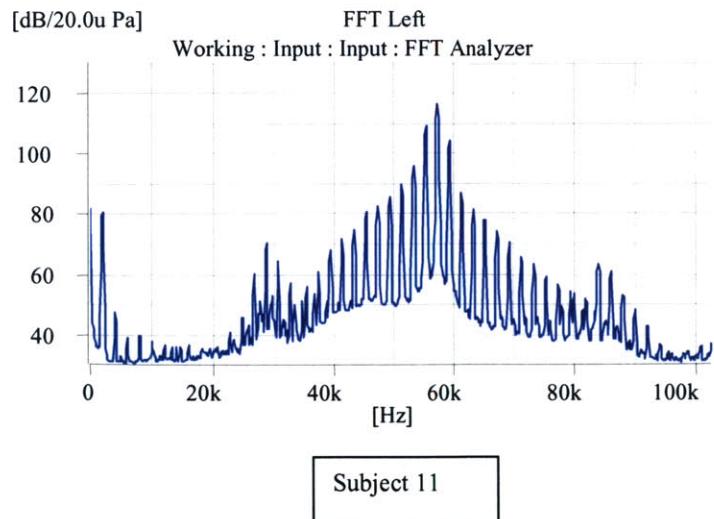
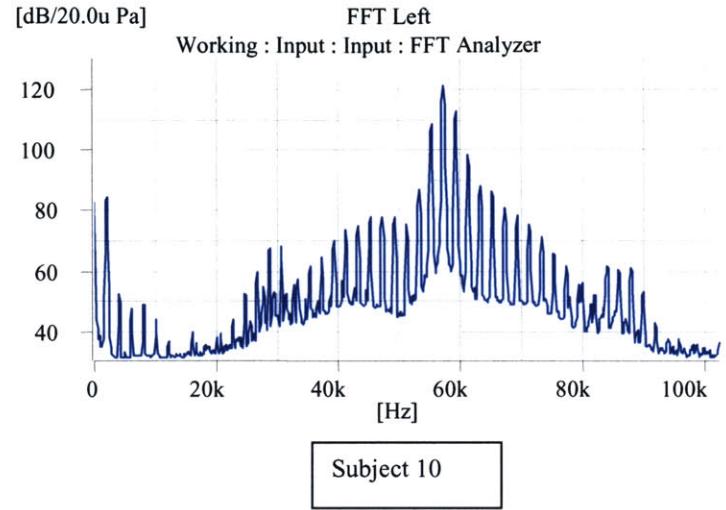
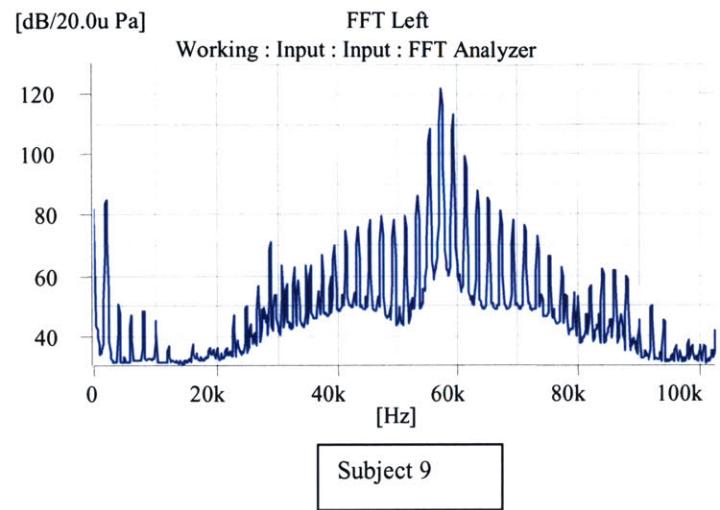
	Audio-Baseline		Ultrasound-Baseline		Ultrasound-Audio	
	Right	Left	Right	Left	Right	Left
Subject #1	1.7	1.7	5	3.3	3.3	1.7
2	3.3	10	3.3	-1.7	0	-11.7
3	-1.7	0	6.7	-1.7	8.3	-1.7
4	-5	3.3	-5	3.3	0	0
5	1.7	1.7	0	5	-1.7	3.3
6	0	0	-3.3	0	-3.3	0
7	-6.7	0	1.7	0	8.3	0
8	-3.3	8.3	-1.7	3.3	1.7	-5
9	0	0	3.3	5	3.3	5
10	-3.3	-6.7	-3.3	-6.7	0	0
11	1.7	-5	3.3	-3.3	1.7	1.7
12	-3.3	3.3	1.7	-3.3	5	-6.7
13	-17	-1.7	3.3	-1.7	5	0
14	11.7	-10	1.7	-5	-10	5
15	-10	0	-10	1.7	0	1.7
16	1.7	1.7	1.7	-1.7	0	-3.3
17	0	0	0	0	0	0
18	-1.7	1.7	-1.7	0	0	-1.7
19	1.7	10	-3.3	11.7	-5	1.7
20	1.7	6.7	1.7	1.7	0	-5
Mean Diff	-0.6	1.3	0.3	0.5	0.8	-0.8
t	0.59	0.29	0.88	1.10	0.54	0.83
p	0.56 (ns)	0.78 (ns)	0.39 (ns)	0.59 (ns)	0.59 (ns)	0.42 (ns)

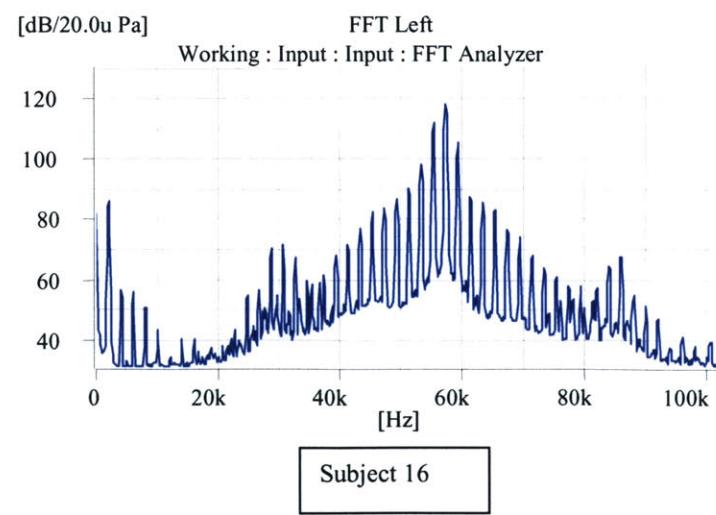
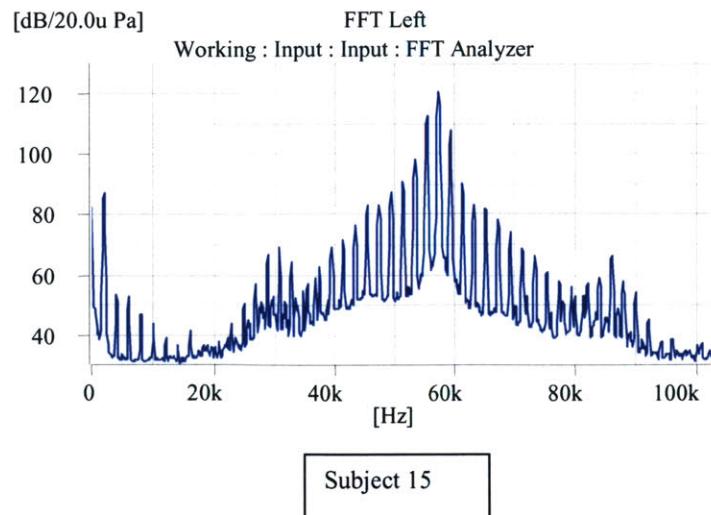
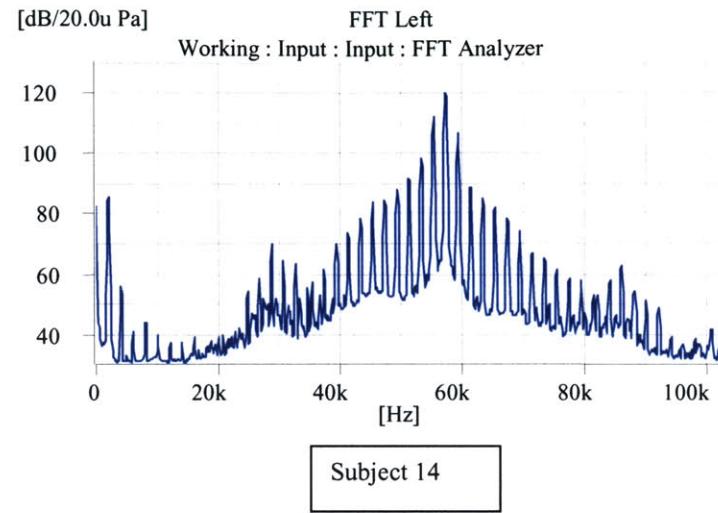
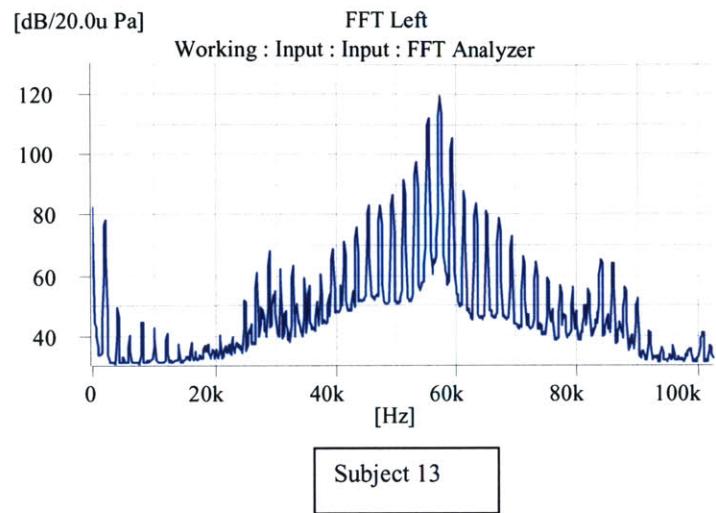
Appendix A.2

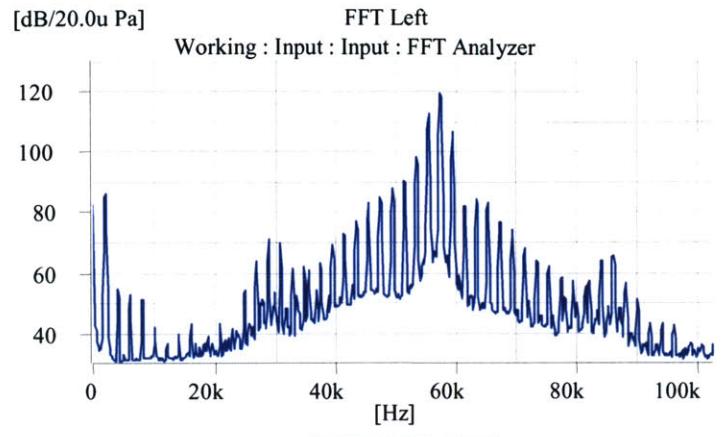
Spectra of Audio Spotlight stimulus for individual subjects.



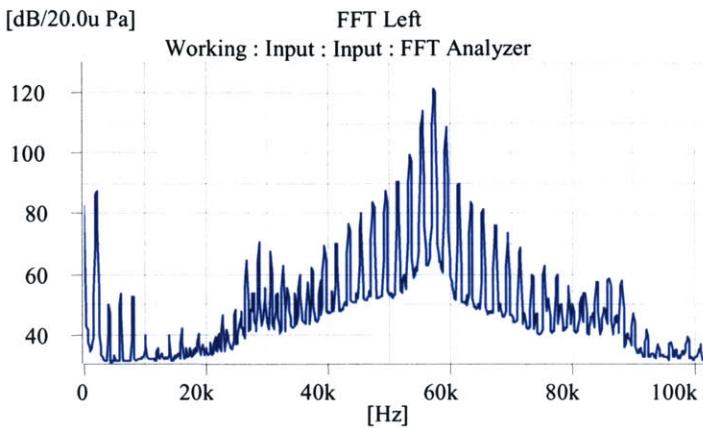




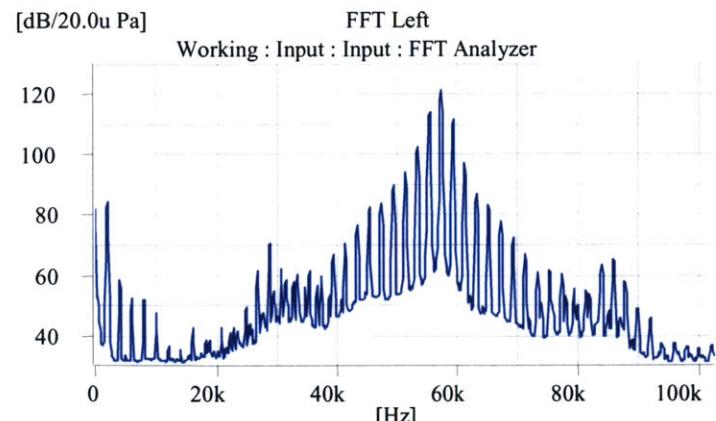




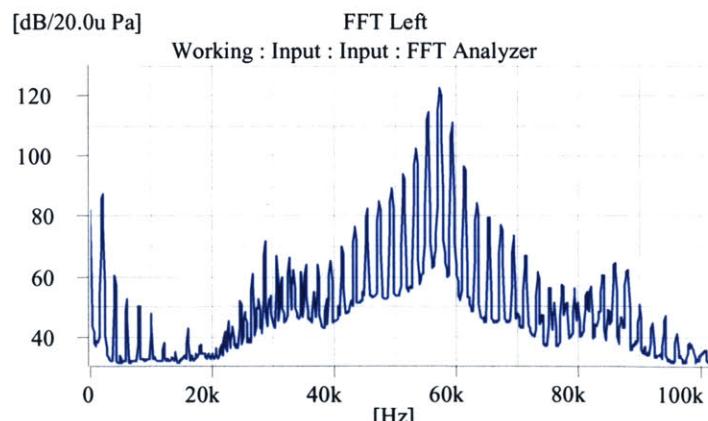
Subject 17



Subject 18



Subject 19



Subject 20

Appendix A.3

OSHA standards for ultrasound exposure.

OSHA & ACGIH Ultrasound Exposure Limits

	Measured in Air in dB re: 20 μPa; Head in Air	Measured in Water in dB re: 1 μPa; Head in Water	
Mid-Frequency of Values	Ceiling Value	8-Hour	Ceiling
third octave band (kHz)		TWA	
10	105 ¹	88	167
12.5	105 ¹	89	167
16	105 ¹	92	167
20	104 ¹	94	167
25	110 ²	--	172
31.5	115 ²	--	177
40	115 ²	--	177
50	115 ²	--	177
63	115 ²	--	177
80	115 ²	--	177
100	115 ²	--	177

¹ subjective annoyance may occur from 75-105 dB; use hearing protection.

² assumes some coupling to water or substrate, if none, limits may be raised by 30 dB with coupling acceleration at the mastoid bone should not exceed 1 g rms. + 15 dB

Martin L. Lenhardt

Professor of Biomedical Engineering
 Medical College of Virginia, Virginia Commonwealth University
lenhardt@hsc.vcu.edu 804.828.9687 fax 827-0065

Education:

Seton Hall University: B.S. Biology (1966)
 M.A. Audiology/Speech Pathology (1968)
Florida State University: Ph.D. Psychoacoustics and Speech Science (1970)
The Johns Hopkins Medical Institutions; Post Doctoral Fellow (1970-71)
 Otolaryngology and Biomedical Engineering
Kirkville College of Osteopathic Medicine: Au.D. (2001)

Licensed Audiologist and Speech Language Pathologist in Virginia

Publications:

- Lenhardt**, M.L.: Effects of Amplitude Modulated Tones in Auditory Averaged Evoked Response. Psychonomic Sci 24:1, 1971
- Lenhardt**, M.L.: Effects of Frequency Modulation on Auditory Evoked Responses. Audiology 10:12-18, 1971
- Lenhardt**, M.L.: A Simple Polarizable Electrode for EEG Audiometry. Audiology 10:365-369, 1971
- Lenhardt**, M.L.: Variabilities in Average Evoked Response Audiometry. J. Comm. Dis. p.551-555, 1972
- Lenhardt**, M.L.: Effects of Verbal Association on Averaged Auditory Evoked Response. Audiology 12:174-179, 1973
- Lenhardt**, M.L.: Vocal Cues in Maternal Recognition of Goat Kids. Applied Animal Ethology, Amsterdam, 3:211-219, 1977
- Lenhardt**, M.L.: Factors in Affective Discrimination of Speech. Neurology, LTTE, 3,28:309-310, March, 1978
- Lenhardt**, M.L.: Incubator Noise and Hearing Loss. Early Human Development, LTTE, 211:89-90, June, 1978
- Lenhardt**, M.L.: Excision of Craniopharyngioma. J. of Neurosurgery, LTTE, 54, 4:577, 1979
- Lenhardt**, M.L.: Pneumatic Whistle for Animal Hearing Tests. Laboratory Animal Science, 812-813, 1979
- Lenhardt**, M.L.: Hearing and Vocalization in Neonatal Gunn Rats. Brief Report, Cry Research Newsletter. 31, 1980
- Lenhardt**, M.L.: Evidence for Auditory Localization Ability in the Turtle. J. of Aud.Res., 21:255-261, 1981
- Lenhardt**, M.L.: Childhood Central Auditory Processing Disorder with Evoked Potential Verification. Archives of Otolaryngology, 107:623-625, 1981
- Lenhardt**, M.L.: Wave V Latency and Chirp (Linear Frequency Ramp): Repetition Rate. Audiology 21:425-432, 1982.

- Lenhardt**, M.L.: Hearing and Vocal Development in Neonatal Gunn Rats. Archives of Otolaryngology, 108:339-342, 1982
- Lenhardt**, M.L.: Evoked Response Abnormalities with Hyperbilirubinemia. J. of Pediatrics 101:1037-1038, 1983
- Lenhardt**, M.L.: Bone Conduction Hearing in Turtles. J. Aud. Res., 22:153-160, 1982
- Fine, M.L. and **Lenhardt**, M.L.: Shallow-water Propagation of the Toadfish Mating Call. Comp. Biochem. Physiol., 1983
- Lenhardt**, M.L., Bellmund, S., Byles, R., Harkins, S. and Musick, J.: Marine Turtle Reception of Bone Conducted Sound. J. Aud. Res., 23:119-125, 1983
- Lenhardt**, M.L.: Effects of Hyperbilirubinemia on Token Test Performance at Six Years of Age. J. Aud. Res., 23:195-204, 1983
- Lenhardt**, M.L., Harkins, S.W.: Turtle Shell as an Auditory Receptor. J. Aud. Res., 23:251-260, 1983
- Lenhardt**, M.L., McArtor, R., Bryant, B., Effects of Neonatal Hyperbilirubinemia on Brainstem Electric Response. J. Pediatrics, 101:281-286, 1984
- Lenhardt**, M.L.: Effects of Spectrum, Background Noise and Stimulation Rate on Auditory Startle Reflex in Hyperbilirubinemic Rats. Oto Head Neck Surg., 93:30-34, 1985
- Lenhardt**, M.L., Shaia, F. and Abedi, E.: Brainstem Evoked Response Waveform Variation Associated with Recurrent Otitis Media. Arch. Otolaryngol., 111:315-316, 1985
- Lenhardt**, M.L.: Tones in Their Bones. The Sciences (NY Academy of Sciences) JULY/AUGUST, p. 40-43, 1985
- Lenhardt**, M.L., Klinger, R.C., Music, J.A.: Marine Turtle Middle Ear Anatomy. J. Aud. Research, 25:66-72, 1985
- Lenhardt**, M.L.: Preyer Reflex in Jaundiced Rats. J. Aud. Research, 25:161-166, 1985
- Lenhardt**, M.L., Clarke, A.M., Harkins, S.W.: High Frequency Hearing in Jaundiced Rats. J. Aud. Research, 26:19-25, 1986
- Dunlap, S., **Lenhardt**, M.L. and Clarke, A.M. High Frequency Skull Vibratory Pattern. Oto Head Neck Sur., 99, 389-391, 1988
- Lenhardt**, M.L., Skellett, R., Wang, P. and Clarke, A.M.: Human Ultrasonic Speech Perception. Science, 253, 1991, 82-85
- Lenhardt**, M.L., Ultrasonic Hearing: Reply Science 1992
- Lenhardt**, M.L. Audible ultrasound revisited. The Hearing Review, 5, 50-52, 1998.
- Moein-Bartol, S., Musick, J.A., **Lenhardt**, M.L. Auditory evoked potentials in loggerhead sea turtles (*C. caretta*) Copeia August, 1999.
- Lenhardt**, M.L. Biomedical Engineering & Speech Language

Pathology. JSHAV, 39, 49-52, 1999

Lenhardt, M.L. What the turtle ear tells us about phylogenetic relationships Proc. Sea Turtle Bio. *in press*

May, 2001

**Preliminary Test of Audio Spotlight™:
1 Hour Exposure, and Exposure to 15 kHz Modulation**

Three subjects were tested in a preliminary look at the effects of exposure to the Audio Spotlight™ for one hour with modulation by music, and for 5 minutes with modulation by a 15 kHz tone. The hour period was chosen because subjective reports of a feeling of fullness in the ears followed exposure of $\frac{1}{2}$ hour to 1 hour with music. The five minutes with 15 kHz modulation was chosen as a worst case of exposure to high frequency audio in addition to ultrasound.

Two issues were addressed in this experiment. The first issue was that the reported feeling of fullness in the ears tends to go away quickly after the sound ceases. Therefore, tympanometry was performed *while the sound was playing*, at the $\frac{1}{2}$ hour and 1 hour points. The second issue was that TTS also tends to decay rapidly with time, with maximum TTS measured at 2 minutes after a sound ceases (TTS₂). Since a full audiogram takes several minutes, hearing in this experiment was only measured at 2 frequencies: 6 kHz in the standard audiometric frequencies, and 14 kHz in the high audio frequencies. The 6 kHz measurement began at 2 minutes after the sound, and the 14 kHz measurement at approximately 3 minutes after the sound. To further reduce the measurement time, only the right ear was measured both for audiometry and tympanometry – this was the ear most directly in the path of the beam.

The following sequence of events was used:

1. Baseline tympanometry and audiometry (right ear only).
2. One hour of music with ultrasonic carrier adjusted to approximately 120 dB SPL (fluctuated with music intensity).
3. Tympanometry at $\frac{1}{2}$ hour and 1 hour, with music still playing and subject still in position 1 meter in front of Audio Spotlight™.
4. Audiometry at 1 hour, beginning exactly 2 minutes after sound turned off.
5. After 5 minute wait, 5 minute presentation of 15 kHz modulated ultrasound, with the ultrasound level at about 121 dB SPL, and the 15 kHz demodulated audio at about 93 dB SPL.
6. Audiometry after 5 minute presentation, beginning exactly 2 minutes after sound turned off.

Results

Tympanometry

The mean change in peak pressure was very small in response to both stimulus conditions (5 - 19 daPa) and indicates no significant pressure effect of the Audio Spotlight™ on the

tympanic membrane. The subjects reported a slight feeling of fullness in the ear, but this went away quickly after the sound ceased.

Audiometry

There was essentially no change in thresholds at 6 kHz during the course of the experiment. At 14 kHz there was a slight *improvement* in thresholds over the course of the experiment; this is likely due to a small change in the criterion for signaling detection of the sound at this high frequency. In no case was there a positive TTS in response to either of the stimulus situations.

Table 1. Mean changes in hearing thresholds and tympanograms (n = 3).

	6 kHz (dB)	14 kHz (dB)	Immitance	Peak Pressure (daPa)
½ Hour-Baseline	na	na	-0.1	19
1 Hour-Baseline	-1.7	-10	-0.1	14
15 kHz-Baseline	-1.7	-11.7	-0.1	19
15 kHz-1 Hour	-1.7	-11.7	-0.0	5

Discussion

From this small sample of 3 subjects, it appears that, despite occasional feelings of fullness in the ears, there is no change with an hour of exposure to the Audio Spotlight at 121 dB in either the tympanogram or the auditory thresholds.

However, the question of the effects of the ultrasonic component is still confounded by the presence of the audio component, although there was no change in either the tympanogram or the audiogram. It would be worth presenting the ultrasound alone.