UNIVERSITY OF SOUTHAMPTON
FACULTY OF ENGINEERING AND APPLIED SCIENCE
INSTITUTE OF SOUND AND VIBRATION RESEARCH

NUMERICAL MODELLING OF THE HEAD-RELATED TRANSFER FUNCTION

Yuvi Kahana

A thesis submitted for the degree of

Doctor of Philosophy

December 2000

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND APPLIED SCIENCE

INSTITUTE OF SOUND AND VIBRATION RESEARCH

Doctor of Philosophy

Numerical Modelling of the Head-Related Transfer Function

By Yuvi Kahana

The localisation of sound sources by humans can be accomplished by the use of acoustical cues only. The Head-Related Transfer Functions (HRTFs) represent the linear, directional transformations of sound signals in free-field, detected in the eardrum of a listener or an artificial head. These functions have a complex structure, especially at frequencies above 5 kHz, due to diffraction, reflection and scattering from the head, torso, and especially the external ear. Traditionally, these functions are measured by using a time consuming and difficult procedure, with expensive apparatus found in well-equipped acoustic laboratories only. As a result, current virtual auditory display systems make use of non-individualised HRTFs, which produce unsatisfactory performance, especially when high fidelity sound quality and localisation of virtual sources in elevation are required.

This thesis investigates various aspects of numerically modelled individualised HRTFs. The computer simulations (undertaken on both a parallel computer and a PC) are based on the exact solution of the wave equation, with the main emphasis on the Boundary Element Method (BEM). The basic features of the HRTF are investigated first with simple geometrical models such as a sphere and an ellipsoid that represent the human head, and a baffled cylinder that represents the concha. Accurate geometric models of two heads and six pinnae are captured by using state-of-the-art 3-D laser scanners and digitisers. These computer models are converted to valid BEM models and their frequency response is simulated. With current hardware technology, and vigilant optimisation of the manipulated mesh models and the solving procedures, baffled pinnae can be investigated up to 20 kHz, and heads with pinnae (but without torso) can be investigated up to 10-15 kHz. High accuracy is obtained when the results of the simulation at the blocked ear canal are compared with measurements made with especially designed and built apparatus in an anechoic chamber, using the same physical head and pinnae used in the simulations.

Once the results of the simulations are validated against measurements, further acoustic features of the external ear are investigated with an emphasis on the 'mode shapes' of the human pinna. Using the Singular Value Decomposition (SVD), the matrix of Green functions relating the acoustic pressure at 'field' points and 'source' points in space is analysed at discrete frequencies. When the field points and the source points are positioned on uniformly sampled spheres, a connection is found between the matrices of the singular vectors and the sampled spherical harmonics. When the method is investigated numerically, and the 'field' points are positioned on different pinnae, their 'mode shapes' are presented, and compared to the classical experiments made by E.A.G. Shaw in the 1970s. The method is investigated further in order to produce 'reduced order' transfer functions by taking into account only the most dominant features of the singular vectors.

Finally, a few examples of numerically modelled sound fields of virtual acoustic imaging systems with various loudspeaker arrangements are given in the frequency and time domains.

ACKNOWLEDGMENTS

Firstly, I would like to thank my supervisor Phil Nelson for all his support throughout this work. I am extremely grateful to him for providing me with the help and freedom to pursue this challenging topic. I am also grateful to Andy Keane, who offered me the position in the Computational Engineering and Design Centre (CEDC), and Maurice Petyt who guided me in the initial stages of the work. I would also like to thank the following for their help:

LMS International who provided invaluable support. In the UK branch I would like to thank Neil Cockrell, Robert Siddall, Lee Sargent, and Colin McCulloch, and in the Leuven branch Luc Cremers, and Peter Seagart. Mike Clark was also very helpful in evaluating the Pre-SYSNOISE software produced by MSC/Nastran. In the ISVR consultancy services Ben Lawton, Mike Lower, and John Fithyan who helped with everything related to the KEMAR artificial head. At Cyber-Site Europe, Guy Hauldren for his patience with scanning all the head and pinnae models. At NCI, Phillip Schwizer who provided the CAD model of the CORTEX artificial head and pinna. At the Hearing and Balance Centre of the ISVR, Graham Brickly who constructed my own ear mould. Mike Stinson of NRC, Richard Duda of San-Jose University, and Victor Sparrow of Penn-State for helpful discussions through personal communication. In Carnegie Mellon University School of Computer Science, Andrew Johnson, for his help with the decimation algorithm. Brian Katz from Penn-State for his help in the initial stages of the work, and Sunghoon Choi from Samsung/SAIT who collaborated on this project part of the time.

For the HRTF rig design and manufacturing, Takashi Takeuchi, Dave Edwards, Allan Sanger and Robert Stansbridge.

Special thanks to the CEDC (Andy Keane and Neil Bresslof) for enabling me to use its parallel computer around the clock for more than two years!

The various sponsors throughout the work: Beatrice Curtey, Anglo Jewish Society, Ian Karten, Samsung/SAIT, Japanse consortium sponsors, and CEDC.

The collegaues I met at the ISVR, Pauli Minnaar, Ole Kirkeby, Ben Cazzolato, Philippe Godano, Emmanuel Fruteau, Douglas Bradshaw, and Boaz Rafaeli.

Finally, I would like to thank my family - without their continual support this thesis would have never been completed.

TABLE OF CONTENTS

LIST OF SYMBOLS

β	Specific characteristic admittance
ω	Angular frequency ($2\pi f$)
ρ_0	Fluid density
η_0	Characteristic impedance of the fluid; $\eta_0 = \rho_0 c_0$
δ_{nk}	Kronecker's symbol: $\delta_{nk} = 1$ for $n = k$; $\delta_{nk} = 0$ for $n \neq k$
θ	Polar angle in spherical co-ordinates
ϕ	Azimuth angle in spherical co-ordinates
σ_i	Diagonal elements of the singular values matrix Σ
Σ	Singular value matrix
γ	Ratio of specific heats
μ	Double layer potential (jump of pressure) in the indirect BEM
σ	Single layer potential (jump of normal derivative of pressure) in the indirect BEM
λ	Wavelength
$C(\mathbf{r})$	Solid angle
c	Speed of sound in fluid
f	Frequency
$g(\mathbf{r} \mid \mathbf{r}_0)$	Free space Green function
H	(Superscript) Hermitian- conjugate transpose
$h_n(kr)$	Spherical Hankel function
j	Imaginary number; $j = \sqrt{-1}$
$j_n(kr)$	Spherical Bessel function
k	Propagation constant, wave number; $k = \omega/c$
$\mathbf{n}$	Unit length normal vector
$n_n(kr)$	Spherical Neumann function
p	Complex pressure
$P_n^m(x)$	Legendre polynomials
Q, q	Source strength
$\mathbf{r}$	Field point position
$\mathbf{r}_0, \hat{\mathbf{r}}$	Source position
$\mathbf{U}$	Singular vector (radiation pattern)
$\mathbf{V}$	Singular vector (source strength pattern)
u_n	Normal velocity
W	Total acoustic power
$Y_n^m(\theta, \phi)$	Complex spherical harmonics
Z	Impedance

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 MOTIVATION

Most current spatial sound reproduction systems are based on the concept of binaural technology. These systems are referred to in the literature as '3-D audio' (Begault, 1994), virtual auditory display (Wenzel, 1993), virtual auditory space (Carlile, 1996), virtual acoustic imaging (Nelson, 1997), and similar variations. The goal of the system designer is to ensure that the reproduced signals at the ears of a listener, through either headphones or loudspeakers, are equivalent to those detected under real listening conditions. In order to manipulate the signals arriving in the eardrum of the listener in a binaural synthesis process, it is required to know the directional characteristics of the physically filtered signals that are encoded in the Head-Related Transfer Function.

It should be noted that there are other approaches to the production of spatial sound, which are not based on the HRTF. For example, 'wave field synthesis' (Berkhout *et al*, 1993) reconstructs the propagating waves in a restricted area using a discrete approximation to the exact Helmholtz-Kirchoff equation, or the 'loudspeakers-walls' system (Ono *et al*, 1998) which recreates an approximation of a desired impulse response of a room. Both methods require a large number of loudspeakers.

In recent years the number of scientific papers, products and applications associated with the HRTF has grown rapidly for two main reasons: (1) the advances in computing power and the possibilities of implementing digital filters with low-cost DSP chips, and

(2) advances in research on the physical, physiological, and psychoacoustical aspects of spatial hearing and the interaction between them.

One of the main limitations of binaural technology is the generalisation of the HRTF of a particular listener or an artificial head for the entire population. When individualised HRTFs are used (i.e. either the recording is done with microphones positioned in the ears of the listener, or monophonic signals are synthesised with the listener's HRTF) and we assume no errors are introduced in any part of the reproduction chain (transducers, acoustic medium, head movements, etc...) we may not need to deal with the complexity of the perception of sound by the auditory system. In practice, errors are inevitable, and *exact* reproduction cannot be achieved. Therefore, psychoacoustical studies must be carried out in order to investigate the physical cues encoded in the HRTF and the perceptual importance of these in the auditory system.

It appears from many psychoacoustical studies published in the last 50 years that the task of localisation of sound is more complex than assumed originally in Lord Rayleigh's duplex theory (Rayleigh, 1907). Although the significance of the external ear is now well recognised as a complex acoustical antenna, it is still not understood how the different cues are combined in the auditory system, and from the neurophysiological view, if all the information detected by the pinna can be encoded by the nervous system.

High fidelity HRTFs are currently required by both the research community and the designers of virtual auditory displays. Traditionally, these databases are acquired by measurements. The procedure of measuring HRTFs is very time consuming, and expensive. These are currently limited to well-equipped acoustic laboratories only, and as a result, most HRTFs are either confidential or restricted to research purposes. There are also many problems encountered when these functions are measured, analysed and compared between different studies. For example it is difficult to define the point at which the microphone

should be positioned in the ear canal, the type of transducers and equalisation techniques that should be used, and methods for dealing with signal to noise ratio problems, etc. In addition, HRTFs are generally measured only at discrete points with a low directional resolution. As a result, any real-time virtual auditory display would need to make use of interpolated functions.

In this study, we suggest an alternative approach to acquiring individualised HRTFs, by using computer simulation techniques rather than by using measurements. Geometrical data derived from an optical image can be converted into its acoustical response, in principle, by solving the wave equation. The idea is not new, as this was stated by Weinrich (1984) who first investigated the response of the human head (without pinnae) using numerical techniques:

> *"The rather complicated geometric shape of the pinna makes a rigorous mathematical treatment very difficult – perhaps impossible"*

and recently also by Shinn-Cunningham and Kulkarni (1996):

> *"Theoretically, it is possible to specify the pressure at the eardrum for a source from any location simply by solving the wave equation.... Needless to say, this is analytically and computationally an intractable problem"*

In this research, we attempt to investigate the feasibility of obtaining accurate HRTFs using computer simulation, and to develop a tool that can be used to investigate the acoustical characteristics of the external ear.

1.2 OBJECTIVES

The focal point of this thesis is to investigate whether it is viable to predict *high frequency* components in the frequency response of the external ear using simulation tools. The objectives of the work and some of the questions addressed throughout the research are given below:

- Investigate the feasibility of using various numerical techniques to compute HRTFs at low to medium frequencies using simple geometrical models. Can we simulate the response of these simple models so that these can be used, for example, in a structural model, such as that proposed by Genuit (1986, 1987)?

- Simulate the HRTFs of accurate geometric models. Can we validate the results with measurements carried out in an anechoic chamber?

- Develop a tool to investigate the acoustical characteristics of the human ear. Can we reduce our problem by substitution of the head with an infinite baffle, and concentrate on the contribution of the external ear alone, independently of the other parts of the human body?

- Identify common characteristics of the external ear by visualising the response at high spatial resolution at different azimuthal and elevation planes. By simulating and measuring a few pinnae under exactly the same controlled conditions, continuous maps of the variation of peaks and notches in the frequency response can be obtained. Are these results comparable with those found in the literature?

- Investigate the acoustical features of the external ear that can be used mathematically to reconstruct individualised HRTFs. In the area of HRTF modelling, low-order parametric functions are required mainly for the implementation of real time virtual auditory displays. Can we find, using simulation tools, common physical patterns that can be used for this purpose?

- Validate the normal mode shapes measured by E.A.G Shaw and published over a period of three decades. To the author's best knowledge, his work summarising the mode shapes of the pinna has not yet been validated nor continued. Can we obtain the same patterns with our simulated pinnae models?

- Visualise sound fields of virtual acoustic imaging systems using loudspeakers. The equalisation zone ('sweet spot') is primarily affected by the loudspeaker arrangement. Can we predict the sound field around the head, while designing an ideal cross-talk cancellation network with the individualised HRTFs modelled at the first step?

1.3 ORIGINAL CONTRIBUTION

The feasibility of using various numerical techniques to simulate the response of HRTFs has been investigated throughout the thesis for the following components of HRTFs: head, torso and shoulders, pinnae, ear canal and eardrum. The following contributions in detail, are summarised below:

- It was found in the research that it is already possible to obtain individualised HRTFs that can be used in the auralisation process. High accuracy HRTFs of the head and pinnae (but without the torso) can be obtained up to 15 kHz for the ipsilateral ear, and up to 10 kHz for the contralateral ear (Chapters 5 and 6).

- If a structural modelling approach is used or specific features of HRTFs are required to be adjusted (such as ITD values), numerical techniques can produce, fairly rapidly, data-bases of simple geometrical models with different sizes and shapes, such as the ellipsoid to represent the human head, or cylinders to represent the shoulders and neck, or the pinnae (Chapter 4).

- Using the principle of reciprocity, it is shown that the problem of interpolating HRTFs is alleviated with the proposed numerical technique. This feature can be used for implementation of real time applications, as well as being used as an ideal tool to investigate the accuracy of interpolation techniques at high frequencies (Chapters 4, 5, and 6).

- The frequency response of six pinnae was simulated and validated with high accuracy against measurements. The simulations of baffled pinnae show high level of agreement with the spectral features found in HRTFs, in the ipsilateral hemisphere (Chapters 5 and 6).

- An automated HRTF measurement apparatus was developed and built to control the accuracy of measurements to a resolution much less than 1°. The efficiency is at its best when *artificial* ears or heads are investigated and the fine structure of the response is visualised (Chapter 6).

- In the investigation of the 'mode shapes' of arbitrary radiating bodies, a number of numerical and analytical simulations have been carried out, and showed the link between the basis functions provided by the Singular Value Decomposition of matrices of acoustic transfer functions, and the basis functions provided by classical acoustical analysis, with particular emphasis on the case of radiation from a sphere using a series of spherical harmonics (Chapter 7).

- The basis functions of a shallow cylinder and accurate pinnae are investigated based on the theory described in Chapter 7. A few examples are given to show how this novel way of predicting the spatial patterns of the pinna can be used in the future for reduced order HRTF modelling based on its physical characteristics (Chapter 8).

- The total sound pressure of sound fields of virtual acoustic imaging systems due to different loudspeaker arrangements is investigated. Once the individualised HRTFs are modelled, it is possible to investigate the effect of scattering around the listener. Multi-channel inversion strategies are used to illustrate the effect of a 2×2 and 4×4

cross-talk cancellation networks. The efficiency of the Stereo Dipole system is demonstrated in the frequency and time domains (Chapter 9).

In addition to the above contributions, further tools and formulations have been developed, some of which are given below:

- Mesh manipulation tools were developed and implemented in order to optimise any arbitrary complex shapes to be used with the BEM. It was found that this was crucial to the successful modelling at high frequencies, when the mesh models are very large, and typically have the order of tens of thousands of vertices and elements on the surface mesh (Chapter 5 and Appendix 1).

- The simulated BEM responses of an ellipsoid were arranged in an interface software that can be used, for example, to repair the response of HRTF measurement at low frequencies due to a poor signal to noise ratio. In addition, the Interaural Time Difference (ITD) is extracted using the interaural group delay differences at low frequencies. The results for the ellipsoid can be advantageous when compared to the classical rigid sphere model (Chapter 4).

- The frequency response of a rigid sphere is derived using the series of spherical harmonics, such that in spherical co-ordinates the response can be calculated anywhere on the sphere from sources positioned at any arbitrary direction (Chapter 7).

1.4 OVERVIEW OF THE THESIS

The research presented in this thesis links two different disciplines: physical cues for sound localisation, and numerical modelling of sound fields of exterior problems.

Chapter 2 introduces the reader to the relevant background of the acoustical characteristics of the external ear. The chapter highlights briefly the relevant aspects that are related to this

work. Some psychoacoustical studies are mentioned to emphasise the significance of analysing the high frequency characteristics of the external ear, and the advantages of binaural synthesis using individualised HRTFs.

Chapter 3 provides the basic formulation of the numerical techniques used in the succeeding chapters. Since our computational problems are very expensive, it is of utmost importance to optimise every aspect of the model. Special formulations are included with a summary of the properties of the different methods employed.

Chapter 4 covers a few basic investigations using the numerical techniques introduced in Chapter 3, in order to explore the feasibility of each technique. The classical sphere, which is used extensively in the literature, is investigated using four different approaches: the Direct Boundary Element Method (DBEM), the Indirect Boundary Element Method (IBEM), the axisymmetric IBEM, and the Infinite-Finite Element Method (IFEM). A Computer Aided Design (CAD) model of the CORTEX artificial head is used to investigate the effects of the torso and shoulders, and these frequency response results are compared to those of a sphere, an ellipsoid, and a head without the torso. At higher frequencies, the pinna plays an important role, and a shallow cylinder can be used to investigate the first quarter wavelength depth mode and its first transverse mode. These basic shapes produce acoustical features that can be found (although with greater complexity) in a full model of the HRTF.

In chapter 5, accurate models are investigated. Two different cases are presented: in the first case, two heads with accurate pinnae (of the KEMAR artificial head and that of the author), and in the second case, six pinnae (five of artificial heads and one of the author) are investigated under baffled conditions. The chapter includes a brief description of the mesh manipulation techniques that were required to convert the original scanned models to BEM models (details on the algorithms are given in Appendix 1). The frequency and time

domain responses are presented, and the main spectral features of the 'blocked meatus' pinnae are analysed and compared with the literature. In addition, the inclusion of an ear canal is briefly investigated, both with a rigid eardrum and with average eardrum admittance values.

The validation of the simulation is given in Chapter 6. A computer controlled rotating arc for HRTF measurements was designed and built. The measurement procedure is described, and a comparison of the response of six baffled pinnae and the HRTFs of KEMAR (without the torso) between simulation and measurement is illustrated and discussed.

Chapters 7 and 8 deal with the mode shapes of arbitrary, complex radiating bodies (or scatterers). In Chapter 7 a theory for finding a set of orthogonal basis functions describing sound radiation and scattering from irregular shaped bodies is presented. It is shown how the basis functions ('mode shapes') of a radiating sphere, described by the complex spherical harmonics, are related to those extracted using the Singular Value Decomposition. The method is also used to investigate numerically the case of an ellipsoid.

In Chapter 8, the theory described above is used to analyse the spatial mode shapes of simple models of the concha (a Helmholtz resonator and a baffled shallow cylinder), and six accurate pinnae using the BEM. It is shown how these patterns can be used to reconstruct a reduced order transfer function. In addition, using the principle of reciprocity, the 'mode shapes' of the pinna are simulated using the same technique used by E.A.G Shaw in his measurements of 'mode shapes'.

In chapter 9, we demonstrate the potential of predicting the sound fields for various loudspeaker arrangements under free-field conditions. The figures include frequency and time domain snapshots of the pressure on the head and the sound field around it, when the inputs to the loudspeakers are filtered with a cross-talk cancellation matrix based on the numerically modelled HRTFs.

In the final chapter (Chapter 10) a discussion and the conclusions of the work are presented with a summary of the computational cost involved in this research. The chapter ends with a proposal for future work.

In Appendix 1, we include a description of the mesh manipulation techniques used in this research. These were an important stage in the process of optimisation of speed, geometrical and acoustical accuracy. The geometrical properties and statistical values of all mesh models are presented.

In Appendix 2, we present the main 'mode shapes' of the pinnae investigated in Chapter 8.

1.5 SOFTWARE TOOLS

This research was based mainly on the vibro-acoustic software package SYSNOISE (LMS international), with all its special formulation modules. Programming was undertaken using the SYSNOISE command language (SCL), and the 'user-defined' subroutines were programmed in Fortran. Mesh manipulation algorithms were developed using C/C++ and included OpenGL modules developed in Carngie Mellon University. Unfortunately, the process of obtaining a valid BEM mesh from the original scanned model is not automatic and requires an iterative process. In addition to the development of mesh decimation tools (that controlled the homogeneous distribution of the vertices), other existing commercial tools were used to optimise the mesh models: CYBERWARE (Cyberware, Inc.) software packages (mainly CYEAT, CYSCULPT), ANSYS pre-processor (Ansys, Inc.), IDEAS (Structural Dynamics Research Corporation, SDRC), HYPERMESH (Altair Engineering, Inc.), PRE-SYSNOISE (LMS and MSC/ Nastran) and AUTOCAD (AutoDesk, Inc.). File formats conversion tools were developed in C/C++. These include two-way conversion of VRML 1.0, WAVEFRONT Obj, ANSYS (ascii cdb format), SYSNOISE free-format, SYSNOISE user-defined format, and SGI inventor. In the measurement set-up, the MLSSA

(DRA Labs) system was used with its macro commands, linked to a PARKER (Parker Hannifin, plc.) automated rotor language to control HRTF measurements and the rotation of the rotating arc.

All analytical simulations involving a rigid sphere (Chapters 4 and 7) were carried out using MATHEMATICA (Wolfram research, Inc.), and all the post-processing, visualisation, and matrix manipulation of the numerically modelled and measurement results were performed with MATLAB (MathWorks, Inc.)

1.6 CO-ORDINATE SYSTEMS AND ANGLE CONVENTIONS

Figure 1-1a presents the co-ordinate systems used when the response of an artificial or a human head was modelled or measured. The Azimuthal angles* are defined as $0° \leq \phi < 360°$. At $\phi = 0°$, the source is in front of the listener, at $\phi = 90°$ the source is to the right ear, $\phi = 180°$ the source is at the rear, and at $\phi = 270°$ the source is on the left. The same convention is used with the baffled pinna, and in this case $0° \leq \phi \leq 180°$.

The elevation angles are defined as $-90° \leq \theta \leq 90°$ where $\theta = 0°$ is the horizontal plane, $\theta = 90°$ is above, and $\theta = -90°$ is below. The same convention is used in the case of the baffle.

The centre of the co-ordinate system in Figure 1-1a is in the centre of the interaural axis, which connects the two ears, and the centre of the co-ordinate system in Figure 1-1b is at the blocked entrance to the ear canal.

The results of the simulations and measurements are given in three planes:

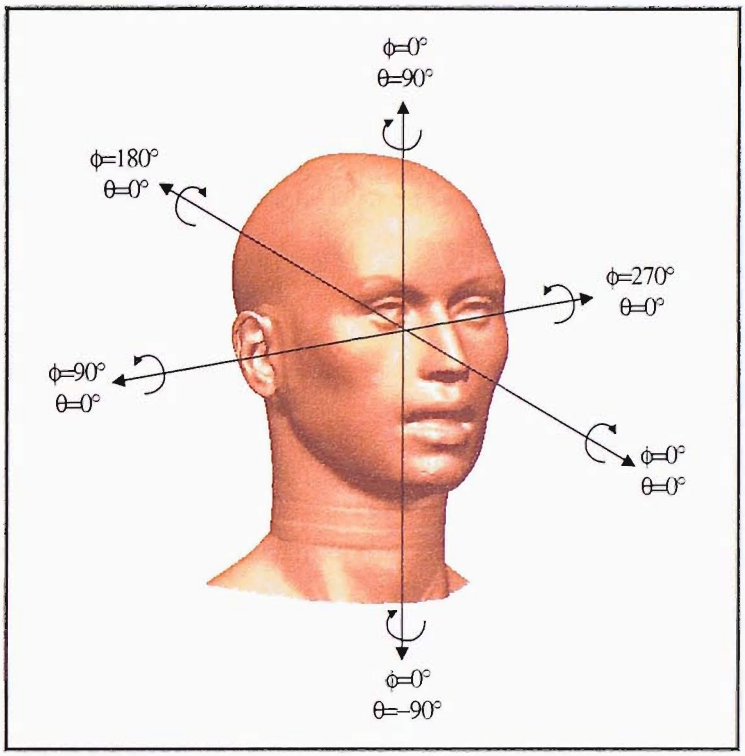
* Note that azimuthal angles are defined with ϕ and not θ. This is mainly for consistency with the spherical co-ordinate system used in Chapter 7.

- Median Vertical Plane. When the full head is investigated, the elevation angles that are symmetrical to both ears are defined as $0° \leq \theta < 360°, \quad \phi = 0°$. When the baffled pinna is investigated, the equivalent plane is 'grazing incidence'[†], with the same angle definitions as above.

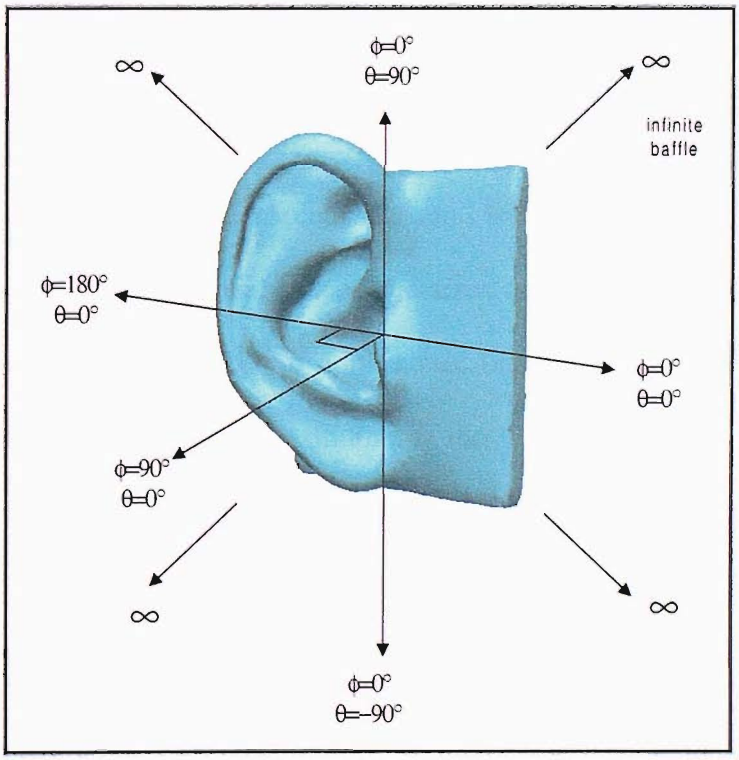
- Horizontal (azimuthal) Plane. For the head $0° \leq \phi < 360°, \quad \theta = 0°$, and for the baffled pinna $0° \leq \phi \leq 180°, \quad \theta = 0°$.

- Lateral Vertical (frontal) Plane. For the head $-90° \leq \theta < 270°, \quad \phi = 0°$, and for the baffled pinna $-90° \leq \theta \leq 90°, \quad \phi = 0°$.

[†] Note that 'grazing incidence' was defined at a normal distance of 5 cm from the baffle, for compatibility with measurements since this was the minimal distance between the centre of the loudspeakers' cones, and the baffle at $\theta = \pm 90°$.



(a)



(b)

Figure 1-1: Co-ordinate systems for simulation and measurement of (a) HRTFs (b) the response of baffled pinnae.

CHAPTER 2

BACKGROUND: THE ACOUSTICAL CHARACTERISTICS OF THE EXTERNAL EAR AND THE HEAD-RELATED TRANSFER FUNCTION

2.1 INTRODUCTION

In the discussion of auditory space perception, questions of sound localisation are tackled from mainly two approaches: physical and psychophysical. It is clear that from the physical point of view, the variation of the eardrum response as a function of source position in space provides the basis for sound localisation. However, in psychophysical studies, despite improved experimental methods in the last few decades, we still have a limited understanding of how the different sound localisation cues are combined in the auditory system.

One of the main difficulties in research into sound localisation and sound reproduction is the high variation in the anatomical shapes and sizes between individuals and in particular the external ear. In our research, we concentrate on the *physical* cues imposed by the external ear, and attempt to provide a simulation tool that can be used for the investigation of different pinnae under controlled parameters.

In this chapter, we summarise briefly the background that is related to our research. For a detailed review of the physics and psychophysics of spatial hearing the reader is referred to the books by Blauert (1997), Gilkey (1996), Carlile (1996) and Begault (1994).

2.2 THE HUMAN EAR

The anatomy of the human ear is generally analysed in three sections: the external ear, the middle ear and the inner ear. The external ear (outer ear) is composed of the pinna and the auditory canal (meatus). The middle ear consists of the eardrum (tympanic membrane), the tympanic cavity, and three small bones, the ossicles, within the cavity. The inner ear includes the organ of Corti, which lies within the cochlea and contains the receptors for the sense of hearing, and the vestibular organs, which contain the receptors for the sense of balance (Blauert, 1997).

When sound travels in air, it is first modified by the shape of the pinna, especially at high frequencies. The pinna acts as a complex acoustical antenna and codes spatial characteristics of the sound field into temporal and spectral attributes. After filtering by the pinna, the sound waves propagate down the ear canal and cause the eardrum to vibrate. These vibrations are transmitted through the middle ear by the ossicles to the cochlea.

In our study, we concentrate mainly on the shape and response of the external ear, with an emphasis on the pinna (the ear canal and the ear drum do not contribute significantly to the localisation of sound, as will be discussed later).

The recognition of the acoustical significance of the external ear is relatively new, and have been appreciated for only a few decades. Its complex effect is summarised by Blauert (1997):

> *"The acoustical effect of the pinna is based upon reflection, shadowing, dispersion, diffraction, interference and resonance."*

A schematic diagram of the external ear is presented in Figure 2-1. The largest hollow in the pinna, the concha, is a broad shallow cavity that is partially divided by the crus helias. The lower part, the cavum, is tightly coupled to the canal whereas the upper part, the cymba, is connected to the fossa of helix (Shaw, 1997). It was found (Teranishi and Shaw,

1968) that these parts have acoustical attributes, whereas the structures extending from the concha, such as the helix, anti-helix and lobule seem to function collectively as a flange.

Different geometrical parameters have been defined to describe the pinna in anthropometric surveys. The work of Burkhard and Sachs (1975) produced the design of an average artificial head (KEMAR) and an average pinna (DB60). The dimensions of this pinna are smaller than the average size of the male pinna. It is designed as an average of the dimensions and response of 24 subjects. These are typical of American and European females as well as Japanese males and females. The ear length and ear breadth (See Figure 2-1 for geometrical definitions) of DB60 are 58.9 mm and 34 mm, respectively, whereas in the work of Alexander and Laubach (1968), where more than 2000 USAF male flying personnel with average age of 28 were investigated, the average length of 67.1 mm with a range of 53.8 mm to 79.7 mm. The average ear breadth was 34.5 mm with a range was 27.4 mm to 42.8 mm (values within these ranges were measured also on individuals by Kuhn, 1983). Later, the DB65 was developed for KEMAR (Maxwell and Burkhard, 1979) with larger dimensions of length and breadth of 66 mm and 37 mm, respectively. It is typical of American and European male pinna sizes.

Note also that the relative rotation of the pinna affects its response. Thus the angle between the pinna and the side of the head, which is between 25° and 45° (Blauert, 1997), and tilt angle with the vertical position of the head, which is between approximately 3° and 8° (Dryfus, 1967) can both influence the acoustical response. However, in our investigation of different pinnae (Chapter 5), all entrances to the ear canal were aligned to be at the origins of the co-ordinate systems, and all rotational angles were aligned to average values (as will be shown later in Figure 5-5).

In the work of Shaw (see the references between 1968 and 1997), the pinna is investigated as a resonator in the frequency domain. It is found that that when the ear canal is blocked,

the cavum concha is responsible for a quarter wavelength resonance. This is of course a function of both the dimensions and the volume of the concha. In the study of Burkhard and Sachs (1975) the volume of the concha of 24 people was in the range of 3.2 cm^3 to 5.8 cm^3. These variations can shift, for example, the first anti-resonance frequency as high as from 8 kHz to 11 kHz (KEMAR, 1978, Chapter 3).

The ear canal is a slightly curved tube with varying cross section (see Stinson and Lawton, 1989, for accurate measured data). It is generally simplified in models to a cylinder with a diameter of 7.5 mm, and length of 22.5 mm and an average volume of 1 cm^3 (Zwislocki, 1970). Average values of the impedance of the eardrum are given by Shaw and Teranishi (1968) and Shaw (1974).

As will be investigated in Chapters 5, 6 and 8, the response of the pinna is very complex at high frequencies and it is very sensitive to the geometric shapes, size and orientation.

2.3 THE BINAURAL TECHNIQUE AND THE HEAD-RELATED TRANSFER FUNCTION

The binaural technique is based on binaural hearing. The technique can be implemented in two closely related approaches: binaural *recording* and binaural *synthesis*. Assume that a listener perceives an auditory event in a free-field environment. In addition, his head is fixed, and the source is stationary. The sound pressure received in both his ears includes the full information of the transformation of sound (it was shown in Blauert, 1997, that bone conducted sound is negligible in the perception of spatial sound). If the two signals, the sound pressure detected at each of the eardrums, are recorded and reproduced exactly (either with headphones or loudspeakers, with the appropriate equalisation), then it is assumed that the complete auditory experience is replicated, including timbre and spatial aspects (Møller, 1992). Generally, the recordings are made with an artificial head (also

termed 'dummy-head' or acoustical mannequin), and as a result, the reproduction is approximated and therefore conflicting cues might occur in a listener.

A more commonly used method is binaural synthesis. In this case, an artificial head is not required, and the monophonic signals can be filtered with the already measured sound transformation, which is now commonly termed the Head-Related Transfer Function (HRTF). The HRTF is a directionally dependent complex valued transfer function (frequency response) that describes the transformations of sound from a stationary source in free space to a point in the ear canal. The source is generally positioned in a three dimensional space, at a constant radius, originating at the centre of the interaural axis (the line that connects the entrances of the two ear canals). The detection point in the ear varies among studies, from a position at an occluded ear canal ('meatus' blocked) and even a few millimetres outside it, and anywhere along the ear canal (Hammershøi and Møller, 1996). In the application of binaural synthesis, HRTF data is generally arranged in the time domain format and is called Head-Related Impulse Response (HRIR). Comprehensive reviews of various studies on the transfer functions of individual anatomical features have been given by Shaw (1975), Kuhn (1983) and Blauert (1997).

The main limitation of binaural recording is that it is based on a recording with a fixed head, and therefore the dynamic cues are missing. In addition, the differences in matching the artificial head pinnae and those of the listener will cause colouration and poor performance in localisation of elevation sources. However, a new trend has appeared in recent years, where binaural synthesis incorporates dynamic cues by using a head tracker, which can compensate for head rotations by updating the filters in real-time (Sandvad, 1996, Wenzel *et al,* 2000, Begault *et al,* 2000 and Mackensen *et al*, 2000).

2.4 SOUND LOCALISATION AND SPACE PERCEPTION

There are clearly different mechanisms in the localisation of real acoustical events, and of virtual sounds when using headphones. Shaw (1982) proposed that headphone studies of spatial imagery be referred to as 'space perception', since the collection of perceptual data, such as segregation of sound, where a single *acoustical event* may give rise to more than one *auditory event*, or ambiguity with respect to whether the sound is externalised (and 'localisation' becomes 'lateralisation'). Most psychoacoustical studies are carried out with headphones, and assume that sound localisation and space perception can be regarded as equivalent.

2.4.1 Binaural cues

Lord Rayleigh's duplex theory (Rayleigh, 1907) was the first to explain how we localise sound: localisation is based on the fact that path lengths are different for the two ears, hence the Interaural Time Difference (ITD), and the head acts as acoustic shadows at higher frequencies producing Interaural Level Difference (ILD). At angles on the median plane or on the 'cone of confusion', additional complex cues are required, and therefore either head movement can resolve ambiguity (Wallach, 1940), or these can be resolved by the filtering of the pinna.

By neglecting the transmission paths within the auditory nervous system, we can assume that binaural cues could be derived by the ratio of the ipsilateral and the contralateral HRTFs in the frequency domain. This ratio produces the ITD and ILD mentioned above. Wightman and Kistler (1997) found that the fact that ITD is frequency dependant and is larger at low frequencies than at high frequencies (Kuhn, 1977) is perceptually irrelevant. Whenever HRTFs are implemented using a minimum phase model (see below), a single value is assigned to the ITD. However, although ITD values are roughly similar among

subjects, the auditory nervous system is very sensitive to changes of the interaural phase or timing. The minimum noticeable difference can be as low as 6 μsec (see Carlile, 1996, section 2.2.1).

The dominance of the ITD cue at low frequencies (below 1.5 kHz) was demonstrated when it was conflicted in a subjective experiment against other localisation cues (Wightman and Kistler, 1992). It was also claimed that the auditory system is sensitive to ITD in the envelopes of high frequency carriers, but this is a less dominant cue.

ILD presentations are very complex, since at high frequencies their dependence with the change of angle is high, and the response varies rapidly between peaks and notches. Their visualisation in different planes reveals some systematic variations, but variations among individuals are high, especially above approximately 8 kHz. It was shown that when ILD are presented in different frequency bands they have similar patterns (Wightman and Kistler, 1997). In addition, similar patterns are noticed for a specific frequency with a change of the elevation angle (Duda, 1997). In the horizontal plane, Middlebrooks and Green (1991) observed that localisation is mainly based on ITD and ILD without pinna cues. However, Musicant and Butler (1984) found that pinna cues indeed helped in resolving front and back confusion, and increased the localisation accuracy when localising sounds within the same quadrant of the horizontal plane.

Another possible binaural cue was suggested by Searle *et al* (1975), as 'binaural pinna disparity'. These authors proposed that the asymmetry between pinnae geometry and acoustical response aids in median localisation performance. This cue is still regarded as of a second order of significance.

2.4.2 Monaural cues: time domain interpretation

Many psychoacoustical studies demonstrated that it is possible to localise reasonably well with one ear plugged, in both horizontal and elevation angles (see Blauert, 1997, Section 4.4, Carlile, 1996, Section 2.2). However, localisation accuracy is dependent on the spectral contents, the frequency bandwidth of the stimuli, and other factors related to practice and context effects.

As the external ear is a linear system, time domain and frequency domain behaviours are related through the Fourier transform. It is assumed that the processing of directional information takes place in one of the two domains. Does the external ear encode the source direction through modulation of time delays or modulation of spectral shape?

Batteau (1962, 1967) was a pioneer in relating localisation in elevation, and the physical cues provided by the external ear. He hypothesised that a simple time domain model, which includes the original signal and two echoes, can give rise to the necessary spectral cues. One echo having a latency of 0-80 μsec varies with the azimuthal position of the source, and a second echo, having a latency of 100-300 μsec varies with the elevation. Some agreement was found by Watkins (1978) and Wright *et al* (1974) when the method was compared with measurements in the lateral vertical plane. Hebrank and Wright (1974) showed that the notches appearing in the frequency domain are matched with the interference of a variable path-length reflection that occurs on the posterior wall of the concha. Hiranaka and Yamasaki (1983) confirmed that major reflections occur within 350 μs after the first arriving sound, and that the delay increases as the source is lowered. In an extensive search for physical cues made with KEMAR, it was shown by Han (1991) that if localisation at high frequencies were based on time delays, it would work only in a very limited region. Since the model of Batteau is based on the physical geometry of the external ear but is said to be too simplistic, a further development was formulated by Chen *et al*

(1992) with the addition of reflected paths. The pinna was modelled using a beam-forming approach. Although this model does not rely on physical principles, it is based on the general geometrical properties of the pinna.

Wightman and Kistler (1997) argued that monaural temporal cues are not likely to be relevant for human sound localisation. Firstly because the HRTF impulse responses are too short to be processed in the auditory system (they are of the order of about 2 msec.), and secondly their previous results (Kistler and Wightman, 1992) suggest that changes in the temporal fine structure of the HRIR do not produce subsequent changes in the apparent positions of sound sources (see also Section 2.7).

2.4.3 Monaural cues: frequency domain interpretation

Although it is accepted now that the pinna acts as a 'frequency domain filter', it is not clear which cues are relevant and necessary for the perception of elevation. Two approaches exist when the spectra of HRTFs are analysed: elevation is perceived through the peaks (the resonance of the pinna), or alternatively, as the notches (the anti-resonance).

It was claimed that for narrow band stimulus, the apparent location is directly related to spectral *peaks* in the subjects HRTFs (Blauert, 1997, Butler, 1997 and Musicant, 1995). However, with regards to vertical localisation, Hebrank and Wright (1974), Butler and Belendiuk (1977), Watkins (1978) and Bloom (1977) have provided strong evidence that with narrow band stimuli, spectral *notches* are responsible for the sensation of source elevation. Shaw (1982) found that in eight out of ten subjects the spectral minima systematically moved along the frequency axis as source elevation varied from high to low. For two subjects the minimum varied in level but not in frequency.

In their work, Hebrank and Wright (1974) used band-pass filters to investigate the frequency range in which the pinna affects localisation. It was concluded that elevation

cues are embedded in a frequency range of 4 kHz to 16 kHz. This conclusion might explain why a large number of reversals occur whenever non-individualised HRTFs are used (Wenzel *et al*, 1993, Møller *et al,* 1996).

The significance of the cues provided by the pinna was demonstrated clearly when its shape was disrupted. Gardner and Gardner (1973) showed that when the pinna is altered by filling its cavities with putty the localisation in the median plane is reduced. Subsequently, more localisation and search for physical cues studies appeared with pinna occlusions (e.g. Oldfield and Parker, 1984, Han, 1991) that supported the significance of the contribution from all parts of the pinna to sound localisation in elevation and also to some extent in the horizontal plane.

2.5 NORMAL MODES

Shaw (1974, 1997) identified six different modes of the human concha. His patterns (1997, pp. 38, also shown in Figure 2-2) were the average among ten subjects. Although the excitation angle, the magnitude at the base of the concha and the resonance frequencies varied between individuals, his identification of monopole, horizontal and vertical dipole patterns provides a very important insight into understanding the physics of the external ear.

These common features have not been used in the construction of individualised response, since it is not straightforward to transform the patterns into HRTFs. To the author's best knowledge his work has not been validated or continued, except the support offered by the work of Middlebrooks (1989), who observed a change in the directional response at 8 kHz and 12 kHz, suggesting correlation with the change of pinna modes from 'vertical' dipoles to 'horizontal' dipoles claimed by Shaw.

In Chapters 7 and 8, we attempt to repeat Shaw's experiment by computer simulation, and to derive a mathematical formulation that can extract the modal characteristics of an individual pinna into its frequency response.

2.6 INDIVIDUALISED AND NON-INDIVIDUALISED HRTFS

Although it has been recognised for a long time that anatomical differences of the head, torso and pinnae exist among individuals and these affect localisation, it was not clear whether these differences require spatial sound reproduction systems to incorporate individualised HRTFs, and in the case of reproduction with headphones, also individualised Headphone Transfer-Functions (HpTFs).

Wenzel *et al* (1993) investigated and compared the localisation of subjects when signals were *synthesised* using their own HRTFs, and with other subjects' HRTFs. When virtual sources in the horizontal plane were investigated, it was concluded that subjects could perform well with non-individualised HRTFs, albeit with the problem of front-back reversals. More errors appeared for 'cone of confusion' elevation angles. It was also suggested that localisation using a non-individualised, but the HRTFs of a 'good localiser', can produce better performance. This idea (of using HRTFs of another person to produce better performance) was investigated also by Morimoto and Ando (1982). They found that subjects could sometimes hear directions in the median plane better using other pinnae, than by localising with their own.

When binaural *recordings* were investigated by Møller *et al* (1996), it was concluded that when localisation of individual recordings (at the blocked entrance of the ear canal) were compared to real life, the performance accuracy was preserved. In a similar manner to the conclusions of Wenzel *et al* (1993), they stated that non-individualised recordings resulted in an increased number of errors for sound in the median plane, and an increased number of

front back reversals. However, the claim that that listeners might localise better with recordings from other individuals then with their own recordings was not supported by their data.

In a recent study, Møller *et al* (1999) investigated non-individualised localisation with various artificial heads. It was concluded that *all* artificial heads produced poorer localisation in listeners, when compared to localisation of real sources. The deterioration was most significant at median plane angles. In addition, it was found that localisation with artificial head HRTFs is comparable to, or poorer than a random human subject. However, they concluded that the design of artificial heads could be improved, since it was found that the number of localisation errors produced by listeners who used non-individualised HRTFs of one of the participating subjects was better than all artificial heads investigated.

Begault *et al* (2000) studied the impact of head tracking, reverberation, and individualised HRTFs on the spatial perception of speech signals. It was concluded that the inclusion of head tracking significantly reduced reversal rates (the cone of confusion is resolved by head motion, see Wallach, 1940). For the horizontal plane, the use of individualised HRTFs did not improve localisation accuracy, or externalisation rates. However it is important to emphasise that speech signals do not include the peaks or notches found in HRTFs due to pinna resonance, so ITD is the primary factor in this study.

2.7 REDUCED ORDER HRTF – MODELLING TECHNIQUES

There are currently a growing number of publications that investigate the performance of reduced cost HRTFs for real-time virtual acoustic systems (for example Wightman and Kistler, 1992, Chen *et al*, 1995, Huopaniemi and Karjalainen, 1997, Evans *et al*, 1998, Gardner, 1999 and Larcher *et al*, 2000). The requirement for efficient coding of HRTFs is for current low-cost real-time applications and cases where multiple virtual sources are

generated simultaneously. In addition, continuous, interpolated HRTFs are required for the generation of moving sources. These interpolations can be achieved in various ways (see for example Christensen *et al*, 1999).

It is suggested that not all the information included in a measured HRIR is required for accurate localisation. For example, as stated in Section 2.4.2 it is well accepted now that modelling the generally non-minimum phase HRTF as a minimum phase transfer function, while preserving the ITD (the interaural group delay at low frequency) results in similar localisation accuracy (Wightman and Kistler, 1989, Middlebrooks and Green, 1990, Kistler and Wightman, 1991, Kulkarni, 1999, Plogsties *et al,* 2000). This result is mainly because the phase information lost in the minimum phase representation is in the high-frequency part of the phase spectrum (a few kHz). This property is very important in the interpolation process in eliminating the comb filtering effect obtained with the original HRTFs. Other simplifications of the magnitude of the spectra generally result in increased errors in elevation and reversals.

The types of HRTF modelling can be classified in two groups: physical models and mathematical models. Most models are currently concentrated on finding a mathematical 'best fit' to empirical HRTF data. These do not attempt to produce individualised HRTF, but to represent the original data in a more compact form.

Shinn-Cunningham and Kulkarni (1996) reviewed the following strategies for HRTF mathematical modelling: eignfunction (Principal Component Analysis), neural-network, Rational function, Pole-Zero (ARMA), and all zero (MA) models.

Recently more mathematical techniques have been suggested. With the inclusion of auditory models, the performances of various models have been investigated subjectively and objectively (Huopaniemi *et al*, 1999). For example, it was suggested that the filters could be described compactly by taking into account the non-uniform resolution of the

auditory system with auditory smoothing and auditory weighting and using frequency warping.

For physical models, only Genuit's structural model (Genuit, 1986, 1987) is a true parametric description that can be used mathematically to construct individualised HRTFs. In his model, an individualised HRTF is constructed from a combination of 16 filter channels and delay elements. Recently, more structural models have been suggested (Brown and Duda, 1998, and Avendano *et al,* 1999). These suggest that elevation cues can be added synthetically based on simplified empirical data (see also Duda, 1997).

In this thesis, we propose an exact physical modelling technique that uses mathematical formulation to extract individualised HRTFs based on individualised basis functions, or spatial patterns of pinna mode shapes (See Chapters 7 and 8). Although the proposed reduced order model is based on physical criteria, its efficiency in the computational cost has not yet been investigated.

2.8 HRTF MEASUREMENT TECHNIQUES

A large body of research has been carried out on HRTF measurement (see Hebrank and Wright, 1974, Mehrgardt and Mellert, 1977, Butler and Belendiuk, 1977, Shaw and Teranishi, 1968, Shaw, 1974, Shaw, 1975, Shaw and Vaillancourt, 1985, Gardner and Martin, 1994, 1995, Møller *et al,* 1995, Hammershøi and Møller, 1996, Carlile and Pralong, 1994, Pralong and Carlile, 1994 and Blauert, 1997).

The external ear transfer functions are measured in the above publications at various places along the ear canal, which gives rise to different responses. However, since up to 12-14 kHz only the longitudinal mode is present in the ear canal, the variation of the position of the microphone does not distort the directionally dependent HRTFs.

Blauert (1997) defines three types of transfer functions: the monaural, the interaural and the free-field transfer function. The monaural transfer function is defined as the measured sound pressure in the ear canal divided by a measurement in the same place, but with the sound source at a reference angle and distance; usually at 0° azimuth and 0° elevation. The interaural transfer function relates sound pressures at equivalent measurements in the two ear canals. The full information of the transformation of sound due to the external ear is preserved when the free-field response is equalised, i.e. the measured response in the ear canal is divided by the measured free field response when the microphone is positioned at the centre of the head, and with the head absent.

2.9 HRTF SIMULATION TECHNIQUES

In the last two decades, various computer simulation techniques have been suggested to model the modification of sound impinging on the human head or parts of the external ear. The growing popularity of the BEM resulted in many papers that studied the steady-state response of simple models (for a review of state-of-the-art papers on the BEM up to 1995, see Ciskowski and Brebbia, Appendix, pp. 261, 1995). However, only recently, problems that are more complex can now be investigated using the BEM (mainly with the IBEM) due to advances in computing power and advanced formulation.

Weinrich (1984) was the first to attempt modelling the response around an 'accurate' geometry of the head. He used analytical and numerical techniques in analysing the response of various parts of the head. He suggested a very simple geometric model of the pinna with a mesh of only 20 elements resembling the shape of the concha. His solution was based on a finite difference approximation method, and the results approximated only roughly the dependence of the first notch with elevation. He also approximated the sound field around an artificial head using the BEM (using a simplified formulation). In this case, the mesh included 212 elements. With a low resolution of elements and without the pinna,

the results were valid only up to approximately 2 kHz. Despite the limited success of his work in modelling realistic HRTFs, Weinrich paved the way for more advanced modelling using the BEM.

Fourteen years later, the work of Katz (1998) is another milestone on the way to acquiring individualised HRTFs using the BEM. Although the idea was similar to the work of Weinrich, the use of BEM models converted from accurate laser-scanned models suggested that although the BEM is associated with 'low-frequency' modelling, we can now solve tens of thousands of simultaneous equations, to predict the response for complex shapes such as the human head. Due to limited computing power, his work was restricted to frequencies below 5 kHz (and this by assuming only four elements per wavelength); as a result, his simulations could not be validated in the high frequency range where pinna resonance and anti-resonance affect the pressure variations (for a discussion about possible geometrical errors see Chapter 5). In addition, the results could not be validated against measurements under the same conditions. His work inspired our research, and especially motivated the question: Can *high frequency* peaks and notches due to the complexity of the external ear be predicted accurately with the BEM?

Ciskowski and Royster (see the work in Ciskowski and Brebbia, 1995) developed a coupled BEM-FEM formulation and investigated the effectiveness of different earplug design configurations. In this work, the ear canal geometry together with the eardrum impedance and the impedance of the earplug were modelled. The results of the response of an open ear canal were compared successfully with measurements, and the attenuation of different earplugs was investigated. This model does not include the concha or the pinna, but suggests the possibilities of investigating the coupling between various earphones and the external ear.

Other numerical approaches were suggested, less computationally demanding. These, however, make use of simplified models and therefore can predict only part of the response of 'real' HRTFs. These are important in the study of the physical origins of the main characteristics of the external ear, found in HRTF measurements. The work of Genuit (1986, 1987) mentioned earlier was based on a structural model. The calculation is based on Kirchoff's diffraction integral. It is reported that this approximate approach yielded results within the tolerance of HRTF measurement of a subject.

Another two examples of such models are given below. Lopez-Poveda and Meddis (1996) presented an approximate physical model of the transfer function of the human concha based on Kirchoff's scalar diffraction theory. They used a simplified model of the concha as a two-dimensional aperture in an infinitely large sound-opaque screen. This model replicated some of the notches (mainly the first one, and partly the third) and its accuracy is limited in the analysis and comparison with the response of an accurate external ear. Speyer (1999) modelled the external ear as three concatenated ducts incorporating the appropriate formulation for the coupling between them. The solution is based on the boundary integral equation, and is proved efficient in replicating the general trends of the first resonance and anti-resonance. As with the work of Lopez-Poveda and Meddis, the simplification of the geometry limits the exact match with real HRTFs, but since the formulation was developed for multiple ducts, it might be possible to improve the accuracy of the prediction with a higher number of ducts (although it has been demonstrated by Shaw (1974, 1975, 1997) that modelling the crus helias is required in matching all the resonance and anti-resonance found in real pinnae).

2.10 ADVANTAGES AND FUTURE APPLICATION OF NUMERICAL MODELLING OF THE HRTF

In recent years, a few HRTF databases have been made available to the research community:

- Average measured responses over 40 years by Shaw and Vaillancourt (1985)
- KEMAR database measured by Gardner and Martin (1994, 1995)
- Human and artificial head measurements in the AUDIS HRTF catalogue (1998)
- Human measurements by Kistler and Wightman (1989)

However, high fidelity HRTFs are still a valuable database that is confined to only well equipped acoustics laboratories. The measurement of individualised HRTF at a high spatial resolution is a very difficult and time-consuming task. Numerical modelling of HRTF might produce cheaper, faster and more accurate HRTFs with a PC in the future. Here we summarise the main advantages:

- There is no need for an anechoic chamber, or any acoustical transducer or apparatus.
- In the simulation, both microphones and sources are ideal with flat and omnidirectional response. 'Real' empirical response could be simulated if required.
- The model and transducers could be accurately positioned anywhere in space, thus alleviating the problem of comparing HRTFs between studies, and highlighting the results which differ only due to anatomical geometry or different boundary conditions.
- The source could be positioned at any distance in the near field or far field.

- The 'microphone' could be positioned at any point along the ear canal (such as directly on the eardrum). It is also advantageous to position a large number of 'microphones' anywhere on the pinna to investigate pressure variation or 'mode shapes'.
- The principle of reciprocity can be used to obtain continuous HRTF anywhere in space with no need to interpolate between discrete data sampled at a low spatial rate of sampling.
- The problem associated with low accuracy of measurement of both magnitude and phase at very low (< 300 Hz), and high (> 15 kHz) frequency due to both transducers' performance and due to poor SNR are both alleviated.
- The exact free-field or diffuse fields can be calculated and visualised in the frequency or time domains. This is possible for both loudspeaker and headphone reproduction.
- It is possible to control the characteristics of air to facilitate higher or lower temperature, humidity and thus characteristic impedance.
- It is possible to control any boundary conditions such as clothing, hair or eardrum impedance and investigate their individual contribution.
- It is possible to generate a 'bank' of HRTFs by modifying and altering the shape of pinnae, head shape and size, neck, torso, etc.
- It might be possible in the future to use a digital camera, already part of a PC, to extract the 3-D geometry of the pinna, and the dimensions of the head, neck and shoulders. This information could be used to generate a reduced order HRTF based on the physical attributes of each component (either with structural modelling or HRTF decomposition based on spatial mode shapes).

- It might be possible in the future to work backwards (through genetic algorithm techniques for example) and to design a specific shape of the pinna, head, torso, etc. for the 'best set' of HRTFs.

2.11 CONCLUSIONS

In this chapter, we reviewed briefly the basic aspects of the physics of the external ear and its relation to sound localisation. The significance of the external ear is now recognised, but it is not clear yet, what information encoded in the HRTF is essential for the auditory system. If individualised HRTFs are used in binaural synthesis, virtual sources are perceived with similar accuracy and authenticity when compared to real sources. However, if non-individualised HRTFs are used, or even reduced-order individualised HRTFs are implemented, not all localisation cues will be present, and therefore the performance of the implemented virtual source will depend on the application, type of stimulus, source position in space, etc. A simulation tool based on the BEM can be used either to obtain individualised HRTFs or to investigate the characteristics of the pinna. The success of this approach depends greatly on its accuracy when compared with measurements. Only when the individual resonance and anti-resonance can be predicted, will the simulation be offered as an alternative to measurement techniques.

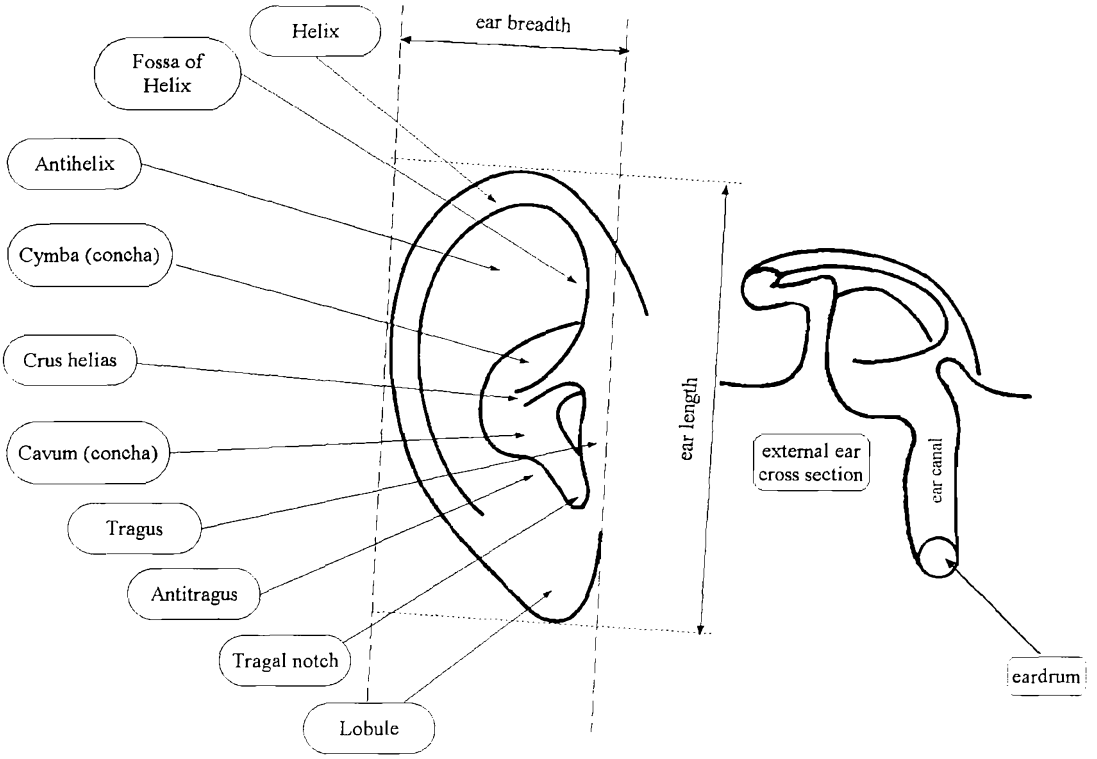


Figure 2-1: The external ear. On the left, the different parts of the pinna, and on the right a schematic cross section with the ear canal and eardrum. After Shaw (1997).

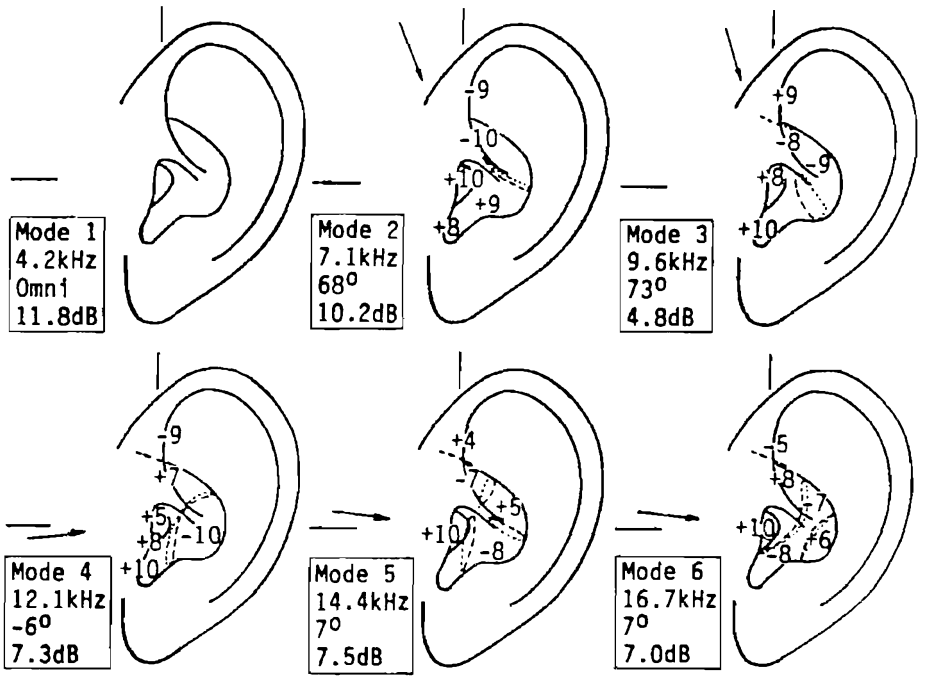


Figure 2-2: Average normal mode shapes of the concha. The taxonomy of the modes for three cases: 'monopole' (mode 1), vertical 'dipoles' (mode 2 and mode 3), and horizontal 'dipoles' (modes 4 to 6). After Shaw (1997).

CHAPTER 3

BACKGROUND - NUMERICAL METHODS FOR EXTERIOR PROBLEMS IN ACOUSTICS

3.1 INTRODUCTION

Among the exact numerical prediction techniques for solving vibro-acoustic problems, the Finite Element Method (FEM) and the Boundary Element Method (BEM) are the most commonly used. They are different from analytical, geometrical, and statistical energy methods in acoustics in that they provide a tool for accurately solving detailed, arbitrary, and complex models.

Both methods are based on the discretisation of the acoustic field and can solve the wave equation, either in the time or frequency domains. As the acoustic domain becomes larger, the number of elements increases rapidly, and therefore these methods are mainly associated with 'low' frequency modelling. FEM is mostly used in structural engineering, and its main limitation in acoustics is the ability to deal only with finite elements (i.e. interior problems). However, in recent years, a new development of FEM has appeared: the Infinite Finite Elements Method (IFEM). The infinite elements alleviate the problem of infinite domains. It is predicted that the popularity of the method will increase in the near future, with significant reduction in the computational effort that is currently required for solving large problems.

However, the main tool for predicting radiation and scattering for exterior problems is still based on Boundary Integral Equations (BIEs), which were already used by the 1800s by

Helmholtz and Rayleigh, but developed only in the 1960s into the BEM, at the same time as computers appeared and proliferated.

In this research, emphasis has been placed on the investigation of 'real' problems that include both large and arbitrary complex models. The main tool was the BEM, although some preliminary investigation of the suitability of the IFEM was also carried out. This chapter summarises briefly only the relevant theory of the numerical methods used in this research. For a thorough summary and mathematical derivation, the reader is referred to the following references: Ciskowsky and Brebbia (1995), which include also a list of state-of-the-art papers in the BEM up to 1995, SYSNOISE (1999) and Desmet (1997), or as otherwise indicated throughout this chapter.

3.2 BASIC THEORY

3.2.1 Governing equations

In a homogeneous, inviscid, compressible fluid, the acoustic pressure field due to radiation or scattering from a rigid body, is governed by the classical scalar wave equation

$$\nabla^2 p(\mathbf{r},t) - \frac{1}{c^2}\frac{\partial^2 p(\mathbf{r},t)}{\partial t^2} = 0 \qquad \mathbf{r} \in \mathrm{V} \tag{3.1}$$

where $p(\mathbf{r},t)$ denotes the instantaneous variation of pressure from the equilibrium that is termed the acoustic pressure, and the spatial vector $\mathbf{r}$ is defined inside the volume V. The speed of sound c is given by $c = \sqrt{(\gamma p_0 / \rho_0)}$, where p_0 is the gas static pressure, ρ_0 is the static density, and γ is the ratio of specific heats.

When time harmonic steady state conditions are considered, the wave equation (3.1) is reduced to the Helmholtz equation

$$\nabla^2 p(\mathbf{r}) + k^2 p(\mathbf{r}) = 0 \qquad \mathbf{r} \in \mathrm{V} \tag{3.2}$$

where the wavenumber k is defined as $k = \omega/c$, ω is the angular frequency, $\omega = 2\pi f$ and f is the frequency.

3.2.2 Boundary conditions

Three types of boundary conditions are associated with the Helmholtz equation: fixed pressure, normal velocity, or acoustic impedance on the surface of the body. These are called Dirichlet, Neumann, and mixed (or Robin) boundary conditions, respectively.

The prescribed normal velocity u_n on the vibrating body is given by

$$u_n = \frac{j}{\omega\rho_0}\nabla p(\mathbf{r}).\mathbf{n} \qquad \mathbf{r} \in S_0 \tag{3.3}$$

where $\mathbf{n}$ is the unit normal on the surface S_0.

An additional boundary condition is needed for radiation problems; the pressure field at infinity must satisfy the Sommerfeld condition

$$\lim_{r\to\infty} r^{\alpha}(\frac{\partial p}{\partial r} + jk\rho_0) = 0 \tag{3.4}$$

where $\alpha = 1/2$, or $\alpha = 1$ for two or three-dimensional domains, respectively. This condition ensures that energy disappears at infinity and only outgoing travelling waves propagate.

3.2.3 Green function

The solution of the wave equation (3.1) is simplified by introducing the acoustic monopole (point) source. The Green function satisfies the inhomogeneous Helmholtz equation

$$(\nabla^2 + k^2)g(\mathbf{r} \mid \mathbf{r}_0) = -\delta(\mathbf{r} - \mathbf{r}_0) \tag{3.5}$$

with $\delta(\mathbf{r} - \mathbf{r}_0)$ the 3-D Dirac function.

The Sommerfeld radiation condition is also applied and this can be written as

$$\lim_{r\to\infty} r\left(\frac{\partial g(\mathbf{r} \mid \mathbf{r}_0)}{\partial r} + jkg(\mathbf{r} \mid \mathbf{r}_0)\right) = 0 \tag{3.6}$$

The free space Green function* represents the free-field pressure in point $\mathbf{r}$ due to an acoustic point source in location $\mathbf{r}_0$ with source strength of a unity

$$g(\mathbf{r} \mid \mathbf{r}_0) = \frac{e^{-jk|\mathbf{r}-\mathbf{r}_0|}}{4\pi \mid \mathbf{r} - \mathbf{r}_0 \mid} \tag{3.7}$$

with $|\mathbf{r} - \mathbf{r}_0|$ the distance between points $\mathbf{r}$ and $\mathbf{r}_0$.

3.3 DIRECT COLLOCATIONAL BOUNDARY ELEMENT METHOD

The direct collocation BEM can be used for solving exterior or interior problems. In our case, we will concentrate on exterior problems in general and on acoustic scattering in particular. The goal of the method is to calculate the pressure and particle velocity in the exterior domain of a body. Once the above acoustic variables have been found, the field variables at any point in the continuum domain can be obtained by using the Boundary Integral Equation (BIE). The three types of boundary conditions described above (Section 3.2.1) can be applied, and the set of equations can be constructed using the direct collocation scheme. The term 'direct' indicates that the pressure and normal velocity distribution on the boundary surface have a direct physical meaning. The method's main advantage is in solving small to medium size problems in which the models have a closed boundary surface (see Chapter 4 and Chapter 5 Section 5.5).

* The notation of the free space Green function is given by $g(\mathbf{r} \mid \mathbf{r}_0)$ since a general Green function $G(\mathbf{r} \mid \mathbf{r}_0)$ is defined and used in Chapter 7.

3.3.1 Direct boundary integral formulation

The direct boundary integral formulation (Schenk, 1968) relates the pressure at any point of an acoustic field to the pressure and normal velocity distribution on the closed boundary surface of the acoustic domain.

For the case of scattering, consider a body with a surface S embedded in an infinite acoustic medium exposed to impinging incident waves with pressure p_i. The total field p_t is a superposition of the incident field p_i and the scattered field p_s. (For the case of radiation, the pressure p_i vanishes). The volume V is the space bounded by S and the surface Σ at infinity, which can be represented by a sphere with an infinite radius (see Figure 3-1). The small sphere σ with a radius $\varepsilon \rightarrow 0$ is defined to deal with the singularity in the Green function when $\mathbf{r} = \mathbf{r}_0$.

Based on Green's third identity that relates the surface integral over S to the volume integral over V bounded by S, it can be shown that the integral formulation for the scattered sound field is given by

$$C(\mathbf{r})p_s(\mathbf{r}) = \int_S [p_s(\mathbf{r}_0)\frac{\partial g(\mathbf{r} \mid \mathbf{r}_0)}{\partial n} - g(\mathbf{r} \mid \mathbf{r}_0)\frac{\partial p_s(\mathbf{r}_0)}{\partial n}]dS(\mathbf{r}_0) \tag{3.8}$$

where the coefficient $C(\mathbf{r})$ represents the solid angle, expressed as a fraction of 4π. This coefficient is 1 for positions inside V, 0 for positions in the interior domain enclosed by the boundary surface S and ½ for positions on the surface where the normal direction is uniquely defined.

For other cases the solid angle can be calculated (see Seybert *et al*, 1985 for the full derivation) for any surface position $\mathbf{r}$ from

$$C(\mathbf{r}) = 1 + \frac{1}{4\pi}\int_S \frac{\partial}{\partial \nu}(\frac{1}{|\mathbf{r} - \mathbf{r}_0|})dS(\mathbf{r}_0) \tag{3.9}$$

where ν is the normal to the closed boundary surface S with a positive orientation into the unbounded domain V ($\nu = -n$).

The total sound field can then be calculated from

$$C(\mathbf{r})p_t(\mathbf{r}) = \int_S [p_t(\mathbf{r}_0)\frac{\partial g(\mathbf{r} \mid \mathbf{r}_0)}{\partial n} - g(\mathbf{r} \mid \mathbf{r}_0)\frac{\partial p_t(\mathbf{r}_0)}{\partial n}]dS(\mathbf{r}_0) + p_i(\mathbf{r}) \tag{3.10}$$

or

$$C(\mathbf{r})p_t(\mathbf{r}) = \int_S [p_t(\mathbf{r}_0)\frac{\partial g(\mathbf{r} \mid \mathbf{r}_0)}{\partial n} + j\rho_0 w u_n(\mathbf{r}_0) g(\mathbf{r} \mid \mathbf{r}_0)]dS(\mathbf{r}_0) + p_i(\mathbf{r}) \tag{3.11}$$

where $p_t(\mathbf{r}_0)$ and $u_n(\mathbf{r}_0)$ are the pressure and normal velocity distribution on the closed boundary surface S. (For radiation problems the term p_i should be omitted).

3.3.2 Numerical implementation

In order to evaluate the integral of Equation (3.11) on the surface S, the body needs to be discretised into E elements that approximate the surface shape and the acoustical variables on the surface. The most commonly used elements are the linear triangular and quadrilateral surface elements.

For the case of linear triangular elements, the global Cartesian co-ordinates x_i $(i = 1,2,3)$, are related to the nodal co-ordinates $x_{i\alpha}$ of each element,

$$x_i(\xi) = \sum_\alpha N_\alpha(\xi) x_{i\alpha} \quad \alpha = 1,2,3 \tag{3.12}$$

and the first order shape functions of the local co-ordinates $(\xi) \equiv (\xi_1, \xi_2, \xi_3)$ and are $N_\alpha(\xi) = \xi_i$ $(i = 1,2,3)$. Each shape function is defined such that it has a unit value at the location of node i and that it is zero at all other node locations (see Figure 3-2).

In each element m, the values of the pressure p and its derivative are approximated by $p_m(\xi)$ and $\frac{\partial p_m}{\partial n}(\xi)$ using the shape functions.

$$p_m(\xi) \cong \sum_{\alpha} N_\alpha(\xi) p_{m\alpha}$$

and

$$\frac{\partial p_m(\xi)}{\partial n} \cong \sum_{\alpha} N_\alpha(\xi) \frac{\partial p_{m\alpha}(\xi)}{\partial n} \tag{3.13}$$

where $p_{m\alpha}(\xi)$ and $\frac{\partial p_{m\alpha}(\xi)}{\partial n}$ are the values of the pressure and its gradient on the element m, respectively.

Hence, after rearranging, the discretised formulation of Equation (3.8) is given by

$$\sum_m \int_{S_m} p_m(\mathbf{r}_0) \frac{\partial g(\mathbf{r} \mid \mathbf{r}_0)}{\partial n} dS(\mathbf{r}_0) - C(\mathbf{r}) p(\mathbf{r}) = \sum_m \int_{S_m} \frac{\partial p_m(\mathbf{r}_0)}{\partial n} g(\mathbf{r} \mid \mathbf{r}_0) dS(\mathbf{r}_0) \tag{3.14}$$

where S_m is the area of the m^{th} element. Substitution of Equations (3.12) and (3.13) into this equation gives

$$\begin{aligned} \sum_m \int_{S_m} \sum_{\alpha} N_\alpha(\xi) p_{m\alpha} \frac{\partial g(\mathbf{r} \mid \mathbf{r}_0)}{\partial n} J(\xi) d\xi - C(\mathbf{r}) p(\mathbf{r}) = \\ \sum_m \int_{S_m} \sum_{\alpha} N_\alpha(\xi) \frac{\partial p_{m\alpha}}{\partial n} g(\mathbf{r} \mid \mathbf{r}_0) J(\xi) d\xi \end{aligned} \tag{3.15}$$

where $J(\xi)$ the Jacobian of the transformation given in Equation (3.12). Equation (3.15) can be rewritten as

$$\begin{aligned} \sum_m \sum_{\alpha} p_{m\alpha} \int_{S_m} N_\alpha(\xi) \frac{\partial g(\mathbf{r} \mid \mathbf{r}_0)}{\partial n} J(\xi) d\xi - C(\mathbf{r}) p(\mathbf{r}) = \\ \sum_m \sum_{\alpha} \frac{\partial p_{m\alpha}}{\partial n} \int_{S_m} N_\alpha(\xi) g(\mathbf{r} \mid \mathbf{r}_0) J(\xi) d\xi \end{aligned} \tag{3.16}$$

Now consider one dedicated point $\mathbf{r}_j$ that is a node on the surface identified by the global number j.

Assume now the following notations:

$p_{j=}p(\mathbf{r}_j)$ is the pressure at node j, $R_j = R(\mathbf{r}_j, \mathbf{r}_0)$ is the distance from node j to any point $\mathbf{r}_0$ on the surface, $g_j = g(\mathbf{r}_j \mid \mathbf{r}_0)$ is the free space Green function between node j and any point $\mathbf{r}_0$ on the surface.

Equation (3.16) can be rewritten as

$$\sum_m \sum_\alpha p_{m\alpha} \cdot a_{mj}^\alpha - p_j[1 + \sum_m C_{mj}] = \sum_m \sum_\alpha \frac{\partial p_{m\alpha}}{\partial n} \cdot b_{mj}^\alpha \tag{3.17}$$

where

$$a_{mj}^\alpha = \int_{S_m} N_\alpha(\xi) \frac{\partial g(\mathbf{r} \mid \mathbf{r}_0)}{\partial n} J(\xi) d\xi \tag{3.18}$$

$$b_{mj}^\alpha = \int_{S_m} N_\alpha(\xi) g(\mathbf{r} \mid \mathbf{r}_0) J(\xi) d\xi \tag{3.19}$$

$C(\mathbf{r})$ is defined in Equation (3.9) and

$$C_{mj} = \frac{1}{4\pi} \int_{S_m} \frac{\partial}{\partial \nu} \frac{1}{R_j} J(\xi) d\xi \tag{3.20}$$

Each m, α combination corresponds to a global node l. However, the global node l may have its origin in different m, α combinations.

By adopting the global scheme we may write

$$\sum_m \sum_\alpha a_{mj}^\alpha \cdot \phi_{m\alpha} \equiv \sum_l \hat{A}_{jl} \phi_l \tag{3.21}$$

$$\sum_m \sum_\alpha b_{mj}^\alpha \cdot \frac{\partial \phi_{m\alpha}}{\partial n} \equiv \sum_l B_{jl} \frac{\partial \phi_l}{\partial n} \tag{3.22}$$

Using Equations (3.21) and (3.22) in Equation (3.17) gives

$$\sum_m \sum_\alpha \hat{A}_{jl} p_l - [1 + \sum_m C_{mj}] p_j = \sum_l B_{jl} \frac{\partial p_l}{\partial n} \tag{3.23}$$

The pressure p_j at node j, can be expressed as $\delta_{jl} p_l$ with δ_{jl} the Kronecker symbol which equals 1 only when $j = l$, therefore we can define

$$A_{jl} \equiv \hat{A}_{jl} - (1 + \sum_m C_{mj}) \delta_{jl} \tag{3.24}$$

Therefore Equation (3.23) becomes

$$\sum_l A_{jl} p_l = \sum_l B_{jl} \frac{\partial p_l}{\partial n} \tag{3.25}$$

At each node l, the pressure and the pressure gradient are related to the same quantites at all other nodes on the surface. For a total number of N nodes , a system of N simultaneous linear algebric equations is created. This can be written in a matrix form as

$$[\mathbf{A}]\{\mathbf{p}\} = [\mathbf{B}]\left\{\frac{\partial \mathbf{p}}{\partial n}\right\} \tag{3.26}$$

where the $N \times N$ matrices $\mathbf{A}$ and $\mathbf{B}$ are made up of combinations of Equations (3.18) to (3.20) . These matrices are fully populated since the pressure and normal velocity at a certain node are related to the values at all other nodes on the surface. The number of simultaneous equations that are solved is the same as the number of nodes on the discretised surface plus the number of source points.

Once the pressure and the normal velocity at each node have been found, the pressure at any point $\mathbf{r}$ in the acoustic domain V (which is not located on S due to singularity) can be obtained from Equation (3.11), which in a matrix form is given by

$$\mathbf{p}(\mathbf{r}) = [\mathbf{C}]\{\mathbf{p}\} + [\mathbf{D}]\{\mathbf{u}_n\} \tag{3.27}$$

where the coefficients $\mathbf{C}, \mathbf{D}$ are integrals evaluated at each node (with a similar procedure to that described in Equation (3.17). These are also given in detail in SYSNOISE, 1999 and Desmet, 1997).

3.3.3 The non-uniqueness problem

A common problem associated with exterior problems solved with the DBEM is that the solution breaks down at certain characteristic frequencies. As proved by Schenck (1968), these frequencies are the eignfrequencies of the corresponding interior problem with modified boundary conditions, i.e. for an exterior Dirichlet problem the frequencies are the eignfrequencies of the interior Neumann problem and *vice versa.* These frequencies appear purely for mathematical and not physical reasons.

These 'irregular frequencies' are removed using the Combined Helmholtz Integral Equation Formulation (CHIEF), proposed by Schenck (1968). This method is based on adding a few equations that collocate the solution at points where $C(\mathbf{r}) = 0$, where it is known that the pressure is zero. These points are also called 'over-determination' points and the over-determined set of equations is solved using the Lagrange multiplier technique. It is required that these points will not be positioned inside the interior cavity on nodal surfaces since they will not add 'real' constraints. As the frequency increases, these nodal surfaces become closely spaced and a careful positioning of these points is required.

Other methods (such as the 'Burton Miller' method which includes additional constraints for the normal derivation of the pressure) exist but are not investigated in the current research.

3.4 INDIRECT VARIATIONAL BOUNDARY ELEMENT METHOD

3.4.1 Indirect boundary integral equation

As stated in the previous section, the DBEM can be used only when the boundary surface S is closed, thus the sound field can be calculated either inside *or* outside the boundary surface. For cases where the domain is open, or includes both closed *and* open boundary surfaces, the Indirect Boundary Element Method (IBEM) is used (see Fillipi, 1977).

The indirect method uses layer potentials that are the differences between the outside and inside values of the pressure and its normal gradient. Along the positive side of the boundary surface these are denoted $p^+(\mathbf{r}_0)$ and $\partial p^+(\mathbf{r}_0)/\partial n$, respectively, and along the negative side of the boundary are denoted $p^-(\mathbf{r}_0)$ and $\partial p^-(\mathbf{r}_0)/\partial n$, respectively (See Figure 3-3a).

The layer potentials are defined as

$$\mu(\mathbf{r}_0) = p^+(\mathbf{r}_0) - p^-(\mathbf{r}_0) \tag{3.28}$$

where μ is generally called the jump of pressure or the *double layer potential.* It represents a distribution of *dipole* sources on the surface, and

$$\sigma(\mathbf{r}_0) = \frac{\partial p^+(\mathbf{r}_0)}{\partial n} - \frac{\partial p^-(\mathbf{r}_0)}{\partial n} \tag{3.29}$$

where σ is generally called the jump of normal derivative of pressure or *single layer potential*, and it represents a distribution of *monopole* sources on the surface.

For open bodies (see Figure 3-3b), it can be shown (Coyette and Lecomte, 1997, SYSNOISE, 1999) that the indirect boundary integral formulation is given by

$$p(\mathbf{r}) = \int_{s^+} \left\{ p^+(\mathbf{r}_0) \frac{\partial g(\mathbf{r} \mid \mathbf{r}_0)}{\partial n_y} - \frac{\partial p^+(\mathbf{r}_0)}{\partial n_y} g(\mathbf{r} \mid \mathbf{r}_0) \right\} dS(\mathbf{r}_0)$$
$$- \int_{s^-} \left\{ p^-(\mathbf{r}_0) \frac{\partial g(\mathbf{r} \mid \mathbf{r}_0)}{\partial n_y} - \frac{\partial p^-(\mathbf{r}_0)}{\partial n_y} g(\mathbf{r} \mid \mathbf{r}_0) \right\} dS(\mathbf{r}_0) \quad (\mathbf{r} \in \mathrm{V}) \tag{3.30}$$

or, in a general form, for open or closed boundary surfaces

$$p(\mathbf{r}) = \int_S [\mu(\mathbf{r}_0) \frac{\partial g(\mathbf{r} \mid \mathbf{r}_0)}{\partial n} - \sigma(\mathbf{r}_0) g(\mathbf{r} \mid \mathbf{r}_0)] dS_a(\mathbf{r}_0) \qquad (\mathbf{r} \in V) \tag{3.31}$$

The term 'indirect' indicates that the boundary variables, i.e. the monopole and dipole distributions on the boundary surface, do not represent any direct physical quantities associated directly with the pressure field.

3.4.2 Variational formulation

The variational formulation is used with the IBEM because it is difficult to apply the collocation method due to the problems of singularities in Equation (3.31).

The single and double layer potentials σ and μ are sought on the three different parts of S according to the boundary conditions: imposed pressure on S_1 ($\mu = 0$ and $\sigma \neq 0$), imposed normal velocity on S_2 ($\sigma = 0$, $\mu \neq 0$ and $\partial p / \partial n = -j\rho_0 \omega u_n$), and imposed normal impedance on S_3 ($\sigma = -jk\beta\mu$, $\mu \neq 0$ and $\partial p / \partial n = -jk\beta p$ where β is the specific admittance function).

Therefore the following boundary conditions (derived in Desmet, 1997) can be reformulated as

$$p(\mathbf{r}) = -\int_{S_1} \sigma(\mathbf{r}_0) g(\mathbf{r} \mid \mathbf{r}_0) dS_1(\mathbf{r}_0) + \int_{S_2} \mu(\mathbf{r}_0) \frac{\partial g(\mathbf{r} \mid \mathbf{r}_0)}{\partial n(\mathbf{r}_0)} dS_2(\mathbf{r}_0)$$
$$+ \int_{S_3} [\mu(\mathbf{r}_0) \frac{\partial g(\mathbf{r} \mid \mathbf{r}_0)}{\partial n(\mathbf{r}_0)} + jk\beta(\mathbf{r}_0) g(\mathbf{r} \mid \mathbf{r}_0)] dS_3(\mathbf{r}_0) \quad (\mathbf{r} \in S_1) \tag{3.32}$$

$$-j\rho_0\omega u_n(\mathbf{r}) = -\int_{S_1}\sigma(\mathbf{r}_0)\frac{\partial g(\mathbf{r}\mid\mathbf{r}_0)}{\partial n(\mathbf{r})}dS_1(\mathbf{r}_0)+\int_{S_2}\mu(\mathbf{r}_0)\frac{\partial^2 g(\mathbf{r}\mid\mathbf{r}_0)}{\partial n(\mathbf{r})\partial n(\mathbf{r}_0)}dS_2(\mathbf{r}_0)$$
$$+\int_{S_3}[\mu(\mathbf{r}_0)\frac{\partial^2 g(\mathbf{r}\mid\mathbf{r}_0)}{\partial n(\mathbf{r})\partial n(\mathbf{r}_0)}+jk\beta(\mathbf{r}_0)\frac{\partial g(\mathbf{r}\mid\mathbf{r}_0)}{\partial n(\mathbf{r})}]dS_3(\mathbf{r}_0)\quad(\mathbf{r}\in S_2) \tag{3.33}$$

$$0 = -\int_{S_1}\sigma(\mathbf{r}_0)\frac{\partial g(\mathbf{r}\mid\mathbf{r}_0)}{\partial n(\mathbf{r})}dS_1(\mathbf{r}_0)+\int_{S_2}\mu(\mathbf{r}_0)\frac{\partial^2 g(\mathbf{r}\mid\mathbf{r}_0)}{\partial n(\mathbf{r})\partial n(\mathbf{r}_0)}dS_2(\mathbf{r}_0)$$
$$+\int_{S_3}[\mu(\mathbf{r}_0)(\frac{\partial^2 g(\mathbf{r}\mid\mathbf{r}_0)}{\partial n(\mathbf{r})\partial n(\mathbf{r}_0)}+jk\beta(\mathbf{r}_0)\frac{\partial g(\mathbf{r}\mid\mathbf{r}_0)}{\partial n(\mathbf{r})}]dS_3(\mathbf{r}_0)$$
$$-\int_{S_1}jk\beta(\mathbf{r})\sigma(\mathbf{r}_0)g(\mathbf{r}\mid\mathbf{r}_0)dS_1(\mathbf{r}_0)+\int_{S_2}jk\beta(\mathbf{r})\mu(\mathbf{r}_0)\frac{\partial g(\mathbf{r}\mid\mathbf{r}_0)}{\partial n(\mathbf{r}_0)}dS_2(\mathbf{r}_0) \tag{3.34}$$
$$+\int_{S_3}[\mu(\mathbf{r}_0)jk\beta(\mathbf{r})\frac{\partial g(\mathbf{r}\mid\mathbf{r}_0)}{\partial n(\mathbf{r}_0)}-k^2\beta(\mathbf{r})\beta(\mathbf{r}_0)g(\mathbf{r}\mid\mathbf{r}_0)]dS_3(\mathbf{r}_0)\quad(\mathbf{r}\in S_3)$$

Equations (3.32) to (3.34) can be written in a compact form as follows

$$\begin{aligned} f_1(\sigma,\mu) &= p && \forall\ \mathbf{r}\in S_1\\ f_2(\sigma,\mu) &= -j\rho\omega u_n && \forall\ \mathbf{r}\in S_2\\ f_3(\sigma,\mu) &= 0 && \forall\ \mathbf{r}\in S_3 \end{aligned} \tag{3.35}$$

These integral equations have to be solved for σ on S_1, and μ on S_2 and S_3. An equivalent variational formulation is given by

$$\int_{S_1} f_1(\sigma,\mu)\delta\sigma dS_1+\int_{S_2} f_2(\sigma,\mu)\delta\mu dS_2+\int_{S_3} f_3(\sigma,\mu)\delta\mu dS_3 =$$
$$\int_{S_1} p\delta\sigma dS_1-\int_{S_2} j\omega u_n\delta\mu dS_2 \qquad \forall(\delta\sigma,\delta\mu) \tag{3.36}$$

The left hand side is a bilinear symmetric form so that the solution (σ,μ) leads to the stationarity of a functional J so that $\delta J = 0$ for all possible variations $(\delta\sigma,\delta\mu)$. The full mathematical derivation of the solution of the variational formulation is shown in detail in Desmet (1997) and SYSNOISE (1999).

3.4.3 Formulation in a discretised form

In a similar way to the discretisation of S in the DBEM, the geometric shape functions are defined and the unknowns σ and μ are expressed in term of their nodal values:

$$\begin{aligned} \sigma(\mathbf{r}) &= \sum \sigma_i N_i(\xi) \\ \mu(\mathbf{r}) &= \sum \mu_i N_i(\xi) \end{aligned} \tag{3.37}$$

It can be shown (see Pierce and Wu, 1983) that the unknown layer potentials are given by

$$\begin{Bmatrix} \boldsymbol{\sigma} \\ \boldsymbol{\mu} \end{Bmatrix} = \begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{C}^T & \mathbf{D} \end{bmatrix}^{-1} \begin{Bmatrix} \mathbf{f}_\sigma \\ \mathbf{f}_\mu \end{Bmatrix} \tag{3.38}$$

where the detailed integral equations of $\mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{f}_\sigma, \mathbf{f}_\mu$ are given in SYSNOISE (1999) and Desmet (1997).

Once both the single and the double layers are known, the BIE can be used to post process the pressure at any field point in the volume which is then given by

$$p(\mathbf{r}) = \sum_e \int_{S^e} [N_\mu(\mathbf{r}_0)\mu \frac{\partial g(\mathbf{r} \mid \mathbf{r}_0)}{\partial n(\mathbf{r}_0)} - N_\sigma(\mathbf{r}_0)\sigma g(\mathbf{r} \mid \mathbf{r}_0)] dS^e(\mathbf{r}_0) \tag{3.39}$$

where the surface S of the body is discretised with e elements such that $S \cong \sum_e S^e$, and S^e is the surface of each element.

3.4.4 The non-uniqueness problem

When the indirect method is used to solve exterior problems with closed domains, the integral evaluation fails at critical frequencies of the interior problem, in a similar way to that which appears with the DBEM. The solution in this case is based on adding interior elements with additional boundary conditions. These are defined as

$$\frac{\partial p^+}{\partial n} + j\rho\omega \mathrm{A}_n p^+ = 0$$

$$\frac{\partial p^-}{\partial n} - j\rho\omega \mathrm{A}_n p^- = 0 \tag{3.40}$$

The absorption, applied as positive real parts of admittance (to increase absorption) can be applied also on the inside layer of the existing closed surface. Further information and practical consideration of implementing these 'singular admittance' elements are given in SYSNOISE (1999).

3.5 SPECIAL FORMULATION OF THE BEM

3.5.1 Symmetry

Baffled structures (Figure 3-4) or structures with symmetry in both geometric and material properties can benefit from a modified formulation which does not require discretisation of the infinite plane. The modified equations are given by

$$\int_{S^H} \left\{ p(\mathbf{r}_0) \frac{\partial g_{\mathbf{H}}(\mathbf{r} \mid \mathbf{r}_0)}{\partial n} - g_{\mathbf{H}} \frac{\partial p(\mathbf{r}_0)}{\partial n} \right\} dS^H(\mathbf{r}_0) = 0$$

$$\int_{S'} \left\{ p(\mathbf{r}_0) \frac{\partial g_{\mathbf{H}}(\mathbf{r} \mid \mathbf{r}_0)}{\partial n} - g_{\mathbf{H}} \frac{\partial p(\mathbf{r}_0)}{\partial n} \right\} dS'(\mathbf{r}_0) = C(\mathbf{r}) p(\mathbf{r}) \tag{3.41}$$

where the vibrating body S' lies in the vicinity of an infinite reflecting baffle S^H. The Green function of a rigid infinite baffle (with zero normal velocity) is $g_{\mathbf{H}} = (e^{-jkR})/R + (e^{-jkR_1})/R_1$ and the Green function of a soft infinite baffle (with zero pressure) is $g_{\mathbf{H}} = (e^{-jkR})/R - (e^{-jkR_1})/R_1$.

3.5.2 Axisymmetric IBEM

In the case of an axisymmtric model, with non-axisymmetric boundary conditions, a reduction in both calculation time and computational storage may be achieved. The 3-D

surface integral of the Helmholtz integral equation can be reduced to a combination of a line integral, and an integral over the angle of revolution, around the z axis (Coyette and Lecomte, 1997). The reference plane for the meridian surface is (r,z) of a cylindrical co-ordinate system (Figure 3-5).

The acoustic pressure can be decomposed into Fourier series along the circumferential direction using

$$p(r_x, z_x, \theta_x) = \sum_{m=-\infty}^{+\infty} p_m(r_x, z_x) e^{jm\theta_x} \tag{3.42}$$

and

$$p_m(r_x, z_x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} p(r_x, z_x, \theta_x) e^{-jm\theta_x} d\theta_x \tag{3.43}$$

After substitution to the Helmholtz equation and following the same mathematical steps that were applied in the 3-D case, the integral representation becomes

$$p_m(\mathbf{r}) = \int_L \left\{ \sigma_m(\mathbf{r}_0) g_m(\mathbf{r} \mid \mathbf{r}_0) - \mu_m(\mathbf{r}_0) \frac{\partial g_m(\mathbf{r} \mid \mathbf{r}_0)}{\partial n(\mathbf{r}_0)} \right\} r(\mathbf{r}) dL(\mathbf{r}_0)$$

where L is the boundary surface generator. The non axisymmetric boundary conditions are handled by their decomposition into Fourier series. For each harmonic m the integral equations are solved with the variational technique described in Section 3.4.2.

3.5.3 Acoustic transparency in the IBEM

The following formulation of the IBEM is a special case of the integral Equation (3.31). Its main use is in problems where the transmissibility of sound by vibro-acoustic interactions can be analysed on both sides of an infinite baffle.

The pressure can be analysed as a superposition of the contribution by the single and double potentials on the three parts of $S: S^l$, the part of S in V^-, S^m, the part of S in the infinite plane, and S^r, the part of S in V^+ (see Figure 3-6).

It is shown in detail by Coyette *et al* (1999), that the total pressure is given by

$$
\begin{aligned}
p(\mathbf{r}) = & \int_{S^m} \left\{ + \frac{\partial p^-(\mathbf{r}_0)}{\partial n(\mathbf{r}_0)} . g(\mathbf{r} \mid \mathbf{r}_0) \right\} dS(\mathbf{r}_0) \\
& + \int_{S^l} \left\{ \mu(\mathbf{r}_0) . \frac{\partial g(\mathbf{r} \mid \mathbf{r}_0)}{\partial n(\mathbf{r}_0)} - \sigma(\mathbf{r}_0) . g(\mathbf{r} \mid \mathbf{r}_0) \right\} dS(\mathbf{r}_0) \qquad \forall\ \mathbf{r} \in V^- \qquad (3.44) \\
& + \int_{T} \left\{ - \frac{\partial p(\mathbf{r}_0)}{\partial n(\mathbf{r}_0)} . g(\mathbf{r} \mid \mathbf{r}_0) \right\} dS(\mathbf{r}_0)
\end{aligned}
$$

and

$$
\begin{aligned}
p(\mathbf{r}) = & \int_{S^m} \left\{ - \frac{\partial p^+(\mathbf{r}_0)}{\partial n(\mathbf{r}_0)} . g(\mathbf{r} \mid \mathbf{r}_0) \right\} dS(\mathbf{r}_0) \\
& + \int_{S^r} \left\{ \mu(\mathbf{r}_0) . \frac{\partial g(\mathbf{r} \mid \mathbf{r}_0)}{\partial n(\mathbf{r}_0)} - \sigma(\mathbf{r}_0) . g(\mathbf{r} \mid \mathbf{r}_0) \right\} dS(\mathbf{r}_0) \qquad \forall\ \mathbf{r} \in V^+ \qquad (3.45) \\
& + \int_{T} \left\{ - \frac{\partial p(\mathbf{r}_0)}{\partial n(\mathbf{r}_0)} . g(\mathbf{r} \mid \mathbf{r}_0) \right\} dS(\mathbf{r}_0)
\end{aligned}
$$

and the variational solution is undertaken by using similar steps to those used with the formulation for the regular 3-D bodies.

3.6 WAVE ENVELOPE INFINITE ELEMENTS

Traditionally the finite element method is used to solve interior acoustic problems, since the numerical implementation requires a finite number of finite elements. However, it is possible to alleviate the problem of infinite domains with various methods (see Givoli, 1992, SYSNOISE, 1999) by modelling the acoustic near field of the radiating body by dividing the unbounded region into an inner region, formed by a mesh of conventional finite elements and an outer region with a single layer of special elements stretching out to infinity. These are used to model the acoustic far field (Figure 3-7). Several approximation methods have been developed, and in our case we use the mapped infinite wave envelope method (Astley *et* al, 1994, 1998 and Cremers *et* al, 1994). The infinite wave envelope element is based on an infinite geometry mapping, extending the element to infinity, and

special shape functions with built-in amplitude decay and a wave-like variation, as shown below.

The formulation of the variable-order, infinite wave envelope element involves the following aspects: the infinite geometry mapping, the shape functions and the weighting functions in a modified Galerkin residual scheme. The first two aspects are discussed below.

3.6.1 Infinite geometry mapping

The infinite geometry mapping consists of the mapping of a unit parent element on to a real element extending to infinity. The mapping is defined by the location of four 'geometric' nodes (in the case of two dimensional element). Figure 3-8 presents an example of the 'parent' and 'mapped' topologies of a two-dimensional wave envelope element.. The corner of the parent element at (1, ±1) is mapped to infinity in the mapped element. The first two geometric nodes 1 and 2 are defined to lie on the envelope (infinite) layer. Nodes 3 and 4 are defined at distances a_1, and a_2 from nodes 1 and 2, respectively along the infinite sides. Mirror nodes 3' and 4' are used as 'virtual sources' and defined at distances a_1, and a_2 from nodes 1 and 2 on the side of the conventional mesh.

This type of geometrical mapping is obtained by introducing a singularity in the radial direction (at ζ=1 or r=∞), yielding an inverse mapping of the form

$$\zeta = 1 - 2\frac{a_i}{r} \qquad (i = 1,2) \tag{3.46}$$

In this equation, r is a radial co-ordinate along the infinite element which is the distance along the 1-3 or 2-4 sides measured from source point 3' or 4'. A detailed description and the formulation in three dimensions can be found in Astley *et* al (1994, 1998) and Cremers *et al* (1994).

3.6.2 Shape functions

The development of the shape functions used within the infinite wave envelope elements is derived from the observation that in spherical co-ordinates, a three-dimensional radiation function $p(r)$ for the exterior region outside a sphere with a radius R can be written in the form of a multipole expansion with an infinite series

$$p(r) = \frac{e^{-jkr}}{r} \sum_{n=0}^{\infty} \frac{F_n(\theta,\varphi,k)}{r^n} \tag{3.47}$$

when r,θ,φ are the spherical co-ordinates relative to the origin r_0 and the functions $F_n(\theta,\varphi,k)$ represent the radiation functions corresponding to acoustic multipoles of increasing order. It has been shown that the series converges absolutely and uniformly when the elements lie outside the smallest sphere circumscribing the structure. Therefore this co-ordinate system is not efficient for long or flat objects. The current variable order infinite wave envelope elements are also formulated in the ellipsoidal co-ordinate system. It is shown that only a finite number of terms (generally less than six) are required to model the amplitude decay of the outgoing propagating waves. The order of the series is thus defined by the number of terms in the $(1/r)$ expansion.

3.7 SUMMARY OF PROPERTIES

- Both the DBEM and the IBEM follow a two-step procedure. In the first step an approximation of the boundary surface variables is determined, which is based on the expansion of the surface geometry and the boundary conditions on the surface in terms of a set of prescribed shape functions which are locally defined on each element. In the second step, a post-processing procedure is undertaken where the field variables at any point are obtained from the boundary integral formulation. The calculation time of the second step is generally much faster than the first step.

- Exterior problems can be solved with either the boundary element method or the infinite element method. The DBEM can be used to solve only for closed bodies where the indirect method can solve the acoustic field on both sides.

- If the boundary surface is closed, the problem can be solved with both methods. The set of equations needed to be solved with the DBEM produces non-symmetric matrices, whereas the IBEM produces symmetric matrices. The total computational effort, however, is determined by the size of the problem (see Chapter 10).

- With the BEM, only the surface is discretised, thus the dimensionality of the problem is reduced from three to two. However, computationally, this property does not result in faster solution times than the FEM, since the matrices are complex, fully populated, frequency dependant, and may be singular. Although the number of nodes and elements in the BEM is much smaller than in the FEM, in large problems, this, again, does not imply greater efficiency in the solution process.

- The non-uniqueness problem appears in both the DBEM and the IBEM. It is a mathematical problem and not a physical one. It occurs only when the body has an enclosed volume. As the size of the model gets larger and the frequency range of interest increases, the higher the number of 'irregular frequencies' that are needed to be removed.

- The advantages of the IFEM are the ability to solve exterior problems both in the frequency domain and the time domain and to alleviate the non-uniqueness problem. Currently, the main limitation is the generation of the complex volume mesh that produces 3-D elements between the discretised 2-D surface elements, and the ellipsoid, or the spheroid infinite layer.

3.8 CONCLUSIONS

This chapter summarised briefly the theory behind the numerical methods used in the research. The formulations of the BEM, currently the main tool for the prediction of acoustic radiation and scattering, have been presented. Since different aspects of HRTFs are investigated numerically, it was required to evaluate different schemes where different methods with different versions of software and hardware produced varying performance. Since a great emphasis of the research was put on high frequency analysis, it was of utmost importance to optimise all the various aspects of modelling, where practical aspects of the implementation and examples of the various methods presented here are given in the following chapters.

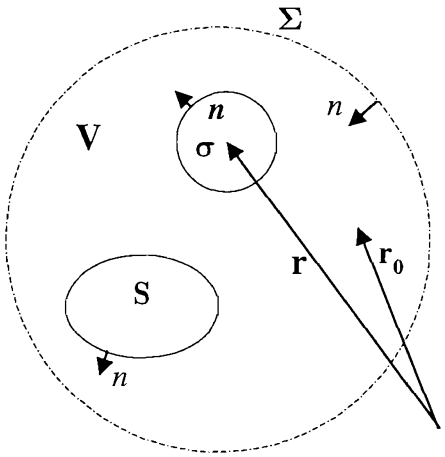


Figure 3-1: Domain definition for an *exterior* point $\mathbf{r}$ in the Direct Boundary Element Method (DBEM).

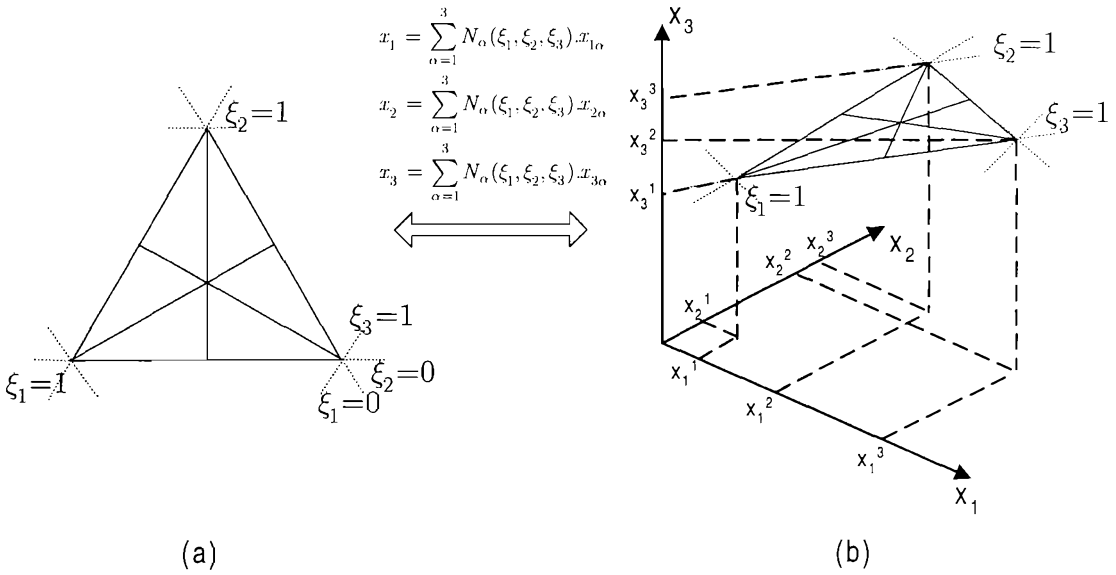


Figure 3-2: Transformation of a planar linear triangular element (a) mapped and (b) parent.

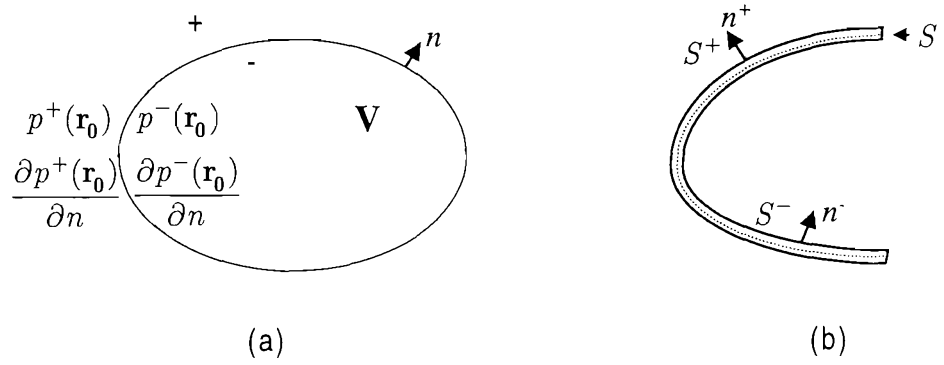


Figure 3-3: IBEM domain definition: (a) closed body (b) thin open body.

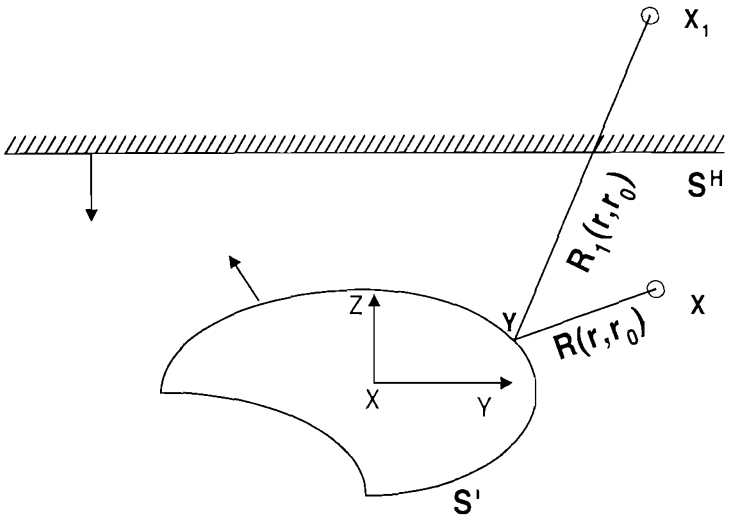


Figure 3-4: Symmetry condition. An Infinite plane (baffle) produces a mirror image and a modified Green function.

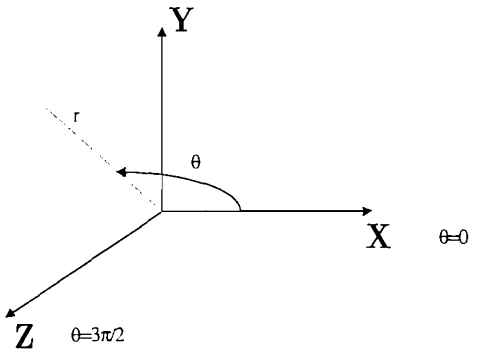


Figure 3-5: Axisymmetric cylindrical co-ordinate system.

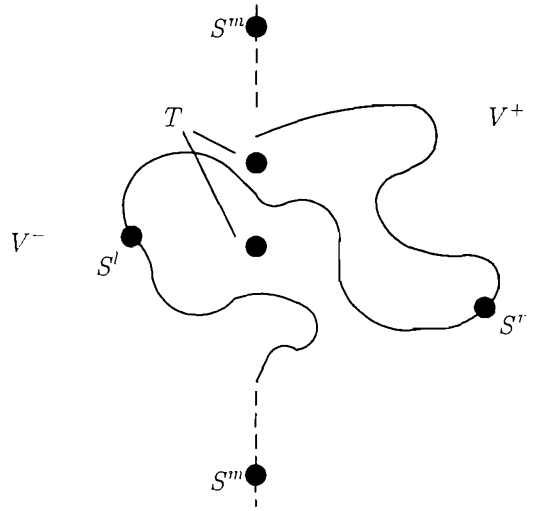


Figure 3-6: Acoustic transparency: a sound wave can propagate from both sides of the baffle (V^- and V^+) through 'transparent' elements (at points *T*) defined at *z*=0. S^m denotes the points on the infinite baffle, and S^r and S^l the points on surface of the body attached to the baffle, on the positive and negative sides of the baffle, respectively.

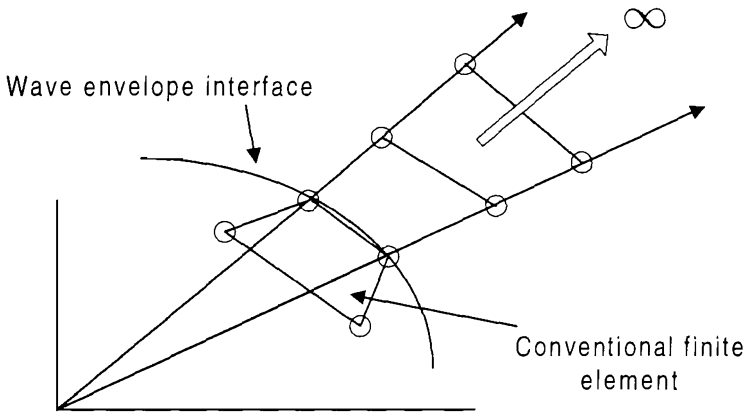


Figure 3-7: Schematic drawing of the IFEM - conventional FEM and infinite layer. The infinite wave envelope element is based on an infinite geometry mapping, extending the element to infinity, and special shape functions with a built-in amplitude decay and a wave-like variation.

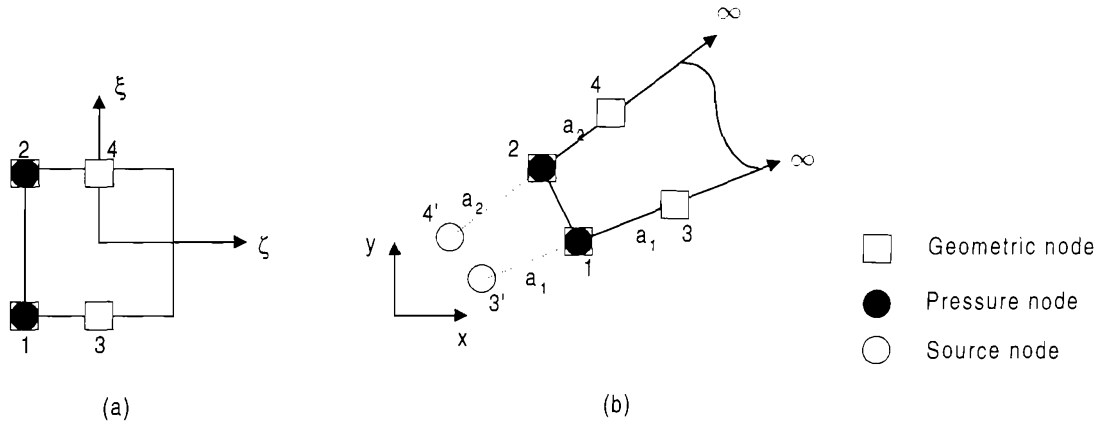


Figure 3-8: Infinite geometry mapping from a two dimensional unit parent element to a real infinite wave envelope element: (a) parent and (b) mapped topologies. The mapping is defined by the location of four nodes:1, 2, 3 and 4. The corners of the parent element at (1, ±1) map to infinity in the mapped element. Nodes 3 and 4, and their mirrors 3' and 4' are defined at distances a_1 and a_2 from nodes 1 and 2. The mapping is given in Equation (3.47).

CHAPTER 4

NUMERICAL COMPUTATION OF THE HRTF USING SIMPLE GEOMETRICAL MODELS

The motivation for investigating the equivalent 'HRTFs' of simple geometrical models is two-fold: (a) to investigate the suitability of different formulations of numerical techniques in simulating the response of small-to-medium size models, and (b) to investigate the effects of the different components in HRTFs which include simplified geometries of the head, torso and the concha.

The investigation of simple geometrical models is important in the analysis of structural modelling of HRTFs, as well as in binaural and monaural physical and psychophysical studies of general attributes of HRTFs. An outline of this chapter is as follows: first, the response of a rigid sphere is investigated analytically and compared to results obtained with various numerical techniques. Subsequently, the responses of an ellipsoid and an artificial head with a torso (but without pinnae) are modelled and compared. Finally, the normal modes of a simplified concha modelled as a baffled cylinder are simulated and analysed. The modelling principles that were investigated and presented in this chapter were implemented throughout the thesis with larger and more complex models.

4.1 THE FREQUENCY RESPONSE OF A RIGID SPHERE

The rigid sphere has been investigated and implemented extensively in research into spatial hearing with various expressions of its magnitude response and effective ITD (e.g. Rayleigh, 1907, Blauert, 1997, Kuhn, 1977, Kahana, 1997a).

4.1.1 Analytical solution

The total complex pressure in the sound field around a rigid sphere due to excitation of a plane wave is given by

$$p_{\text{tot}} = p_0 \exp(j\omega t) \sum_{m=0}^{\infty} (-j)^{(m+1)}(2m+1) P_m(\cos\phi) \cdot \cdot \frac{j_m(kr)\mathrm{n}'_m(ka) - \mathrm{n}_m(kr)j'_m(ka)}{j'_m(ka) - j\mathrm{n}'_m(ka)} \tag{4.1}$$

where p_0 is the amplitude of the pressure of the incident wave, ω is the angular frequency, k is the wave number, a is the sphere radius, r is the distance from the centre of the sphere, $P_m(\cos\phi)$ are the Legendre Polynomials, ϕ is the angle of incidence in the horizontal plane*, n_m is the spherical Neumann function and finally, j_m is the spherical Bessel function.

On the surface of the sphere this equation reduces to

$$p_{\text{tot}} = p_0 \exp(j\omega t) \sum_{m=0}^{\infty} \frac{(-j)^{(m+1)}(2m+1)(ka)^{-2} P_m(\cos\phi)}{j'_m(ka) - j\mathrm{n}'_m(ka)} \tag{4.2}$$

The full derivation of the above expressions is given in Kahana (1997a). Both Equations (4.1) and (4.2) are derived for the horizontal plane only. A general formulation of Equation (4.1), which includes also dependency on elevation angles, will be presented in Chapter 7.

* Note that azimuthal angles are defined with ϕ and not θ. This is mainly for consistency of the spherical co-ordinate system used in Chapter 7.

4.1.2 Numerical solution

In principle, numerical modelling of sound scattering by a sphere should not impose great difficulties. In practice the sphere 'emphasises' the problem of singularities in both the DBEM and the IBEM. Due to its symmetry in every plane, it is found that removing 'irregular frequencies', especially at high frequencies, is a difficult task. The following worked-out examples of modelling the response of a sphere and the total sound field around it are investigated using four approaches: the DBEM, the IBEM, the IFEM, and the axisymmetric IBEM.

4.1.2.1 The non-uniqueness problem

The response of a rigid sphere was investigated at two points on its surface: one in the front ($\phi = 0°$) opposite the source, and one at the rear ($\phi = 180°$), in the shadow zone. A mesh of a sphere with a radius of 0.1 m including 1266 nodes and 1264 quadrilateral linear elements was used (see Figure 4-1). A plane wave source was positioned in the far field, at a distance of 1 m and $\phi = 0°$.

As stated in Chapter 3, only in the DBEM and the IBEM does the problem of the non-uniqueness of the solution exist. In the DBEM, over-determination points have been applied. Since these points 'regularise' the solution of the set of equations using the Lagrange multiplier technique, the higher the ratio between over-determination points and real nodes of the elements, the larger the error. At high frequencies, 'irregular frequencies' will be removed with many over-determination points, and for improved accuracy, increased mesh resolution is required. Practically, the solution procedure starts with no over-determination points at all. Then, when the first 'irregular frequency' is found[†] (around 1.7 kHz for a sphere with a radius of 0.1 m, which corresponds to the characteristic

frequency $k = \pi a$) more points are added. Figure 4-2 demonstrates that a single point will not entirely remove the 'irregular frequency' but four points are necessary (note that as a result the results in the frequency range of 1.6 kHz to 2 kHz are distorted). For large models, the addition of 15-30 over-determination points was found to give no deterioration in the results and also to remove severe peaks when the size of the elements corresponds to six elements per wavelength.

For the case of the IBEM (Figure 4-3), 25 'singular admittance' elements (arranged in a plane with 5×5 elements) are added inside the cavity of the sphere, and the inside of the elements (with negative normal vector) are applied with absorption boundary conditions. As frequency increases it is difficult to remove the singularity. A higher frequency range was investigated in this case and it is demonstrated that as frequency increases, larger errors were obtained with higher amplitude and increased affected frequency bandwidth. Note that a smoother result was obtained although a residual effect remained.

4.1.2.2 The principle of reciprocity

In both simulation and measurement, the HRTF database is obtained by changing the angular position of the source on an imaginary sphere, generally with a radius of 1.4 m to 2 m. The disadvantage of this method is that it is time consuming, and also the database is limited to the discrete positions of the measured/simulated sources. An alternative method is suggested by using the principle of reciprocity (Kinsler *et al*, 1982, pp. 165-168, see Figure 4-4): the acoustic pressure p_1 produced at point B in a fluid by a source at another point A in the fluid where $\mathbf{u}_1$ is the velocity of the radiating element of the source, is the same as the pressure p_2 produced at point A by the same source located at point B where

† 'Irregular frequencies' are the eignfrequencies of the corresponding interior problem with modified boundary conditions. i.e. for an exterior Dirichlet problem the frequencies are the eignfrequencies of the interior Neumann problem.

$\mathbf{u}_2$ is the velocity of the radiating element of the source, irrespective of the presence of boundaries of normal impedance.

It is shown by Kinsler *et* al (1982) that

$$\int_{S_A} p_2 \mathbf{u}_1 \cdot \mathbf{n}\, dS = \int_{S_B} p_1 \mathbf{u}_2 \cdot \mathbf{n}\, dS \tag{4.3}$$

If the sources are small with respect to the wavelength and several wavelengths apart then the pressure is uniform over each source so that

$$\frac{1}{p_1}\int_{S_A} \mathbf{u}_1 \cdot \mathbf{n}\, dS = \frac{1}{p_2}\int_{S_B} \mathbf{u}_2 \cdot \mathbf{n}\, dS \tag{4.4}$$

It therefore follows that

$$\frac{Q_1}{p_1(r)} = \frac{Q_2}{p_2(r)} \tag{4.5}$$

where Q is the volume velocity of each of the monopole sources, and $p(r)$ is the pressure at a distance r from the source. If we model $Q_1 = Q_2 = \text{const.}$, then the pressure simulated anywhere in space, when a monopole source is positioned close to the point of interest (e.g. at the entrance to the ear canal), is equivalent to the pressure produced at this point due to a source with the same volume velocity anywhere in space.

In practice, a few approaches for numerical implementation of the principle of reciprocity exist, as follows:

- Positioning the source on a specific node of the surface and compensating for the loss of energy, radiated into the interior cavity. This method results in 'irregular frequencies' with severe peaks which are very difficult to remove (see a worked-out example below).

- Positioning the source in the exterior domain, very close to the surface. This will ensure radiation will occur only in the exterior domain. The proximity of the source to the surface requires a local refinement of the mesh (Figure 4-1). Note that some excitation of the interior might still exist, but this is not as severe as in the previous case.

- Simulating the source using a vibrating piston with an equivalent volume velocity by applying a velocity boundary condition to one side of the element. This method will be inaccurate if the resolution of the mesh is low with large elements.

Figures 4-5a to 4-5f summarise the difference in simulating the principle of reciprocity using the first two approaches. The sphere presented in Figure 4-1 (but with a reduced radius to *r=6.85 cm* in order to shift the 'irregular frequencies' upwards, thus eliminating errors of this kind in this investigation) is excited by a monopole source with a source strength of unity, positioned at 1.4 m away. The pressure was calculated on the surface of the sphere, at two points: the front ($\phi = 0°$), and at the rear ($\phi = 180°$). The responses shown in Figures 4-5a and 4-5b were calculated using the DBEM. As can be seen in the figures, the direct response and the response calculated using the principle of reciprocity agree to within a tolerance of 0.5 dB. It should be noted that at higher frequencies the expected 'irregular frequencies' appear and needs to be removed using CHIEF over-determination points.

When the IBEM is investigated, Figure 4-5c and Figure 4-5d demonstrate the effects of positioning the source on the surface of the sphere. Since the source is now radiating in both directions, the energy should be increased proportionally to the local solid angle (see Equation 3.9 in Chapter 3). If the local point was positioned on an infinite plane, then the solid angle is ½ (see Equation 3.8), and the radiation will be distributed evenly for both

sides of the plane. The local solid angle for the node on which the source was positioned was calculated to be 2.4π, hence the source strength needs to be multiplied by a factor of $4\pi / 2.4\pi = 1.67$.

It can be seen that the results are in good agreement, except for a severe peak – an 'irregular frequency' at around 1.7 kHz. This is mainly because of the real 'physical' excitation of the interior cavity. Note that when the source is positioned in the far field no 'irregular frequencies' appear, as expected.

The results of the simulation based on the second approach are presented in Figures 4-5e and 4-5f. As before, the IBEM was used for the modelling. However, in this case the source was positioned in the exterior domain. As a 'rule of thumb' the position of the source close to the surface should not exceed one normal edge length (this is due to the integration sensitivity as $r \rightarrow 0$). So in order to position the source very close to the surface, a local refinement is required (see typical refinements in Figure 4-1 and in Figure 4-11). As shown in the figures, hardly any numerical excitation is observed, and in principle these can be minimised by applying absorption to the inside elements of the mesh. The accuracy at the rear could be improved even further by using quadratic elements instead of the linear elements used in this case.

This method was used throughout the research with both the DBEM and the IBEM and found to be reliable, efficient and accurate.

4.1.2.3 The BEM with symmetric models

The formulation of the boundary integral equation (Equation 3.11 and 3.23 for the DBEM and the IBEM, respectively) can be used to solve arbitrary, complex geometries. As demonstrated in Section 3.4.1 it is possible to reduce the computational CPU time when the acoustic field and the model are symmetric with respect to one, two or three axes.

The procedure to solve symmetric models comprises two stages (see Figure 4-6):

- The problem is run with the symmetric boundary conditions. In the case of a single source in the original problem (as shown in the figure), this will result in duplication of the geometry and also any boundary conditions, including the acoustic source. The duplication is for both the amplitude and the phase of the source. Note that rigid boundary conditions ($u_n = 0$) should also be applied to elements that lie on the plane of symmetry.

- The problem is run again, this time with antisymmetric boundary conditions. In this case the duplicated source has a similar amplitude but negative phase. Note also that pressure release boundary condition ($p = 0$) should be applied to the elements that lie on the plane of asymmetry.

- The results are superposed and divided by a factor of two. This is clearly illustrated in Figure 4-6. The contributions from the mirror sources are cancelled since these have opposite signs, whereas the real sources are superposed, which is equivalent to a single monopole source with double the volume velocity.

This property is used together with the principle of reciprocity to optimise the CPU time whenever the DBEM and the IBEM are used (as will be shown later in Figure 4-15 to

Figure 4-17). These properties will be further investigated in Chapter 5, Section 5.4.2 for the case of modelling HRTF of accurate representation of real heads, and also in Chapter 9, Section 9.5 where the sound field is modelled around two KEMAR artificial heads arranged with four sources in a symmetrical arrangement.

Using the principle of reciprocity, the overall CPU time is reduced by a factor of 4: doubling the number of nodes ($n \rightarrow 2n$) increases the running time by a factor of 8

($(2n)^3 / n^3$), but the process has to be repeated twice (see a summary of the computational cost in Chapter 10, Section 10.2.8).

4.1.2.4 The IBEM with Axisymmetric models

The formulation of the axisymmetric IBEM is given in Section 3.5.2. For the case of a sphere, or an ellipsoid, this method is far superior to other methods as a result of the following factors:

- Only a very simple mesh is required with a single curve consisting of line elements.
- The solution is very fast compared to the 3-D BEM.
- Severe 'irregular frequencies' do not appear in this case (although they still do exist), due to the different mathematical formulation of the problem (as presented in Chapter 3, Section 3.5.2).

In the following example, the mesh consists of 50 line elements (Figure 4-7a). When a harmonic order of 7 was used (i.e. 2^7 elements are generated for each element in the revolving line mesh), this resulted in a 3-D mesh presented in Figure 4-7b (which can be investigated up to approximately 10 kHz, or $ka \approx 20$ based on $a = 0.1$). In practice it was found that this formulation requires more Gaussian points used in the integration, when compared to the DBEM and the IBEM, especially for points at high frequencies in the far field[‡]. The results obtained using this technique are discussed in Section 4.1.2.6.

4.1.2.5 IFEM modelling

The response of a sphere can be modelled very efficiently also with the wave envelope method (IFEM)[*]. Figure 4-8 shows 1/8 of an IFEM mesh model of a sphere. It includes

[‡] In current version of SYSNOISE 5.4, this formulation suffers from errors due to singularities on the surface, and therefore only the scattered sound field was used in the analysis.

[*] Currently, in SYSNOISE 5.4, two methods for solving infinite element models exist: the wave envelope method (Astley, 1994, 1998, Cremers *et al*, 1994) and the prolate and oblate formulation given by Burnett (1994, 1998a, 1998b). However, the latter formulation still suffers from implementation problems, and therefore the results are not included here.

1638 nodes and 1522 linear hexahedron elements. It was found that the solution is optimised when hexahedron elements are used and not tetrahedron elements (these affect the bandwidth of the matrices). The number of layers of the conventional FEM mesh is frequency dependent, and three layers were sufficient for accurate modelling.

As shown in Section 3.6, the pressure is composed of a multipole expansion with an infinite series. The response in the front and at the rear of the sphere are presented in Figures 4-9a and 4-9b, respectively, when only a few terms are used. It is demonstrated that three terms of the series are sufficient for accurate modelling. At the rear, no improvement was obtained when the number of terms was increased to 10. The deviation from the analytical solution is a result of the lower resolution mesh than required, and probably the use of linear elements and not quadratic elements.

4.1.2.6 Sound field simulation

All the above methods were investigated with the modelling of the sound field around a sphere. The excitation was a plane wave (from the right side), the area around the sphere was 0.5 m × 0.5 m (with a grid of 400 × 400 points), and the frequency was 2 kHz ($ka \approx 4$).

Errors of less than 2% were obtained when the results are compared with the analytical solution in the front, and an increase of the error up to 7%, in the shadow zone. A higher mesh resolution (in the order of 20 elements per wavelength), a higher quadrature and quadratic elements will minimise the errors even further.

It was concluded that under these conditions the axisymmetric IBEM and the IFEM are superior to conventional DBEM and IBEM both from the computational speed and from the fact that 'irregular frequencies' do not exist (in IFEM) or have little effect (in axisymmetric IBEM). According to Burnett (1994) the relative efficiency of the IFEM is at its greatest when large models are investigated at high frequencies. Comparison at higher

frequencies imposes great difficulties with the use of conventional BEM models, both due to many 'irregular frequencies' and the large size.

4.2 THE FREQUENCY RESPONSE OF A RIGID ELLIPSOID

Different formulations of the frequency response of prolate spheroids have been previously derived (Spence and Granger, 1951, Sugiyama *et al,* 1991, Novy, 1998). These are much more complex than those of the sphere and also do not address the ITD. As a result they are rarely used in binaural synthesis.

Recently, Duda *et al* (1999) developed a simple approximate solution for calculating the ITD of an ellipsoid that is based on the path lengths to the two 'ears'. His results confirmed a better match to an ITD of a listener than the sphere at elevation angles.

In the following example, the response of an ellipsoid was modelled using the DBEM, using the symmetric properties. Figure 4-11a shows the geometry of a sphere which is the average of the CORTEX head, with a radius of 8.75 cm, and Figure 4-11b shows the geometry of a stretched sphere (in three directions) to match the size of the head with r_x=9.6 cm, r_y=7.9 cm and r_z=11.6 cm. Both the sphere and the ellipsoid models included 5493 nodes and 10824 elements. The 'ears' were positioned as close as possible to the original positions of the ears of the CORTEX head. The position found for the right ear (that was used in the simulation) was: $\phi = 100°$ and $\theta = 260°$ (in the y-axis 0.86 cm below the origin of the ellipsoid).

Examples of the variation of the magnitude of the right 'ear' of the ellipsoid are presented in Figure 4-12. The 3-D surface plot in Figure 4-12a shows the response in the horizontal plane. The angle varies from $\phi = 0°$ (front), through $\phi = 90°$ (right) to $\phi = 355°$ in steps of 5°. The ipsilateral response is characterised with gradual increases up to almost a factor of 2 at high frequency (5 kHz) when the source is positioned at 90°. The 2-D plot in

Figure 4-12b shows the same data: at $\phi = 90°$ an increase of almost 6 dB is noticed, and the contralateral response is characterised with sharp notches. They are profound at high frequencies, at $\phi \approx 265°$ and $\phi \approx 295°$ (Note the asymmetry of angles along the 'interaural axis' due to the lower and rear position of the 'ear' detection point on the ellipsoid). These notches are clearly seen in Figure 4-12b on a logarithmic scale, and increase the dynamic range up to 40 dB in this frequency range. Figure 4-12c presents the response at the right 'ear' due to sources positioned in elevation ($\theta = 45°$). ϕ varies again from 0° to 355° in step of 5°. The gain at the ipsilateral ear is now slightly less than 6 dB, and the minima are less dramatic with changing the angular position of the source in the contralateral side. These characteristics are demonstrated also in the 2-D plot in Figure 4-12d.

Recently, Minnaar *et al* (2000) used this data to compare the ITD of a sphere (with a radius of 82 mm and the ears diametrically disposed on the sides of the head), the ellipsoid presented above and average ITD values of 70 people.

He concluded that the sphere produces satisfactory results for the horizontal plane, but a better match by the ellipsoid is produced, for example in the lateral vertical plane (see Figure 4-13). A better fit could have been obtained with further adjustments of the positions to the 'ears'. The ITD was calculated by comparing the gradient of the phase (the group delay) of the two ears, at low frequencies.

The magnitude of the response of the ellipsoid in the horizontal plane and in elevation is compared to the sphere in Figure 4-15, 4-17 and 4-18.

4.3 THE EFFECT OF THE SHOULDERS AND TORSO

Previous measurements of the effect of the torso (Shaw, 1974, Burkhard and Sachs, 1975, Kuhn, 1983) revealed that the main contribution was in the frequency region up to 3 kHz

depending on the angle of incidence (although the reflections from the torso result in a 'comb' filtering which affect higher frequencies as well). Recently Avendano *et al* (1999) investigated a head and torso model for the evaluation of the localisation of elevated sources at frequencies below 3 kHz. In their psychophysical test of the structural model, they concluded that elevation could be synthesized at low frequencies, where the pinna has very little effect.

A CAD model of the CORTEX artificial head (Figure 4-14a) was provided by NCI (CORTEX, 1996). The original model includes the artificial head C1 and the torso TO1. This model was designed according to IEC959 (1990). It was converted to a valid BEM model by operating mesh decimation techniques, which are described in Chapter 5 and in Appendix 1. The decimated BEM model is shown in Figure 4-14b. Its half model includes 5596 nodes and 10926 linear triangular elements. Assuming six elements per wavelength this model could be investigated only up to 2.5 kHz. This large model highlights the difficulties associated with the 'non-uniqueness' problem: it was found that removing 'irregular frequencies' at high frequencies was very difficult. The larger the model, the higher number of modes that exist in the cavity as frequency increases. The number of 'irregular frequencies' was significantly reduced (but still remained problematic) when the model was divided into two identical parts and the symmetric formulation was used.

It was concluded that modelling the torso with the BEM should be limited to the frequency range where it contributes the most (up to 3 kHz).

The results of the response of the sphere, ellipsoid, the C1 artificial head (without pinnae) and the C1+TO1 artificial head and torso are presented in Figure 4-15 to 4-18.

Figures 4-15a to 4-15d show the response in the horizontal plane, and Figures 4-15e to 4-15h show the response in elevation.

The modulation of the response of the torso observed in the figures is mainly due to a delayed path in addition to the direct path (with different path lengths in the order of approximately 14-18 cm). The modulation is angularly dependent with the maximum variation noticed between 800 Hz and 2 kHz. The same trend is noticed for elevation angles. The figures also show the similarity between the ellipsoid and the CORTEX head without pinnae.

Another way of comparing the responses is by comparing the Interaural Transfer Function (ITF) of these models. Figure 4-16 shows the ILD of the ellipsoid for an arbitrarily chosen angle of elevation ($\phi = 135°, \theta = 45°$). Between 500 Hz and 2 kHz the sphere produces an underestimated ILD. The ellipsoid produces a very similar response when compared to the CORTEX C1, and C1+TO1 up to 1.5 kHz.

Figure 4-17 shows the phase difference of the ITF for the source excitation from the same direction, which can be used to calculate the ITD. The errors of the phase are calculated with respect to the C1 head. The ellipsoid shows very small errors up to 2 kHz and the torso is characterised by the modulation of the phase. The sphere shows larger errors. This is also supported when the ITD of the sphere and the ellipsoid were compared by Minnaar *et al* (2000), and also by Duda *et al* (1999) who compared the ITD of an ellipsoid with the ITD of a single subject.

4.4 A CYLINDER IN A BAFFLE AS A SIMPLIFIED CONCHA

In the previous sections we investigated the effect of HRTFs of simple head models. In a similar manner to the case of a sphere and ellipsoid where the response can demonstrate the basic physical effect of the head as a 'shadowing obstacle', it is possible to investigate the basic features of the pinna. Teranishi and Shaw (1968) investigated the response of a baffled cylinder to reveal the fundamental resonance frequencies.

In this section we simulate the response at the bottom of a baffled cylinder (at a point which is 3 mm offset to the centre of the base of the cylinder). A plane wave source was positioned at grazing incidence (the plane of the baffle). This problem was solved with the DBEM where the surrounding of the cylinder has been smoothly raised from the baffle plane. A more efficient method is implemented in Chapter 8 with the introduction of the 'IBEM transparency' formulation.

The variation of the amplitude of the response is presented in Figure 4-18 on a linear scale. Two resonance frequencies are detected: the first at 4.2 kHz is excited almost uniformly from every angle (the slight offset of the detection point at the bottom does not affect significantly this resonance mode). However, the second resonance, a transverse mode appearing at 10.4 kHz is angularly dependent, as at certain angles an interference causes the pressure to be cancelled, and at others to be superposed.

The variation of pressure at the bottom of the cylinder and its surrounding is shown in Figure 4-19 with the source being positioned on the right side. The two resonance frequencies detected in Figure 4-18 are investigated, and the first quarter wavelength at 4.2 kHz is presented in Figure 4-19a. The amplification of the pressure is almost uniformly distributed, and with very little variation of the phase (Figure 4-19b). The first transverse mode is found at 10.4 kHz. In this case a zero pressure nodal line occurs at the centre of the bottom cylinder and maxima at both sides. The phase variation reveals a negative phase on the right side and a positive phase on the left side. These mode shapes will be investigated further in Chapter 8 as they are fundamental in the analysis of the characteristics of the external ear with a blocked ear canal.

4.5 CONCLUSIONS

Different formulations of the DBEM, the IBEM and the IFEM have been used in the investigation of simple geometrical models. It is clear that the efficiency of each numerical technique depends of the type of problem. For the case of a sphere four methods which are based on different formulations have been used and the results have been computed with high accuracy compared with the analytical solutions. With the closed models investigated in this chapter, it was found that better control of accuracy and eliminating the problem of 'irregular frequencies' was achieved with the DBEM. However the method is inefficient when the models become large (see a discussion in Chapter 10). The investigation with the IFEM demonstrated its efficiency. The results do not suffer from 'irregular frequencies' and in this relatively small problem the computational speed was equivalent to the axisymmetric IBEM. The latter is advantageous compared to all other methods if the model is axisymmetric, due to a simple mesh, faster calculation time, and avoiding the problems of 'irregular frequencies'. The investigation of the torso is limited to frequencies below 2.5 kHz mainly because of the large number of 'irregular frequencies' in the large cavity. The main advantages of computing the response of the ellipsoid and the torso are the ability to produce the most important features of HRTFs at low frequencies that can be used in the synthesis of virtual sources in azimuth or elevation.

The analysis of the baffled cylinder showed two basic features: (1) the frequency response at ~4.2 kHz is excited almost unidirectionally from any angle of incidence, and it corresponds to a quarter of a wavelength resonance. (2) The first transverse mode appeared around 10.4 kHz. These results are in agreement with the measurements by Shaw and Teranishi (1968) and Teranishi and Shaw (1968) and motivated further investigation of the cylinder and pinna mode shapes, that will be presented in Chapter 8.

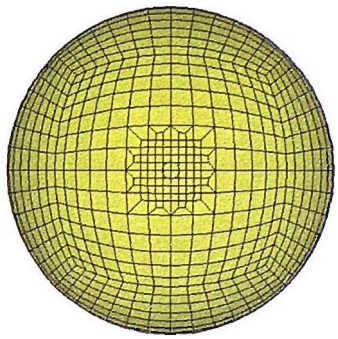


Figure 4-1: A sphere with a radius of *r=0.1 m* was used for modelling the scattered sound field. The local refinement is used in Section 4.1.2.2, when the principle of reciprocity is implemented. The mesh comprises 1266 nodes and 1264 quadrilateral linear elements. Assuming six elements per wavelength, the maximum frequency is 5.8 kHz.

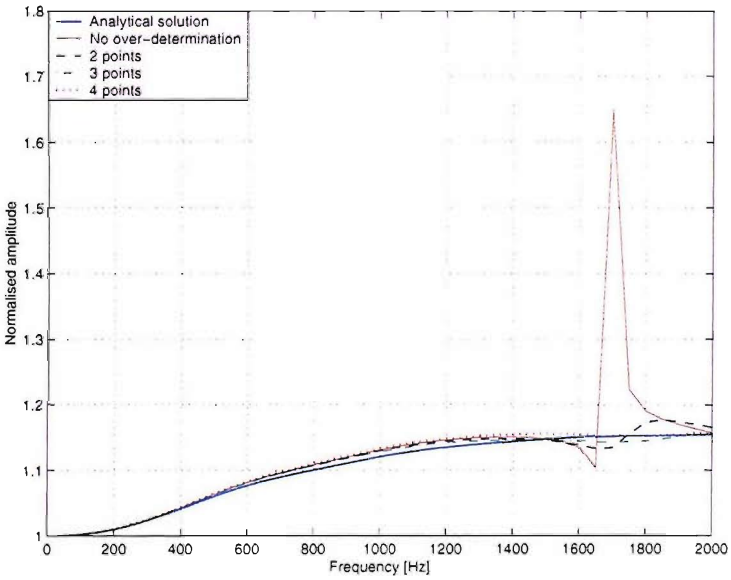


Figure 4-2: The frequency response of a rigid sphere (with a radius of *r=0.1 m*) modelled with the DBEM. The point on the sphere is positioned at the rear, at $\phi = 180°$. The effects of adding over-determination points is presented and compared with the analytical solution.

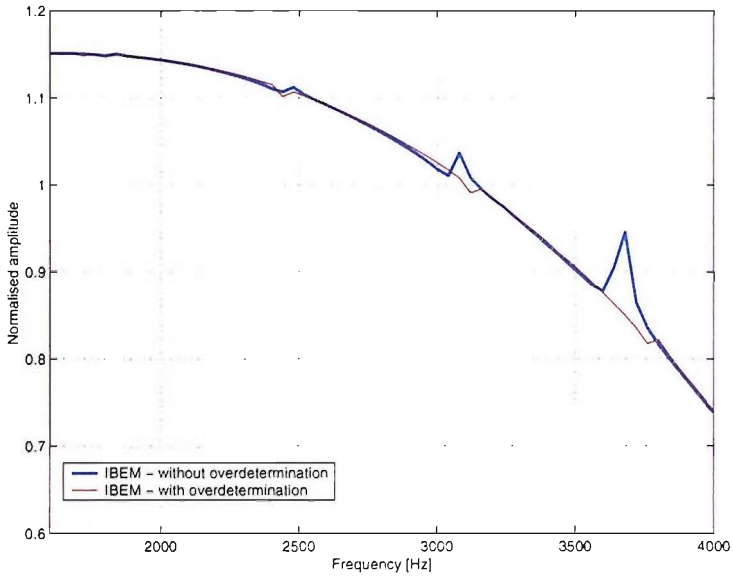


Figure 4-3: The frequency response of a rigid sphere (with a radius of *r=0.1 m*) modelled with the IBEM. The point on the sphere is positioned at the rear, at $\phi = 180^{\circ}$. 'Irregular frequencies' are smoothed by adding 25 (5 × 5) special 'singular admittance' elements with large, real admittance values inside the cavity.

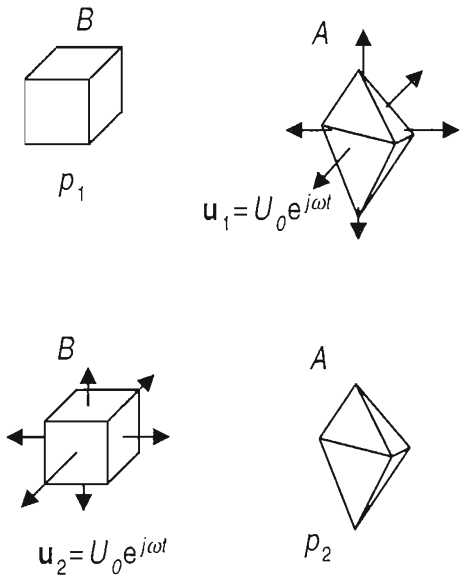
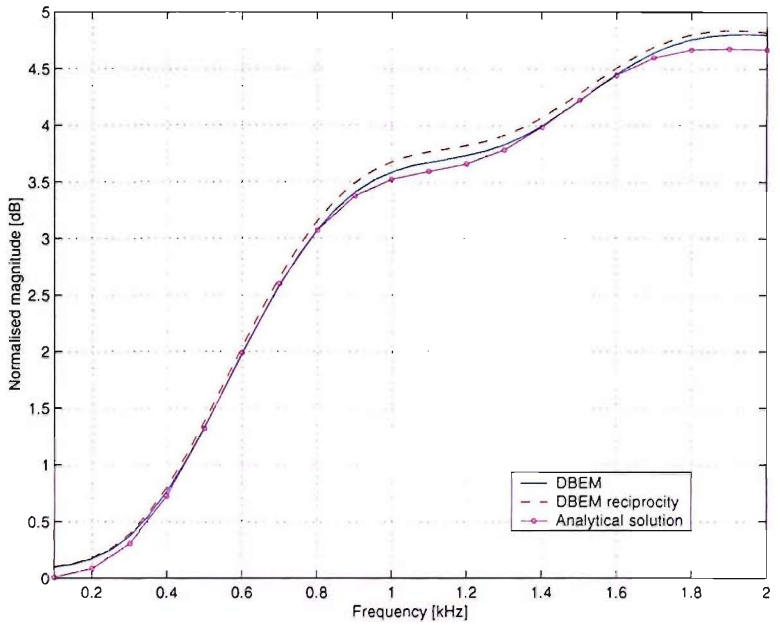
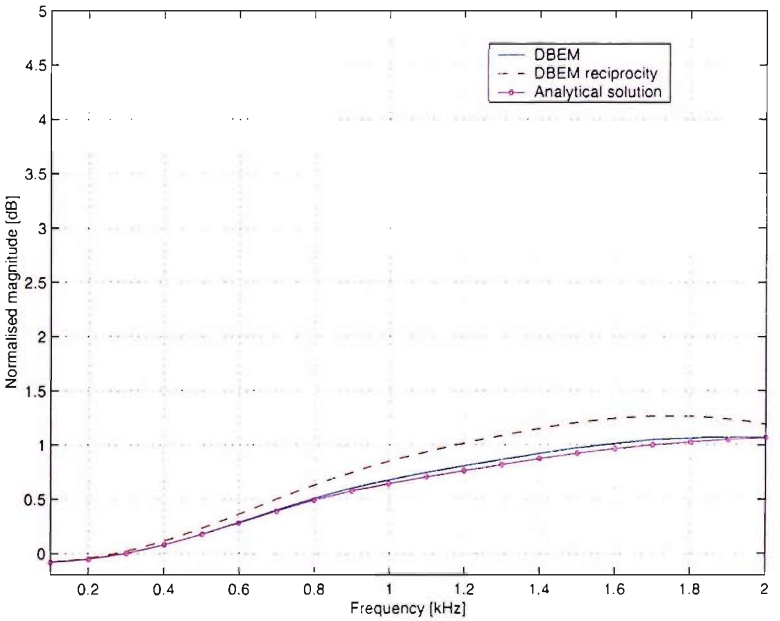


Figure 4-4: The principle of reciprocity: the acoustic pressure p_1 produced at point *B* in a fluid by a source at another point *A* in the fluid is the same as the pressure p_2 produced at point *A* by the same source located at point *B* (after Kinsler *et al*, 1982).

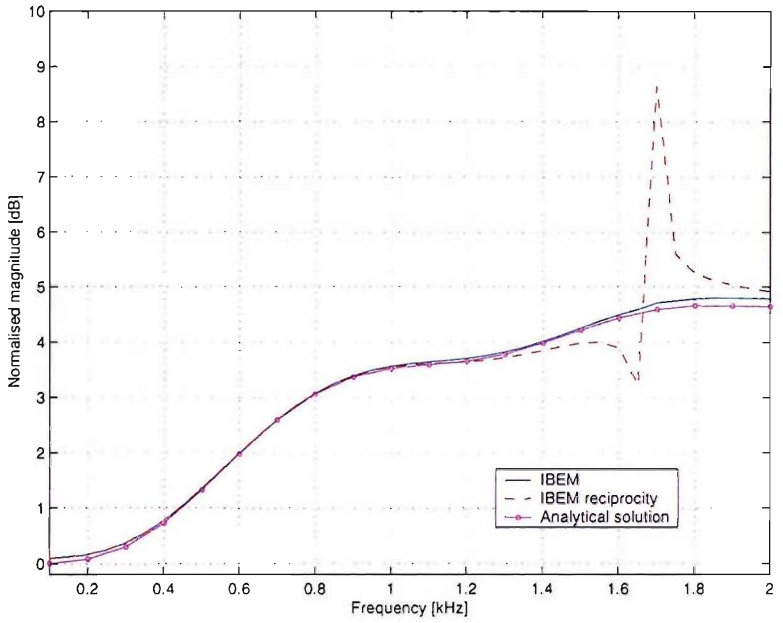


(a)

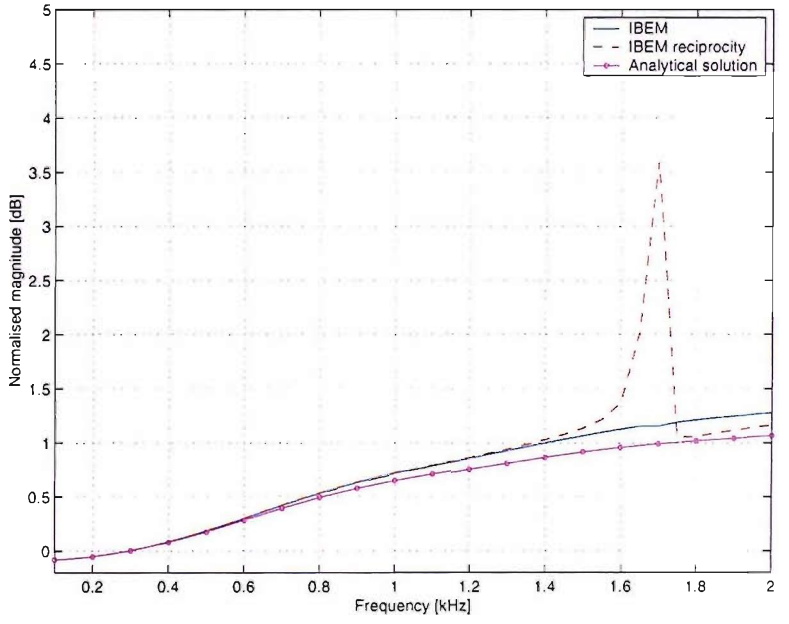


(b)

Figure 4-5: Three approaches for numerical implementation of the principle of reciprocity. The DBEM with the source positioned on the surface of the sphere (a) frontal angle $\phi = 0°, \theta = 0°$ (b) rear angle $\phi = 180°, \theta = 0°$. In these cases a smaller radius was used (*r=6.85 cm)* in order to avoid the treatment of over-determination points.

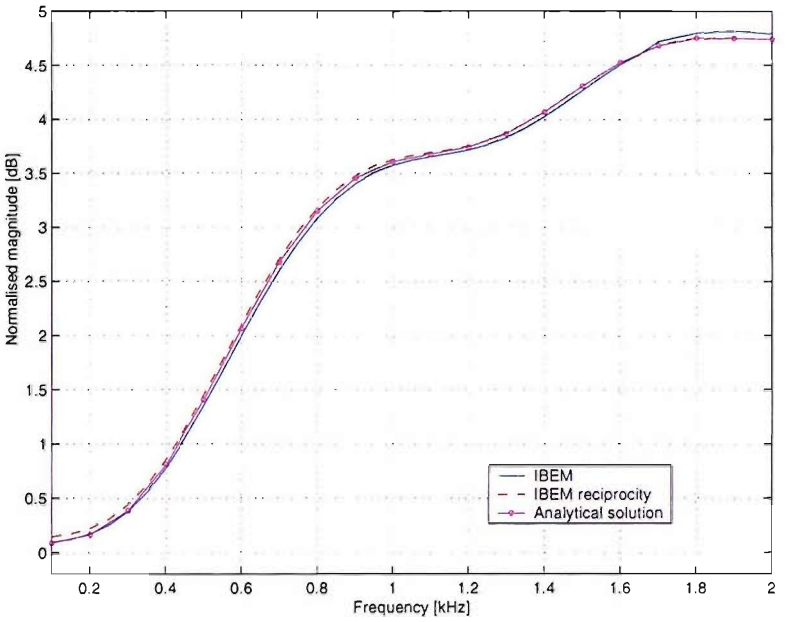


(c)

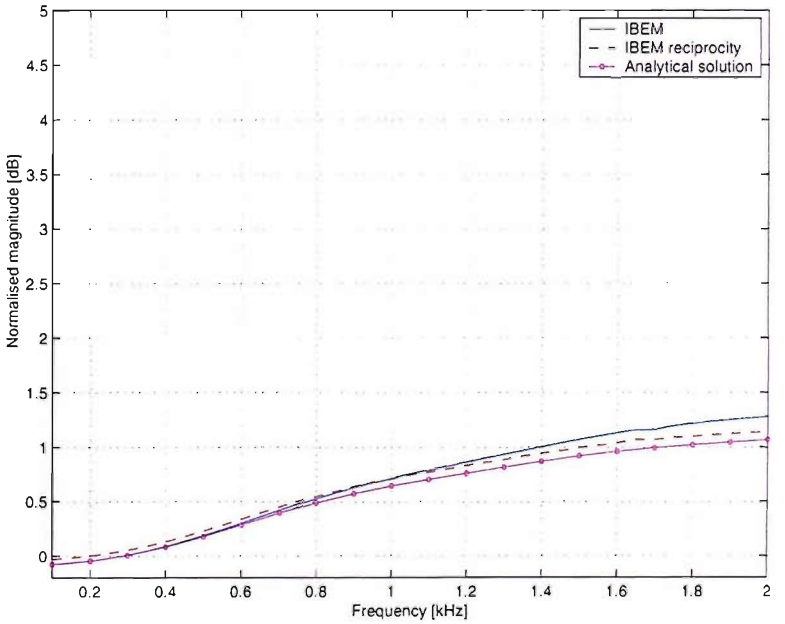


(d)

Figure 4-5 (cont.): The IBEM with the source positioned on the surface of the sphere (c) frontal angle $\phi = 0°, \theta = 0°$ (d) rear angle $\phi = 180°, \theta = 0°$.



(e)



(f)

Figure 4-5 (cont.): The IBEM is used with a mesh which is refined locally and the source is positioned close to the surface in the exterior domain. (e) frontal angle $\varphi = 0°, \theta = 0°$ (f) rear angle $\phi = 180°, \theta = 0°$.

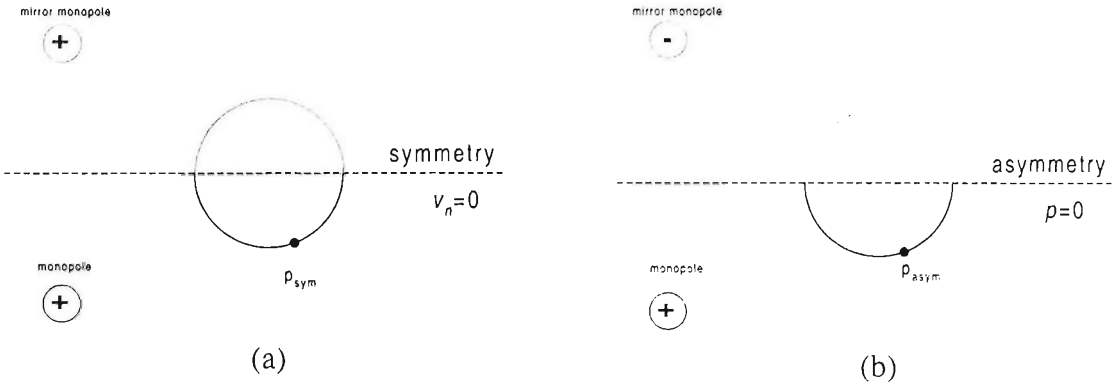


Figure 4-6: For symmetric models, the total sound pressure is the superposition of symmetric and asymmetric models (and division by a factor of two).

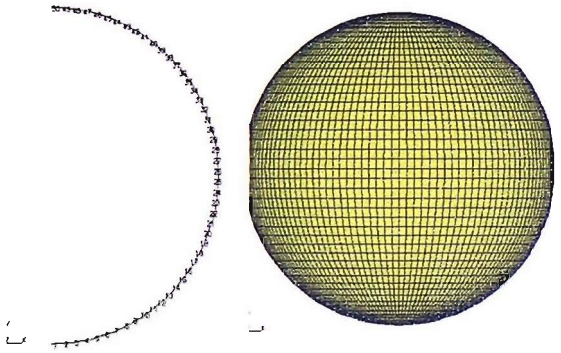


Figure 4-7: Mesh models of a rigid sphere, *r=0.1 m* using the IBEM axisymmetric formulation (a) 2-D rotated arc with 50 line elements, and (b) The resulting IBEM mesh with harmonic of order 7, equivalent to a 3-D mesh with 6450 elements.

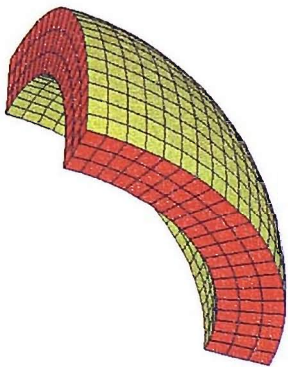
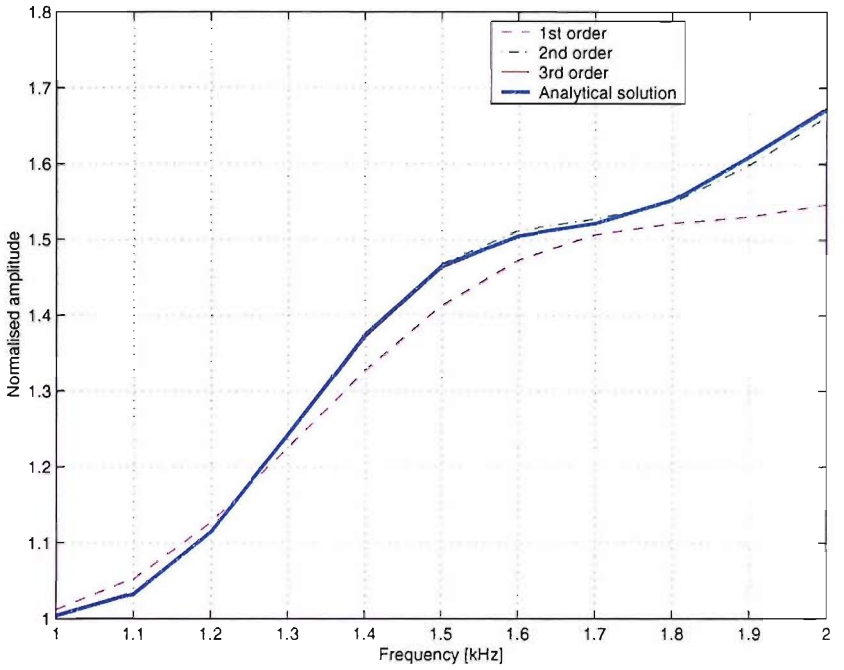
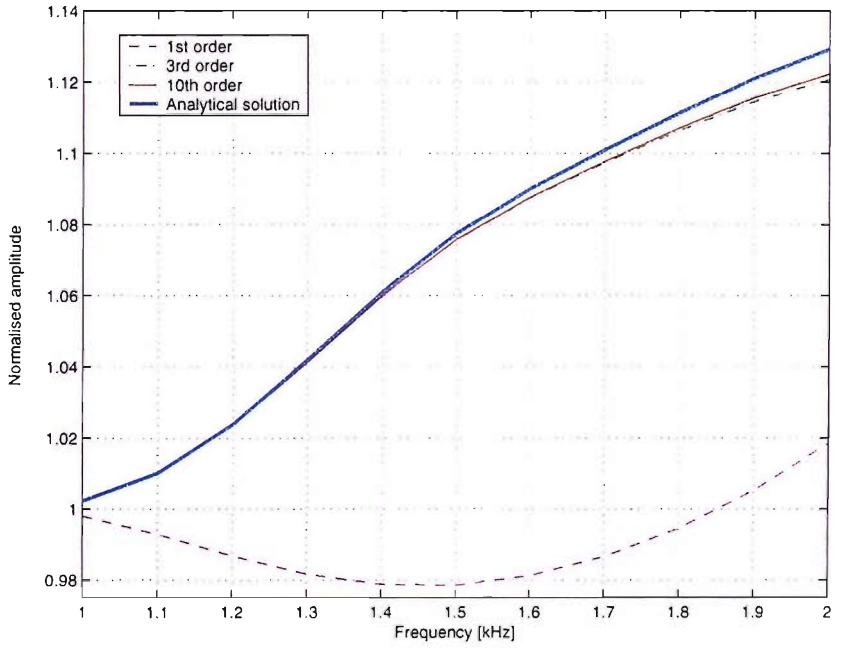


Figure 4-8: 1/8 of a sphere IFEM mesh. This mesh model comprises 1522 nodes and 1638 linear hexahedron elements.



(a)



(b)

Figure 4-9: The effect of the order of the series implemented with the IFEM for scattering around a rigid sphere. The results (a) in the front ($\phi = 0°$) (b) at the rear, $\phi = 180°$, are compared to the analytical solution.

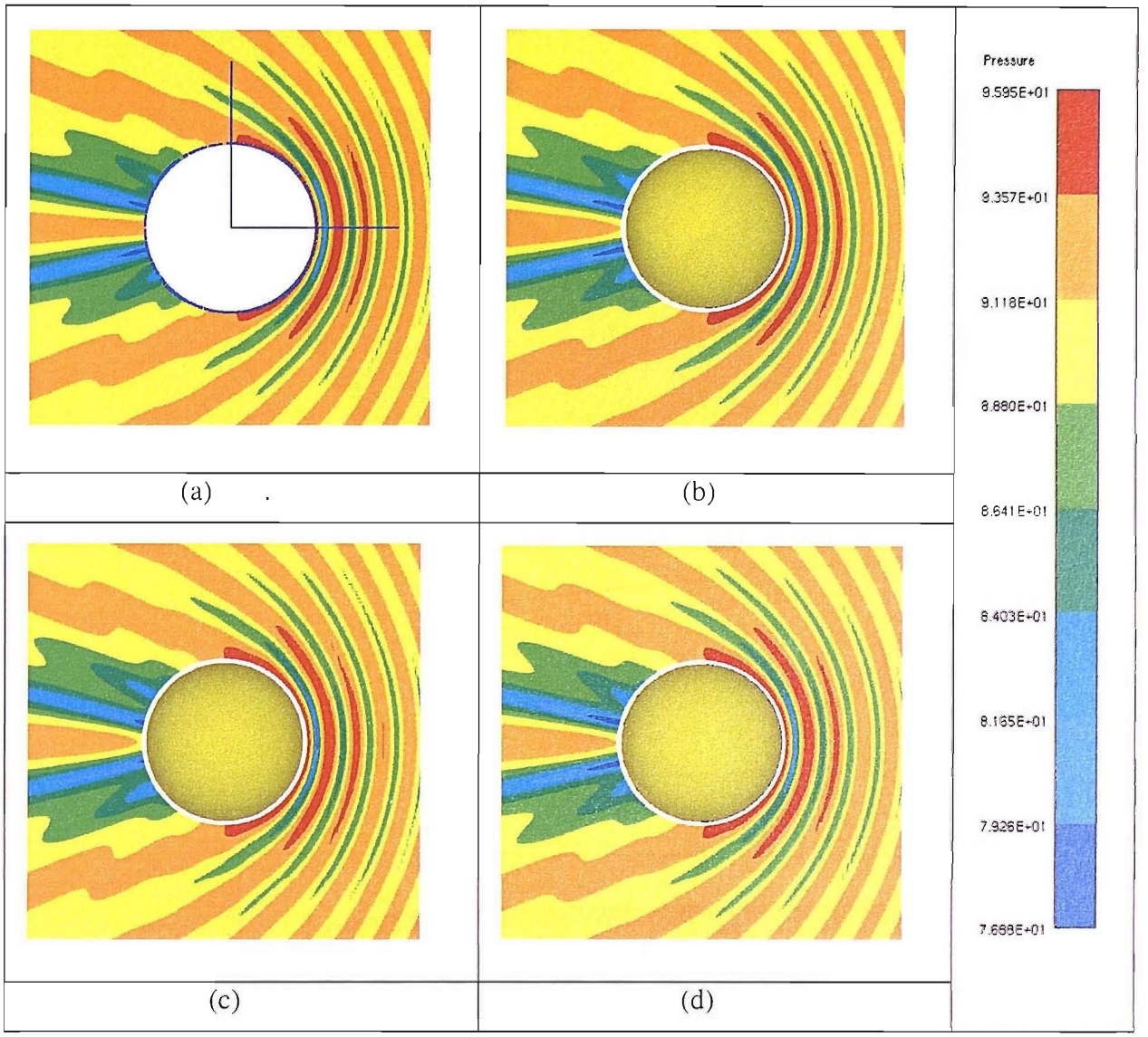


Figure 4-10: The magnitude (in dB, not equalised) of the total sound pressure at 0.5 x 0.5 m around a sphere with a radius of *r=0.1 m* excited with a plane wave source on the right at 2 kHz (~*ka=4*) (a) with the IFEM (b) with the IBEM (c) with the DBEM (d) with the axisymmetric IBEM.

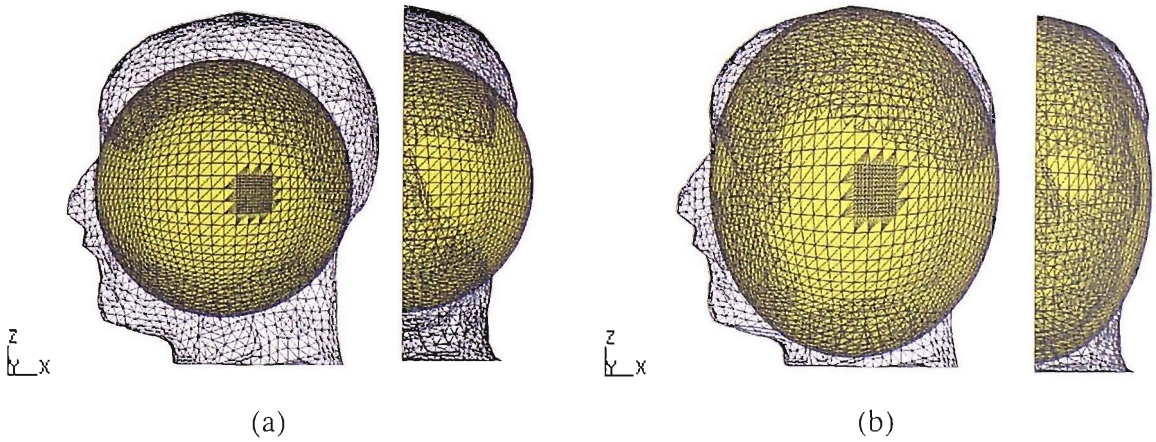
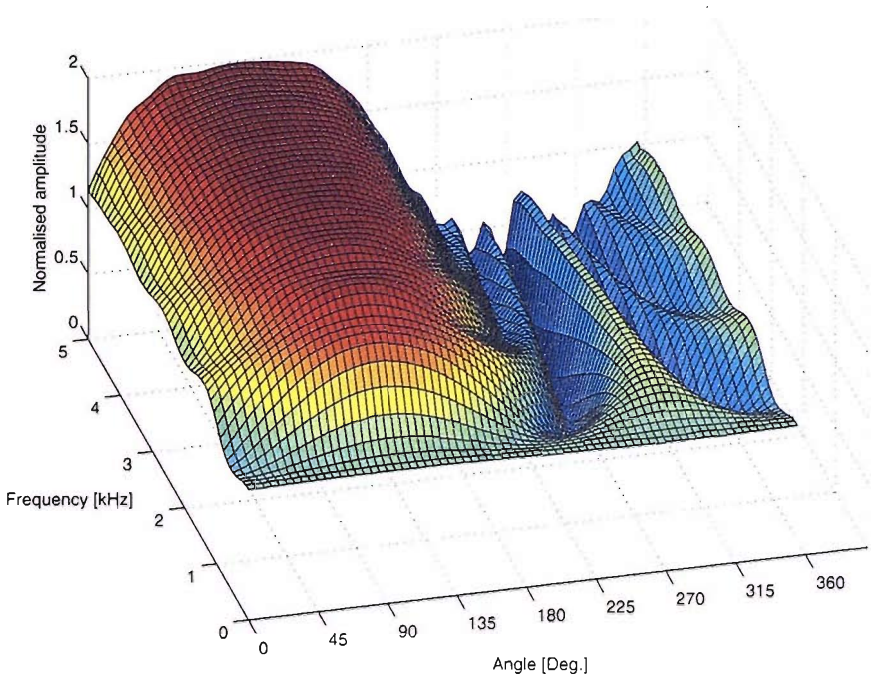
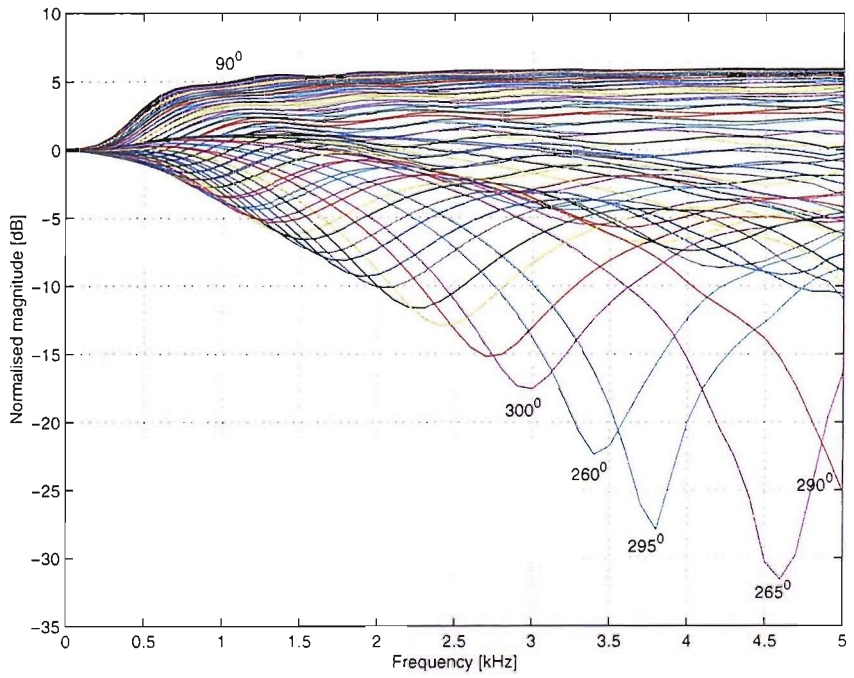


Figure 4-11: Two simplified half mesh models: The head of the CORTEX artificial head (without pinnae) converted from CAD model, fitted with (a) a sphere with r=8.75 cm and (b) an ellipsoid with: r_x=9.6, r_y=7.9, r_z=11.6 cm

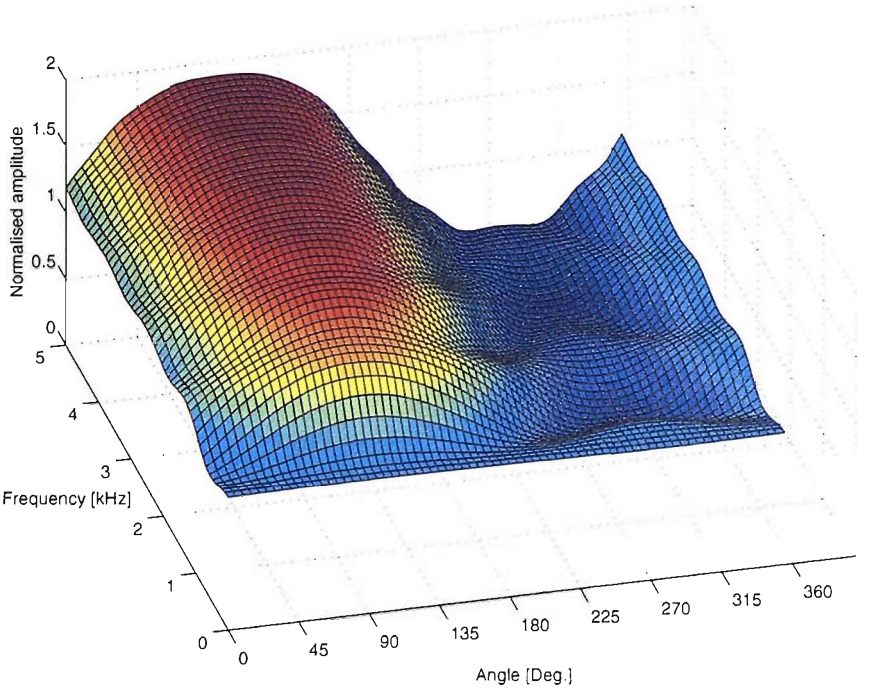


(a)

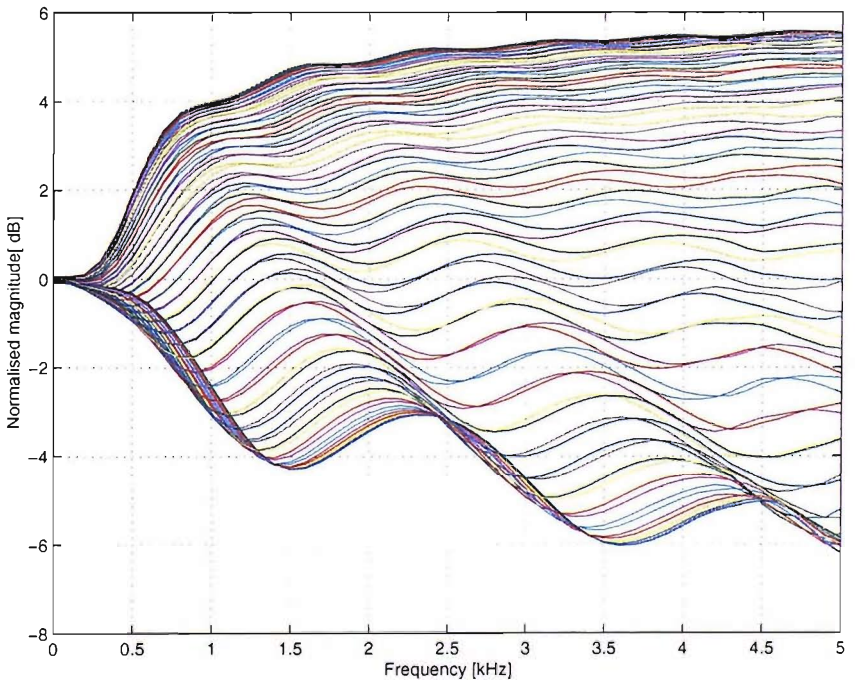


(b)

Figure 4-12: Plots of the amplitude and magnitude of the HRTFs of an ellipsoid (a) 3-D plot of the linear normalised amplitude in the horizontal plane (b) 2-D plot of the same data in [dB]. Simulation was undertaken using the principle of reciprocity and the points, in resolution of 5° are at a distance of 10 m.



(c)



(d)

Figure 4-12 (cont.): (c) as in (a) but the plane is defined at 45° of elevation (d) as in (b) the plane is defined at 45° of elevation.

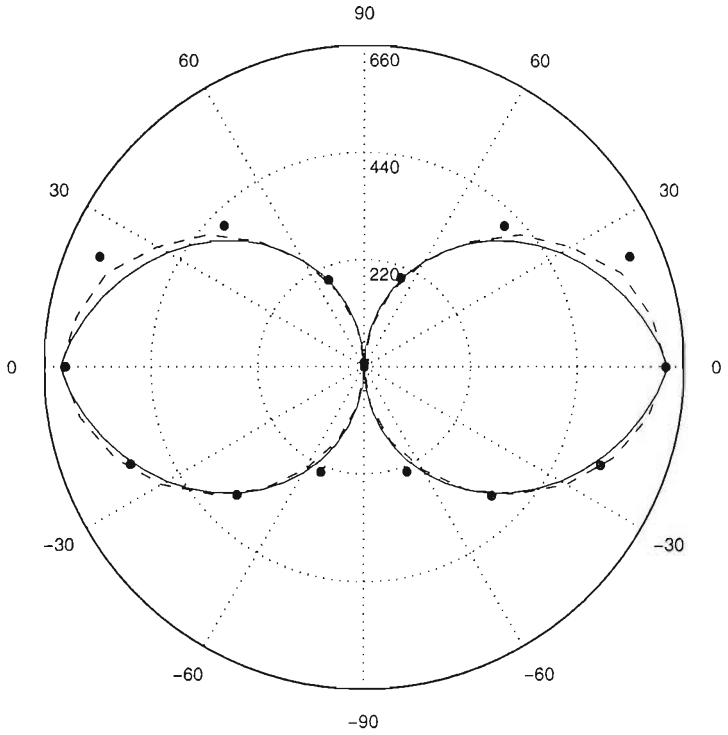
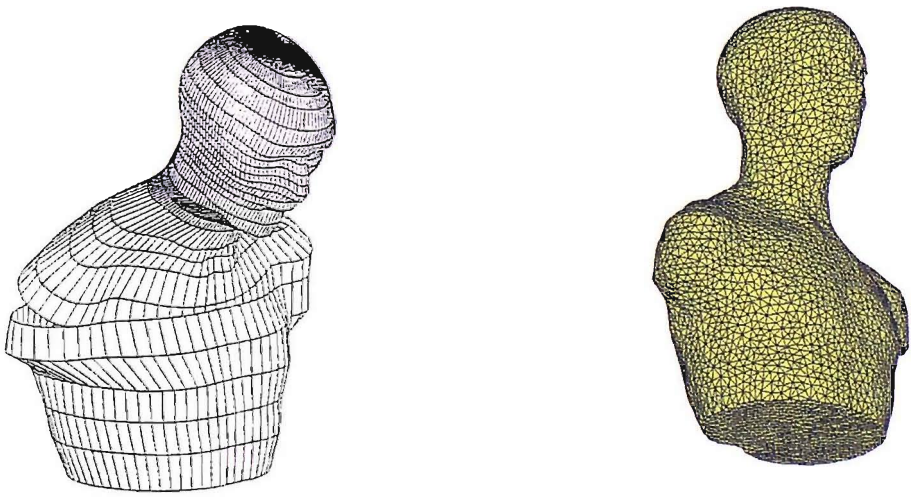


Figure 4-13: A comparison of ITD in the lateral vertical plane: with a sphere (solid line), an ellipsoid (dash line, data from above), and average of 70 people (After Minnaar *et al*, 2000). The polar axis shows the angles in degrees, and the vertical axis shows the ITD in μsec.



(a) (b)

Figure 4-14: The models of CORTEX artificial head (a) original CAD model (b) Decimated BEM model for head without pinna (a half model was used with 5596 nodes and 10926 linear triangular elements).

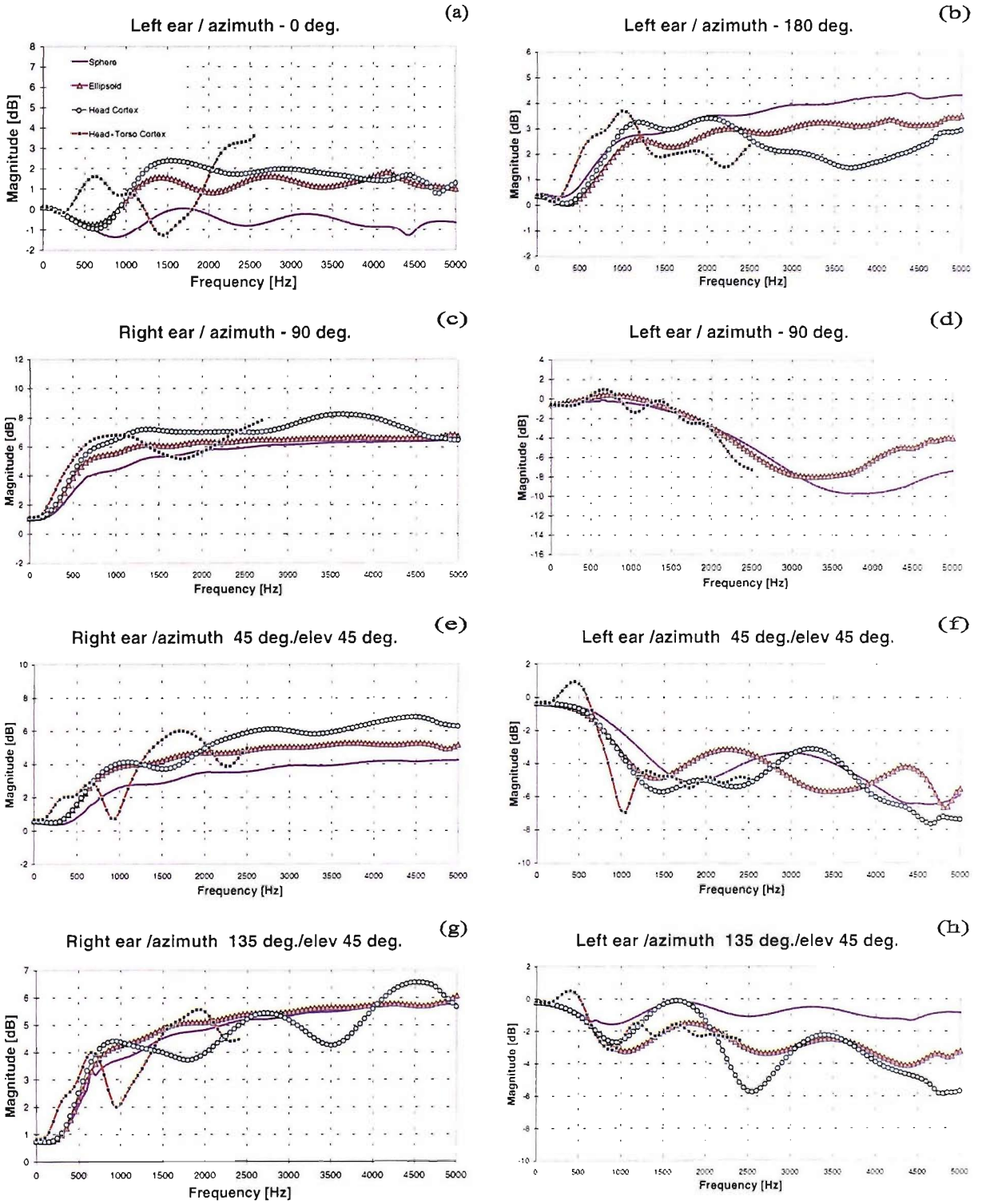


Figure 4-15: The magnitude of the transfer functions of a sphere, an approximation of an ellipsoid, the head of the CORTEX (C1), and the head with a torso (C1 and TO1), at various angles of incidence. The source is positioned at a radius of 1.4 m from the centre of the models, and the 'ears' of the models are positioned as close as possible to the original positions at the entrance to the ear canal.

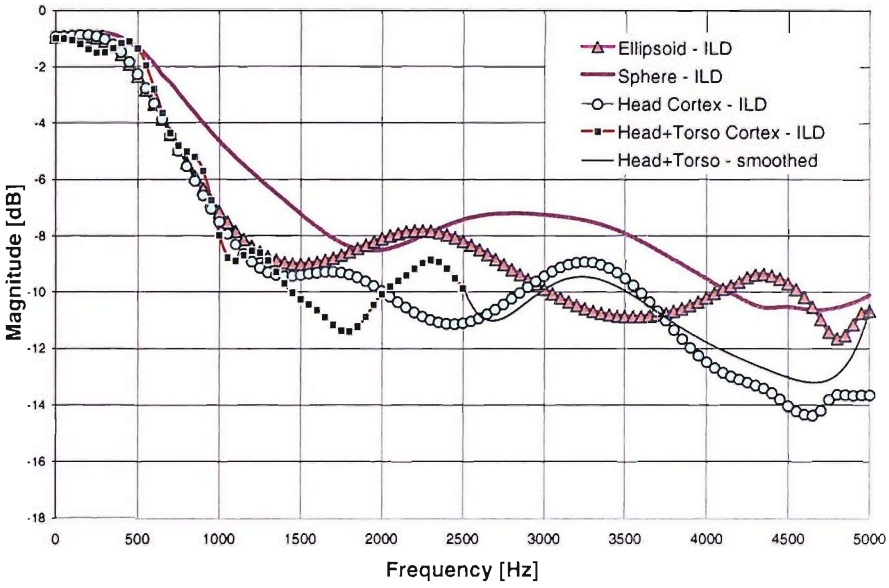


Figure 4-16: The Interaural Level Difference (ILD) for an arbitrarily chosen elevation angle (at $\phi = 135°, \theta = 45°$) for the above models.

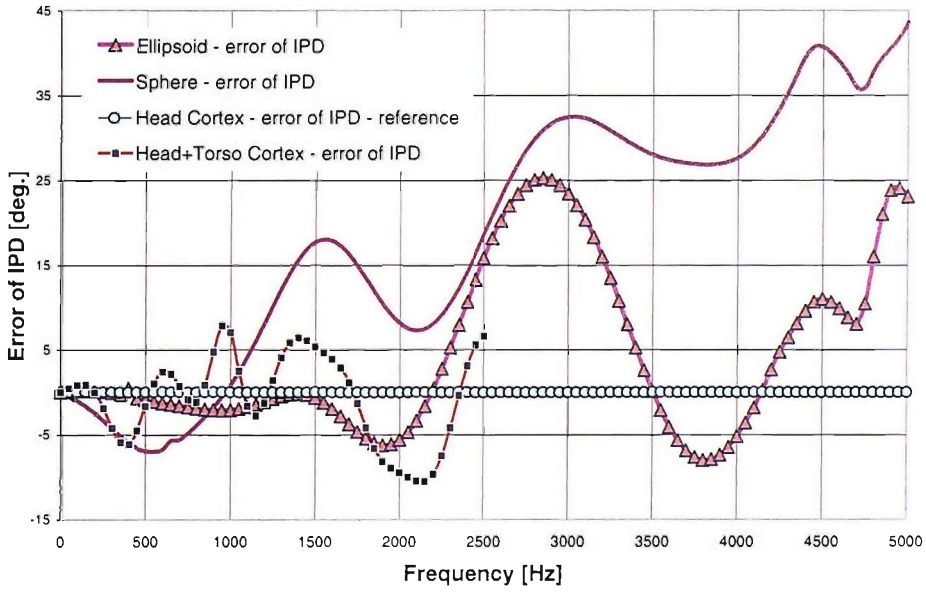


Figure 4-17: The errors of the Interaural Phase difference (IPD) at $\phi = 135°, \theta = 45°$ when the IPD of the above models is subtracted from the reference IPD of the CORTEX head (C1).

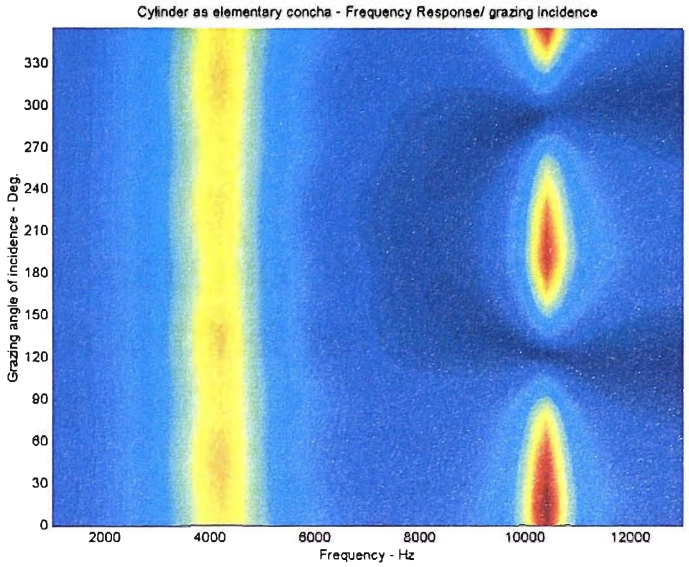


Figure 4-18: Frequency response of a baffled cylinder excited at grazing incidence. The first 'omnidirectional' mode (at 4.2 kHz) is almost independent of the angle of excitation, whereas the first transverse mode pattern (at 10.4 kHz) shifts with angle. The scale is similar to the scale presented in Figure 4-19(c). The simulation point at the base of the cylinder at which the above results were computed was offset by 3 mm from the centre.

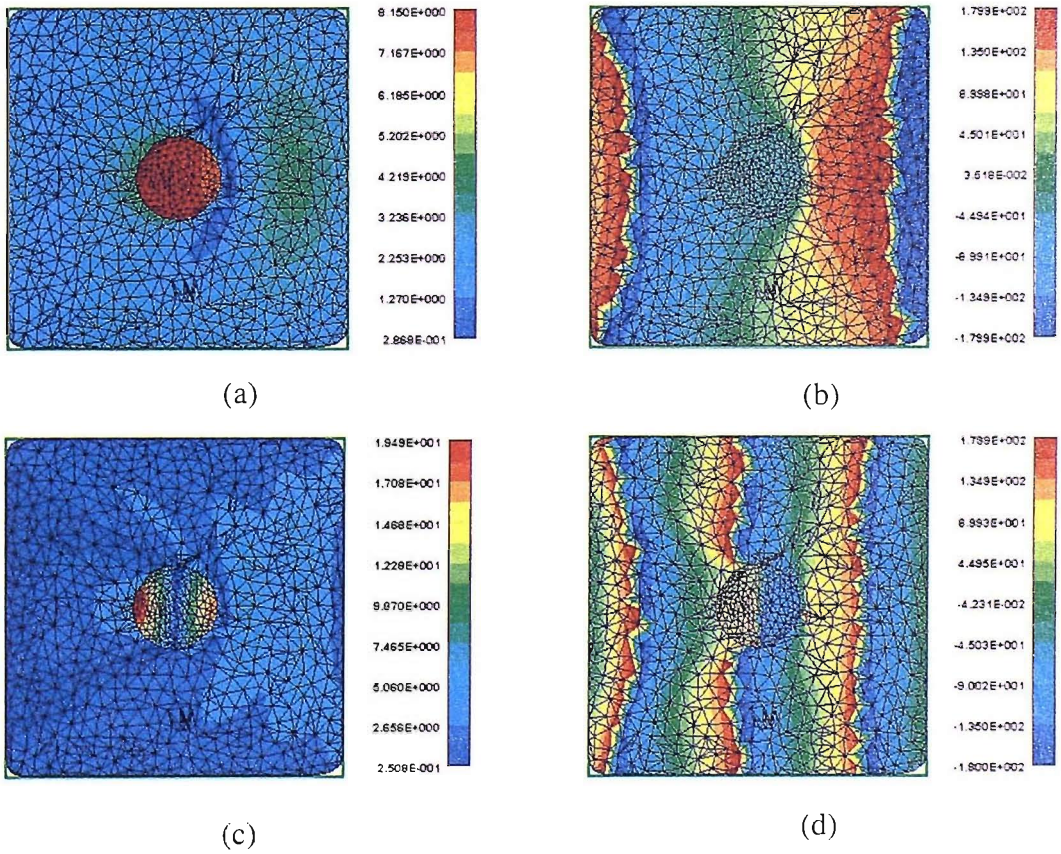


Figure 4-19: The pressure variation in a baffled cylinder. These normal modes have been excited by a plane wave source at grazing incidence. (a) The amplitude of the pressure at 4.2 kHz (b) the phase of the pressure at 4.2 kHz (c) the amplitude of the pressure at 10.4 kHz (d) the phase of the pressure at 10.4 kHz.

CHAPTER 5

NUMERICAL MODELLING OF THE RESPONSE OF THE EXTERNAL EAR AND THE HRTF USING ACCURATE MESH MODELS

5.1 INTRODUCTION

In the previous chapters we presented simple geometrical models such as a rigid sphere or a baffled cylinder which can be validated against analytical solutions, and also we introduced more complex models such as the ellipsoid that is more difficult to investigate analytically, or the head with torso that cannot be investigated analytically. In this chapter we attempt to simulate the response of 'real' and complex shapes such as the human head and pinnae, and to extend the frequency range. In practice, many technical hitches make this problem far from being trivial. Therefore, a few key questions that have emerged during this part of the research are summarised as follows:

- What method of capturing the geometry should be used? Can we obtain a high accuracy geometric model of the head, and especially the pinna? Can the computer model include all the folded parts of the helix, fossa of helix, and antihelix? Can the entire volume of the cymba and cavum concha be captured correctly?
- Is it feasible to obtain a mesh model that includes both the precise representation of the pinna, and also has a small size (i.e. small number of triangles/elements), so that it can be handled by the computing hardware and the BEM software? What techniques are required to be developed in order to manipulate the mesh model?

- Do numerical singularities and other artefacts of using the BEM distort the results?

This chapter attempts to answer these question. An outline of this chapter is as follows: first, the acquisition of the mesh models and their manipulation for valid BEM models are described. Subsequently, procedures of modelling of individualised HRTFs and the response of baffled pinnae are described and the results are analysed. Examples of modelled impulse response are shown, and the inclusion of an ear canal for the case of a baffled pinna is investigated.

5.2 MESH MODELS

5.2.1 Acquiring the computer models

The initial assumption made during this work was that the highest resolution possible is required for the mesh models of the head and pinnae. It was not clear at that stage what accuracy is required and how sensitive will be the modelling of the acoustical response to geometry approximations. There are currently a few techniques available to obtain a computer model by scanning a physical model. These include: Computed Tomography (CT), Magnetic Resonance Imaging (MRI), 3-D ultrasonic imaging, etc. These are generally used for internal scanning for medical purposes. The main advantage of the 3-D laser scanner technique used in this research is that it can produce fairly quickly an accurate mesh of the surface made out of triangles*.

Simpler methods for obtaining 3-D mesh models based on a still image or multiple images have been suggested. For example Kyriakis and Holman (1997) proposed that the grey levels in the pictures are used to define key features of the structure of the external ear.

* Note that in principle, the use of quadrilateral elements can produce a higher accuracy of the simulation compared with triangular elements, but due to the format of the original data, quarilateral elements were not used for scanned models.

Approaches like this might be very useful in obtaining an individualised HRTFs without the need for complex and expensive laser scanners, such as those used in this research. However, the accuracy and the feasibility of using such simple methods is yet to be investigated.

5.2.1.1 3-D scanned models

In the initial work, the Cyberware 'Head and Face' colour 3-D scanner suggested by Katz (1998), had been used. This scanner includes the 3030/RGB digitising head and the PS motion system. It is mainly used for medical applications, anthropometry, human interface, and in the film industry. The geometry is captured by means of an optical range-finding system that produces around a half of a million interconnected triangles and their vertices in approximately 15 seconds. We refer to this scanner as having a 'low-resolution'. Further details are given in Appendix 1.

To begin with, it was found in the scanning stage, that the accuracy of the pinna, which is our most important part of the head, was poor, since during the motion of the scanner the laser beam detects only unhidden parts. In the initial trials using the KEMAR head, the rear part of the pinnae were significantly distorted, and more importantly, the resulting concha was much shallower then the original rubber ear of KEMAR, and without the details of the cavum and cymba concha, helix, antihelix and fossa of helix. However, only after a set of simulations, it was found that another scanning technique was required for high accuracy modelling of the pinna.

Although still based on the technology of *laser* scanning, the Cyberware 'Mini model' 3-D scanner is based on the high-resolution 3030RGB/HIREZ scan head with a mid-size high-resolution motion system. It moves slowly from side to side in the horizontal plane in a straight line, parallel to the object. The principle of operation is similar to the 'head and face' scanner, but with software controlled, the data is accumulated through repeated scans

at different angles of the pinna. In this way, almost every curvature can be captured correctly. Even the ear canal geometry can be obtained, by providing an additional model with the original's cross-section. Then, the software matches the 'internal' data with its own 'external' data by matching overlapping sections. The duration of this scanning procedure is much longer, of the order of a few hours; before any post-processing is applied to fix connectivity, rough surface, holes, etc.

A few tools were developed and used to integrate the two scans into a single mesh model for both KEMAR head and YK head. Description of the procedure is given in Appendix 1.

5.2.1.2 CAD models

The original CAD model of the CORTEX MK1 and the decimated BEM model were presented in Chapter 4 (see Figure 4-14a and Figure 4-14b). The main difficulty with the original format of this model (IGES standard, with slices) is that the elements and the vertices are not necessarily connected. Also the CAD model of the CORTEX pinna (designed according to IEC 959, 1990) was also provided, but due to the low resolution of this format, only the head and torso model have been manipulated. Once the mesh was repaired interactively by using various tools, it was decimated in order to produce a maximal number of homogeneously distributed vertices that can be solved in-core with the IBEM. Due to symmetry properties, a half model was created and used.

5.2.1.3 Pinna moulding

The external ear of the author was moulded in the Hearing and Balance Centre of the ISVR. The moulding was created in two stages, to make sure that both the internal part of the external ear (concha, and 1 cm inside the ear canal) as well as the outside part around the pinna is captured (see Figure 5.1).

Two plaster models were created from the negative mould. With one was cut along the long side of the ear and included the cross section of part of the ear canal that was moulded. This

cross section was required in the scanning process, to ensure all the curved parts of the pinna are captured. The frame of the plaster model was shaped and attached to an approximate rectangular frame, in a similar way to which the KEMAR pinnae are produced, so that it can be attached to the baffle in the anechoic chamber for measurements (see Chapter 6).

5.2.2 Mesh decimation

The original scan produced a polygonal mesh that describes the surface geometry of the head or the pinna. A polygonal mesh is defined as a collection of edges, vertices and triangular polygons. Each edge is shared by at most two polygons. An edge connects two vertices, and a polygon is a closed sequence of edges (Foley *et al*, 1990).

In the literature, curves and surface mesh simplification algorithms are investigated from different perspectives in fields such as cartography, virtual reality, computer vision, computer graphics, scientific visualisation, computer-aided geometric design, and finite element related methods. Since the CPU time of the BEM increases drastically with the number of nodes (see Chapter 10 for a discussion on the computational cost), it is crucial to optimise the size of the mesh. It is well known that the maximum frequency in the BEM/FEM/IFEM corresponds to the longest edge in the mesh. Any alteration to this global limit will distort the overall results. Therefore a homogeneous distribution of the nodes and elements is required. Research into this topic has revealed that a commercial package for this specific task does not exist. More than 100 algorithms have been developed for decimating mesh models (Heckbert and Garland, 1997), mainly concentrating on preserving the accuracy of the rendered model but these do not necessarily operate with the same restrictions that acoustic BEM would require, i.e. controlling the resolution. Preserving the shape while reducing the number of elements in the model are conflicting requirements in mesh decimation algorithms. In fact, it is impossible to make the distance

between vertices exactly the same for all vertices in the mesh while still describing accurately the shape of the object.

The main algorithm used in this research has been developed by Johnson and Hebert (1997). It is described briefly below (with more details given in Appendix 1). Our modification to the original algorithm includes the addition of controlling separate areas within a model, so that different mesh resolution can be applied to different parts in the head (i.e. the pinna). The main advantage of the algorithm is in successfully handling the two forces in mesh decimation: preserving the shape by limiting a defined maximum 'global shape error' and distributing the vertices homogeneously by local operators.

The goal of this method is to adjust the resolution of the original mesh to a desired value while minimising the edge length spread of the model's histogram. The principle of the algorithm used here is similar to other mesh simplification algorithms in that it iteratively changes the mesh by applying *local* mesh operators. First, all the edges of the model are ordered in a priority queue, with the order being determined by each edge length and a shape change measure. Two operations are iteratively applied to the edges to achieve the desired resolution: edge-split is used to remove long edges, and edge-collapse is used to remove short edges. During edge split an edge is divided in the middle and produces a new vertex, two new edges and two elements, without changing the accuracy of the model. However, the operation of edge-collapse does change the accuracy. An edge is reduced to a point: an edge and two elements are eliminated, and two vertices are replaced by a new vertex. In addition to these local operations, a maximum allowable shape change for the mesh is defined to prevent the mesh from changing too much. As a result, the low priority edges are treated first, and the programme is stopped only if the *global* shape change criteria are met.

Another approach is based on 'edge collapse' only. In this case the original mesh is used where the vertices are already distributed homogeneously (with a resolution of 1 mm for the head scans, and 0.35 mm for the pinnae scans). The priority queue and the shape change measures are similar to those used in the previous method. In this case all vertices in the resulted decimated mesh are found also in the original mesh.

5.2.3 Additional mesh manipulation techniques

Although mesh decimation is the most important tool, essential to convert scanned models to BEM models, more tools are required such as:

- Closing the volume of the head. The scanner produces two large holes: at the top of the head, and around the neck. It is essential to produce a closed volume otherwise the DBEM cannot be used at all, and the IBEM will produce erroneous results since the double layer potential (see Chapter 3) will take into account pressure inside the head and outside it.

- Mesh refinement. This is the opposite operation of decimation, where more vertices and edges are added without changing the topology (such as 'edge split'). This is essential when the principle of reciprocity is used, and source is positioned very close to the blocked ear canal. The cavum concha must be refined. This is also discussed in Section 5.4.2.

- Slicing the model into half. Assuming the head is symmetric and the variation between the ears does not affect the response at only one of them. Slicing the model also at other positions is required in order to compare the degradation of accuracy with different mesh resolutions.

- Reorientation of model. The IBEM and the DBEM symmetry work only with respect to x, y, or z co-ordinates. For the 'IBEM transparency' the infinite baffle must be defined at z=0.

- Substitution of low-resolution pinna with the high resolution one. This involved reorientation of the new pinna in six degrees of freedom, and matching specific patterns of it with patterns of the pinna on the head.

All of these techniques are described in Appendix1.

5.3 BEM MODELS

5.3.1 Heads

Two heads were scanned: the KEMAR artificial head (without the torso) and the head of the author (referred to as the 'YK head'). The DB60 rubber pinna and the YK plaster pinna were scanned separately, and were integrated to the KEMAR and YK head models, respectively, replacing the original 'low-resolution' pinnae.

Two types of decimated KEMAR models are presented in Figure 5-2. In both cases the original data included more than 400000 triangles (around 200000 vertices). The target was to obtain a suitable BEM mesh that could be used to modelling at the maximum frequency possible with the IBEM in-core solver[†]. The mesh on the left (Figures 5-2a, 5-2b and 5-2c showing the vertices, elements, and rendered model, respectively) was decimated using our proposed algorithm that produces homogeneously distributed vertices, thus optimising its size, geometry, and maximum frequency. This resulted in approximately 23000 elements that can be used up to 15 kHz if four elements per wavelength are assumed (for the

[†] SYSNOISE 5.4 is compiled with a 32-bit compiler. This limits the maximum RAM to be used with the in-core solver to approximately 1.2 Gb. Although more memory was available, it could not be used for larger models.

ipsilateral ear), and 10 kHz if six elements per wavelength are assumed (for the contralateral ear).

When a conventional mesh decimation algorithm was used (the mesh on the right, optimised for computer graphics applications), and the number of elements was limited to 23000 elements, a mesh with non-uniform distribution of vertices was obtained (Figure 5-2d), where planar areas were described with less triangles (Figure 5-2e), and complex areas retained a higher density of triangles to preserve the geometry. Note that with both decimation algorithms the rendered images (Figure 5-2f and 5-2c) are very similar. This mesh (on the right column) could be investigated using the BEM reliably only up to 1 kHz (!). This emphasises the significance of optimising the mesh distribution while retaining the accuracy of complex shapes such as the pinna.

It was found very useful to reorient the head in space so its ears are at the same height (both entrances to the ear canal are adjusted to be at y=0). The centre of the line connecting these two points was defined as the centre of the head with the origin of the co-ordinate axis. The head was then divided into two identical parts through a vertical axis that intersects the head along the nose. All the vertices near the slicing plane were snapped to x=0. The new model was then decimated to a few mesh resolutions. Figure 5-3 shows four different model resolutions of half model of KEMAR[‡].

When analysis at low frequencies is required, there is no need to use a very detailed model and in principle the frequency range of interest is divided to 'band-pass' regions to optimise the size of the model and consequently its maximum frequency. This solving procedure, with mesh hierarchies, also reduces significantly the need for treatment of 'irregular frequencies'. This issue is discussed in Section 5.4.2.

[‡] These figures show the decimated KEMAR with the 'low-resolution' pinnae to emphasise the concept of 'mesh hierarchies'. In modelling, these pinnae were replaced with decimated high resolution pinnae.

Since we are interested eventually in modelling *individualised* HRTFs, the YK head was scanned. The main problem in this case was the elimination of hair, which was achieved using a 'shower cap', as seen in Figure 5-4a. The final manipulated BEM model is shown in Figure 5-4b, where the shape of the scalp was smoothed, the original 'low-resolution' pinna was replaced with the accurate pinna, and the resolution optimised for frequencies in the region of 10 kHz[§].

5.3.2 Pinnae

Seven pinnae were scanned and investigated. Four pinnae of KEMAR (the right pinnae: DB60, DB65, DB90 and DB95[**]), B&K, CORTEX, and YK. Since DB95 is not used to represent a typical human ear, and includes a large fitting space in the cavum concha for containing hearing aids, it was found that the resulting response was not representative of a typical ear, and it was decided not to include it in the following analysis of the results.

All pinnae were scanned with the 'high-resolution' scanner. The original models included approximately 150000 triangles (75000 vertices), and include three-dimensional information of the pinna (including its frame and base). Two strategies of model manipulation were used as follows:

- Fitting the pinnae into the head (DB60 and YK pinna only). In this case the surrounding of the pinna was removed, and only the outside, thin shell surface remained. This model was later adjusted further to match the curvature of its boundaries and the surrounding of the pinna in the head.
- Smoothing the boundaries to a specific plane (z=0).

[§] Due to larger dimensions of this head compared with KEMAR, for a given maximum number of elements that could be handled with the IBEM in-core solver, lower resolution with larger elements were obtained, hence slightly reduced frequency range compared to KEMAR.

[**] Note that DB90 and DB95 have larger ear canal openings. Their geometry is mainly suited for the use of earmoulds.

For the latter case, two methods were used; first, since the upper boundaries of the pinna do not lie on a single plane, a gradual slope was added from the higher level to the lowest level (approximately a height of 7-10 mm, as in the sides of artificial head pinnae). This resulted in relatively large models (see Figure 5-5/DB60 and Figure 5-6). Later, the bottom level of the pinna was snapped to a single plane at z=0. The acoustic response was investigated to demonstrate that the errors were negligible compared to the previous case (see Figure 5-5/YK pinna and the pinnae used in Chapter 8). Six smoothed pinnae are presented in Figure 5-5. These were investigated with mesh resolutions corresponding to a maximum frequency of more than 20 kHz. All models include a refined area around the blocked entrance to the ear canal for using the principle of reciprocity (detailed geometrical properties and statistical values of these pinnae are presented in Appendix 1).

Although we concentrated in our study on the 'blocked meatus' response, the addition of an ear canal is not impossible. Stinson and Lawton (1989) showed large variations in ear canal shapes and sizes: the ear canal cross section area can vary between 40 mm^2 and 90 mm^2 with an average of 65 mm^2, also due to eardrum impedance variations, the calculated transformation of the pressure from the ear canal entrance to the innermost point in the ear canal vary substantially, where up to 8 kHz the variations can be within a 4 dB from the average, and above this frequency range the variations can be of 20 dB and more. Therefore, we chose to add a typical cylinder with the dimensions of 22.5 mm length and 7.5 mm in diameter (as used by Zwislocki, 1970). The difficulty in the process of this mesh is the special treatment required with the 'IBEM transparency' module. Since the baffle lies above the ear canal and below the pinna, and sound waves are required to propagate 'through' the entrance to the ear canal, the pinna must be oriented such that the ear canal entrance is exactly at z=0. Then the entrance must be closed with a planar surface with elements at z=0 and the edges at the boundary of the circular surface should be connected to

edges of the canal below and edges of elements of the pinna above. These added elements will be defined as 'transparent' elements in the solving procedure.

It was found that the formulation of the 'IBEM transparency' is very inflexible to deviation of the mesh from these requirements. An example of the cross section of the baffled DB60 and the simplified ear canal (without the transparent elements) is given in Figure 5-6. Note that the addition of the ear canal to the model of the *head* is a much simpler process, since the addition of transparent elements is not required, and the entire mesh model remains a closed volume.

The frequency response of the 'low-resolution' DB60 was investigated as described in Section 5.5.1. Figure 5-7a, b shows two views of the *original* cut-out scan of the pinna, before any decimation was applied. Cleary seen are the lack of details, the distortion (in the form of extrusion) in the back of the posterior wall of the concha, and the shallow cavity of the concha. This model will be investigated in order to study the degradation of the acoustical response when compared with the results obtained with the high accuracy mesh model of the DB60 pinna.

5.4 NUMERICAL MODELLING OF INDIVIDUALISED HRTFs

5.4.1 Limitations of the model

When HRTFs are measured with subjects in the anechoic chamber the sound transformation detected very close to the eardrum includes all the contributions form the body, torso, neck, head, pinnae, ear canals and eardrums, as well as the effects of clothing, and hair. In our simulations we limited our model to include only a rigid head and pinnae that are blocked at the entrance to the ear canals. We therefore expect the following discrepancies when our model is compared with an average measurement:

- **Blocked ear canal**. The valuable positions of the 'microphone' at the blocked ear canal was suggested initially by Yamaguchi and Sushi (1956), investigated extensively by Shaw (1974), and evaluated with the binaural technology by Møller *et al* (1995) and Hammershøi and Møller (1996). It should be born in mind that the response detected at this position equalised with the free-field response (without the head) is the only directional component in the HRTF since the transformation in the ear canal is insensitive to direction up to around 12 kHz (Shaw and Ternaishi, 1968). The transformation therefore needs to be compensated for the ear canal response.

- **Missing torso**. The effects of the neck and torso were investigated by Burkhard and Sachs (1975), Preves (KEMAR, 1978) and Kuhn (1977). Male and female have on average different neck lengths. These differences can shift the frequency of the first pressure minimum (see Figures 4-15e to 4-15h and Burkhard and Sachs, 1975), detected at the eardrum for sources in the front from 1.4 kHz to 1.2 kHz. As a result KEMAR includes three neck sizes. The torso also contributes mainly up to 2-3 kHz. At 1.2 kHz a clothed torso can alter the pressure by 3 dB compared with the head only.

- **Boundary conditions**. The effects of simulating flesh, hair and clothing were also investigated by Shaw (1974) and Burkhard and Sachs (1975). Hair in the form of a wig makes the pressure minima at the eardrum at around 10 kHz less deep but has very little effect at lower frequencies. The effect of neglecting the impedance of the skin is found to be negligible (<1 dB difference below 8 kHz). Kuhn (1977) found that ITD and ILD of KEMAR are different depending whether or not the torso is bare or clothed. Difference of ILD can be up to 3-5 dB up to 3 kHz, and ITD can be reduced significantly (from $\sim 600\ \mu\text{sec}$ to $\sim 500\ \mu\text{sec}$) below 700 Hz if the

torso is clothed. Modelling the impedance values of the hair was carried out by Katz (1998). He found that the hair contributes up to 6 dB at frequencies below 5 kHz. This was higher than the conclusions of Shaw (1974) who stated that the contribution of hair is probably less than 3-5 dB (the effect of clothing) at very high frequencies.

5.4.2 Problem optimisation with the IBEM

The use of the symmetry property described in the previous chapter is investigated here with the IBEM[††], for the case of KEMAR. Two cases are compared: a full model with 30000 elements, and a half model with 15000 elements. The response is compared between the two cases at the blocked ear canal due to a monopole source positioned at $\phi = 90°, \theta = 0°$. It is shown in Figure 5-8 that the errors are less than 0.2 dB up to 10 kHz[‡‡]. As shown in Chapter 4, Section 4.1.2.3, the CPU calculation time is four times faster than when the full head is modelled.

The principle of reciprocity (Section 4.1.2.2) is an essential tool in numerical modelling of HRTFs. Once the frequency response was calculated when the source is positioned very close to the entrance to the ear canal, it is possible to calculate the pressure at any angular position and distance in space at greater speeds than the calculation time of the conventional method (see Section 10.2.8). Figure 5-9 shows the response at the blocked ear canal of a baffled pinna (DB60 with 6887 nodes and 13488 elements). Two cases are investigated:

[††] All the simulations undertaken in this research assumed linear elements, the speed of sound is $c = 340$ m/sec, the fluid density is $\rho = 1.21$ m/s^3 and the boundary conditions are defined for rigid body ($u_n = 0$) except otherwise described.

[‡‡] The models used in this comparison were with the 'low-resolution' pinnae. Therefore the main resonance is shifted to 7 kHz. However, the motivation here is to validate the superposition of the pressure with the symmetric properties.

- The source is positioned in the far field, and the frequency response is detected 1 mm away from the surface of the blocked entrance of the ear canal.
- The source is positioned 1 mm away from the surface, and the pressure is detected in the far field at the position that was previously that of the source.

It is shown that the results are nearly identical providing that enough over-determination points (30) are inserted in the cavity between the pinna and the baffle.

Visualisation of HRTFs could be presented in many ways, since the complex structure at high frequency varies significantly in both frequency and time domains, and with angular directions. In Figure 5-10, 3-D colour maps show how the magnitude of the pressure in the near field (0.5 m) at the reciprocal points, changes with an angular position. The response is shown at discrete frequencies. As the frequency increases, the dynamic range obtained becomes higher. At 200 Hz the variation is almost omnidirectional, and variations of ±3 dB are only due to proximity to the head, and equalisation with respect to its centre. At 1 kHz the head still has similar characteristics to a sphere, with equal contours of magnitude in both the ipsilateral and contralateral ears. Note that the minimum pressure occurs at angles that are not directly opposite the right ear, due to superposition of waves with equal path lengths. At 2 kHz, slight variation can be seen between positive and negative elevation angles. Also the response at the contralateral becomes direction dependent. The pinna still has very little effect at this frequency. At 5 kHz the boost is mainly due to the first resonance of the concha (17 dB at this distance), and the attenuation is mainly due to the shadowing of the head (-39 dB). We can expect even more complex variations in the contralateral ear as frequency increases.

These figures demonstrate the advantage of obtaining continuous maps of HRTFs. As frequency increases the complex patterns show that interpolation of HRTFs based on low-

resolution sampling of measurements will produce large errors, especially for the contralateral ear. In the next sections we investigate the variation of HRTFs with a higher frequency range at different planes.

5.4.3 Results

The results presented below and also in Section 5.5.3 are analysed and compared with the literature, with further simulation results compared with measurements in Chapter 6.

5.4.3.1 Median vertical plane

The modelled amplitude and magnitude of the HRTFs of the YK head are presented in three different planes (The co-ordinate system is defined in Chapter 1, Section 1.6). In all cases the response is detected at the blocked ear canal of the right ear, and the sources are positioned on a sphere with a radius of 2 m. Figures 5-11a and 5-11b present the response due to sources in the median vertical plane. Since the torso is absent the angles below -40° and above 220° are meaningless and therefore are excluded from the figures. The figures are presented on both linear and logarithmic scales. The following observations can be made and compared with the literature:

The first resonance is clearly seen around 4 kHz. It almost does not change with the angle of elevation, but a maximum amplification is obtained when the sound is in the front and above, between $\theta = 0^\circ$ and $\theta = 40^\circ$ (see Shaw, 1997 for similar trends although the pinna is baffled). A similar pattern was obtained with the baffled cylinder (see Figure 4-18). At all angles the first resonance causes an amplification of 10-17 dB (as in Shaw and Teranishi, 1968, Møller *et al*, 1995, Hammershøi and Møller, 1996).

It was noted by a few researchers that a possible frontal cue could be the low pass notch that moves from 6 kHz to 10 kHz as elevation increases from $\theta = -40^\circ$ below the horizontal plane and up to $\theta = +60^\circ$ above the plane (Shaw 1974, Butler and Belendiuk,

1977 and Mehergart and Mellert, 1977). The low pass filter for these angles is clearly seen in the figures. This notch can be explained by the fact that for higher elevations, the reflected path length is shorter, so that the notch in the spectrum is at an increased frequency (Hebrank and Wright, 1974).

Blauert studied the perception of elevation in three sectors: frontal ($-15° < \theta < 45°$), above ($45° < \theta < 135°$), rear ($135° < \theta < 200°$). A 'boosted band' appears around 8 kHz when the source is overhead (Blauert, 1996, pp. 110, Figure 2.47). Hebrank and Wright (1974) also noted that sources above are characterised by a ¼ octave peak between 7 kHz and 9 kHz. In our case, the maximum amplification shown in Figure 5-11b occurs at an angle of $\theta = 80° \sim 90°$ overhead and the frequency is 7.8 kHz.

5.4.3.2 Lateral vertical plane

With the simulation of HRTFs in the lateral vertical plane (presented in Figure 5-12) a larger dynamic range is noticed due to the shadowing of the head. The structure of the notches at contralateral angles is very complex. The general structure of this plane can be compared to the data measured on 40 subjects by Møller *et al* (1995) in this plane: the frequency of the notch at angle $\phi = 90°, \theta = -40°$ starts at 6 kHz and increases with angles up to 10 kHz at $\phi = 90°, \theta = +30°$. The same phenomenon appears in their results (Møller *et al*, 1995, Figure 15), and also in Shaw and Teranishi (1968, pp. 248).

The maximum pressure is obtained at $\phi = 90°, \theta = 0°$ at 5.5 kHz with a boost of 15 dB. The attenuation of angles in the shadow zone reaches levels of -33 dB, and the overall dynamic range of HRTFs in this defined plane is almost 50 dB. Therefore the features of the HRTFs cannot be characterised under this scale. Also, the reliability of the large attenuation in the shadow zone is probably not very high due to lack of mesh resolution, and this is investigated against measurements in Chapter 6 (for the case of KEMAR).

5.4.3.3 Horizontal plane

The variation of HRTFs in the horizontal plane (presented in Figure 5-13) is mainly characterised by the deep notch whose frequency - 8.5 kHz, is fairly constant. Similar features exist in Shaw, 1997, p. 29 (although the minima there occur at 9.5 kHz, see also Section 5.5.3.3). In his averages from 12 studies (more than 100 subjects), the measurement is close to the eardrum and averaged. However it was shown in Shaw and Teranishi (1968, Figures 4 and 10) that the minimum is present in both cases and that it is a result of an interference effect. They concluded that the minimum around 8 kHz is relatively independent of the angle of incidence in the horizontal plane as is shown in our case, although this minimum is not present in the contralateral ear response (as can be seen in Figure 5-13).

5.5 NUMERICAL MODELLING OF THE RESPONSE OF BAFFLED PINNAE

5.5.1 The response of a 'low-resolution' DB60

In the initial stage of the work, the 'low-resolution' scanner was used. At this stage, the full HRTF database (of the head with the pinna) was simulated. Due to the difficulties in obtaining reliable results at high frequencies, it was not clear to what extent the accuracy of the distorted geometry affects the acoustical response. In order to verify that the resonance frequencies are due to the geometry of the pinna, and not due to numerical errors that might arise with a large model, the original data of the pinna was 'cut' from the head and transformed to a baffled plane. The geometry of this pinna model is depicted in Figure 5-7a and Figure 5-7b. In this case, the original model comprises 5535 nodes and 10844 elements. Virtually the same acoustical characteristics appeared with the baffle compared with the full head for certain angles (see below). The mesh was not decimated to ensure that no additional geometrical distortions would occur, but only smoothed at its boundaries

to the plane of the baffle. The response of the 'low-resolution' DB60 for grazing incidence is presented in Figure 5-14a. The response of the accurate DB60 pinna for the same plane under similar conditions is presented in Figure 5-14b.

It can clearly be seen that the visually distorted parts result in acoustical errors across angles and frequencies. The general trend that appears in both pinnae is two main resonance frequencies with large amplification at frontal angles, whereas rear and bottom angles have very little amplification compared with the front and top angles. However two significant errors appear:

- Due to the shallow concha all resonance frequencies shift upwards. In this case 6.5 kHz and 11.5 kHz obtained with the 'low-resolution' pinna, compared with 4.9 kHz and 10.3 kHz obtained with the accurate pinna.
- Although the magnitude of the response is of the same order the overall change of the response with angle and frequency is not as distinctive as with the accurate pinna. This is expected since the variation of the folded parts of the pinna is also not very detailed.

It is concluded that the 3-D mesh with accurate geometry is essential for modelling the response at high frequencies.

5.5.2 The effect of the baffle on the response

The pinna is attached to the side of the head at different positions and angles among individuals (see Alexander and Laubach, 1968, and Bukhard and Sachs, 1975). By attaching the pinna to a baffle we inevitably change the transformation of sound detected at the eardrum, or blocked entrance to the ear canal. However, since we investigate only the ipsilateral side, we assume that differences between the baffle and the head, will not be

significant, if we limit our expectations to the general structures of peaks and notches, varying across frequency and directions.

In Figures 5-15a to 5-15f, we compare the response of two pinnae (DB60 and YK pinna) for the baffled case and when the pinna is attached to the head. The positions of the sources in space were aligned approximately with the same angular directions since the co-ordinate systems are different in both cases. The plots on the left compare the response at grazing incidence angles. These are very similar, and the slight differences can be attributed to the fact that the source and microphone arrangement in the two cases is different and therefore the angle of excitation is slightly misaligned. However, both the frequencies of the peaks and notches as well as the magnitudes represent the response obtained with the full head.

On the right, three source positions in the horizontal plane have been investigated with KEMAR and DB60. As the source moves to the rear, the error increases due to interaction with the baffle[§§]. In this case we can still study the general trend of the variation of the response but the accuracy is frequency and direction dependent, in particular if all pinnae under investigation have the same baffled conditions.

5.5.3 Results

5.5.3.1 The response at grazing incidence

A few authors emphasised the significance of the notch that varies between 6 kHz and 10 kHz according to the angle of sound source (Shaw and Teranishi, 1968, Blauert 1972, Bloom, 1977 and Butler and Belendiuk, 1977). The results presented in Figure 5-16 can be compared with the results obtained by Shaw (1997, pp. 34-35) since he investigated the response of baffled pinnae. Only the general trends should be analysed since probably the dimensions of the pinnae are different in these studies, and he used a point source

[§§] In the figures on the right the response of KEMAR was 6 dB higher than presented due to free-field equalisation.

positioned in the near-field (8 cm from the entrance to the ear canal), whereas, in this case a monopole was simulated in the far field using the principle of reciprocity. In his figures, the response was presented in a 2-D format: magnitude *versus* frequency. In this form, the trends of the most significance frequencies are highlighted, and also certain features such the 'low-pass' observed with parallel sloping lines between 5 and 11 kHz. However, it is difficult to observe the variation of magnitude with angle, especially at high frequencies.

In Figure 5-16 we present, on a linear scale, the response of six pinnae to excitation at grazing incidence. The response is normalised with the response detected at the centre of the baffle, without the pinna. At DC the amplitude of the response is unity (or 0 dB). In our cases, it is found that pinna amplification can be as high as 4.5 (increase of 13 dB). The figures show the variation of maxima and minima as a function of frequency and angle. We can conclude with the following observations:

- In a similar way to the case of the median plane angles to the full head (see Section 5.4.3.1) the notch for all six pinnae starts at around $\theta = -90^\circ$ at 6 kHz and increases up to $\theta = +40^\circ$ at 10 kHz.

- In general, the peaks in the figures are arranged in vertical lines, i.e. at certain resonance frequencies, the pinna is excited with different efficiencies depending on the angle of excitation. There are no more than six resonance frequencies, as Shaw found, but in some pinnae only five resonance frequencies were found.

- In the work of Shaw, maximum amplification is always observed at the first quarter wavelength (between 4 and 5 kHz). An almost similar amplification level can be found at higher frequencies as well (e.g. subjects A, D, and J, in his figures, Shaw, 1997, pp. 34-35). This trend is evident in our case, with the exception that, with some ears, higher amplification at high frequencies is noticed compared with the

first resonance. This can be explained by the fact that the source is positioned 5 cm away from the baffle, and amplification values at high frequencies are very sensitive to its position in the vicinity of the baffle. In addition, due to approximation made using the principle of reciprocity where the source is not positioned on the surface but 1-2 mm away, some variation of amplification can be obtained at frequencies above 10 kHz.

- The range of frequencies around 4 kHz, corresponding to one of Blauert's 'v'-bands (1997), was found to play an important role in front-back discrimination. This frequency is attenuated for rear sources and boosted for frontal sources due to destructive and constructive interference respectively between direct and reflected sound. In all figures the frontal angles ($-30° \leq \theta \leq +30°$) are characterised by a boost, and angles at the rear (above and below) are characterised by smaller amplification values.

For grazing incidence excitation, a comparison of all resonance frequencies found by Shaw, and this work is given in Table 5-1. The first resonance frequency is very similar in all cases except for a shift in the first resonance of DB60. This increase is expected due to smaller dimensions of this particular pinna (the larger version of KEMAR, the DB65 has a high level of agreement with the averages of Shaw). In two cases not all of the six resonance frequencies were found. This was investigated further with measurements in Chapter 6 and also with mode shape analysis in Chapter 8. The reliability of the last mode is doubtful due to the large sensitivity to source position, mesh geometry, and alignment, although in general a peak was always found in this range.

DB60	DB65	DB90	YK	CORTEX	B&K	Shaw
4.9	4.2	4.2	4.1	4.2	4.1	4.2
7.8	7.2	7.2	7.6	7.2	7.7	7.1
10.3	9.5	9.6	-	9.6	10.5	9.6
-	11.6	11.8	11.2	11.8	12.2	12.2
14.0	14.8	14.7	14.0	14.1	15.3	14.4
17.0	18	18.4	17.8	17.4	18.0	16.7

Table 5-1: Resonance frequencies in kHz of six pinnae modelled with the BEM. The frequency corresponds to the frequency at which the maximum amplification is reached in the resonance frequency range. The results of Shaw (1997) show the average of 10 pinnae.

5.5.3.2 The response in the lateral vertical plane

The main features appearing in Figure 5-17 can be summarised as follows: two main peaks occur at different frequencies: at 4-5 kHz with the strongest excitation at normal direction ($\phi = 90°, \theta = 0°$), and around 10 kHz at upper angles ($\phi = 90°$, $40° \leq \theta \leq 80°$). Note the similarity with the resonance frequencies of the cylinder. The behaviour of the YK pinna is slightly different with the second resonance occurring at 8 kHz instead of 10 kHz. A progressive notch is observed from 6 kHz, down at $\phi = 90°, \theta = -90°$, and increases with frequencies up to approximately 12 kHz at approximately $\phi = 90°, \theta = +40°$. Variations above 10 kHz have no consistent structure.

5.5.3.3 The response in the horizontal plane

The most striking feature in the horizontal plane (Figure 5-18) is the simpler variation of the first few peaks and notches. The first resonance shifts only slightly from 4 kHz to 6 kHz for angular positions of sources that shift from the front to $\phi = 130°$ at the rear. In all pinnae, from $\phi = 130°$ to $\phi = 180°$ the peak in this frequency range is substituted with a notch. In fact this notch starts at frontal angles (from $\phi = 20°$ at around 1 kHz) and progress monotonically with frequency up to $\phi = 130°$ at 5 kHz, and then the attenuation is noticed at all angles. The deepest notch at around 9 kHz hardly changes as a function of the angle. This notch is found problematic in every HRTF database in the horizontal plane

(see Gardner and Martin, 1994, and AUDIS, 1996) since it requires a large dynamic range in the equalisation process and also since it has a narrow bandwidth, equalisation may result in a boost at a slightly shifted frequency. As before, other spectral features do not repeat systematically among all pinnae above 10 kHz.

5.5.4 Examples of modelled impulse responses

An analysis of measured impulse responses for baffled pinnae could not be found in the literature. The following impulse responses were calculated by applying the Inverse Fast Fourier Transform (IFFT) to frequency responses calculated between 1 kHz and 20 kHz, and therefore the sampling frequency is 40 kHz. It should be noted that at the Nyquist frequency the imaginary part of the pressure should be zero (or the phase). Therefore, a simple linear correction was applied to the complex pressure values at high frequencies (between 15 kHz and 20 kHz, with gradual linear decrease of the phase down to 0°), without changing the magnitude.

The impulse response at grazing incidence for DB60 and DB65 are given in Figure 5-19a and Figure 5-19b. These responses are characterised by two adjacent peaks for angles from $\phi = -50°$ up to $\phi = +30°$. The same trend can be seen in Hiranaka and Yamasaki (1983, Figs. 3 and 4) although they investigated the impulse response with the effect of the head.

For the lateral vertical plane, the DB90 is investigated, with a clear trend of a delayed secondary reflection as the source is lowered. Hiranaka and Yamasaki (1983) confirmed that major reflections occur within 350 μsec after the first arriving sound, and that the delay increases as the source is lowered.

5.5.5 Baffled pinna with a cylindrical ear canal

Møller (1992) defined the following variables for sound transmission: p_1 is the pressure at centre position of head, p_2 is the pressure at entrance to blocked ear canal, p_3 is the pressure at entrance to the open ear canal, and p_4 is the pressure at the eardrum. Therefore the free-field equalised HRTF can be expressed as

$$\frac{p_4}{p_1} = \frac{p_2}{p_1}\frac{p_3}{p_2}\frac{p_4}{p_3} \tag{5.1}$$

It has been shown (Møller, 1992, Møller *et al*, 1995, and Hammershøi and Møller, 1996) That the ratio p_2 / p_1 is direction dependent, whereas p_3 / p_2 and p_4 / p_3 are direction independent. It has also been shown that measurements at the blocked ear produce a lower standard deviation between subjects than measurements undertaken at the open entrance and the eardrum and this point is suitable for binaural recordings. The significance of this approach is the simplification of measuring HRTFs (see also the SNAPSHOT system used by Wenzel *et al* (2000) that is designed to measure the response at the blocked ear canal of subjects). However the quality of the sound will be different due to the variation in resonance between blocked and open ear canal (see Shaw and Teranishi, 1968).

As stated in Section 5.3.2 the motivation of inclusion of a simple cylindrical canal has been raised due to variations among individuals for the canal shape, size and eardrum boundary conditions. Therefore we include an average ear canal with dimensions based on the Zwislocki coupler (Zwislocki, 1970), and the boundary conditions based on average data published by Shaw (1974).

5.5.5.1 Boundary conditions

Using the 'IBEM transparency' formulation (Chapter 3, Section 3.5.3), the following boundary conditions were applied:

- 'Transparency' for elements defined on the plane of the entrance to the ear canal, at z=0. This is to ensure that waves can propagate through this 'opening', where the 3-D mesh is attached to an infinite baffle.

- Monopole source excitation. Using again the principle of reciprocity, the source was positioned 1 mm above the bottom of the canal (at –21.5mm). The mesh includes 8203 nodes, 16118 elements, of which 85 elements are transparent at z=0, and 666 elements comprise the eardrum and are applied with the impedance boundary conditions.

- Frequency dependent admittance values have been assigned to the elements of the circular area at the bottom of the cylinder. The values are based on the average impedance values (resistance and reactance) given by Shaw (1974).

In the frequency range of up to 8 kHz these values are similar to the response of the Zwislocki ear canal simulator (Zwislocki, 1970). The impedance values, measured in acoustic Ohms, were converted to specific impedance by multiplying with the area of the bottom of the canal and converted to specific admittance values in Rayls. It should be noted that this simulation provides only a rough approximation to the response that would have been detected by the Zwislocki simulator DB100. The impedance values were measured by Zwislocki at the four branches (see KEMAR, 1978), and not at the bottom. However, it is possible in principle to calibrate the response of the canal, by applying admittance values obtained from empirical data.

5.5.5.2 Results

The results presented in Figure 5-20 include simulation with two models: blocked ear canal (See Figure 5-5), and the current model with the inclusion of the ear canal. For the latter two cases were investigated: rigid boundary conditions at the eardrum, and frequency dependent complex admittance values. The general characteristics of the three curves are in

agreement with the results of Shaw and Teranishi (1968). A strong resonance appears at 3 kHz due to the canal resonance. Amplification is in the order of 25 dB compared to the response of the baffle. The admittance values reduce the amplification to 15-17 dB. The blocked ear canal creates a peak around 4.5-5.5 kHz depends on the angle of excitation. In grazing incidence the amplification is 10-12 dB (see also Shaw, 1997). Notice that the sharp minima around 8 kHz is present in both blocked pinnae types as stated by Shaw and Teranishi (1968).

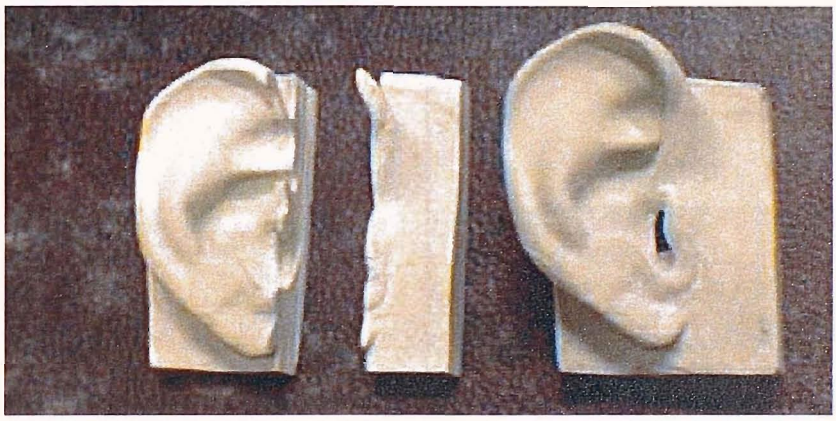
It should be noted that this feasibility study requires further investigation of the properties of boundary conditions at high frequencies, and more accurate modelling, for example of the four-branch Zwislocki coupler.

5.6 CONCLUSIONS

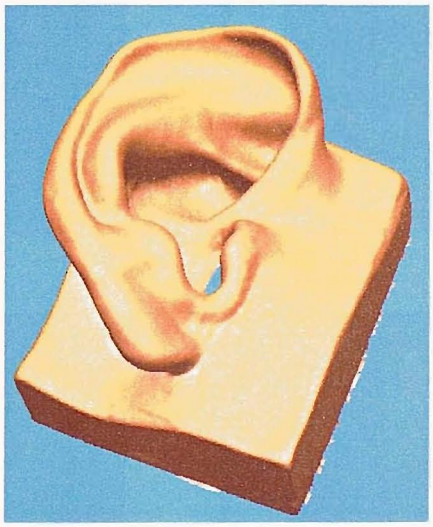
The method of capturing and manipulating scanned surface models has been described. Since the BEM is very inefficient when large mesh models are investigated, the optimisation of various stages in the simulation is found to be crucial: the vertices of the mesh should be distributed homogeneously, but the overall shape should not be distorted. the pinna scanned with an accurate laser scanner needs to replace the low-resolution pinna. the mesh needs to be cut to half and reoriented in space so that the symmetry option in the BEM can be used. Local refinement around the blocked ear canal is required when the source is positioned very close for simulation using the principle of reciprocity. This was used to demonstrate the spatial variation of the pressure at discrete frequencies. The visualisation in three dimensions demonstrated the high variation of the pressure in the shadow zone (the contralateral side).

The simulation results of individualised HRTF are presented in the median, lateral vertical and horizontal planes. The results are compared to previously published observations of spectral features of the HRTF.

Due to the high complexity of the pinna shape, and also its frequency response at high frequencies, it was found that high accuracy is required in the original scanned model. In order to obtain good agreement at frequencies in the region of the first quarter wavelength resonance, at least the shape, size and volume of the concha should be close to the physical model. The simulation of baffled pinnae reveals similar peaks and notches found in the simulation of the full head (for the ipsilateral ear only, of course). At high frequencies the dominance of the pinna is evident compared with diffraction effects around the head. The peaks and notches of six pinnae have been analysed and compared with the literature. Continuous maps can reveal the similarities and differences between pinnae. Large variations appear above approximately 10 kHz even when all pinnae are baffled and oriented to have the same tilt angles with the baffle. It is expected that these high variations will occur at lower frequencies, around 6-8 kHz if the HRTFs include also the effect of the head.



(a)



(b)

Figure 5-1: YK plaster pinna model (a) still picture of two identical moulded plaster pinnae. One was cut vertically for the detection of the curved surface by the laser scanner (b) the resulting computer 3-D mesh model with approximately 112000 vertices and 225000 triangles.

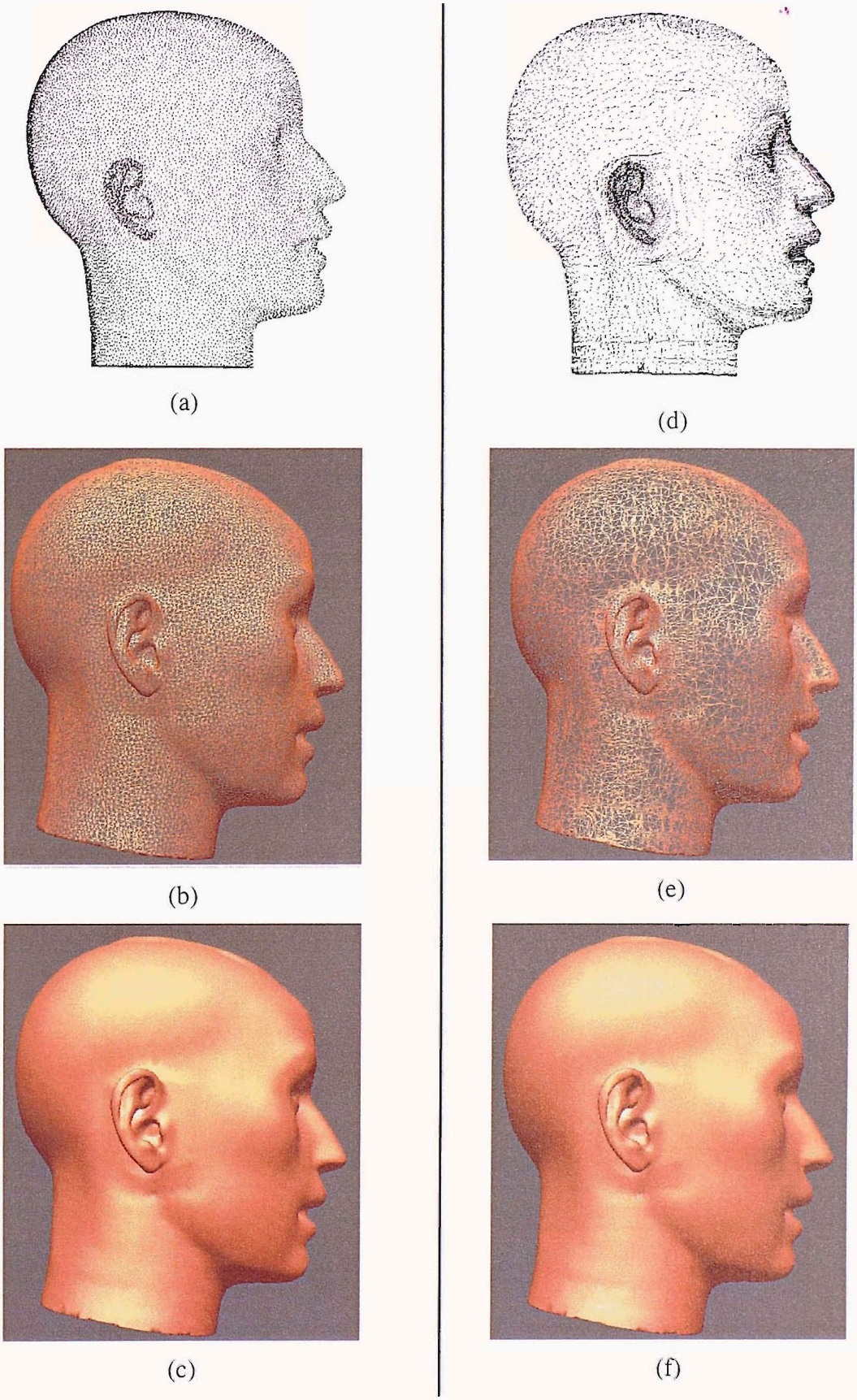


Figure 5-2: Decimated mesh with vertices (top), elements (middle) and rendered models (bottom) of KEMAR. In both cases the model comprises 23000 elements and 11500 vertices. (a, b, c) homogeneous mesh decimation optimised for the BEM. (d, e, f) conventional mesh decimation algorithm optimised for computer graphics.

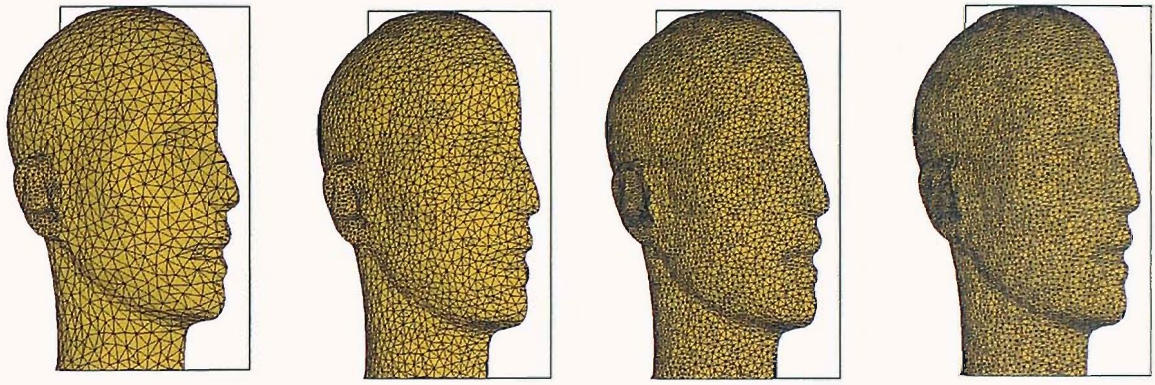
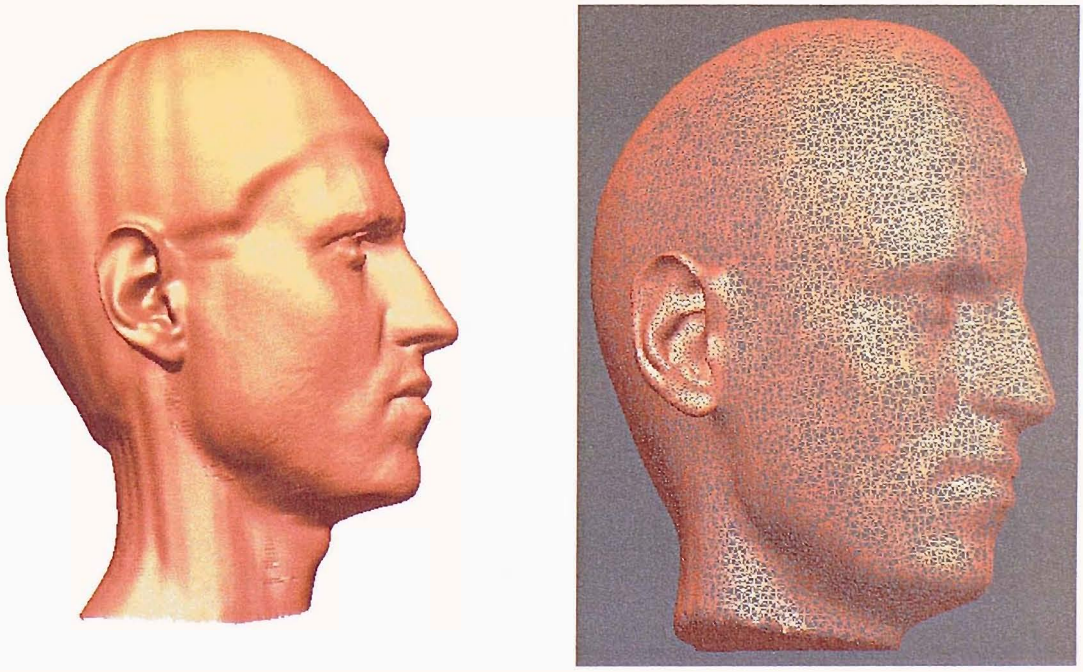
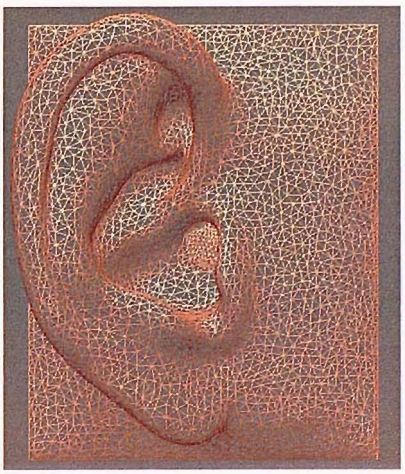


Figure 5-3: Mesh hierarchy of half models of KEMAR with the following approximate number of elements (a) 2500 (b) 5000 (c) 10000 (d) 15000. The pinna of the models is decimated separately with a higher resolution to enable the positioning of the source close to the entrance of the blocked ear canal, when the simulation is undertaken with the principle of reciprocity. Also shown is the plane of symmetry for each model.

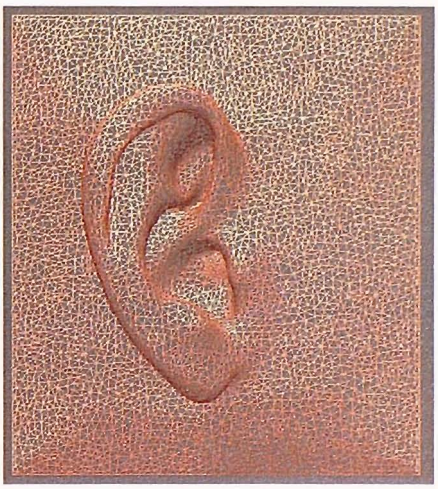


(a) (b)

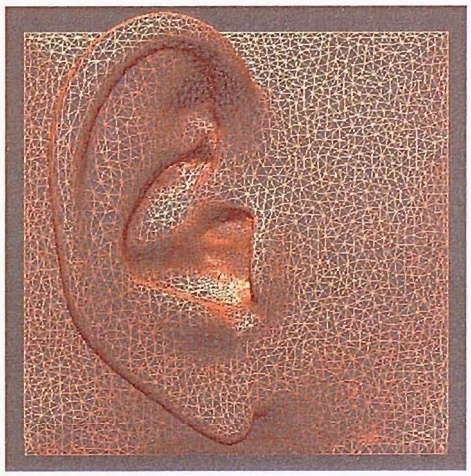
Figure 5-4: YK head model (a) Original rendered model with a 'shower cap' used to conceal the hair. The model consists of 418000 elements (b) Hybrid model includes decimated head and pinnae captured with two types of scanners.



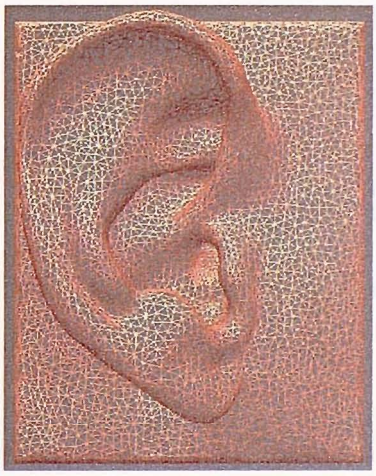
DB90



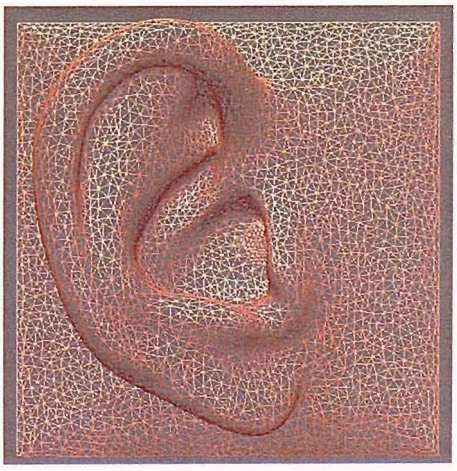
DB60



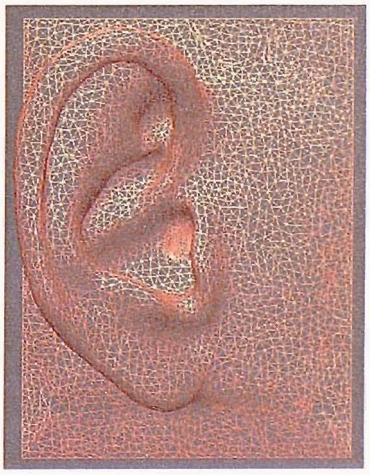
CORTEX



YK



B&K



DB65

Figure 5-5: BEM pinnae models: decimated, aligned and smoothed into a rectangular frame lying at z=0. Distortion of the geometries of the pinnae was kept minimal.

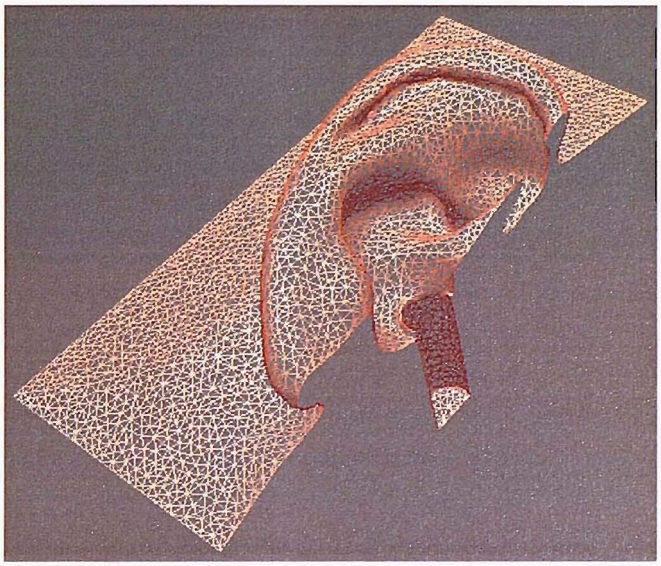
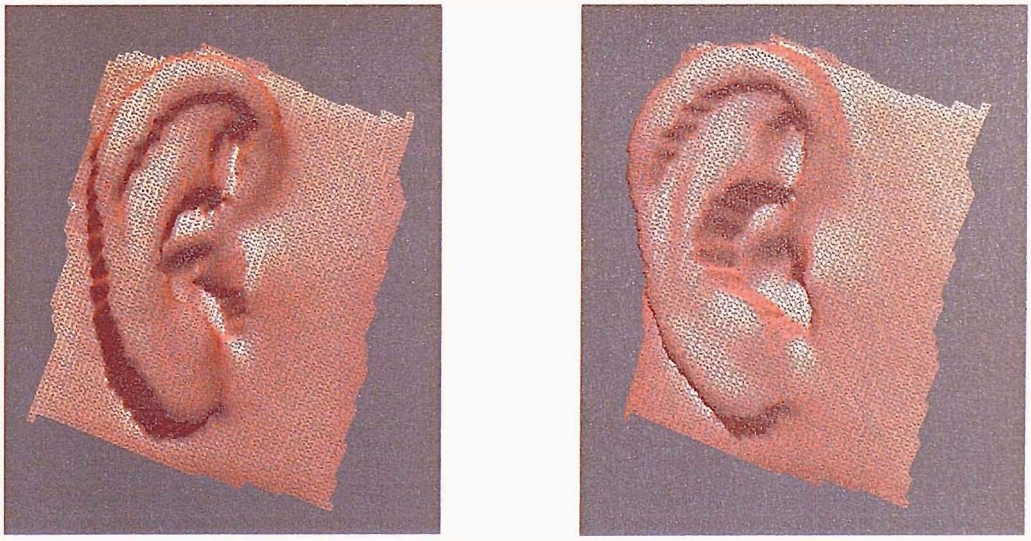


Figure 5-6: Decimated BEM model of DB60 with the addition of a cylindrical ear canal. Only the cross section is shown, without the 'transparent' elements positioned at the entrance to the canal. The entire model consists of 8189 nodes and 16113 elements.



(a) (b)

Figure 5-7: Original 'low resolution' DB60 mesh model with approximately 6000 nodes and 12000 elements. Two views are shown to demonstrate the coarse representation of the pinna (a) a 'solid' rear and shallow cavum concha (b) distorted antihelix and the posterior wall of the cavum concha.

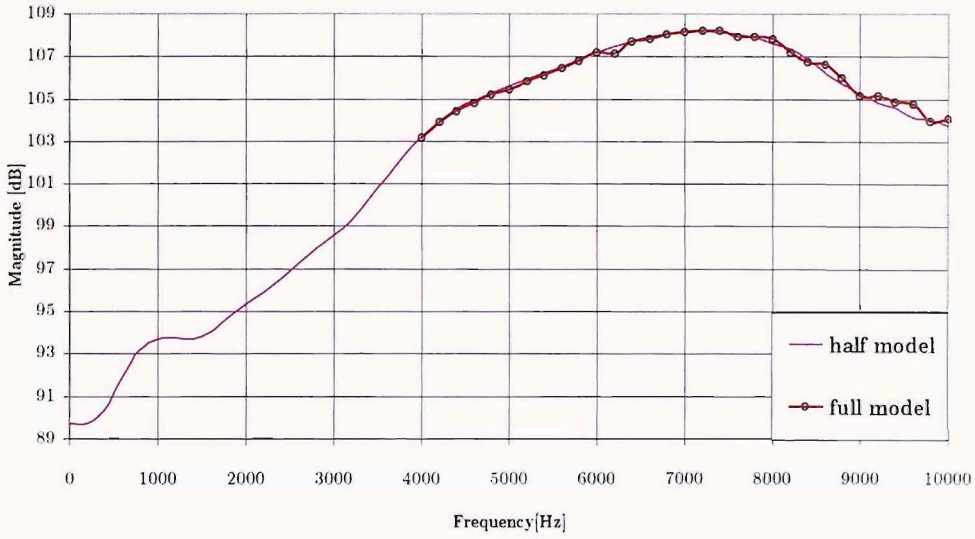


Figure 5-8: Validation of the use of the property of symmetry. When using a half of a model the error is less than 0.1 dB, and the CPU time is improved by a factor of 4.

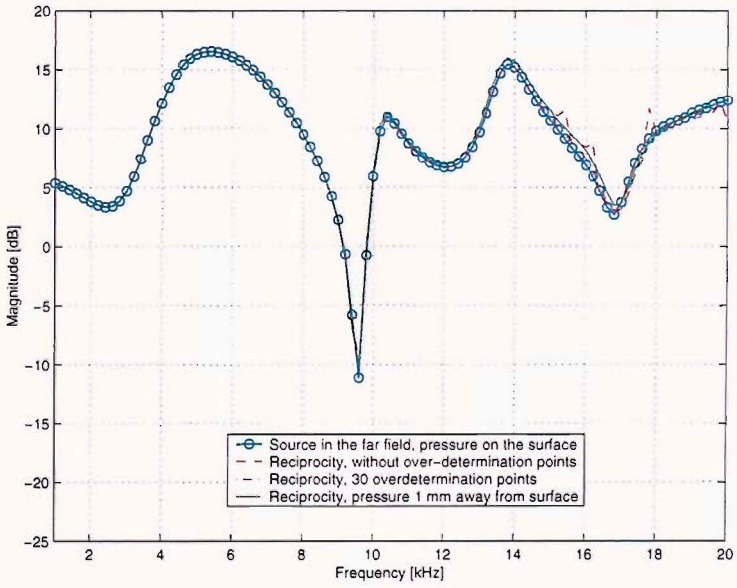


Figure 5-9: Validation of the principle of reciprocity using a baffled DB60 pinna. The model consists of 6887 nodes and 13488 elements. The curves shown are the frequency response at the entrance to the ear canal due to a source at normal angle, $\phi = 90^{\circ}$. Reciprocity is checked when the pressure is investigated on the surface of the pinna due the source in the far field and also at the position of the source which is 1 mm away from the surface.

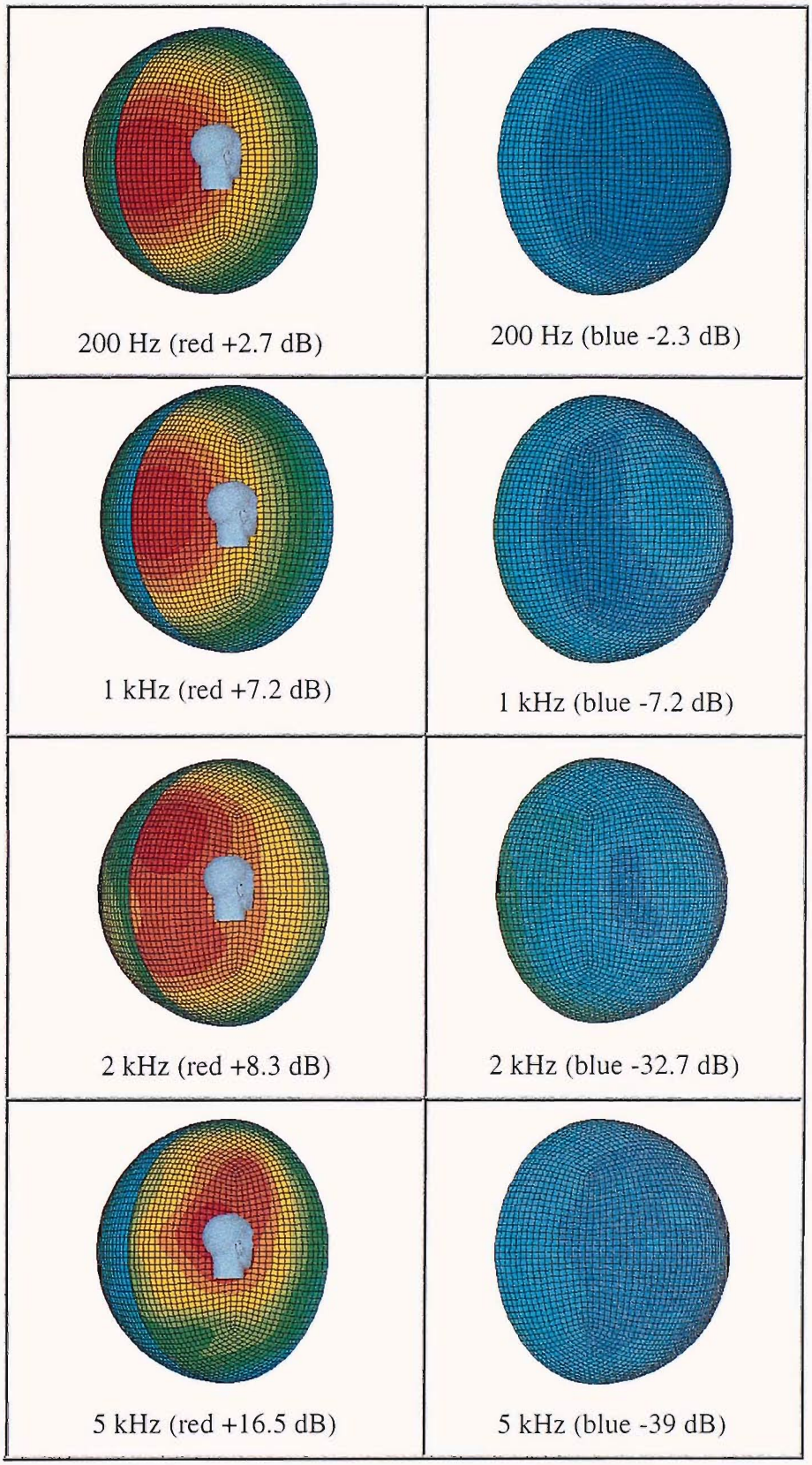
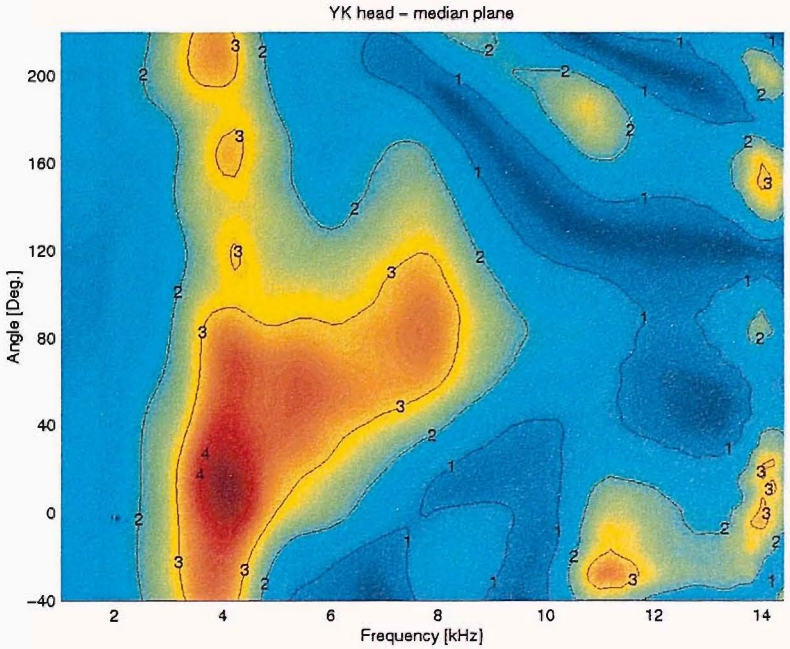
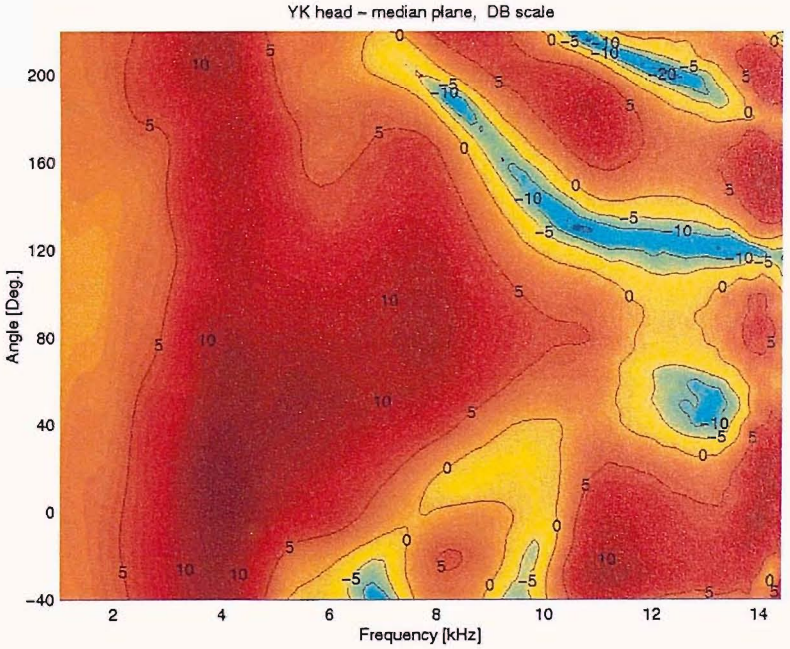


Figure 5-10: Spatial colour maps of HRTFs of KEMAR using the principle of reciprocity. In this case the response was calculated at 0.5 m away from the head. Maximum magnitude variations are given at each frequency (blue corresponds to minimum, and red for maximum). The colour corresponds to the magnitude of the pressure detected at the left ear due to a source at that position.

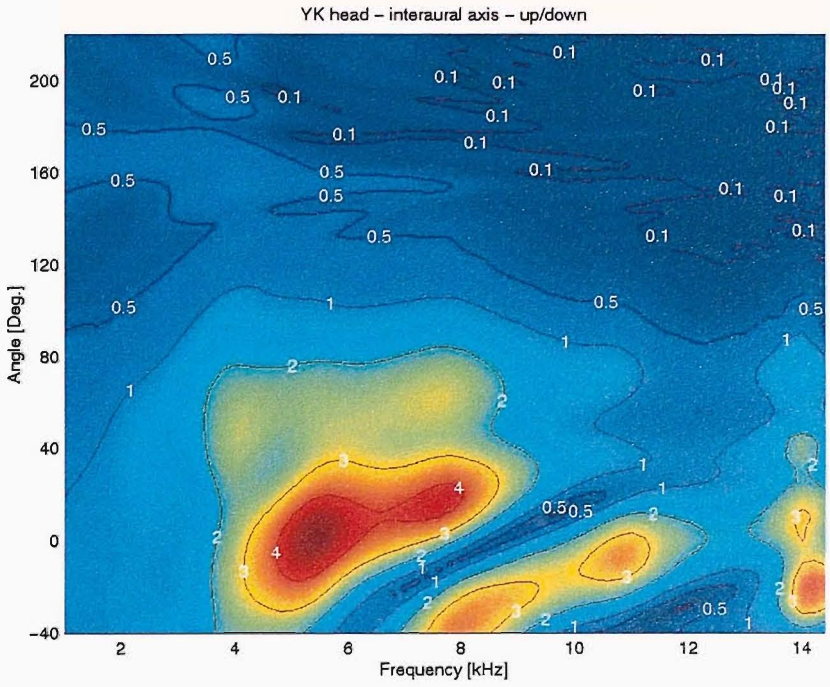


(a)

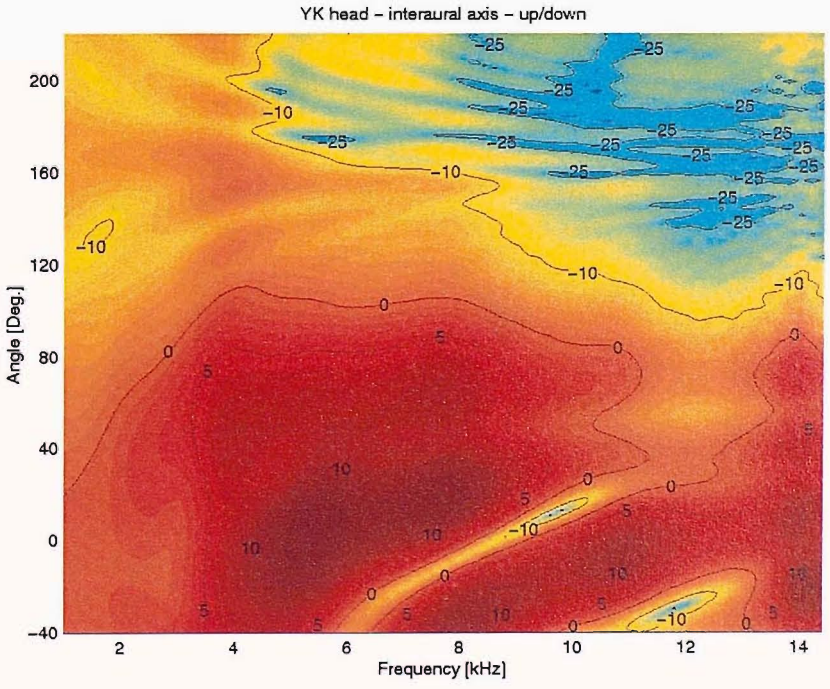


(b)

Figure 5-11: Normalised median vertical plane HRTFs for the YK head (the pressure detected at the blocked ear canal and divided by the pressure detected at the centre of the head and the head was absent). (a) linear amplitude (b) magnitude in dB. Simulation undertaken at a resolution of 1°, and steps of 200 Hz (71 frequencies).

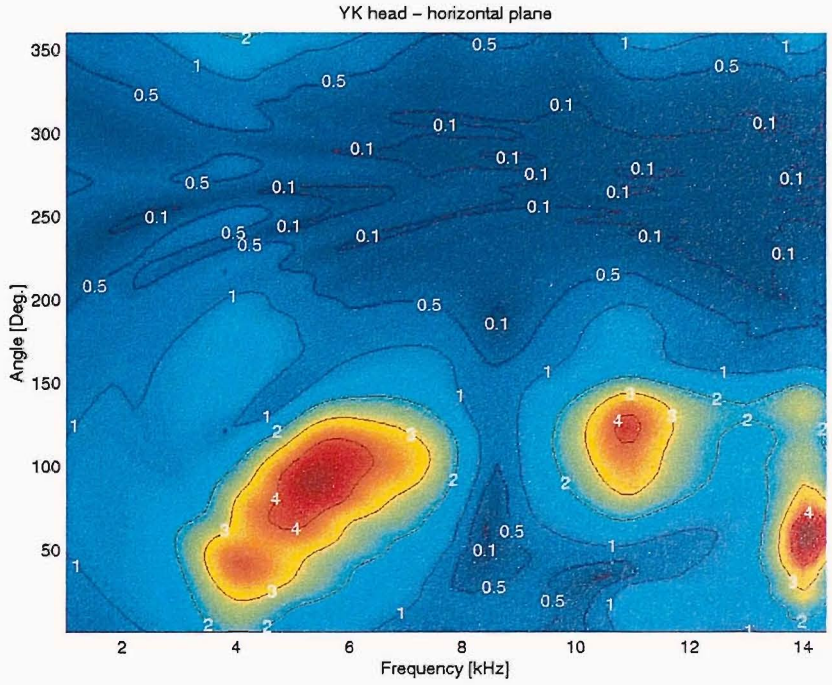


(a)

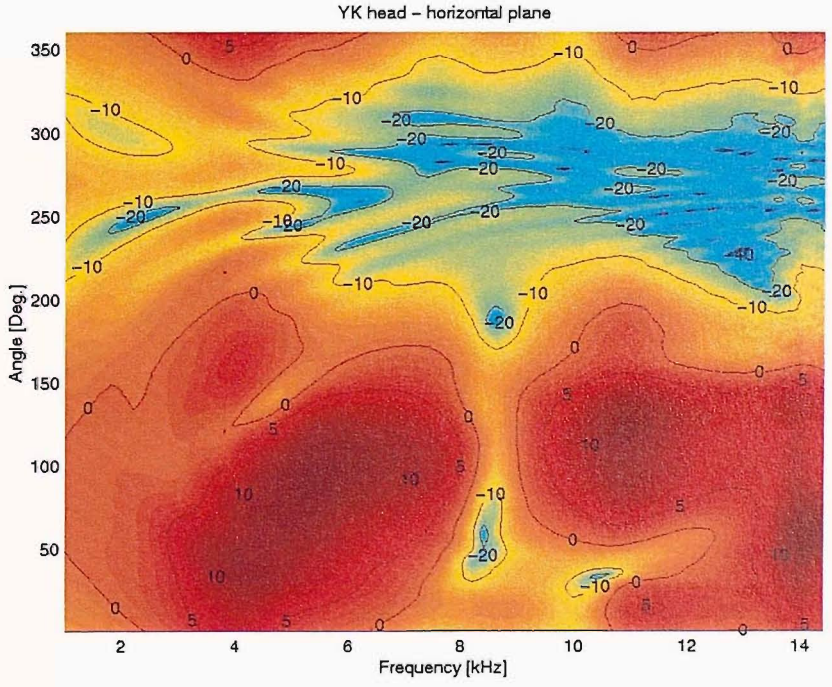


(b)

Figure 5-12: Normalised lateral vertical plane HRTFs of the YK head (a) linear amplitude (b) magnitude in dB. Parameters as in Figure 5-11.

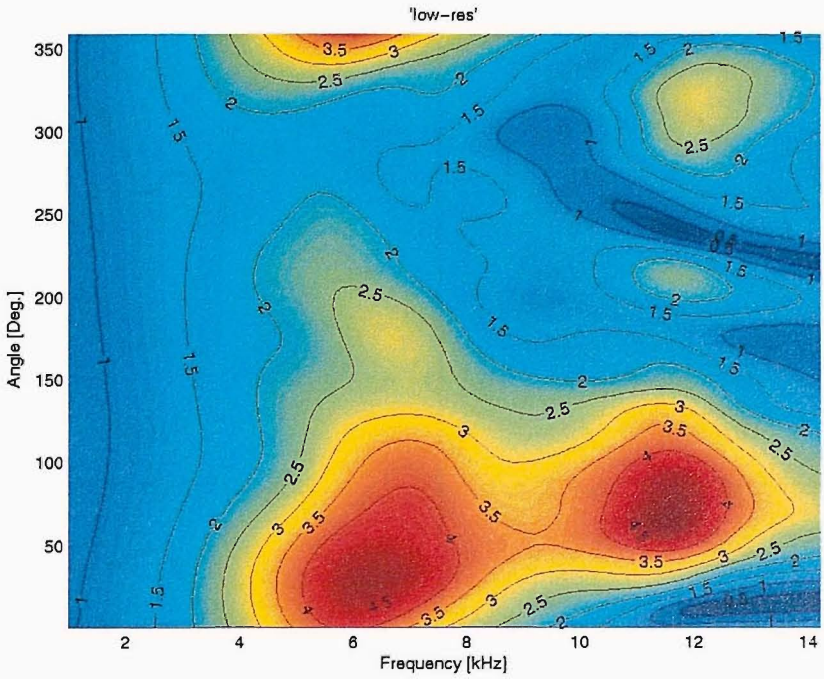


(a)

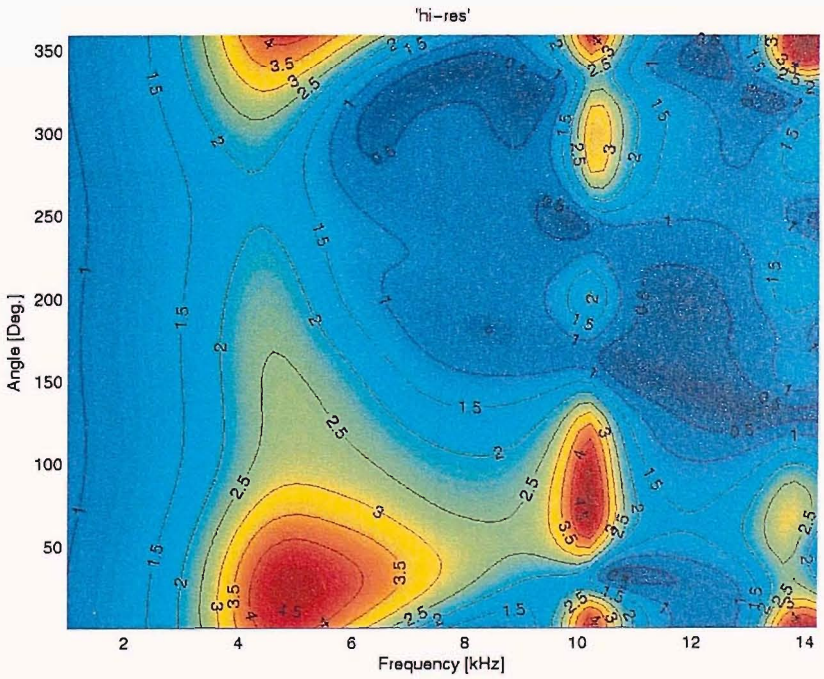


(b)

Figure 5-13: Normalised horizontal plane HRTFs of the YK head (a) linear amplitude (b) magnitude in dB. Parameters as in Figure 5-11.

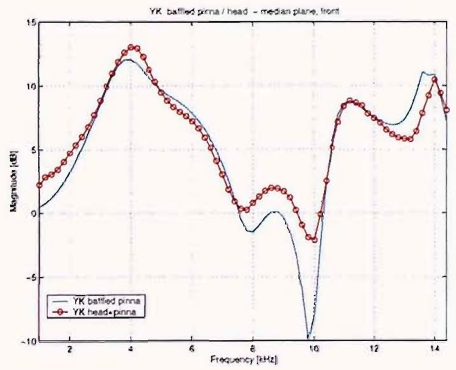


(a)

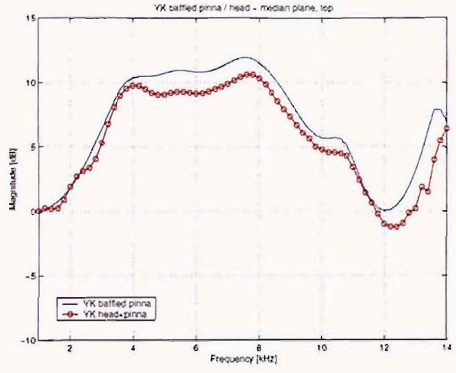


(b)

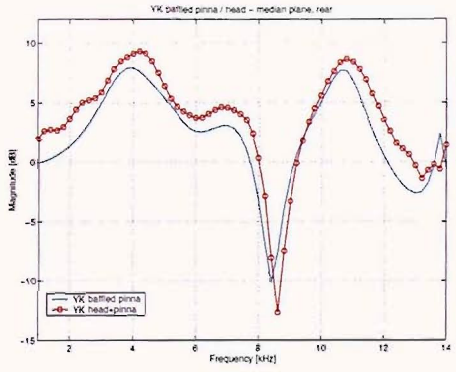
Figure 5-14: Comparison of the normalised frequency response of (a) 'low resolution' DB60 and (b) accurate DB60 at grazing incidence. Contours represent amplification, after the response was equalised with the response detected at the centre of the baffle without the pinna.



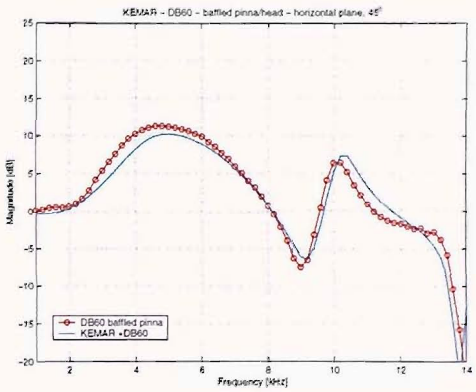
(a)



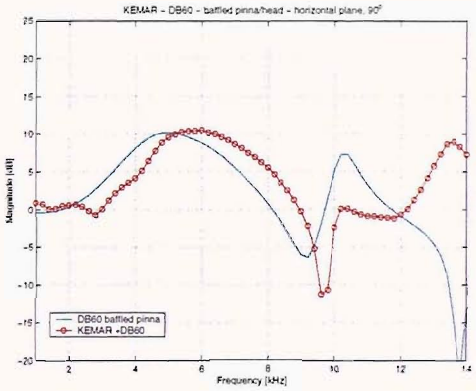
(b)



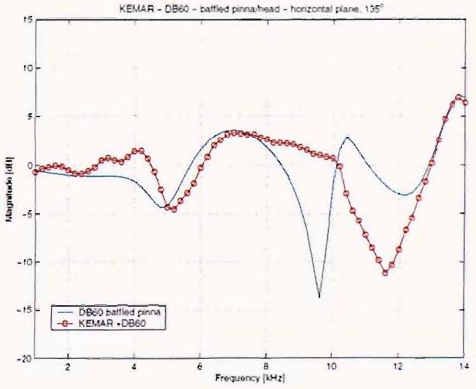
(c)



(d)

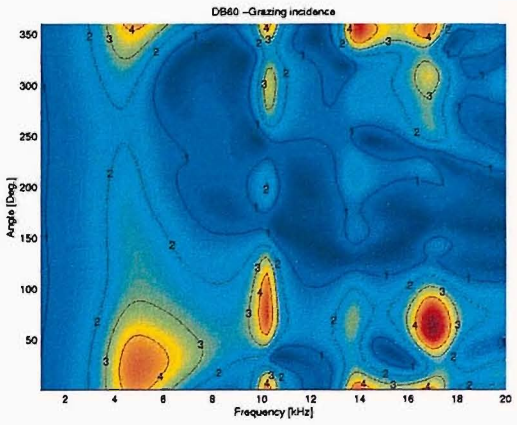


(e)

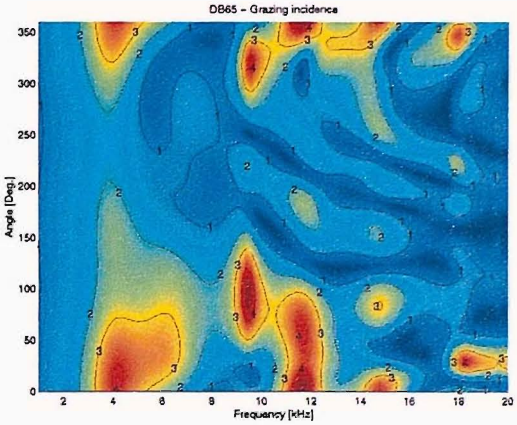


(f)

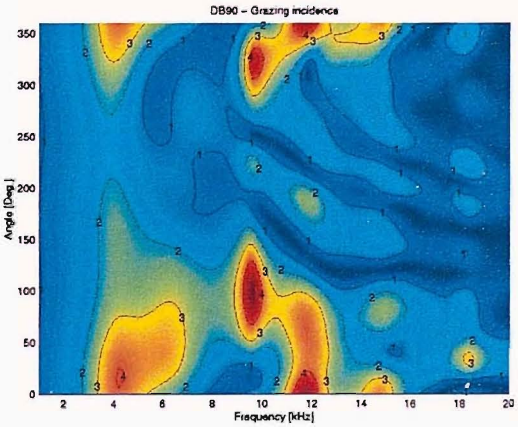
Figure 5-15: Comparison of the simulated normalised frequency response of baffled pinnae, and HRTF (head and pinnae) (a) YK median plane: front $\phi = 0°; \theta = 0°$ (b) YK median plane: above $\phi = 0°; \theta = 90°$ (c) YK median plane: rear $\phi = 0°; \theta = 180°$ (d) KEMAR horizontal plane: above $\phi = 45°; \theta = 0°$ (e) KEMAR horizontal plane: normal angle $\phi = 90°; \theta = 0°$ (f) KEMAR horizontal plane: rear $\phi = 135°; \theta = 0°$.



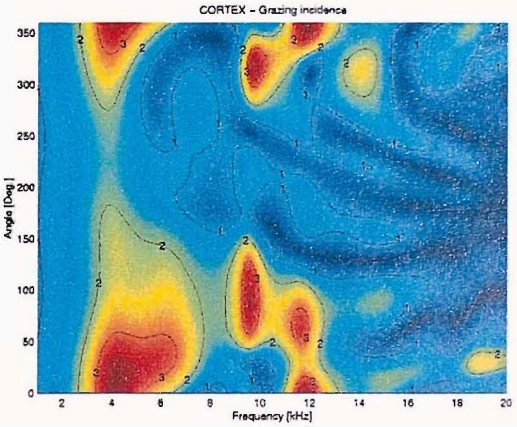
DB60



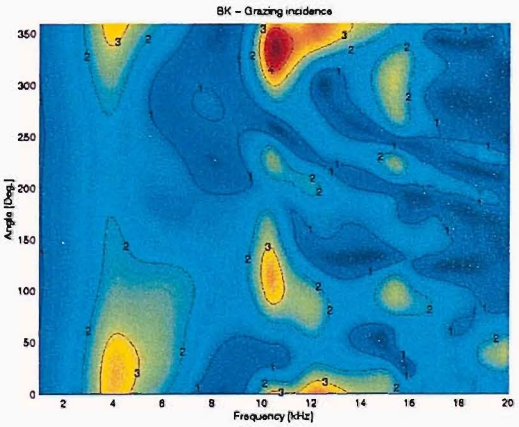
DB65



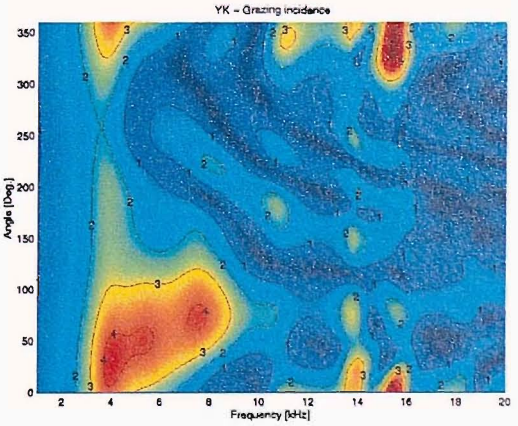
DB90



CORTEX

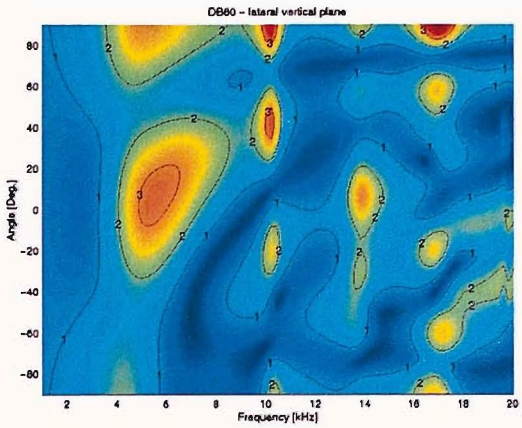


B&K

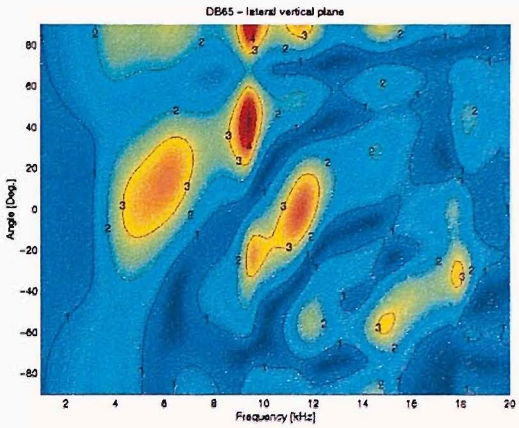


YK

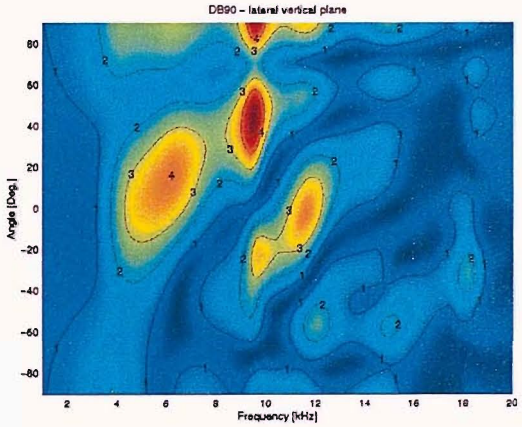
Figure 5-16: The normalised response of six baffled pinnae in grazing incidence at a resolution of 1° and steps of 200 Hz.



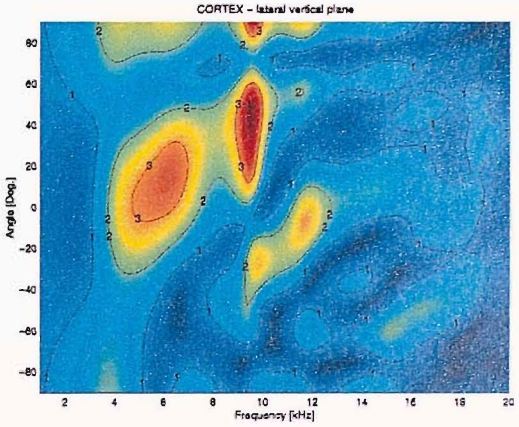
DB60



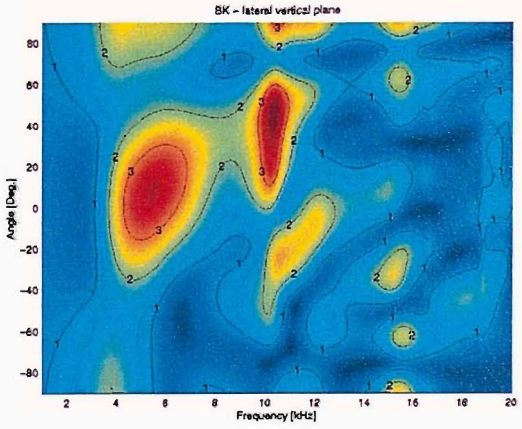
DB65



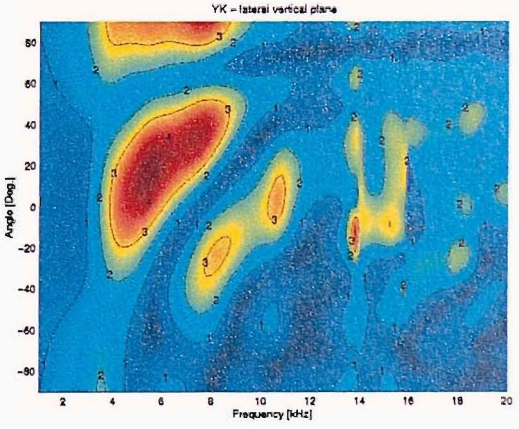
DB90



CORTEX

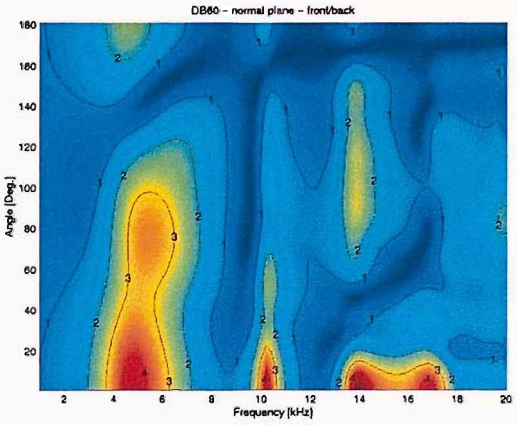


B&K

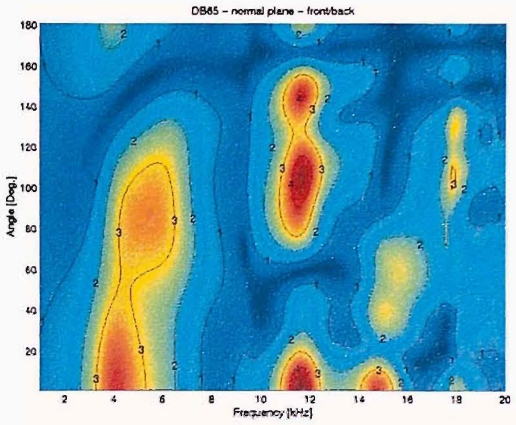


YK

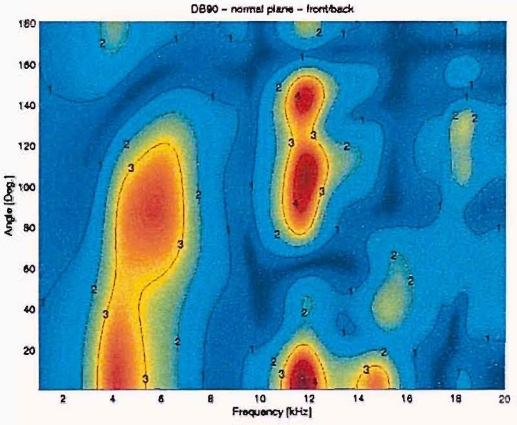
Figure 5-17: The normalised response of six baffled pinnae in the lateral vertical plane ($\phi = 90°, \ -90° \leq \theta \leq +90°$).



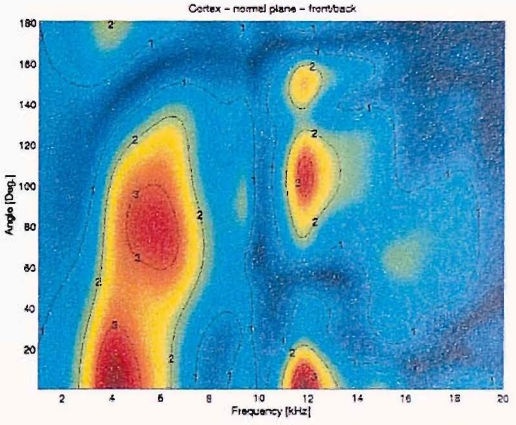
DB60



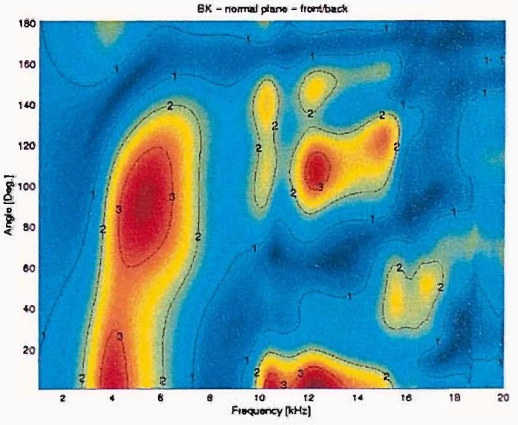
DB65



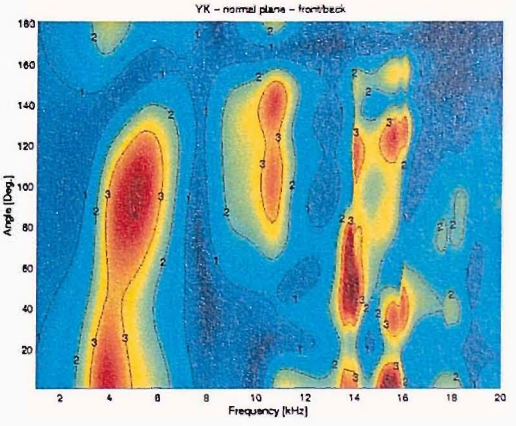
DB90



CORTEX

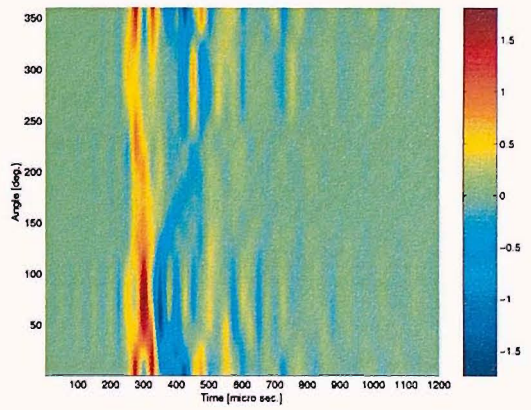


B&K

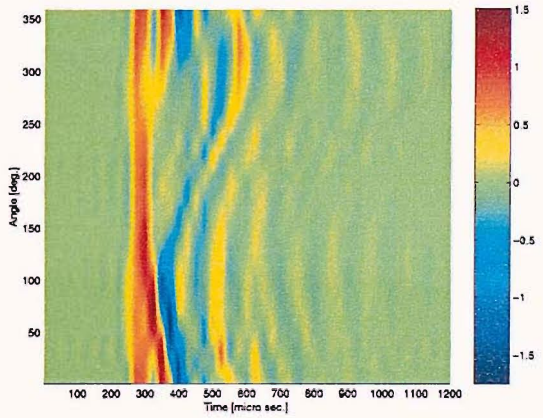


YK

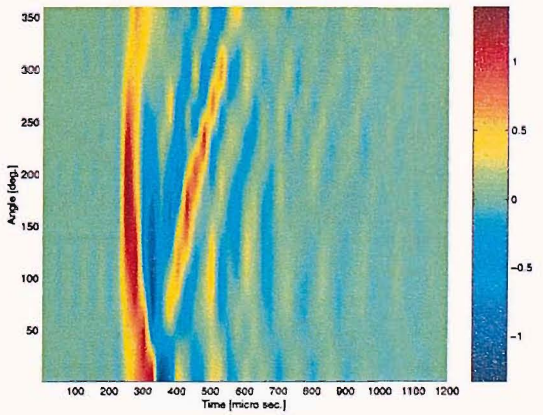
Figure 5-18: The normalised response of six baffled pinnae in the horizontal plane ($0° \leq \phi \leq 180°, \theta = 0°$).



(a)



(b)



(c)

Figure 5-19: Simulation of the impulse response of baffled pinnae (a) DB60 - grazing incidence (b) DB90 - grazing incidence (c) DB90 - lateral vertical plane. The Nyquist frequency is 20 kHz. Impulse responses were obtained by applying Inverse FFT to the responses presented in the previous figures, and applying linear phase correction at high frequencies.

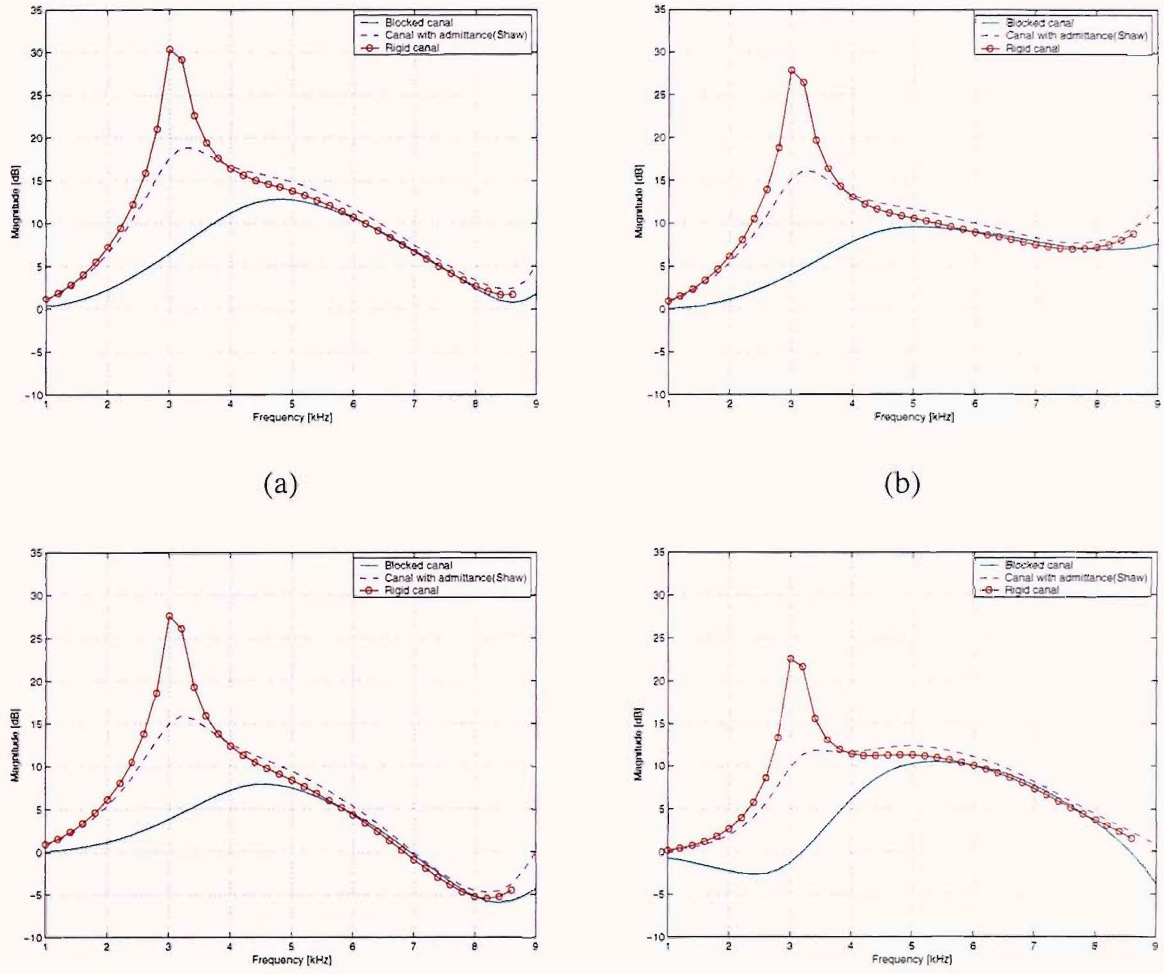


Figure 5-20: The response of DB60 in three conditions: Blocked ear canal, with rigid eardrum, and eardrum with averaged impedance boundary conditions by Shaw (1974). (a) front $\phi = 0°, \theta = 0°$ (b) above $\phi = 0°, \theta = 90°$ (c) normal $\phi = 90°, \theta = 0°$ (d) rear $\phi = 180°, \theta = 0°$.

CHAPTER 6

MEASUREMENTS OF THE RESPONSE OF THE EXTERNAL EAR AND THE HRTF

6.1 INTRODUCTION

Many publications have dealt with measurements and analysis of HRTFs in the last 50 years (For reviews of the most significant publications see Blauert, 1997, Møller *et al*, 1995, Hammershøi and Møller, 1996, and Carlile, 1996). The definitions, methods and equalisation techniques vary among the studies, and a technical standard for HRTF measurements does not exist. Higher accuracy and better signal-to-noise-ratio (SNR) are now obtained with the use of computers and improved stimuli, such as maximum-length sequences (Rife and Vanderkooy, 1987) or Golay codes (Zhou *et al*, 1992), compared with pure tones (used in the early days by Shaw and Teranishi, 1968 and Searle *et* al, 1975). In addition, the inclusion of video cameras and a turn-table (Møller *et al*, 1995), or an electro-magnetic head tracker (Middlebrooks *et al*, 1989), enables faster procedure with full control of the position of the source relative to the listener. However, the positioning of the microphone, either a condenser probe microphone (Shaw and Teranishi 1968, Blauert, 1997, Møller *et al*, 1995), or a miniature Electret microphone (Wightman and Kistler, 1989, Carlile and Pralong, 1994, and Møller *et al*, 1995) still remains problematic, as the ideal HRTF should be measured at the entrance to the auditory system, i.e. the eardrum, which cannot be used in practice for safety reasons. HRTFs described in the literature have been measured at four main positions along the ear canal: deep in the canal, in the middle of the canal, at the entrance to the canal, and at the blocked entrance of the ear canal. As described

in Chapter 5, we choose the latter option since it has been shown (Hammershøi and Møller, 1996) that measurements undertaken at the blocked entrance to the ear canal include the full spatial information encoded in HRTFs, and also since the geometrical model obtained with the laser scanner does not include any internal information such as the ear canal. However, the inclusion of average ear canal geometry and average eardrum admittance values should not impose great difficulties, as demonstrated in Section 5.5.5.

The goal of the current set of measurements is mainly to validate the results obtained with the simulations. As presented in the previous chapter, we concentrate on two different cases: (a) the simulation of HRTFs of the head and pinnae, without the torso, and (b) the response of blocked meatus pinnae, attached to an 'infinite' baffle.

In view of the high sensitivity of the measured and simulated response at high frequencies both due to source location and individual geometry of the pinna, it was of utmost importance to control parameters such as the geometry of the pinnae and their orientation, and the position of the transducers, and make them as similar as possible in both physical and simulation environments. Other secondary factors such as the speed of sound, or 'noise' generated either electronically or acoustically in the measurement procedure, or numerically in the simulation procedure were observed and analysed.

6.2 METHOD

Measurements were carried out in the large anechoic chamber of the ISVR. Only an artificial head (KEMAR) and artificial pinnae were used, to avoid alignment problems encountered with subjects. In addition, non rigid boundary conditions such as hair, clothing, eardrum impedance, and contribution to the total pressure due to torso, shoulders and knees were avoided in both simulations and measurements. Measurements were repeated four

times, on different days (with a time gap up to six months), to ensure high repeatability with measurements.

6.2.1 Motorised rotating arc

Two arcs were designed and built: a smaller one with a radius of 1.15 m and a larger one with a radius of 1.6 m. Since generating HRTFs for binaural synthesis is *not* our main goal in this set of measurements, we used the small arc since its assembly is simpler and faster. The resulting distance between the cone of the loudspeaker and the microphone positioned at the entrance to the blocked ear canal of the pinna is exactly 1 m. 19 loudspeakers are mounted at an equal spacing of 10° between each other (see Figure 6-1 and Figure 6-2). It is estimated that the positional error due to the radius of the arc and its curvature is less than 1 cm, and due to the spacing between the speakers is less than 0.5°.

The rotation of the arc is achieved by a transmission chain between the arc and a step motor. When the system was calibrated it was found that every pulse to the step motor would rotate the arc by 1/10932 of 1°. (i.e. in order to rotate the arc by 180°, 1456789 pulses are required). The rotation of the arc in this case takes only 4 minutes.

The effect of reflections due to adjacent loudspeakers mounted on the arc was investigated by covering all loudspeakers, except the one used for the measurement, with absorbent material, and also when the same loudspeaker was used for measurement of baffled pinnae in the anechoic chamber without the arc and all the loudspeakers. Differences at frequencies below 10 kHz were less than 1 dB, and variations up to 2 dB were noticed at higher frequencies.

6.2.2 The baffle

A large baffle was constructed out of thick plywood. Its large dimensions (with a maximum width of 2.3 m, and a maximum height of 2.4 m, see Figure 6-2) ensured that the

measurements of the acoustical response of the pinna are not distorted due to diffraction and reflection from the boundaries of the baffle. In the centre of the baffle a circular aluminium plate was positioned, able to rotate around its centre, and a rectangular opening enables the positioning of the artificial pinnae (see Figure 6-3a). The response of human pinnae can also be measured when subjects sit behind the baffle but these are not included in this study. An additional circular plate included an opening for the holding Electret microphone, flush mounted with the baffle plane (see Figure 6-3b). This was used for equalisation of the transducers for each source position in space.

6.2.3 MLSSA system

The Maximum Length Sequence System Analyzer (MLSSA) has been used widely in recent years in HRTF measurements. The MLS method offers a number of advantages compared to traditional frequency and time domain techniques. A detailed review of the MLS method is given by Rife and Vanderkooy (1989). With the accurate positioning capability of the transducers with respect to the head/pinnae in our measurement set-up, a single channel measurement system is sufficient. Consecutive measurements with all loudspeakers and two microphones (in the case of KEMAR) were used to generate a matrix of impulse responses capturing the space with a dense resolution.

6.2.4 Transducers and amplifiers

A miniature microphone, Sennheiser KE 4-211-2 was used. This microphone has been used previously for HRTF measurement (Møller *et al*, 1995) and its main advantages are the cylindrical shape with small dimensions (diameter of 4.75 mm and height of 4.2 mm), and fairly flat frequency response (for its typical frequency response see Møller *et al*, 1995, Figure 4).

The loudspeakers (by Fostex, Inc.) are based on a single unit-drive with a diameter of 70 mm, which are designed to produce a reasonably flat response (±2 dB between 1 kHz and 15 kHz). Since the main goal of the measurement was to capture the contributions of the pinna at high frequencies, and the pressure changes only slightly below 1 kHz, it was not required to pay additional attention to low frequencies in the design of the loudspeaker. Nevertheless, reliable responses were obtained down to 300 Hz.

An Electret microphone amplifier that was built at the ISVR was used and its magnitude response is ±0.4 dB over the entire frequency range. The input signals to the loudspeakers were amplified by a single channel of a Yamaha P2160 power amplifier.

6.2.5 Electronic switching box and software

The activation of each of the speakers mounted on the arc was undertaken through an electronic multiplexer switching box. The communication with the MLSSA system was carried out using a single pulse coding system that is controlled through the macro command language of the MLSSA system. The operation of the multiplexer is described as follows: the circuit is controlled by 10 msec. pulses from MLSSA. As each pulse is received, the circuit steps to the next function. A switch is provided to connect one, two or four microphones to the MLSSA analogue input. The MLSSA analogue output is fed via a power amplifier to one of up to 36 loudspeakers. On receipt of a pulse, the circuit advances to the next selected microphone. A pulse detector monitors the pulse rate. If it fails to detect a pulse for a nominal 5 seconds, it interprets this as the end of the set of measurements at that angle. The circuit resets to the first microphone and the first loudspeaker and a pulse is sent to the motor controller to adjust the angle of the arc. When the motor controller has completed its programmed operation, it sends a pulse back to the circuit and on to MLSSA to proceed with the acquisition program. All logic functions are performed by CMOS gates and counters. The microphones are switched by an analogue multiplexer integrated circuit.

The loudspeakers are switched by optically coupled MOS solid-state relays for silent operation and low cross talk. LEDs are connected in series with each drive to indicate which loudspeaker is in operation. Level shifting is required between the 5V CMOS logic circuits and the 24V logic used by the motor controller. All software commands (Parker Hanifin, Inc) were run in parallel to the MLSSA system (version 10W) in the Microsoft Windows environment.

6.3 INDIVIDUALISED HRTF MEASUREMENT

6.3.1 KEMAR head

The head of KEMAR with the 'small' pinna – DB60 was used to validate the HRTF of a 'full head'. The response of YK head was measured, but it is not included here, since the contribution of the torso and other errors prevent comparison directly of the response with the simulations (which do not include the torso and the model is rigid).

Since the torso was absent in the simulation process, a special attachment was used (see Figure 6-4a) to ensure that the centre of the rotation axis of the head was at the middle of the interaural axis. The main discrepancy between the geometry of the head in the simulation and measurement was the extension of the neck (see Figure 6-4b). However, we assume that it does not contribute significantly to the variations between simulation and measurements. Two different types of DB60 pinnae were investigated (one 10 years old, and the other new) and these produced slight differences in performance at high frequencies (it should be noted that the flexibility of the pinnae was very different and fine details were also different). The old DB60 pinna that was scanned was used in both simulation and measurements.

By using a rigid head without a torso, we minimise the causes for deviations between simulation and measurements, where the geometry of the pinna plays the most important role.

6.3.2 Measurement set-up and equalisation

The MLSSA system was set with the following parameters:

- The sampling frequency was 48.193 kHz. This is the nearest sampling frequency to 48 kHz the internal clock of MLSSA can provide*. Since the purpose of these measurements was not for binaural synthesis, our main consideration here is to ensure that the Nyquist frequency is above 20 kHz.

- The dynamic range of the system was set to 'on' to ensure the approximate 65 dB SNR of KEMAR (Gardner and Martin, 1995) is available. The binary file structure of MLSSA (*.tim) includes the impulse response data and the amplification used for each measurement, thus enabling the obtaining of the absolute response. This is crucial, especially when the response is measured in the contralateral ear, and the SNR in this case is much lower than for measurements taken in the ipsilateral ear.

- The free-field sound pressure level was approximately 75 dB(A) at a distance of 1 m. A higher level was not required to improve SNR, and it might have caused non-linear effects in the measurements.

- The highest order of MLS with MLSSA was used (14) which is sufficient for our measurements. This sequence results in 65535 points. A length of 4096 points was much more than was required for both measurement arrangements, and the data was reduced and windowed in a post-processing procedure.

* Ideally, HRTF data should be measured using a sampling frequency of 160 kHz; this results in resolution of 6 μsec, which is the noticeable difference of ITD when localizing low frequency sounds.

- To avoid frequency aliasing a 20 kHz Chebyshev low-pass filter in the MLSSA hardware was used.
- The measurements were undertaken using 16 pre-averages that ensured the highest SNR.
- The temperature in the anechoic chamber was recorded and varied between 16.8 °C and 19.3 °C throughout the measurements. These result in variation of the speed of sound between 340 m/sec and 342 m/sec (all simulations with SYSNOISE were undertaken with $c = 340$ m/sec).
- The alignment of source positions in space was achieved by the following steps:
    - Alignment with a spirit level that was attached to the arc. The fine-tuning was achieved by sending only a few tens of pulses to the step motor.
    - Impulse response measurements were undertaken at the maximum sampling frequency, 160 kHz. For example, to ensure that the head was positioned exactly between the right and the left loudspeakers (No. 1 and No. 19, respectively), the two impulse responses should result in equal delays. In a similar way, the loudspeaker on the median plane (No. 10) should produce similar delays in both ears of the KEMAR head.
    - The measured response was immediately equalised and compared with the simulation results. Variations of angular positions of the measurements up to 4° were found better to match simulation results due to slight misalignment between the co-ordinate systems of the measurements and simulations.

6.3.3 Results

Free-field HRTF measurements require the equalisation of the transducers and the microphone at the centre of the head when the head is absent. A few measurements were undertaken for each of the 19 loudspeakers, all positioned in the horizontal plane ($\phi = 0°$). Since the microphone is not completely omnidirectional, this method of equalisation still results in some errors.

The results for measurements in the median plane are presented in Figure 6-5a with a continuous 3-D map of the simulation in the range of $-40° \leq \theta \leq 220°$ and in Figures 6-6a to 6-6e with 2-D plots of the magnitude at a few angles. All general peaks and notches appear in both simulations and measurements. Contours are presented on a linear scale as amplification with the reference of the free-field response with no interpolations applied. The main characteristic of this figure is a smoother variation in the simulation results compared to measurements. The same simulated and measured responses are shown at discrete angles in Figure 6-6 at $\phi = 0°, \theta = -40°, 0°, 40°, 90°, 130°, 180°$. Figures 6-6a, 6-6b and 6-6d show that errors of <1 dB appear up to 15 kHz. At other angles where sharper notches are found (Figures 6-6c, 6-6e and 6-6f) errors increase locally at the position of the notch. It is not clear if the discrepancies are due to errors in measurements, such as positioning or SNR, or due to a higher accuracy of the simulation techniques. However, it is concluded that simulation techniques of HRTFs can be used up to 15 kHz. It should be noted that median plane simulation and measurement do not impose great difficulty from the SNR point of view, since there are no effects of the shadowing due to the head, which have strong attenuation at high frequencies.

Figure 6-7 compares the simulation and measurement in the worst case from the point of view of SNR, since it presents the ILD (the difference between the magnitude of the right ear and the left ear in dB). As the frequency increases the sensitivity of the modelling and

measurement of the contralateral ear is shown. The contour maps show the ILD of the HRTFs of KEMAR at the lateral vertical plane, at angles of $\phi = 90°, -40° \leq \theta \leq 90°$. At $\theta = 90°$ (above the head) the ILD is 0 dB over the entire frequency range (we did not measure in this case the response of the DB61 pinna, which is the left ear of KEMAR and produces a different response at high frequencies than the DB60 pinna. In addition we assume a complete symmetry, in both the head and the pinna). At low frequency (in the region of 1 kHz) the ILD is on average 6 dB and increases with frequency and with angles approaching the horizontal plane and below. In this region a boost of approximately 30 dB is noticed between 5 and 8 kHz. This is mainly due to the broad peak of concha resonance and increasing attenuation with frequency in the contralateral ear. In addition, a sharp notch is noticed and increases with frequency (from 7 kHz to 13 kHz) as elevation angle increases (in the range of $-40° \leq \theta \leq +20°$). Good agreement between simulation and measurement is noticed between 90° and 40° up to 15 kHz. Our simulation model seems not to produce continuous variation above 10 kHz as expected and found in the measurements. The inaccuracy of the model can be explained by the low resolution of the mesh, assuming only four elements per wavelength. In all our simulation results presented in Chapter 4 and 5, at least six elements per wavelength are required for a reasonable accuracy in the shadow zone. However, the improvement of the mesh model and the whole procedure of simulation are straightforward, if more memory (RAM) is available for solving the problem with the IBEM symmetry, with the in-core solver.

However, all the peaks to the right of this notch have poor reliability in our simulation. This is due to the following reasons.

- The complex structure of peaks and notches at high frequency is a result of shadowing and pinna resonance. A much higher mesh resolution is required for

obtaining accurate results (in this case assuming six elements per wavelength the maximum frequency that can be investigated reliably is around 10 kHz).

- Results are very sensitive to source position and geometry. Even the extension of the neck can shift or alter the anti-resonance in the contralateral ear.

6.4 BAFFLED PINNAE MEASUREMENT

The equalisation of the measurements of baffled pinnae was undertaken by repeating all measurements twice: once when the microphone is positioned at the blocked entrance to the ear canal, and a second time when the microphone is positioned at the centre of the baffle without the pinna. The responses are then divided in the frequency domain. A typical pair of measured frequency responses shown in Figure 6-8 (in this case of the YK pinna, with excitation from normal direction at $\phi = 90°$). The response in both cases is similar up to approximately 2 kHz, the frequency at which the pinna starts to produce its broad resonance due to the quarter wavelength of the concha. At higher frequencies the peaks and notches are mainly affected by the resonance and anti-resonance of the pinna. All measurements presented below are after the equalisation of the transducers' response. The same principle applied with simulation results but in this case the transducers' frequency response are constant over the frequency range and only the phase changes with frequency due to the inherent delay which is a result of the distance between the positions of the source and the baffle (1 m).

In the simulation process the edges of the pinnae were smoothed to the plane of the baffle (z-plane). However, both pinnae and the KEMAR head do not have flat edges. Therefore, the pinnae were supported in the slot to produce smooth continuation with the baffle. This resulted in a very similar (although not exact) geometry to that which was used in the simulations.

Measurements of six pinnae (four of KEMAR, one of Brüel and Kjær 4127 Head and Torso Simulator (HATS), and a replica of YK pinna) were undertaken and repeated a few times to ensure repeatability. First, measurements were carried out before the rotating arc assembly was assembled, and only a single loudspeaker was used. The loudspeaker was positioned manually on a stand in the horizontal plane on a radius of 1 m, at 19 positions, every 10°. The aluminium plate was rotated manually around its axis 18 times, every 10°, to produce the full hemisphere response. The main advantage of this method is that it includes only a single loudspeaker, and therefore the equalisation is fast and accurate. However, the main disadvantage is that the whole process is manual and time consuming. Later, all measurements were run automatically where a full scan of the entire sphere in a resolution of 2° took approximately one hour.

A large amount of data was collected for all six pinnae, with frequency response measured at high resolution sampling of the hemisphere. The figures shown include only a few quantitative comparisons at single angles, and also a single continuous map which provides an insight to the trends of the variation of peaks and notches at a resolution of 1° in the lateral vertical plane.

6.4.1 Measurement set-up and equalisation

- The frequency response was measured with a few sampling frequencies: 48.193 kHz, 60.60 kHz, 75 kHz and 160 kHz. Although the high sampling frequencies include a higher resolution of the captured impulse responses, the frequency response remained the same since all measurements were undertaken with the antialiasing Chebyshev filter at 20 kHz.

- The response at each location of loudspeaker was equalised with the microphone positioned in the centre of the baffle without the pinna. Due to the high accuracy

achieved with the positioning of the arc in both measurements, it was possible to ensure that most significant imperfections of the rig would be cancelled.

6.4.2 Results

Figure 6-9 shows a comparison between the simulated and measured response of DB60. The measurements were undertaken with a single loudspeaker (without the arc). A systematic error appears at around 2 kHz that is due to a poor response of the specific loudspeaker used in this set of measurements. Nevertheless all peaks and notches are replicated by simulation. Note especially the high accuracy of the prediction of the notch at 9.5 kHz for grazing incidence of $\theta = 120^{\circ}$ in Figure 6-9c. At other angles the errors are within 2 dB up to 15 kHz with maximum errors appearing at deep notches either due to magnitude difference or slight shift of resonance frequency. All the following comparisons include measurements undertaken with the aid of the rotating arc. Figure 6-10 shows a comparison between the simulated and measured response of DB65. An almost perfect match is obtained up to 10 kHz. Accuracy remains high for higher frequencies but it depends on the angle of excitation. The highest errors are achieved as before where deep notches occur. These are very sensitive to source positions.

Figure 6-11 shows a comparison between the simulated and measured response of the B&K pinna. As before, errors are within a range of 2 dB up to 13-15 kHz. It should be noted that the variability in repeated measurements with this frequency range remained low (up to 2 dB) although variations were as high as 5 dB above 15 kHz.

Finally we present the response of pinnae in which the identical positioning of the microphone in both simulations and measurements is not possible as in the previous cases due to irregular curvature of the opening at the entrance to the ear canal. Figure 6-12 shows a comparison between the simulated and measured response of DB90, Figure 6-13 shows a

comparison between the simulated and measured response of the YK pinna, and Figure 6-14 shows a comparison between the simulated and measured response of DB95. In all cases a good agreement is obtained only up to 7 kHz, but the general trends are followed up to a higher frequency. The worst case was for the DB95 with a sharp notch obtained with simulation between 7 kHz and 11 kHz. It should be noted that this pinna has a large hole at the entrance to the ear canal that is dedicated for ear moulds, and therefore the model is not accurate for our purposes of both measurements and simulations.

Figure 6-15 shows an example of a spatial map of continuous variation of spectral peaks and notches of the baffled DB60 pinna in the lateral vertical plane. Both simulations and measurements were undertaken at a resolution of 1°. The contours show amplification in steps of 1, 2, 3 and 4 times the response detected at the equalisation point at the centre of the baffle (without the pinna). It is interesting to observe how the general trends are replicated for all angles and all frequencies albeit with amplitude differences at the higher frequency range.

In particular, discrepancies were found at the amplitude of peaks at frequencies higher than the first quarter wavelength. These differences are likely to be as a result of geometry differences and sensitivity due to the proximity of the sources to the baffle. These 'boosted' peaks were not found, for example, when the pinnae were modelled with the full head.

6.5 CONCLUSIONS

The numerical modelling of the response of the external ear was validated with a high level of accuracy. Two cases were studied: the free-field equalised response of HRTFs and the response of baffled pinnae. The accuracy is significantly degraded for modelling the response of the contralateral ear where low mesh resolution is available. With all

optimisation of numerical techniques and mesh models, reliable results can be achieved up to 10 kHz.

The responses of baffled pinnae show many of the spectral characteristics found in the response of the ipsilateral HRTFs. The addition of the baffle does not alter the performance at grazing incidence, and in other angles the general trend is preserved. The response is very sensitive to the geometry and source positions in space, and high agreement was found with pinnae for which the position of the microphone was accurate. In the case of DB90, DB95 and YK pinna, the positioning of the microphone was not as accurate as with the other pinnae due to irregular shape at the entrance to the ear canal, and as a result a slightly higher deviation between simulation and measurement was obtained.

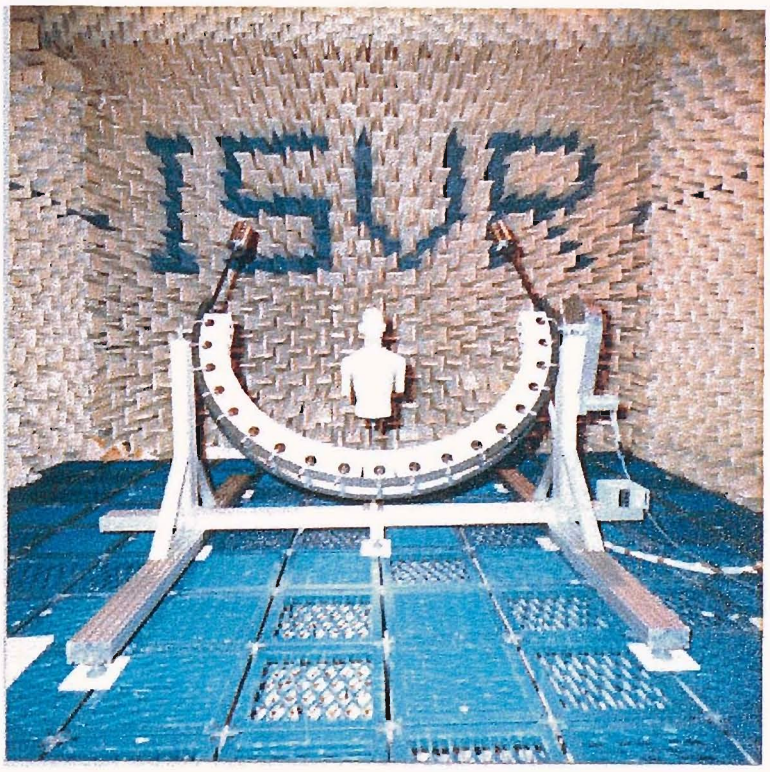


Figure 6-1: Apparatus for HRTF measurement (the measurement results presented in this chapter are based on KEMAR without torso, as shown on Figure 6-4a).

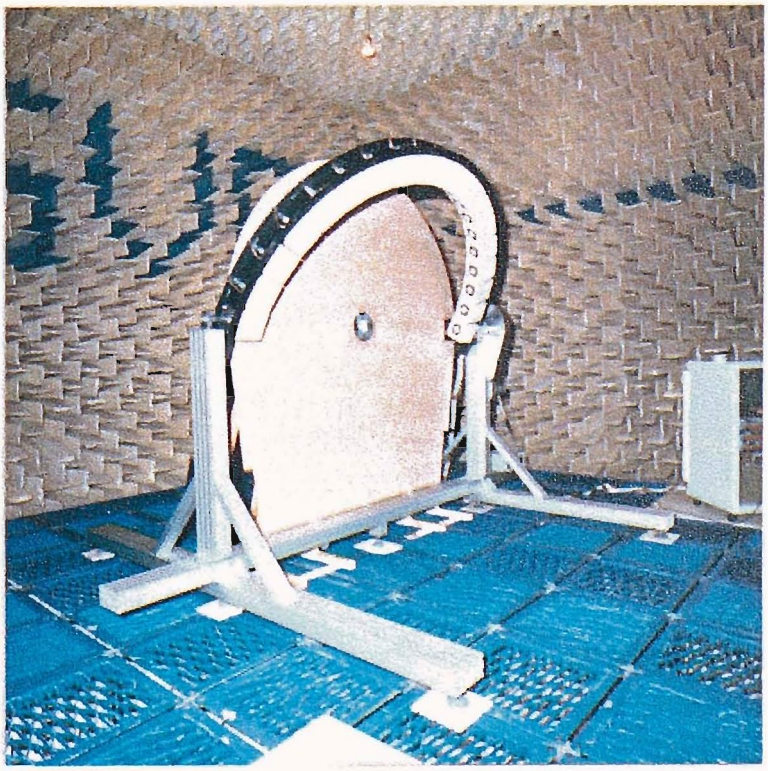
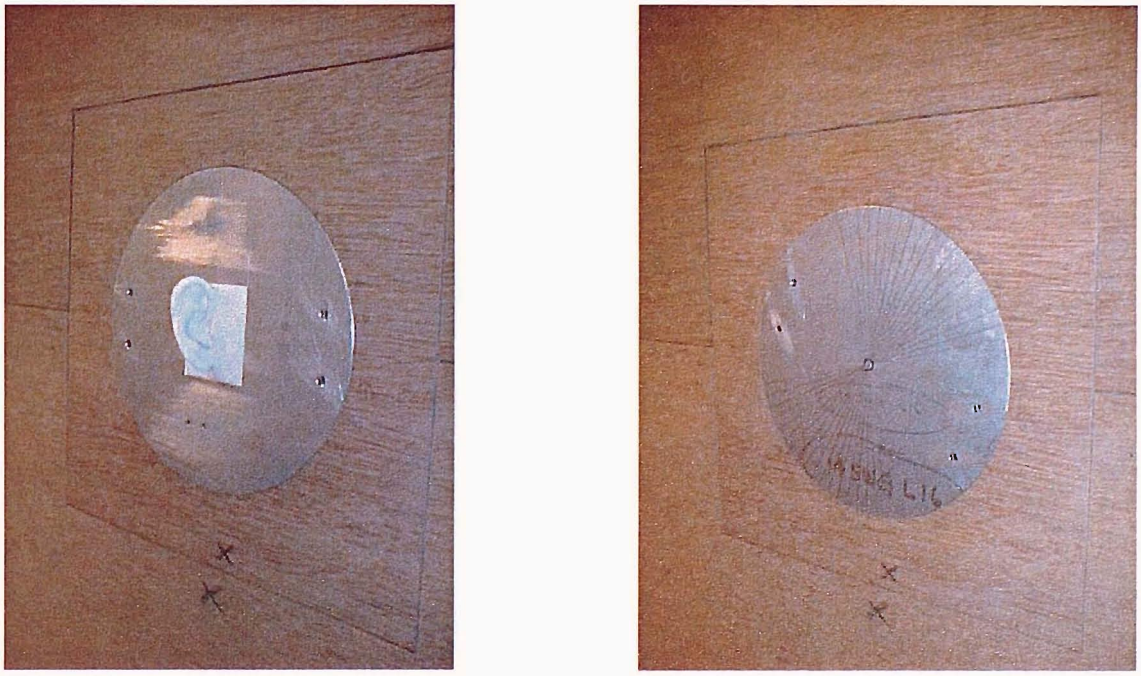
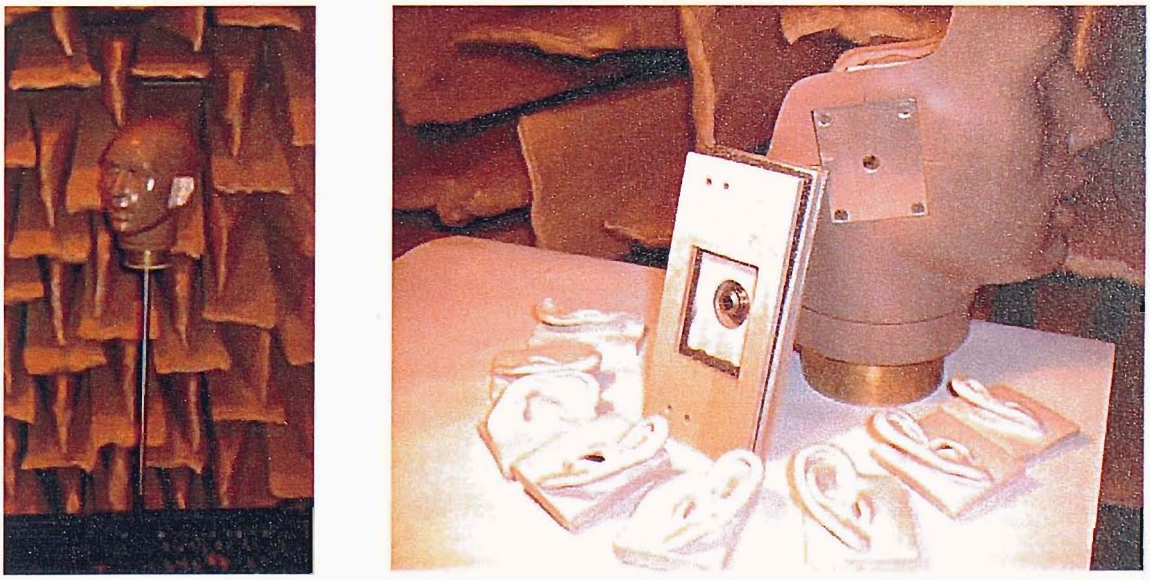


Figure 6-2: Apparatus for the measurement of the response of baffled pinnae.



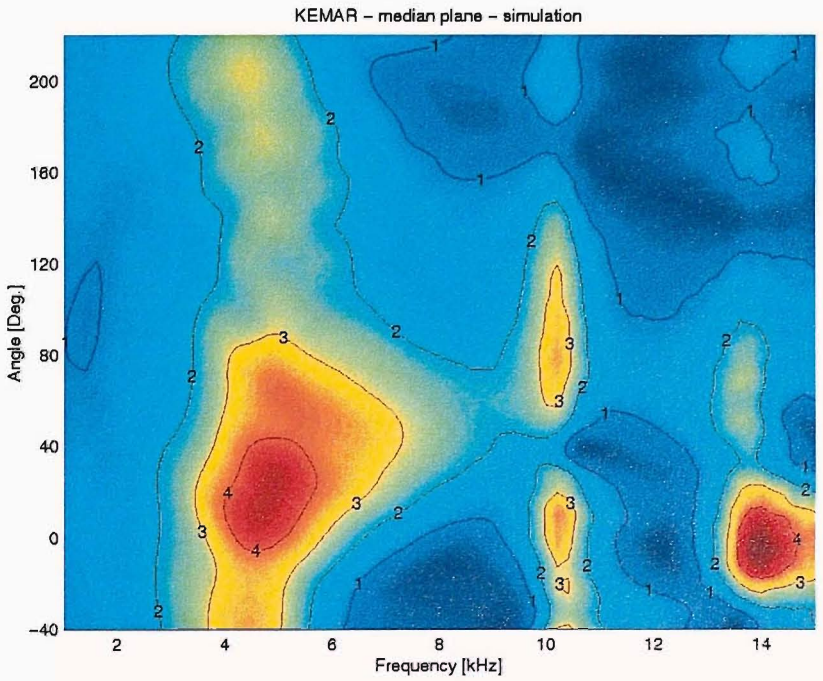
(a) (b)

Figure 6-3: An aluminium rotating plate at the centre of the baffle used for (a) attaching replaceable rubber and plaster pinnae (b) attaching Electret microphone for its calibration (Sennheiser KE 4-212-2).

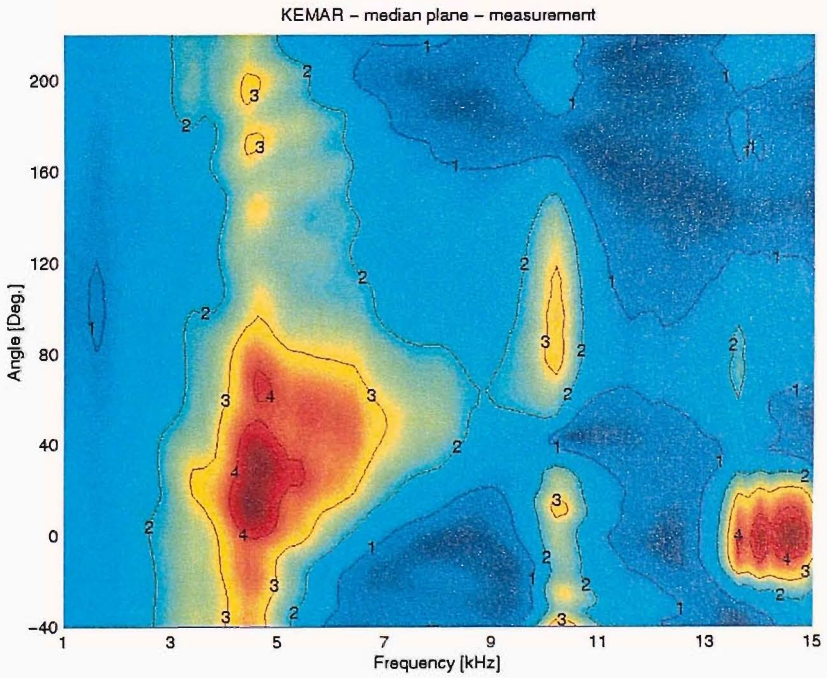


(a) (b)

Figure 6-4: Special modification and adjustments for measurements (a) attachment for KEMAR (without torso) to a turn-table or the bottom frame of the arc (b) head and baffle attachments and the various pinnae investigated.

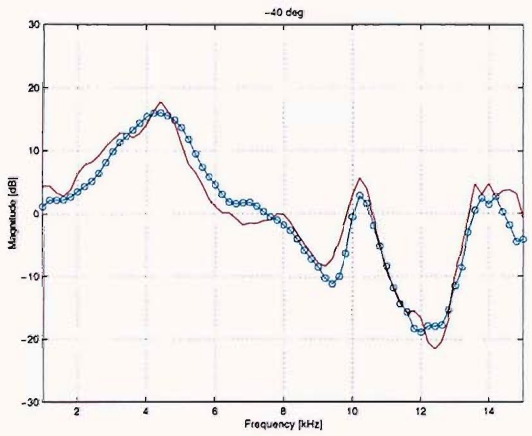


(a)

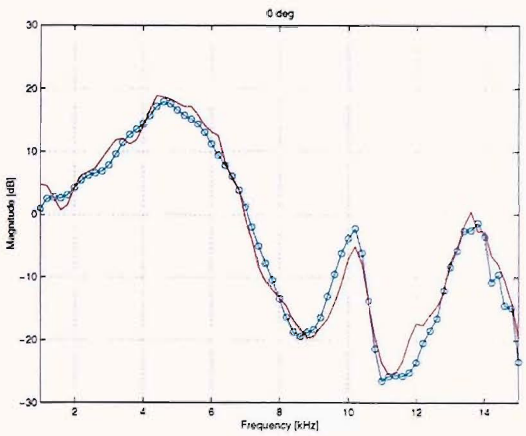


(b)

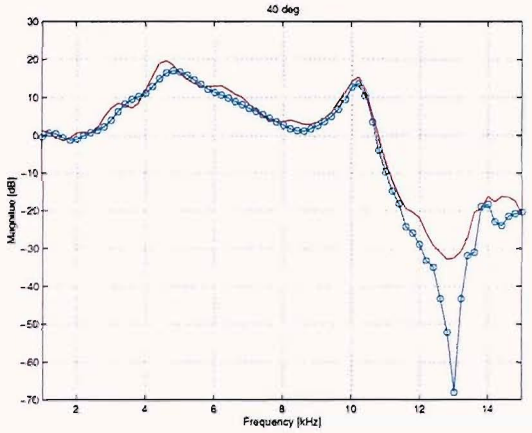
Figure 6-5: Comparison of the HRTFs of KEMAR on the median plane (a) simulations (b) measurements. In both cases the spatial resolution is 1° and no interpolations are made.



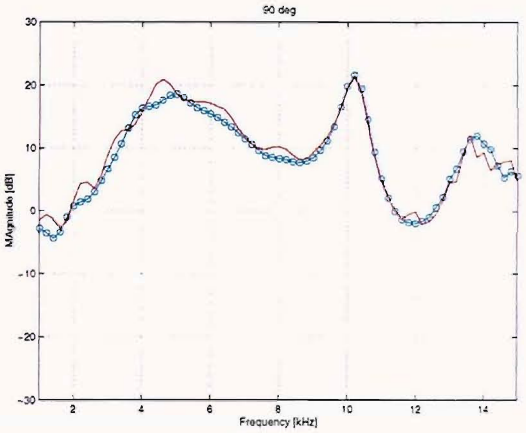
(a)



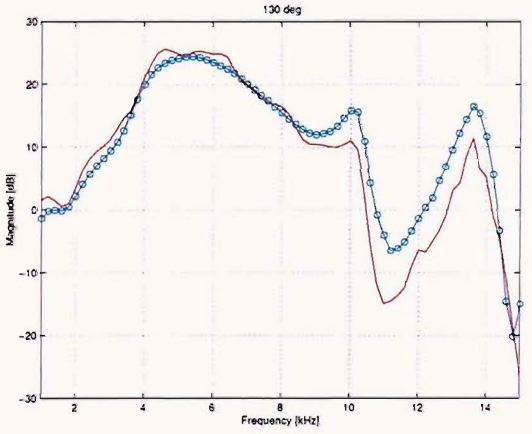
(b)



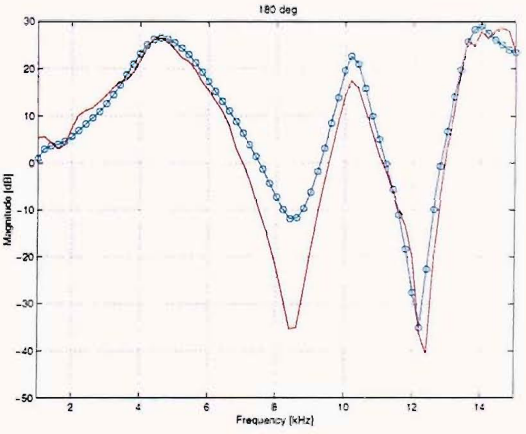
(c)



(d)

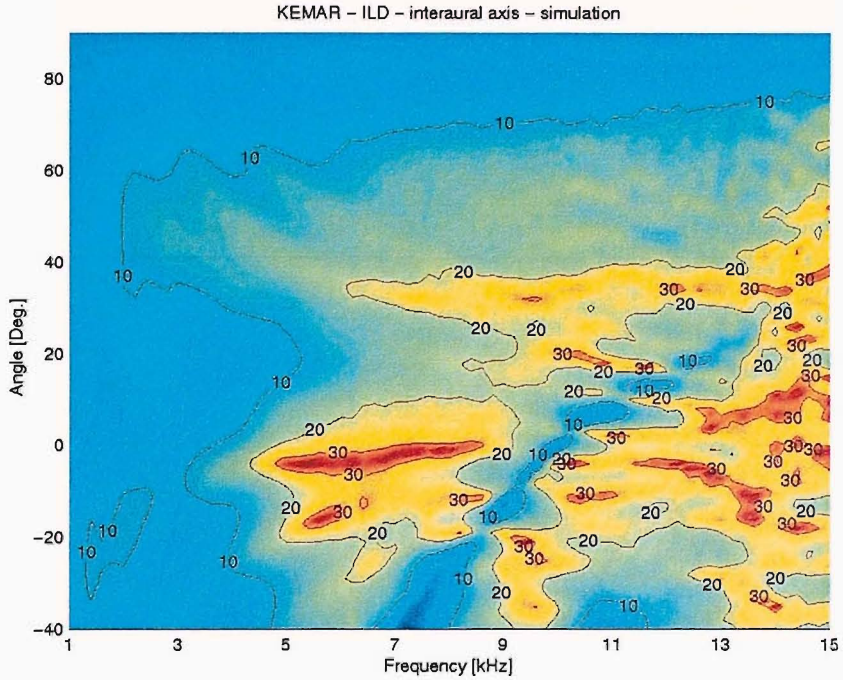


(e)

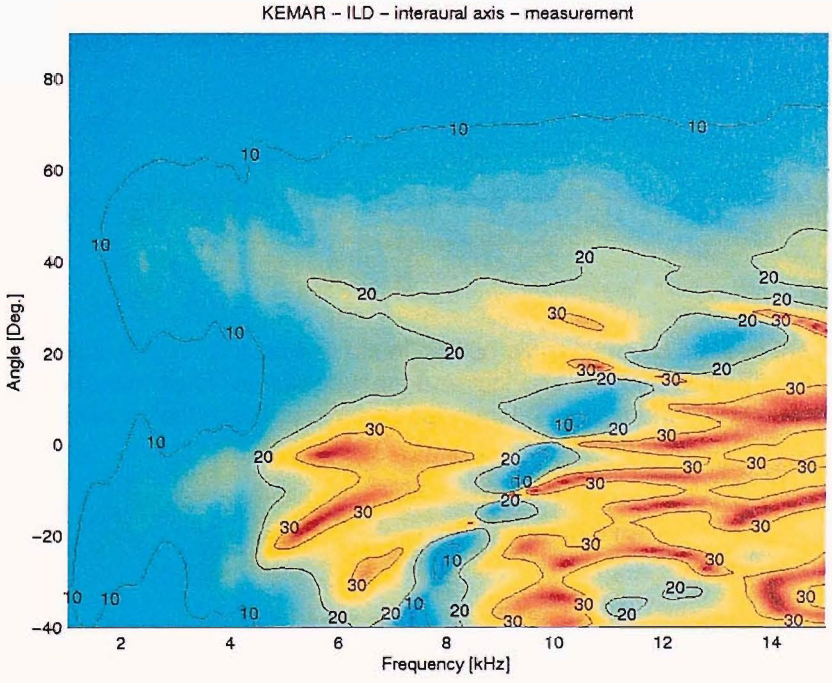


(f)

Figure 6-6: Comparison of simulations and measurements of HRTFs of KEMAR on the median plane at the following angles (a) $\theta = -40°$ (b) $\theta = 0°$ (c) $\theta = 40°$ (d) $\theta = 90°$ (e) $\theta = 130°$ (f) $\theta = 180°$. The simulation results are shown in red, and the measurements are shown in blue.



(a)



(b)

Figure 6-7: Comparison of the ILD of the HRTFs of KEMAR on the lateral vertical plane (a) simulations (b) measurements. In both cases the spatial resolution is 1° and no interpolations are made.

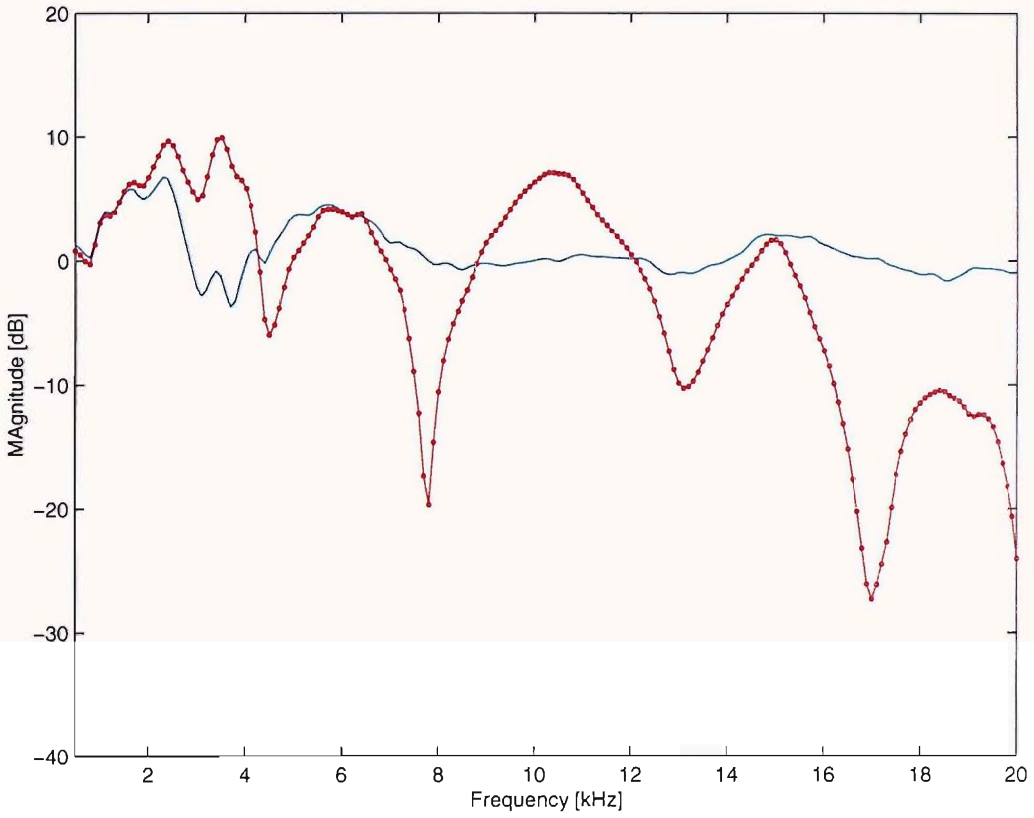
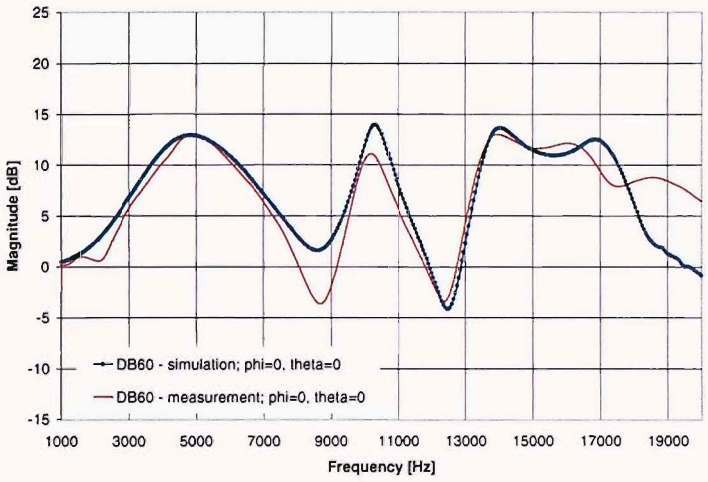
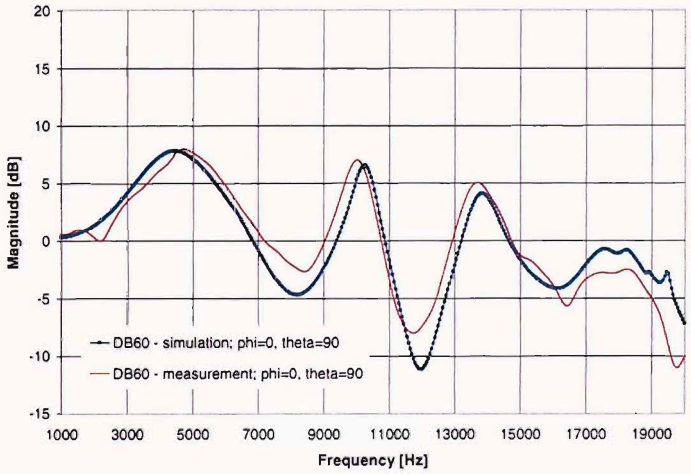


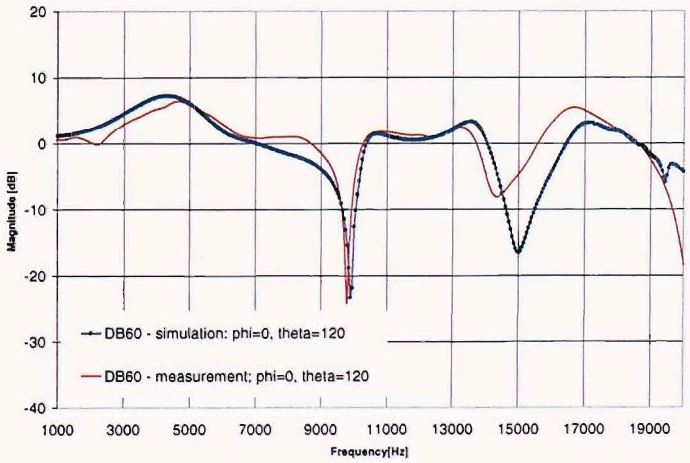
Figure 6-8: A typical pair of measurements of baffled pinnae (in this case of the YK pinna). The blue curve is the response of the transducers (microphone position as in Figure 6-3b). The red curve is the response detected at the blocked ear canal. The response is identical up to 1.5 kHz.



(a)

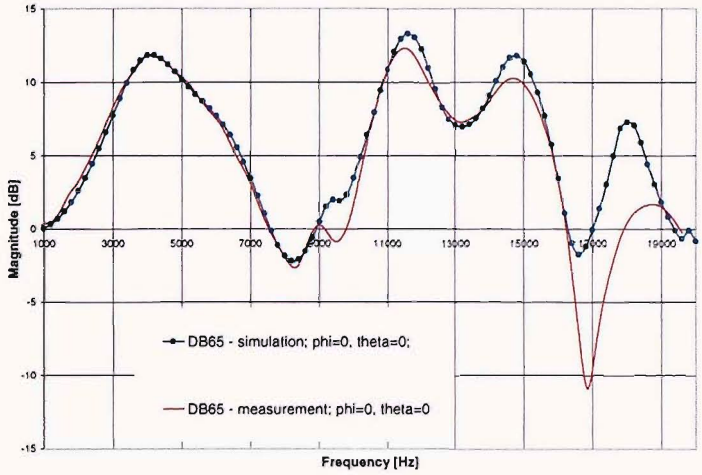


(b)

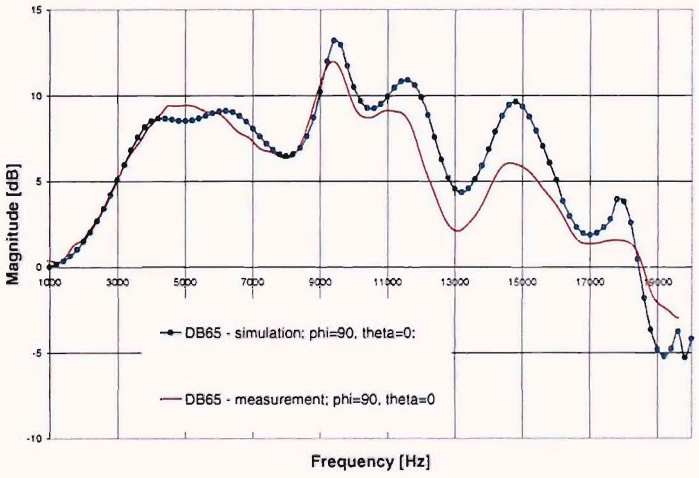


(c)

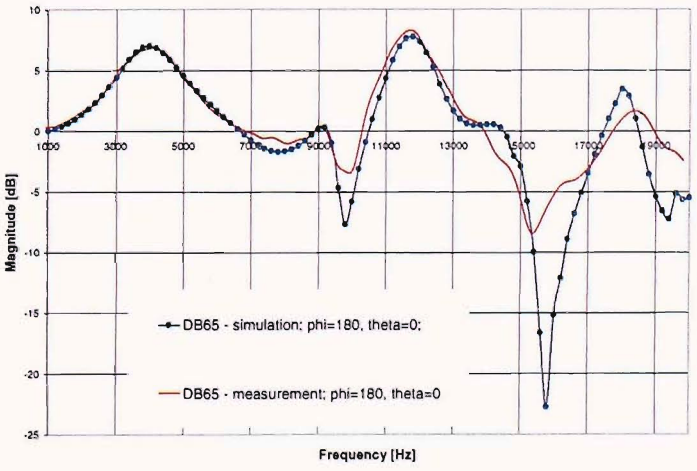
Figure 6-9: Comparison of the simulated and measured response of baffled DB60 (a) front $\theta = 0°, \phi = 0°$ (b) above $\phi = 0°, \theta = 90°$ (c) rear section $\phi = 0°, \theta = 120°$.



(a)

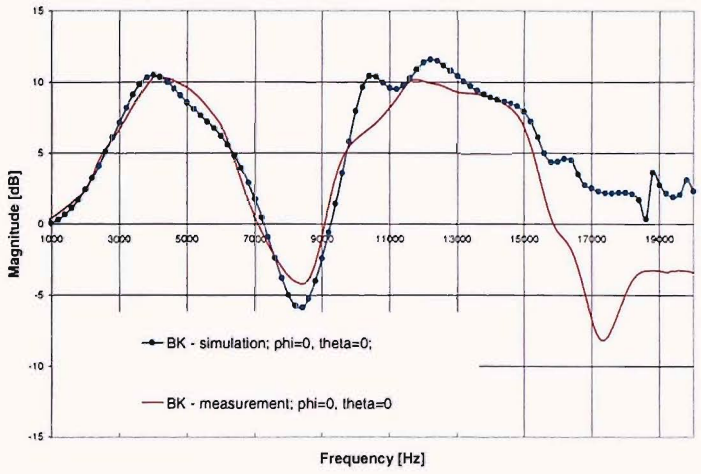


(b)

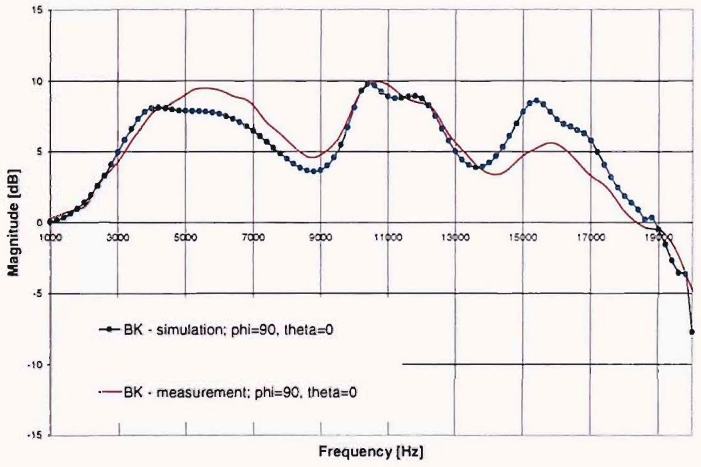


(c)

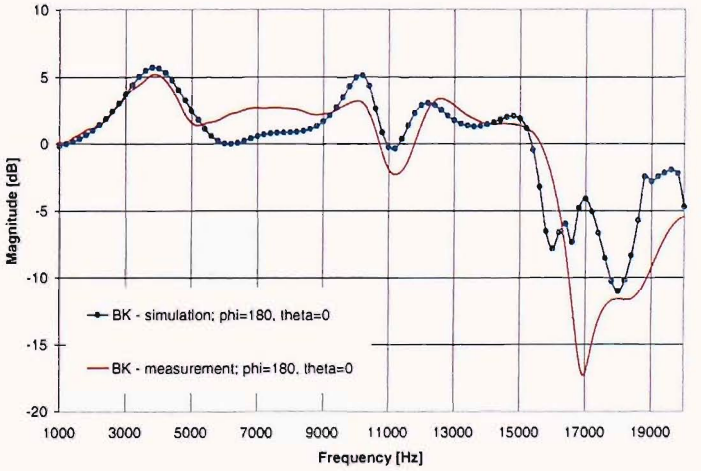
Figure 6-10: Comparison of the simulated and measured response of baffled DB65 (a) front $\phi = 0°, \theta = 0°$ (b) normal $\phi = 90°, \theta = 0°$ (c) rear $\phi = 180°, \theta = 0°$.



(a)



(b)



(c)

Figure 6-11: Comparison of the simulated and measured response of baffled DB65 (a) front $\phi = 0°, \theta = 0°$ (b) normal $\phi = 90°, \theta = 0°$ (c) rear $\phi = 180°, \theta = 0°$.

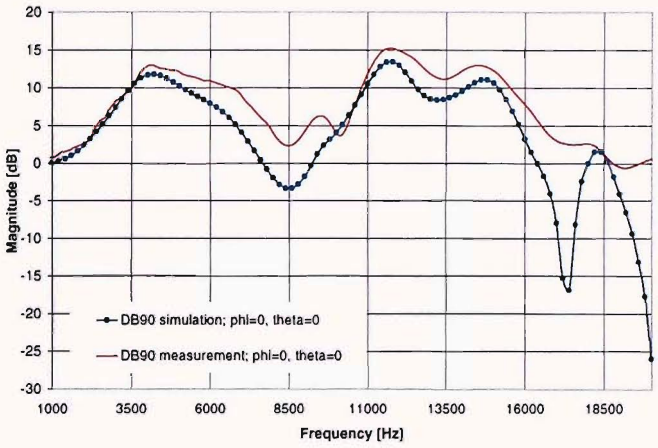


Figure 6-12: Comparison of the simulated and measured response of baffled DB90 in the front, at $\phi = 0°, \theta = 0°$.

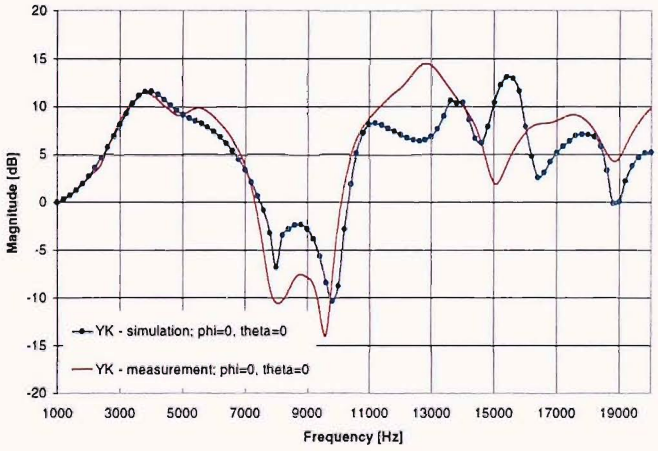


Figure 6-13: Comparison of the simulated and measured response of baffled YK pinna in the front, at $\phi = 0°, \theta = 0°$.

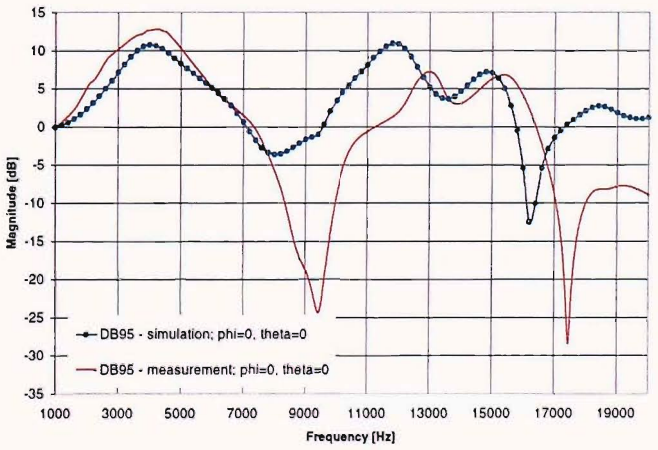
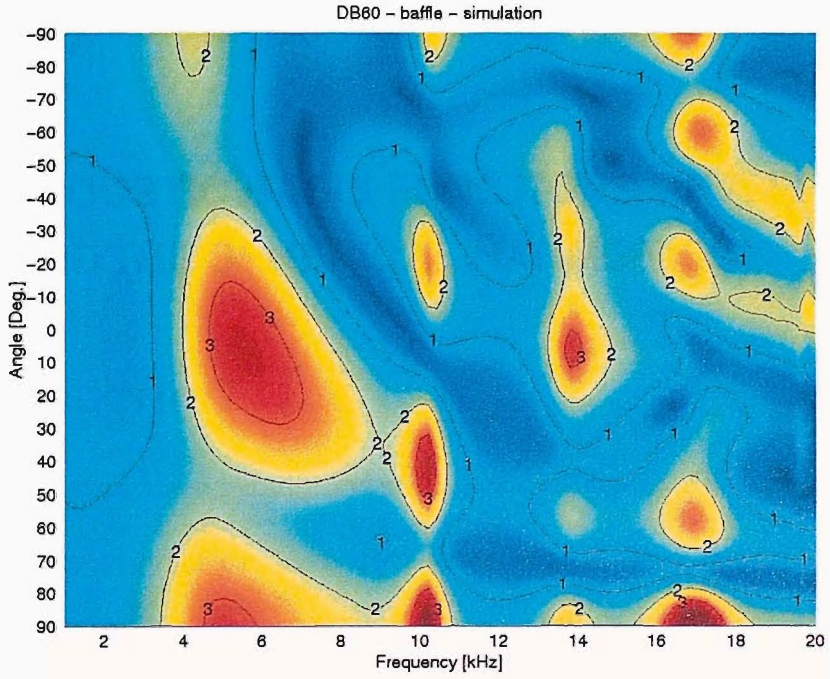
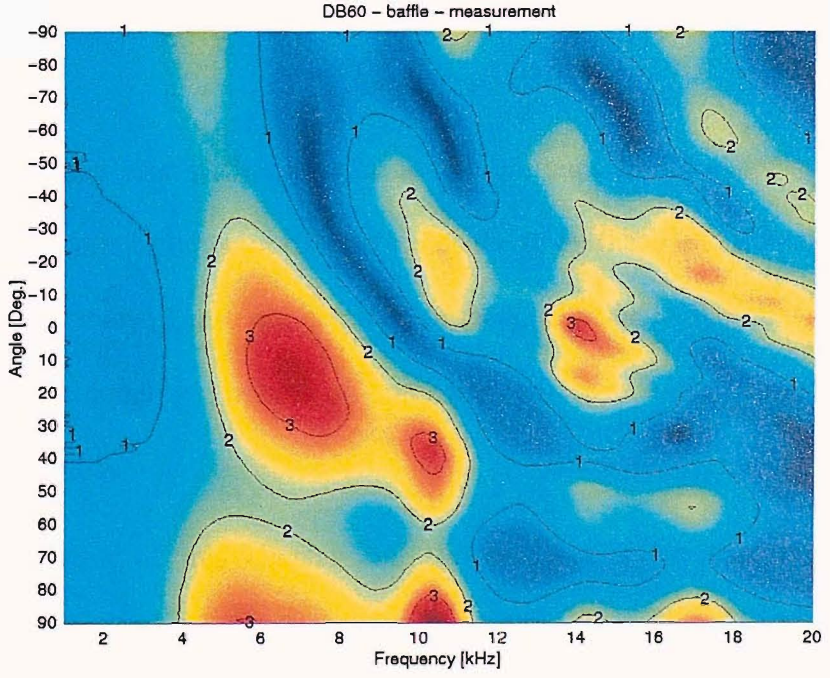


Figure 6-14: Comparison of the simulated and measured response of baffled DB95 in the front, at $\phi = 0°, \theta = 0°$.



(a)



(b)

Figure 6-15: Spatial map and contour plot of a comparison between (a) the simulated, and (b) the measured response of baffled DB60 in the lateral vertical plane ($\phi = 90°$, $-90° \leq \theta \leq +90°$) with a resolution of 1°.

CHAPTER 7

SPATIAL ACOUSTIC BASIS FUNCTIONS OF A RIGID SPHERE AND THE SPHERICAL HARMONICS

7.1 INTRODUCTION

In the previous chapters, the reliability and efficiency of using the BEM in the prediction of the response of the external ear have been demonstrated. Two main properties of the simulations are in particular advantageous:

- The calculation of the variation of the response with angular position is simplified using the principle of reciprocity, thus scattering can be replaced with radiation and *vice versa.*
- It is assumed that if the pressure at the entrance to the ear canal is predicted accurately, other points on the surface of the body (head and pinna) should have a similar accuracy. These are obtained immediately due to the inter-connectivity of the BEM equation where *all* pressure values on the surface are solved simultaneously.

The main motivation of this chapter and the next one is to analyse the acoustical characteristics of the head (simplified as a sphere) and various models of pinnae, using a 'modal' representation, with a mathematical formulation based on the Singular Value Decomposition (SVD). It is hypothesised that the extraction of the 'mode shapes' of the

head and pinna, will provide an insight into their physical characteristics, and that the decomposition of HRTFs would be possible based on these main features*.

SVD is a powerful method for analysing complex matrices with a wide variety of applications in mathematics and engineering. In acoustics, it has been used over the last decade as a mathematical tool in studies of sound radiation and scattering. In estimating the strength of acoustic sources using Near Acoustic Holography (NAH), Veronesi and Maynard (1989) used SVD on a matrix of transfer functions relating the strengths of discrete sources and acoustic pressure at a number of field points. The general formulation relating sources and field positions was used subsequently by Borgiotti (1990), who showed that at a given frequency, the 'radiation operator' relating the velocity on the surface of a vibrating body to the sound pressure produced over a surface in the radiated field can be used in analysing the complex singular vectors of the velocity and radiation patterns. Photiadis (1990) investigated a Green function matrix relating the pressure in the far field and the sources on a body. The relationships between the singular vectors of the 'source modes' and the associated 'radiated modes' have been analysed using the SVD. These are related at each frequency through the scalar singular values. Using a similar approach, the radiation modes of a plate have been investigated by Elliot and Johnson (1993), and Currey and Cunefare (1995). These authors concluded that the source 'mode shapes' contribute orthogonally to the total sound power over the surface considered. The SVD is also used also in acoustic inverse problems of radiation and scattering (see for example Fillipi *et al*, 1988, Kim and Lee, 1990, Grace *et al*, 1996, Nelson and Yoon, 2000 and Nelson, 1999) since it provides a least square solution when the matrices are underdetermined.

* It should be noted briefly that the SVD formulation (Gardner, 1999, Larcher *et al*, 2000) and the use of the series of spherical harmonics (Evans *et al*, 1997, Jot *et al*, 1999) have been used independently in the context of efficient coding scheme of HRTFs. In all of these cases the dimensions of the matrix analysed was the number of sources in the far field and the number of frequencies in the HRTF function, whereas in our case another dimension is added (the number of 'microphones' on the body), it is used in a different context and therefore the operation of the SVD has a different meaning .

Orthogonal basis functions have been used for a long time in the analysis of acoustic fields (Morse and Feshbach, 1953, Morse and Ingard, 1968 and Arfken, 1970). In his paper, Photiadis (1990) compared the radiation from a vibrating sphere using a set of orthonormal spherical harmonics with the SVD and stated that the source and radiation modes in both cases are similar apart form the phase. He also concluded that the multipole expansion embodies more information than is contained by the SVD.

In this chapter, we attempt to investigate the relationship between the basis functions of classical acoustics and the SVD through the scattering from a rigid sphere. An outline of this chapter is as follows: firstly, the response of a sphere is derived using an infinite series of complex spherical harmonics. This formulation is used in the construction of a Green function matrix relating a number of field points in the far field (on a large sphere) and a number of sources on the surface (of a small sphere). The analogy of the 'mode shapes' and the spherical harmonics is described. In addition, the deviation from ideal conditions such as non-uniform sampling of either the surface or the far field, as well as stretching the sphere to an ellipsoid and using numerical techniques for the solution are discussed. Finally, the frequency response reconstruction based on the SVD formulation is presented and analysed.

7.2 THEORY

7.2.1 Solutions of the wave equation in spherical co-ordinates

The scalar Helmholtz equation governing the behaviour of the complex acoustic pressure $p(\mathbf{r})$ is given in Equation (3.2). In spherical co-ordinates (r, θ, ϕ) this equation becomes:

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial p}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial p}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2 p}{\partial\phi^2}+k^2p=0 \qquad (7.1)$$

As indicated by Morse and Ingard (1968) and Temkin (1981) this equation is separable such that the solution is expressed as a functional of the radial, polar and azimuthal* dependence $p(\mathbf{r}) = F(r)G(\theta)H(\phi)$. Therefore three ordinary differential equations are derived

$$r^2 \frac{d^2F}{dr^2} + 2r\frac{dF}{dr} + (k^2r^2 - C^2)F = 0 \tag{7.2}$$

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{dG}{d\theta}\right) + \left(C^2 - \frac{m^2}{\sin^2\theta}\right)G = 0 \tag{7.3}$$

$$\frac{d^2H}{d\phi^2} + m^2H = 0 \tag{7.4}$$

where k^2, C^2 and m^2 are the separation constants. The solutions of Equation (7.4) are given by $e^{\pm jm\phi}$ where $j = \sqrt{-1}$, but that m must be a positive or negative integer (m = 0, ± 1, ± 2 . .) if the pressure is to be a single valued function of ϕ (i.e. to ensure that $p(r,\theta,\phi + 2\pi) = p(r,\theta,\phi)$). It was shown by Temkin (1981) that for Equation (7.3) to have solutions that are finite at $\theta = 0$ and $\theta = \pi$ then $C = n(n+1)$ where n must be a positive integer (n = 0, 1, 2, . . .). The solutions of Equation (7.3) when m = 0 are given by the Legendre polynomials (sometimes called also the Legendre functions of the first kind) $P_n(\cos\theta)$. For non-zero m, the solution is given by the associated Legendre polynomials $P_n^m(\cos\theta)$ defined by

* Some care must be taken in identifying the notational convention being used. In this entry, θ is taken as the polar (colatitudinal) coordinate with $\theta \in [0,\pi]$, and ϕ as the azimuthal (longitudinal) coordinate with $\phi \in [0,2\pi]$. This is the convention normally used in physics, as described by Arfken (1970) and in Mathematica (Wolfram Research, Inc.). In the co-ordinate system used for describing HRTFs, θ usually denotes the longitudinal coordinate and ϕ the colatitudinal coordinate.

$$P_n^m(x) = \frac{1}{2^n n!}(1-x^2)^{m/2}\frac{d^{m+n}}{dx^{m+n}}(x^2-1)^n, \qquad -n \le m \le n \tag{7.5}$$

Note also that the Condon-Shortley phase term $(-1)^m$ is omitted in this notation[†], although it was shown by Arfken (1970) that since m can be negative

$$P_n^{-m}(x) = (-1)^m \frac{(n-m)!}{(n+m)!}P_n^m(x) \tag{7.6}$$

The polar and azimuthal dependences of the pressure are combined in the definition of the spherical harmonics, given by Arfken (1970)[‡]

$$Y_n^m(\theta,\phi) = (-1)^m \sqrt{\frac{2n+1}{4\pi}\frac{(n-m)!}{(n+m)!}}P_n^m(\cos\theta)e^{jm\phi} \tag{7.7}$$

These spherical harmonics satisfy the orthogonality condition

$$\int_0^{2\pi}\int_0^{\pi} Y_n^{m*}(\theta,\phi)Y_k^l(\theta,\phi)\sin\theta d\theta d\phi = \delta_{nk}\delta_{ml} \tag{7.8}$$

where δ_{nk} is the Kronecker delta with $\delta_{nk} = 1$ for $n = k$ and $\delta_{nk} = 0$ for $n \neq k$ and δ_{ml} is analogously defined. Finally, the functions describing the radial dependence of the acoustic pressure that satisfy Equation (7.2) are given by the spherical Hankel functions of order n defined by

$$h_n^{(1)}(kr) = j_n(kr) + jn_n(kr) \tag{7.9}$$

$$h_n^{(2)}(kr) = j_n(kr) - jn_n(kr) \tag{7.10}$$

[†] The definitions of the Legendre Polynomials and the Spherical harmonics vary in the literature. It should be mentioned that the definitions in Mathematica (Wolfram research, Inc.) are different from the above formulation and the Condon-Shortley phase is included. Therefore the computational model presented in Section 7.3.1 has been compensated for this discrepancy.

[‡] Note that in some cases the spherical harmonics are defined differently. For example Morse and Ingard (1968), defined the spherical harmonics as three sets of scalar functions.

where $j_n(kr)$ and $n_n(kr)$ are respectively the spherical Bessel and Neumann functions of order n. These are in turn related to the Bessel function of order (n + 0.5) and Neumann function of order (n + 0.5) by

$$j_n(kr) = \sqrt{\frac{\pi}{2kr}} J_{n+0.5}(kr) \tag{7.11}$$

$$n_n(kr) = \sqrt{\frac{\pi}{2kr}} N_{n+0.5}(kr) \tag{7.12}$$

The spherical Hankel functions can also be deduced from Arfken (1970)

$$h_n^{(1)}(x) = -j(-1)^n x^n \left(\frac{d}{xdx}\right)^n \left(\frac{e^{jx}}{x}\right) \tag{7.13}$$

$$h_n^{(2)}(x) = j(-1)^n x^n \left(\frac{d}{xdx}\right)^n \left(\frac{e^{-jx}}{x}\right) \tag{7.14}$$

It therefore follows that for n = 0 for example

$$h_0^{(1)}(kr) = -je^{jkr} / kr \quad , \quad h_0^{(2)}(kr) = je^{-jkr} / kr \tag{7.15a, b}$$

which define either incoming or outgoing waves depending upon the choice of time convention.

7.2.2 Radiation from a point source on a rigid sphere

In the following sections the formulations derived are for the case of radiation, and not scattering as required from the point of view of representing HRTFs. However, the radiation problem from a point source on a rigid sphere into a point in the far field is analogous to the resultant sound pressure on the sphere due to a point source in the far field due to the principle of reciprocity.

The sound field produced by a point source situated at a vector position $\mathbf{r}_0$ in an unbounded medium is described by the free space Green function $g(\mathbf{r} \mid \mathbf{r}_0)$ which is a solution of

$$(\nabla^2 + k^2)g(\mathbf{r} \mid \mathbf{r}_0) = -\delta(\mathbf{r} - \mathbf{r}_0) \tag{7.16}$$

When a harmonic time dependence of $e^{j\omega t}$ is assumed, the free space Green function describing waves propagating outwards from $\mathbf{r}_0$ is given by

$$g(\mathbf{r} \mid \mathbf{r}_0) = \frac{e^{-jk|\mathbf{r}-\mathbf{r}_0|}}{4\pi \mid \mathbf{r} - \mathbf{r}_0 \mid} \tag{7.17}$$

This function can be expressed in terms of a series expansion of the spherical harmonics described in Equation (7.7) (see, for example, Arfken, 1970, Chapter 16, p. 768, Morse and Feshbach, 1953, Chapter 11, p. 1466, or Morse and Ingard, 1968, Chapter 7, p. 352. This series expansion can be written as

$$g(\mathbf{r} \mid \mathbf{r}_0) = -jk\sum_{n=0}^{\infty} j_n(kr_0)h_n^{(2)}(kr)\sum_{m=-n}^{n} Y_n^m(\theta,\phi)Y_n^{m*}(\theta_0,\phi_0) \quad (r > r_0) \tag{7.18}$$

where $r = \mid \mathbf{r} \mid$ and $r_0 = \mid \mathbf{r}_0 \mid$. This series expansion differs slightly from that given by Morse and Ingard (1968) since an $e^{j\omega t}$ time dependence has been assumed and thus $h_n^{(2)}(kr)$ has been used in place of $h_n^{(1)}(kr)$. As shown by Morse and Ingard (1968, Chapter 7, p. 355) this series expansion can be used to construct the Green function $G(\mathbf{r} \mid \mathbf{r}_0)$ describing radiation from a point source on the surface of a rigid sphere. It is convenient to modify the free space Green function $g(\mathbf{r} \mid \mathbf{r}_0)$ by adding a free standing wave $\chi(\mathbf{r})$ which satisfies the boundary conditions on the surface (that the radial gradient of the sum of $g(\mathbf{r} \mid \mathbf{r}_0)$ and $\chi(\mathbf{r})$ is zero on the surface of the sphere) through the

homogeneous Helmholtz equation $(\nabla^2 + k^2)\chi(\mathbf{r}) = 0$. Based on Morse and Ingard (1968) the Green function satisfying these conditions can be written as

$$G(\mathbf{r} \mid \mathbf{r}_0) = -jk\sum_{n=0}^{\infty}\left[\frac{j_n(kr_0)h_n^{(2)'}(ka) - j_n{}'(ka)h_n^{(2)}(kr_0)}{h_n^{(2)'}(ka)}\right]h_n^{(2)}(kr) \times \sum_{m=-n}^{n} Y_n^m(\theta,\phi)Y_n^{m*}(\theta_0,\phi_0) \tag{7.19}$$

where a is the radius of the sphere and the prime denotes differentiation with respect to the argument.

The general solution of the inhomogeneous Helmholtz equation is given by

$$(\nabla^2 + k^2)p(\mathbf{r}) = -Q_{\text{vol}}(\mathbf{r}_0) \tag{7.20}$$

where $Q_{\text{vol}}(\mathbf{r}_0)$ represents some volume source distribution, and the pressure can be expressed as

$$p(\mathbf{r}) = \int_V Q_{\text{vol}}G(\mathbf{r} \mid \mathbf{r}_0)dV + \int_S [G(\mathbf{r} \mid \mathbf{r}_0)\nabla_0 p(\mathbf{r}_0) - p(\mathbf{r}_0)\nabla_0 G(\mathbf{r} \mid \mathbf{r}_0)].\mathbf{n}\,dS \tag{7.21}$$

where S is the surface with unit outward normal vector $\mathbf{n}$ that bounds the volume V containing the field point $\mathbf{r}$, and ∇_0 is the gradient operator with respect to the co-ordinates defined by $\mathbf{r}_0$. Since $G(\mathbf{r} \mid \mathbf{r}_0)$ given by Equation (7.19) has been chosen to satisfy $\nabla_0 G(\mathbf{r} \mid \mathbf{r}_0).\mathbf{n} = 0$ on the surface of the sphere, and if there are no other sources within V, then

$$p(\mathbf{r}) = \int_S G(\mathbf{r} \mid \mathbf{r}_0)\nabla_0 p(\mathbf{r}_0).\mathbf{n}\,dS \tag{7.22}$$

where the integration is carried out only over the surface of the sphere. (The other part of S is that part of the surface bounding the volume V outside the sphere and thus lies at infinity; the contribution to the surface integral from this part of S can be shown to be zero by

satisfying the Sommerfeld radiation condition. See Chapter 3 and also Morse and Ingard, 1968). If the sphere is now assumed to have an arbitrary radial velocity distribution $\nabla_0 p(\mathbf{r}_0).\mathbf{n} = j\omega\rho_0 u_n(\theta_0,\phi_0)$ then it follows that

$$p(\mathbf{r}) = j\omega\rho_0 a^2 \int_0^{2\pi}\int_0^{\pi} G(\mathbf{r} \mid \mathbf{r}_0) u_n(\theta_0,\phi_0)\sin\theta_0 d\theta_0 d\phi_0 \tag{7.23}$$

It therefore follows that the expression for the pressure field generated by an arbitrary surface velocity distribution can be written as

$$p(\mathbf{r}) = \rho_0 c_0 k^2 a^2 \sum_{n=0}^{\infty} \frac{[j_n(ka)h_n^{(2)'}(ka) - j_n{}'(ka)h_n^{(2)}(ka)]}{h_n^{(2)'}(ka)} h_n^{(2)}(kr) \times \sum_{m=-n}^{n} U_n^m Y_n^m(\theta,\phi) \tag{7.24}$$

where the factor U_n^m accounts for the degree to which a given spherical harmonic is driven by a given velocity distribution and is defined by

$$U_n^m = \int_0^{2\pi}\int_0^{\pi} u_n(\theta_0,\phi_0) Y_n^{m*}(\theta_0,\phi_0)\sin\theta_0 d\theta_0 d\phi_0 \tag{7.25}$$

In this case we apply a point source with strength $q(\hat{\mathbf{r}})$ at $(\hat{\theta},\hat{\phi})$ instead of the general velocity distribution given by

$$u_n(\theta_0,\phi_0) = -q\frac{\delta(\theta_0 - \hat{\theta})\delta(\phi_0 - \hat{\phi})}{\sin\theta_0} \tag{7.26}$$

then U_n^m is simply given by $-qY_n^{m*}(\hat{\theta},\hat{\phi})$.

In addition, Arfken, 1970, p. 529, shows that

$$j_n(x)n_n{}'(x) - j_n{}'(x)n_n(x) = (1/x^2) \tag{7.27}$$

So it can be shown that, the term in the square brackets in Equation (7.24) reduces to $-j/k^2a^2$ and the resulting pressure is given by

$$p(\mathbf{r}) = j\rho_0 c_0 q(\hat{\mathbf{r}}) \sum_{n=0}^{\infty} \frac{h_n^{(2)}(kr)}{h_n^{(2)'}(ka)} \sum_{m=-n}^{n} Y_n^m(\theta,\phi) Y_n^{m*}(\hat{\theta},\hat{\phi}) \tag{7.28}$$

It should be noted here that Equation (7.28) shows that the frequency dependence of the radiated sound within the terms $h_n^{(2)}(kr)/h_n^{(2)}(ka)$ is independent of θ and ϕ.

7.2.3 The singular value decomposition

For a given linear acoustic field we can specify a matrix $\mathbf{G}(\mathbf{r} \mid \hat{\mathbf{r}})$ of Green functions relating the pressures produced at a number of field points defined by the position vector $\mathbf{r}$ to the strengths of a number of point sources at positions defined by the position vector $\hat{\mathbf{r}}$. Therefore, the pressure is given by

$$\mathbf{p}(\mathbf{r}) = \mathbf{G}(\mathbf{r} \mid \hat{\mathbf{r}})\mathbf{q}(\hat{\mathbf{r}}) \tag{7.29}$$

where $\mathbf{p}(\mathbf{r})$ is the vector whose elements define the field pressure values and $\mathbf{q}(\hat{\mathbf{r}})$ is the vector whose elements define the source strengths. By using the singular value decomposition (SVD) of the matrix $\mathbf{G}(\mathbf{r} \mid \hat{\mathbf{r}})$ the characteristics of radiation and scattering problems are investigated. First we start, with the sphere - a 'separable' problem, and then we show the effectiveness of the method with 'non-separable' problems.

By using the SVD, arbitrary complex matrix $\mathbf{G}(\mathbf{r} \mid \hat{\mathbf{r}})$ of order $K \times L$ to be expressed as

$$\mathbf{G}(\mathbf{r} \mid \hat{\mathbf{r}}) = \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^{\mathrm{H}} \tag{7.30}$$

where $\boldsymbol{\Sigma}$ is the $K \times L$ matrix whose entries are zero apart from the diagonal elements σ_i which consist of the singular values of $\mathbf{G}(\mathbf{r} \mid \hat{\mathbf{r}})$. The superscript H denotes conjugate transpose. If R is the rank of $\mathbf{G}(\mathbf{r} \mid \hat{\mathbf{r}})$ then the singular values satisfy

$$\sigma_1 \geq \sigma_2 \geq . \ . \ . \ \sigma_R > 0 \ , \ \sigma_{R+1} = . \ . \ . \ . = \sigma_p = 0 \quad p = \min(K, L) \tag{7.31}$$

The matrices $\mathbf{U}$ and $\mathbf{V}$ are respectively of dimension $K \times K$ and $L \times L$ and are unitary matrices having the orthogonality properties

$$\mathbf{U}^{\mathrm{H}}\mathbf{U} = \mathbf{U}\mathbf{U}^{\mathrm{H}} = \mathbf{I} \quad \mathbf{V}^{\mathrm{H}}\mathbf{V} = \mathbf{V}\mathbf{V}^{\mathrm{H}} = \mathbf{I} \tag{7.32 a, b}$$

The columns $\mathbf{u}_i$ of the matrix $\mathbf{U}$ and the columns $\mathbf{v}_i$ of the matrix $\mathbf{V}$ respectively define the left and right singular vectors of $\mathbf{G}(\mathbf{r} \mid \hat{\mathbf{r}})$. The significance of the singular vectors is that they provide sets of orthogonal basis functions for describing the spatial variation in radiated pressure and their relationship to spatial variations in source strength. Specifically, it follows from Equations (7.29) and (7.30) that

$$\mathbf{p}(\mathbf{r}) = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{H}}\mathbf{q}(\hat{\mathbf{r}}) \tag{7.33}$$

and since by virtue of Equation (32a), $\mathbf{U}^{-1} = \mathbf{U}^{\mathrm{H}}$, then this expression may be written as

$$\mathbf{U}^{\mathrm{H}}\mathbf{p}(\mathbf{r}) = \boldsymbol{\Sigma}\mathbf{V}^{\mathrm{H}}\mathbf{q}(\hat{\mathbf{r}}) \tag{7.34}$$

Therefore, for all $i \leq p$, $(p = \min(K, L))$ we may write

$$\mathbf{u}_i^{\mathrm{H}}\mathbf{p}(\mathbf{r}) = \sigma_i \mathbf{v}_i^{\mathrm{H}}\mathbf{q}(\hat{\mathbf{r}}) \ . \tag{7.35}$$

This demonstrates that a specific spatial pattern in the radiated field defined by $\mathbf{u}_i^{\mathrm{H}}\mathbf{p}(\mathbf{r})$ is linearly related to a specific spatial pattern of source strength distribution defined by $\mathbf{v}_i^{\mathrm{H}}\mathbf{q}(\hat{\mathbf{r}})$. These two patterns are related by the singular value σ_i.

7.2.4 Alternative forms of the Green function matrix

There is also a close relationship between the singular value decomposition and the series expansions used in the analyses of acoustical problems in separable co-ordinate systems. In

particular, this close relationship can be demonstrated for the case of spherical radiation expressed in terms of a series of spherical harmonics. The SVD of the Green function matrix can be expressed in terms of the left and right singular vectors such that

$$\mathbf{G}(\mathbf{r} \mid \hat{\mathbf{r}}) = \sum_{i=1}^{P} \sigma_i \mathbf{u}_i \mathbf{v}_i^{\mathrm{H}} \quad (P = \min(K, L)) \tag{7.36}$$

where the matrix is shown to consist of a linear superposition of P component matrices each defined by the outer product $\mathbf{u}_i \mathbf{v}_i^{\mathrm{H}}$ and weighted in the summation by the singular value σ_i. Now note that we may use Equation (7.28) to define the elements of the Green function matrix $\mathbf{G}(\mathbf{r} \mid \hat{\mathbf{r}})$ relating the acoustic pressure at K points in the sound field to the source strength at L points on the surface of a sphere. This matrix can be written in the form

$$\mathbf{G}(\mathbf{r} \mid \hat{\mathbf{r}}) = \begin{bmatrix} \sum_{n=0}^{\infty} f_n \sum_{m=-n}^{n} Y_n^m(\theta_1, \phi_1) Y_n^{m*}(\hat{\theta}_1, \hat{\phi}_1) & \cdots & \sum_{n=0}^{\infty} f_n \sum_{m=-n}^{n} Y_n^m(\theta_1, \phi_1) Y_n^{m*}(\hat{\theta}_L, \hat{\phi}_L) \\ \sum_{n=0}^{\infty} f_n \sum_{m=-n}^{n} Y_n^m(\theta_2, \phi_2) Y_n^{m*}(\hat{\theta}_1, \hat{\phi}_1) & \cdots & \sum_{n=0}^{\infty} f_n \sum_{m=-n}^{n} Y_n^m(\theta_2, \phi_2) Y_n^{m*}(\hat{\theta}_L, \hat{\phi}_L) \\ \sum_{n=0}^{\infty} f_n \sum_{m=-n}^{n} Y_n^m(\theta_K, \phi_K) Y_n^{m*}(\hat{\theta}_1, \hat{\phi}_1) & \cdots & \sum_{n=0}^{\infty} f_n \sum_{m=-n}^{n} Y_n^m(\theta_K, \phi_K) Y_n^{m*}(\hat{\theta}_L, \hat{\phi}_L) \end{bmatrix} \tag{7.37}$$

where it follows from Equation (7.28) that

$$f_n = j\rho_0 c_0 h_n^{(2)}(kr) / h_n^{(2)'}(ka) \tag{7.38}$$

Since each term in the series comprising each element of the matrix is weighted by the same factor f_n, it is possible to write the matrix as a linear superposition of matrices having the form

$$\mathbf{G}(\mathbf{r} \mid \hat{\mathbf{r}}) = \sum_{n=0}^{\infty} f_n \sum_{m=-n}^{n} \mathbf{y}_n^m(\mathbf{r}_k) \mathbf{y}_n^{m\mathrm{H}}(\hat{\mathbf{r}}_l) \tag{7.39}$$

where the following vectors have been defined

$$\mathbf{y}_n^{m\mathrm{T}}(\mathbf{r}_k) = [Y_n^m(\theta_1,\phi_1)\ Y_n^m(\theta_2,\phi_2)\ \ .\ .\ .\ .\ .\ .\ Y_n^m(\theta_K,\phi_K)] \tag{7.40}$$

$$\mathbf{y}_n^{m\mathrm{T}}(\mathbf{r}_l) = [Y_n^m(\hat{\theta}_1,\hat{\phi}_1)\ Y_n^m(\hat{\theta}_2,\hat{\phi}_2)\ \ .\ .\ .\ .\ .\ .\ Y_n^m(\hat{\theta}_L,\hat{\phi}_L)] \tag{7.41}$$

There are some obvious similarities between the singular value decomposition expressed in the form of Equation (7.36) and the matrix series given by Equation (7.39). The vector $\mathbf{y}_n^m(\mathbf{r}_k)$ has elements which correspond to the values of the (n, m)'th spherical harmonic at each of the points in the radiated field at which we evaluate the pressure. The vector $\mathbf{y}_n^m(\hat{\mathbf{r}}_l)$ has corresponding elements at each of the points on the surface of the sphere at which we evaluate the source strength. The vectors $\mathbf{y}_n^m(\mathbf{r}_k)$ thus define the basis functions for the pressure field in much the same way as the vectors $\mathbf{u}_i$. Similarly the vectors $\mathbf{y}_n^m(\mathbf{r}_k)$ define the basis functions for the source strength distribution in an analogous manner to the vectors $\mathbf{v}_i$. One should note however, that one would not necessarily expect an exact correspondence between the matrix series given by Equation (7.39) and the SVD in Equation (7.36). Most obviously, the weighting factors differ in that the singular values σ_i are purely real whilst the coefficients f_n are in general complex. Additionally, the matrix series in Equation (7.39) is infinite, while the matrix series representation of the SVD is finite and consists of $p = \min(K, L)$ terms.

Nevertheless, there is a connection between the two descriptions of the Green function matrix. Note that if we choose to represent the Green functions in Equation (7.37) as a sum up to a maximum number of $n = N$ spherical harmonics, then the truncated Green function matrix $\mathbf{G}_N(\mathbf{r} \mid \hat{\mathbf{r}})$ can be expressed as the matrix product

$$\mathbf{G}_N(\mathbf{r} \mid \hat{\mathbf{r}}) = \mathbf{Y}(\mathbf{r}_k)\mathbf{F}\mathbf{Y}^{\mathrm{H}}(\hat{\mathbf{r}}_l) \tag{7.42}$$

where the columns of $\mathbf{Y}(\mathbf{r}_k)$ are given by the vectors $\mathbf{y}_n^m(\mathbf{r}_k)$ and the columns of $\mathbf{Y}(\hat{\mathbf{r}}_l)$ are given by the vectors $\mathbf{y}_n^m(\hat{\mathbf{r}}_l)$. The matrix $\mathbf{F}$ is a diagonal matrix consisting of the coefficients f_n. To make clear the structure of these matrices, in the case N=1, K=5 and L=3, the Green function matrix can be expanded as

$$\mathbf{G}_N(\mathbf{r}|\hat{\mathbf{r}}) = \begin{bmatrix} Y_0^0(\theta_1,\phi_1) & Y_1^{-1}(\theta_1,\phi_1) & Y_1^0(\theta_1,\phi_1) & Y_1^1(\theta_1,\phi_1) \\ Y_0^0(\theta_2,\phi_2) & Y_1^{-1}(\theta_2,\phi_2) & Y_1^0(\theta_2,\phi_2) & Y_1^1(\theta_2,\phi_2) \\ Y_0^0(\theta_3,\phi_3) & Y_1^{-1}(\theta_3,\phi_3) & Y_1^0(\theta_3,\phi_3) & Y_1^1(\theta_3,\phi_3) \\ Y_0^0(\theta_4,\phi_4) & Y_1^{-1}(\theta_4,\phi_4) & Y_1^0(\theta_4,\phi_4) & Y_1^1(\theta_4,\phi_4) \\ Y_0^0(\theta_5,\phi_5) & Y_1^{-1}(\theta_5,\phi_5) & Y_1^0(\theta_5,\phi_5) & Y_1^1(\theta_5,\phi_5) \end{bmatrix} \begin{bmatrix} f_0 & 0 & 0 & 0 \\ 0 & f_1 & 0 & 0 \\ 0 & 0 & f_1 & 0 \\ 0 & 0 & 0 & f_1 \end{bmatrix} \times \begin{bmatrix} Y_0^{0*}(\hat{\theta}_1,\hat{\phi}_1) & Y_0^{0*}(\hat{\theta}_2,\hat{\phi}_2) & Y_0^{0*}(\hat{\theta}_3,\hat{\phi}_3) \\ Y_1^{-1*}(\hat{\theta}_1,\hat{\phi}_1) & Y_1^{-1*}(\hat{\theta}_2,\hat{\phi}_2) & Y_1^{-1*}(\hat{\theta}_3,\hat{\phi}_3) \\ Y_1^{0*}(\hat{\theta}_1,\hat{\phi}_1) & Y_1^{0*}(\hat{\theta}_2,\hat{\phi}_2) & Y_1^{0*}(\hat{\theta}_2,\hat{\phi}_2) \\ Y_1^{1*}(\hat{\theta}_1,\hat{\phi}_1) & Y_1^{1*}(\hat{\theta}_2,\hat{\phi}_2) & Y_1^{1*}(\hat{\theta}_2,\hat{\phi}_2) \end{bmatrix} \tag{7.43}$$

Now note that it may be shown that as $K,\ L \rightarrow \infty$, then the matrices $\mathbf{Y}(\mathbf{r}_k)$ and $\mathbf{Y}(\hat{\mathbf{r}}_l)$ become unitary matrices. For example, the diagonal elements of the matrix $\mathbf{Y}^{\mathrm{H}}(\mathbf{r}_k)\mathbf{Y}(\mathbf{r}_k)$ consists of terms given by

$$\mathbf{y}_n^{m\mathrm{H}}(\mathbf{r}_k)\mathbf{y}_n^m(\mathbf{r}_k) = \left|Y_n^m(\theta_1,\phi_1)\right|^2 + \left|Y_n^m(\theta_2,\phi_2)\right|^2 \ldots \left|Y_n^m(\theta_K,\phi_K)\right|^2 = \sum_{k=1}^{K}\left|Y_n^m(\theta_k,\phi_k)\right|^2 \tag{7.44}$$

The orthogonality property of the spherical harmonics given by Equation (7.8) shows that

$$\int_0^{2\pi}\int_0^{\pi}\left|Y_n^m(\theta,\phi)\right|^2 \sin\theta d\theta d\phi = 1 \tag{7.45}$$

The term $\sin\theta d\theta d\phi$ can be regarded as an element of area of a sphere of unit radius, and if the K points at which the spherical harmonics are sampled are at the centres of segments of equal area, the integral in Equation (7.45) can be expressed as the summation

$$\sum_{k=1}^{K} |Y_n^m(\theta,\phi)|^2 \, \Delta S_k = 1 \tag{7.46}$$

Since $\Delta S_k = 4\pi / K$, then it follows that the summation in Equation (7.44) will tend to a value of $K/4\pi$ as $K \rightarrow \infty$. By an exactly analogous argument, it is evident that the orthogonality property of the spherical harmonics results in the off-diagonal terms of the matrix $\mathbf{Y}^{\mathrm{H}}(\mathbf{r}_k)\mathbf{Y}(\mathbf{r}_k)$ tending to zero as $K \rightarrow \infty$. It is thus concluded that

$$\mathbf{Y}^{\mathrm{H}}(\mathbf{r}_k)\mathbf{Y}(\mathbf{r}_k) = (K/4\pi)\mathbf{I} \qquad \text{as } K \rightarrow \infty \tag{7.47}$$

and similarly that

$$\mathbf{Y}^{\mathrm{H}}(\hat{\mathbf{r}}_l)\mathbf{Y}(\hat{\mathbf{r}}_l) = (L/4\pi)\mathbf{I} \qquad \text{as } L \rightarrow \infty \tag{7.48}$$

7.2.5 The singular value decomposition and the spherical harmonics

It will be demonstrated by the numerical simulations presented below that there is indeed, under certain circumstances, a direct connection between the results of the singular value decomposition of the Green function matrix and the matrices $\mathbf{Y}(\mathbf{r}_k)$ and $\mathbf{Y}(\hat{\mathbf{r}}_l)$ of sampled spherical harmonics. Specifically it will be demonstrated that when $\mathbf{Y}(\hat{\mathbf{r}}_l)$ and $\mathbf{Y}(\mathbf{r}_k)$ can be regarded as unitary matrices, such that Equations (7.47) and (7.48) hold to a good approximation, then we may write

$$\mathbf{U}_N = \mathbf{Y}(\mathbf{r}_k)\mathbf{T}(\mathbf{r}_k) \;, \quad \mathbf{V}_N = \mathbf{Y}(\hat{\mathbf{r}}_l)\mathbf{T}(\hat{\mathbf{r}}_l) \tag{7.49a, b}$$

The matrices $\mathbf{U}_N$ and $\mathbf{V}_N$ comprise the left and right singular vectors associated with the first N singular values which, if the SVD is taken of the truncated Green function matrix $\mathbf{G}_N(\mathbf{r} \mid \hat{\mathbf{r}})$, are the only non-zero singular values. The matrices $\mathbf{U}_N$ and $\mathbf{V}_N$ are found to be linear combinations of the sampled spherical harmonics. The matrices $\mathbf{T}(\mathbf{r}_k)$ and $\mathbf{T}(\hat{\mathbf{r}}_k)$

effectively specify the combinations of the columns of $\mathbf{Y}(\mathbf{r}_k)$ and $\mathbf{Y}(\hat{\mathbf{r}}_l)$ respectively that must be added to produce the columns of $\mathbf{U}_N$ and $\mathbf{V}_N$. In fact the left and right singular vectors are found to be linear combinations of sampled spherical harmonics of a certain order n, and this results in the matrices $\mathbf{T}(\mathbf{r}_k)$ and $\mathbf{T}(\hat{\mathbf{r}}_l)$ having a certain block diagonal structure. Furthermore, since $\mathbf{U}_N$ and $\mathbf{V}_N$ are unitary, and when $\mathbf{Y}(\mathbf{r}_k)$ and $\mathbf{Y}(\mathbf{r}_l)$ are unitary, then both $\mathbf{T}(r_k)$ and $\mathbf{T}(\hat{\mathbf{r}}_l)$ must be unitary. This follows since from Equations (7.49a) and (7.32a)

$$\mathbf{U}_N^{\mathrm{H}}\mathbf{U}_N = \mathbf{T}^{\mathrm{H}}(\mathbf{r}_k)\mathbf{Y}^{\mathrm{H}}(\mathbf{r}_k)\mathbf{Y}(\mathbf{r}_k)\mathbf{T}(\mathbf{r}_k) = \mathbf{I} \tag{7.50}$$

and when $\mathbf{Y}^{\mathrm{H}}(\mathbf{r}_k)\mathbf{Y}(\mathbf{r}_k) = (K/4\pi)\mathbf{I}$, then it follows that $\mathbf{T}^{\mathrm{H}}(\mathbf{r}_k)\mathbf{T}(\mathbf{r}_k) = (4\pi/K)\mathbf{I}$. Similarly, $\mathbf{T}^{\mathrm{H}}(\hat{\mathbf{r}}_l)\mathbf{T}(\hat{\mathbf{r}}_l) = (4\pi/L)\mathbf{I}$. Thus it follows from Equations (7.49a, b) that the SVD of the Green function matrix given by Equation (7.30) can be written as

$$\mathbf{G}_N(\mathbf{r} \mid \hat{\mathbf{r}}) = \mathbf{Y}(\mathbf{r}_k)\mathbf{T}(\mathbf{r}_k)\boldsymbol{\Sigma}_N\,\mathbf{T}^{\mathrm{H}}(\hat{\mathbf{r}}_l)\mathbf{Y}^{\mathrm{H}}(\hat{\mathbf{r}}_l) \tag{7.51}$$

where $\boldsymbol{\Sigma}_N$ is the diagonal matrix of the N non-zero real singular values. It is also evident from Equation (7.43) therefore that the diagonal matrix $\mathbf{F}$ of the complex amplitudes of the spherical harmonics is given by

$$\mathbf{F} = \mathbf{T}(\mathbf{r}_k)\boldsymbol{\Sigma}_N\,\mathbf{T}^{\mathrm{H}}(\hat{\mathbf{r}}_l) \tag{7.52}$$

These findings are confirmed by the results of the numerical simulations presented below. It will also be shown that for these results to hold, that it is vital that the points on the surface of the sphere and the points on the surrounding spherical surface are sampled at the centres of segments of equal area. This ensures the orthogonality of the columns of the matrices $\mathbf{Y}(\mathbf{r}_k)$ and $\mathbf{Y}(\hat{\mathbf{r}}_l)$.

7.3 NUMERICAL SIMULATION BASED ON THE ANALYTICAL MODEL OF RADIATION FROM A SPHERE

7.3.1 Computational model

The Green function given by Equation (7.28) was used in order to construct the matrix of Green functions relating sources placed at 32 points on the surface of a rigid sphere (forming an Icosahedron) of radius 0.1 m to 32 points on a surrounding far field spherical surface of radius 10 m. The number of points chosen enabled both spherical surfaces to be divided into 60 segments of equal area as illustrated in Figure 7-1. In Equation 7.28, the first 5 terms were used in the series expansion (n=0 to n=4). As a result, the first $N = 25$ terms were included in the series. It was assumed that the characteristic impedance $\eta_0 = \rho_0 c_0 = 411.4$ Rayls. The calculation is undertaken at a low frequency (ka=0.1, where a=0.1 m). The SVD of this matrix was undertaken using the 'SVD' routine in MATHEMATICA (Wolfram Research, Inc.) and yielded only 25 significant singular values σ_i, as expected. The number of singular values found was thus equal exactly to the total number of terms in the spherical harmonic expansion (i.e. one term corresponding to n = 0, three terms corresponding to n = 1, five terms corresponding to n=2, seven terms corresponding to n = 3 and so forth). The remaining singular values were smaller than a factor of 10^{-10} times the smallest significant singular value at the given low frequency. A plot of the variation of the dominant 16 non-zero singular values as a function of ka is shown in Figure 7-2. It is clear that the singular values are grouped in spherical harmonic order with the largest singular value corresponding to n = 0, the next three corresponding to n = 1, the next five corresponding to n = 2 and the next seven corresponding to n = 3. The real parts of the singular vector corresponding to n = 0, 1 and 2 which are given by the columns $\mathbf{u}_i$ are shown as surface colour maps in Figure 7-3a. These plots clearly show the 'monopole-like', 'dipole-like' and 'quadrupole-like' nature of the singular vectors associated

respectively with spherical harmonic orders corresponding to $n = 0, 1, 2$. (Note however, as discussed in detail by Morse and Ingard, 1968, Chapter 7, p. 346, that the sound fields associated with the multipole expansion are not simply related to the spherical harmonics; the quadrupole field for example consisting of a combination of spherical harmonics of order $n = 2$ and a monopole field). For comparison, the real and imaginary parts of the spherical harmonics corresponding to $n = 0$, 1 and 2 are shown in Figures 7-3b and 7-3c respectively. It was observed at this point that clearly similar patterns exist, but the rotation of the shapes cannot be identified intuitively.

7.3.2 Calculation of the unitary transformation matrices

A check was undertaken on the unitary nature of the matrices $\mathbf{Y}(\mathbf{r}_k)$ and $\mathbf{Y}(\hat{\mathbf{r}}_L)$ as defined in Equation (7.42) at a value of $ka = 0.1$. These matrices of sampled spherical harmonics were found, for the sampling geometry illustrated in Figure 7-1, to be unitary to an extremely good approximation. The matrices $\mathbf{Y}^{\mathrm{H}}(\mathbf{r}_k)\mathbf{Y}(\mathbf{r}_k)$ and $\mathbf{Y}^{\mathrm{H}}(\hat{\mathbf{r}}_l)\mathbf{Y}(\hat{\mathbf{r}}_l)$ were found to be given by $(K/4\pi)\mathbf{I}$ and $(L/4\pi)\mathbf{I}$ respectively to a very good accuracy although the maximum diagonal term exceeded the theoretical value by 3.7%. The maximum off-diagonal term was 3% of the theoretical value of the diagonal terms. The elements of the matrix $\mathbf{Y}^{\mathrm{H}}(\mathbf{Y}_k)\mathbf{Y}(\mathbf{r}_k)$ are shown on a grey scale plot in Figure 7-4. It is interesting to note, however, that the sum of the diagonal terms in these matrices was *exactly* equal to the theoretical value. Similarly the sum of the off-diagonal terms was a factor 10^{-5} smaller than the theoretical value of the diagonal terms.

The matrices $\mathbf{T}(\mathbf{r}_k)$ and $\mathbf{T}(\hat{\mathbf{r}}_l)$ were then calculated. Since there were only $n = 4$ terms in the spherical harmonic series expansion and thus only 16 non-zero singular values, only the

first $N = 16$ columns of the matrices $\mathbf{U}$ and $\mathbf{V}$ were used in SVD expansion of the matrix $\mathbf{G}_N(\mathbf{r} \mid \hat{\mathbf{r}})$. Thus

$$\mathbf{G}_N(\mathbf{r} \mid \hat{\mathbf{r}}) = \mathbf{U}_N \, \boldsymbol{\Sigma}_N \, \mathbf{V}_N^{\mathrm{H}} \tag{7.53}$$

and the unitary transformation matrices are defined using Equations (7.49a, b). Pre-multiplication of these equations respectively by $\mathbf{Y}^{\mathrm{H}}(\mathbf{r}_k)$ and $\mathbf{Y}^{\mathrm{H}}(\hat{\mathbf{r}}_l)$ then shows that

$$\mathbf{T}(\mathbf{r}_k) = (4\pi / K)\mathbf{Y}^{\mathrm{H}}(\mathbf{r}_k)\mathbf{U}_N \quad , \quad \mathbf{T}(\hat{\mathbf{r}}_l) = (4\pi / L)\mathbf{Y}^{\mathrm{H}}(\hat{\mathbf{r}}_l)\mathbf{V}_N \tag{7.54a, b}$$

The results of these computations are shown in Figures 7-5 and 7-6 which show both the real and imaginary parts of these two matrices and their block diagonal structure. Note that the blocks of terms in these matrices are arranged in accordance with spherical harmonic order n, with submatrices of dimension 1×1, 3×3, 5×5 and 7×7 appearing in diagonal blocks. This in turn implies that the first left and right singular vectors are equal to the first columns of the matrices $\mathbf{Y}(\mathbf{r}_k)$ and $\mathbf{Y}(\hat{\mathbf{r}}_l)$ respectively (i.e. corresponding to the $n = 0$ spherical harmonic) whilst the second, third and fourth singular vectors are linear combinations of the second, third and fourth columns of $\mathbf{Y}(\mathbf{r}_k)$ and $\mathbf{Y}(\hat{\mathbf{r}}_l)$ corresponding to the $n = 1$ spherical harmonics, and so forth. A plot of the elements of the matrix $\mathbf{T}(\mathbf{r}_k)^{\mathrm{H}}\mathbf{T}(\mathbf{r}_k)$ is shown in Figure 7-7 which confirms the unitary structure of the matrix $\mathbf{T}(\mathbf{r}_k)$.

Finally, a check was undertaken to ensure that the matrix $\boldsymbol{\Sigma}_N$ of singular values was related to the matrix $\mathbf{F}$ of complex spherical harmonic amplitudes through the relationship $\mathbf{F} = \mathbf{T}(\mathbf{r}_k)\boldsymbol{\Sigma}_N \, \mathbf{T}^{\mathrm{H}}(\hat{\mathbf{r}}_l)$. It was found that this relationship holds to an excellent approximation. Figure 7-8 shows on a logarithmic scale the elements of $\mathrm{Re}\{\mathbf{T}(\mathbf{r}_k)\boldsymbol{\Sigma}_N \, \mathbf{T}^{\mathrm{H}}(\mathbf{r}_l)\}$ computed from this expression. The results are in agreement with

those deduced from the analytical solution to within the precision of the numerical calculation.

7.3.3 Sensitivity to choice of mesh

A further set of numerical simulations were undertaken that were based on the above analytical model, but with non-uniformly sampled spherical surfaces. The surface mesh used is illustrated in Figure 7-9. This form of sampling was applied both to the surface of the rigid sphere and on the far field surface. In this case, the matrices $\mathbf{Y}(\mathbf{r}_k)$ and $\mathbf{Y}(\hat{\mathbf{r}}_l)$ were no longer found to be unitary. A plot showing the real part of $\mathbf{Y}^{\mathrm{H}}(\mathbf{r}_k)\mathbf{Y}(\mathbf{r}_k)$ is shown in Figure 7-10 which demonstrates that although the diagonal terms are still dominant, the diagonal terms are a maximum of 19.1% above the theoretical value, although their sum is again exactly equal to the theoretical value. The off-diagonal terms are a maximum of 24.7% of the theoretically predicted diagonal terms. However, the *sum* of the off diagonal terms was 0.05% of the theoretically predicted values of the diagonal terms. Similarly, plots of the real and imaginary parts of $\mathbf{T}(\mathbf{r}_k)$ for this case are shown in Figure 7-11 and a plot of the real part of $\mathbf{T}(\mathbf{r}_k)^{\mathrm{H}}\mathbf{T}(\mathbf{r}_k)$ is shown in Figure 7-12. Also a plot of the real part of the matrix product $\mathbf{T}(\mathbf{r}_k)\sum_N \mathbf{T}^{\mathrm{H}}(\hat{\mathbf{r}}_l)$ is shown in Figure 7-13. In this case the agreement with the results for the analytical solution for $\mathrm{Re}\{\mathbf{F}\}$ is far less good.

A plot of the 16 dominant singular values is shown in this case as a function of *ka* in Figure 7-14. Again 5 terms (up to n=4) were used in the spherical harmonic series and 25 dominant singular values were yielded by the SVD. The smallest singular values were again found to be a factor 10^{-9} times the smallest dominant singular value. However, it is clear that the singular values are no longer grouped in spherical harmonic order as convincingly as the case with perfectly uniform sampling of the field. This therefore

illustrates the sensitivity of the choice of mesh to the relationship between the 'source modes' and 'field modes' comprising the columns of **V** and **U** and which is quantified by the relevant singular value.

7.4 Numerical simulation of a sphere and ellipsoid using the BEM

The numerical simulations undertaken below were undertaken using the DBEM. As an initial verification of the software, a numerical solution was computed for the case of a sphere whose surface was sampled as illustrated in Figure 7-15 with 152 vertices. The SVD of the resulting Green function matrix was then undertaken and calculated using MATLAB (MathWorks, Inc.). The results are illustrated in Figure 7-16. The agreement with the results for the sphere sampled as shown in Figure 7-9 was found to be excellent (compare Figure 7-14 with Figure 7-16). Although a different number of nodes in each sphere was used in the analytical and numerical calculations, the deviation of the curves from the case where the nodes are distributed uniformly (Figure 7-2) is similar.

The singular vectors presented in Figure 7-3 were calculated at low frequencies. Similar patterns and scales were obtained for a higher frequency (f=1 kHz, $ka=1.8$). It is not clear at this stage if at higher frequencies, the patterns are distorted or not. Due to long computational times the calculation was executed for low frequencies only.

7.4.1 The source and field mode shapes of a rigid ellipsoid

As a further illustration of the utility of the SVD in analysing numerical solutions, a further simulation was undertaken numerically in order to deduce the 'field mode shapes' and 'source mode shapes' associated with an ellipsoid and a far field spherical surface. The ellipsoidal geometry illustrated in Figure 7-17 was used and the far field sphere was sampled using the same geometry as that illustrated in Figure 7-15.

The behaviour of the singular values is shown in Figure 7-18, where there were subtle departures from the behaviour seen in the case of a sphere (Figure 7-14 and Figure 7-16). Most notably, the groups of spherical harmonic orders become more spread as the individual singular values depart more from the frequency dependence found with the sphere. In view of the evident sensitivity of the results shown above to the choice of mesh, it is difficult at this stage to be sure to what extent the basic change of geometry influences the results of the SVD. The mode shapes associated with both source and field are presented in Figures 7-19 to 7-28. The calculations were undertaken at two frequencies in order to investigate the consistency of the patterns and values with increasing frequencies and the extent to which the singular vectors associated with lower order singular value are changed. Figure 7-19 and 7-20 show the 'monopole-like' real and imaginary parts of the singular vectors at 500 Hz (ka=0.9) and 1 kHz (ka=1.8), respectively. No attempt was made to fix the scale and variation of colours, so although specific patterns are seen, the interpretation should be in accordance to the scale. The real part of the singular vector on the surface of the ellipsoid is negative with only a small variation (9.7%). The corresponding imaginary value is approximately 500 times smaller, and is in agreement with the results obtained in the case of radiation from a sphere (see Figure 7-3. Only the real part of the left singular vectors of the analytically derived 32 × 32 Green function matrix is shown since the imaginary values are zero). In fact, in all the following figures the imaginary part of the singular vectors on the surface of the ellipsoids are smaller by a factor of between 50 and 2500 than the real part of the same singular vectors. Also, the real and imaginary singular vectors in the far field on the surface of the sphere shown in Figure 7-19 are almost omnidirectional where in both cases only positive values appear with variation in the order of 8% from the maximum values. These variations are a result of the 'distortion' of the shape of the sphere.

As frequency increases, similar patterns appear but these are not frequency independent (the scales have different values). Note also the change of the sign of the imaginary part of the singular vectors at 1 kHz compared with the equivalent singular vectors at 500 Hz. This is a result of the change of the phase with frequency for a given distance between the sources and the points on the surface.

Figures 7-21 to 7-26 present the three 'dipole' shapes. In all figures the variation of the positive to negative values occur at the same direction of the dipole axis as found in Figure 7-3 for the case of the sphere. Note that all scales are symmetrical with respect to the zero values (i.e. in each scale the absolute value of the maximum and minimum is identical). The variation in the scale for the real part of the singular vectors for the cases of 500 Hz and 100 Hz is in the order 5-10% and is different for each dipole order since the axes of the ellipsoid have different lengths (see Figure 7-17). Since the imaginary part of the singular vectors have much smaller values, the change of their amplitudes with frequency is higher.

Finally, the first 'quadrupole' singular vectors are presented in Figures 7-27 and 7-28. Even for the fifth singular values, the mode shapes are distorted only slightly with the scales. As before the phase changes but all patterns are similar to the quadrupole patterns obtained with the SVD operation.

The main difference between radiation from an ellipsoid and radiation from a sphere is that for the ellipsoid, the singular vectors (the 'mode shapes') are frequency dependent. More complex cases will be investigated in Chapter 8 which have more distinctive resonance frequencies.

7.5 EXTRACTION OF HRTFs BASED ON THE SVD METHOD

Note that it is possible using the SVD formulation to reconstruct the frequency response from a source in a particular position p_n and a field point in another position q_m. Based on Equation (7.33), the vector $\mathbf{p}$ can be given by the product

$$\begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \\ \vdots \\ p_N \end{bmatrix} = \begin{bmatrix} u_{11} & u_{21} & \cdots & u_{n1} & \cdots & u_{N1} \\ u_{12} & u_{22} & \cdots & u_{n2} & \cdots & u_{N2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ u_{1n} & u_{2n} & \cdots & u_{nn} & \cdots & u_{Nn} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ u_{1N} & u_{2n} & \cdots & u_{nN} & \cdots & u_{NN} \end{bmatrix} \begin{bmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_n & & \\ & & & & \ddots & \\ & & & & & \sigma_N \end{bmatrix}$$

$$\times \begin{bmatrix} v_{11}^* & v_{12}^* & \cdots & v_{1n}^* & \cdots & v_{1N}^* \\ v_{21}^* & v_{22}^* & \cdots & v_{2n}^* & \cdots & v_{2N}^* \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ v_{m1}^* & v_{m2}^* & \cdots & v_{mn}^* & \cdots & v_{mN}^* \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ v_{M1}^* & v_{M2}^* & \cdots & v_{MN}^* & \cdots & v_{MN}^* \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_m \\ \vdots \\ q_M \end{bmatrix} \tag{7.55}$$

For the case of an excitation due to a specific source, we assume only q_m is non-zero, then

$$\begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \\ \vdots \\ p_N \end{bmatrix} = \begin{bmatrix} u_{11} & u_{21} & \cdots & u_{n1} & \cdots & u_{N1} \\ u_{12} & u_{22} & \cdots & u_{n2} & \cdots & u_{N2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ u_{1n} & u_{2n} & \cdots & u_{nn} & \cdots & u_{Nn} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ u_{1N} & u_{2n} & \cdots & u_{nN} & \cdots & u_{NN} \end{bmatrix} \begin{bmatrix} \sigma_1 v_{1m}^* q_m \\ \sigma_2 v_{1m}^* q_m \\ \vdots \\ \sigma_n v_{nm}^* q_m \\ \vdots \\ \sigma_N v_{Nm}^* q_m \end{bmatrix} \tag{7.56}$$

or, in a compact form

$$p_n = \sum_{n=1}^{N} \sigma_n u_{nn} v_{nm}^* q_m \tag{7.57}$$

Thus the summation can be limited to a number which is less than N, if the last terms of the series of σ_n are much smaller than the first terms.

In figure 7-29a and 7-29b, the frequency responses detected in the front ($\phi = 0°, \theta = 0°$) and at the rear ($\phi = 180°, \theta = 0°$) of a sphere, when the source is positioned at a distance of 3 m at $\phi = 0°, \theta = 0°$ are presented, respectively. Since the variations between the largest and the smallest singular values is decreasing as frequency increases, the contributions from low order singular values is not negligible. The first 20 terms in the summation are sufficient if a maximum of ka=1 (f≈550 Hz, at a=0.1 m) is required, For higher frequencies, up to ka=2, 50 terms are required.

For this case, the reconstruction of the frequency response when a large frequency range is required, using the SVD method, is clearly not efficient. As will be demonstrated in the next chapter, the method is efficient when the radiating or scattering body is characterised by a strong resonance.

7.6 Conclusions

A number of numerical simulations have been undertaken in order to examine the connection between the basis functions provided by the singular value decomposition of matrices of acoustic transfer functions and the basis functions provided by classical acoustical analyses. In particular, it has been found that for radiation from the surface of a sphere to a spherical surface in the far field, the left and right singular vectors associated with the SVD are related to the sampled spherical harmonics by a unitary transformation. This relationship is dependent on the source and field being sampled by dividing the spherical surfaces involved into segments of equal area. The relationship is found not to hold for non-uniform sampling of the surfaces involved.

When the SVD is investigated on a more complex geometry such an ellipsoid, the deviation of the results of the singular values from the spherical harmonics is noticeable, and the corresponding singular vectors become frequency dependent.

Since the singular values are not characterised for these two cases with a resonance at a particular frequency, the reconstruction of a frequency response for a given source position and a point on the sphere is successful only at low frequencies when a limited number of singular values are used in the summation.

In the next chapter the method is implemented numerically on complex geometries (such as the human pinna) which are known to have clear resonance.

Figure 7-1: A spherical surface sampled at 32 uniformly distributed points (forming an Icosahedron). Each vertex is used to position sources on the surface, and field points in the far field.

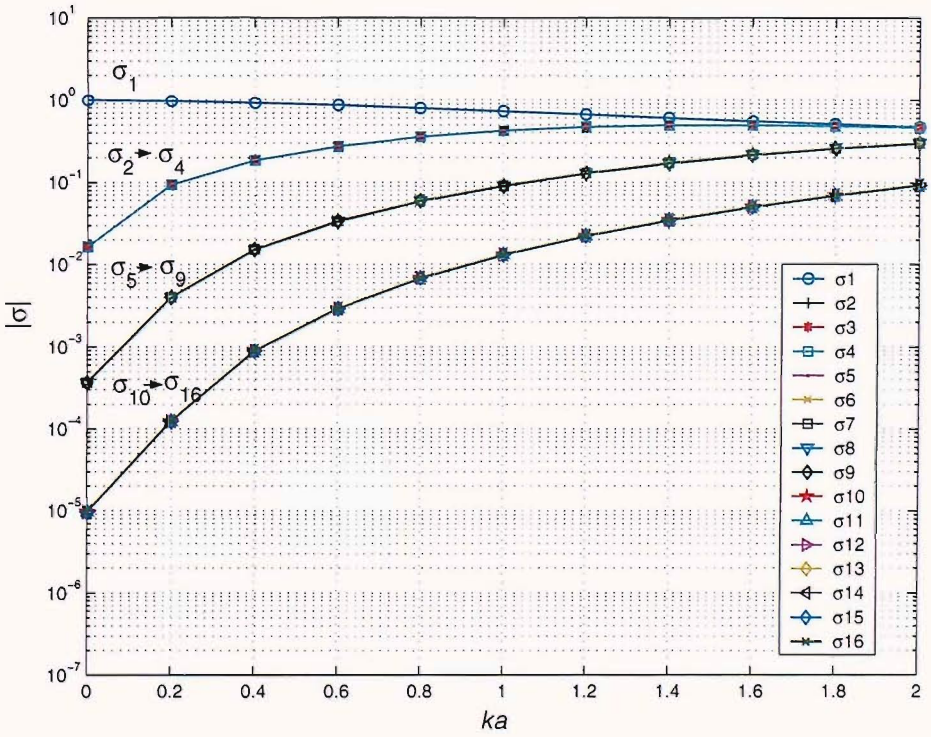


Figure 7-2: The singular values of the analytically generated 32 × 32 Green function matrix relating points on the surface of a rigid sphere of radius 0.1 m to points on the surface of a far field surface of radius 10 m. Both spherical surfaces are based on the mesh illustrated in Figure 7-1.

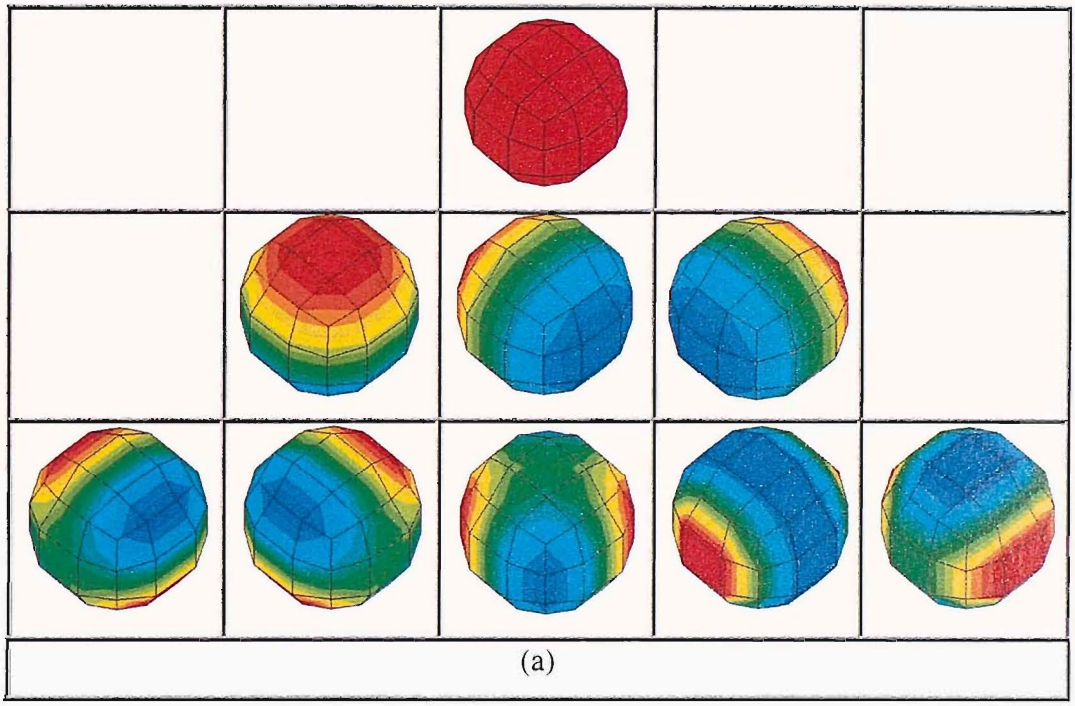
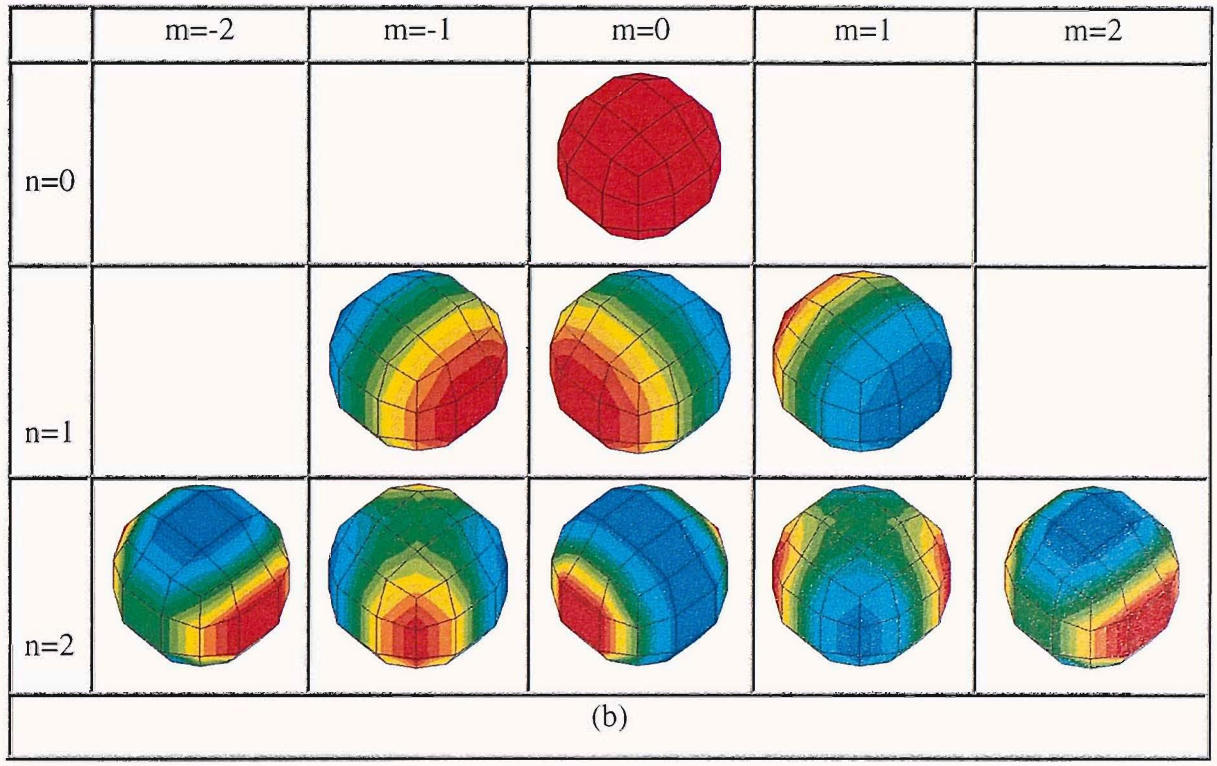


Figure 7-3: Three dimensional colour maps illustrating (a) the variation over the far field sphere of the normalised real part of the left singular vectors of the analytically generated 32 × 32 Green function matrix (the plots shown correspond to the nine most dominant singular values). Values in red correspond to +1, and values in blue correspond to -1.



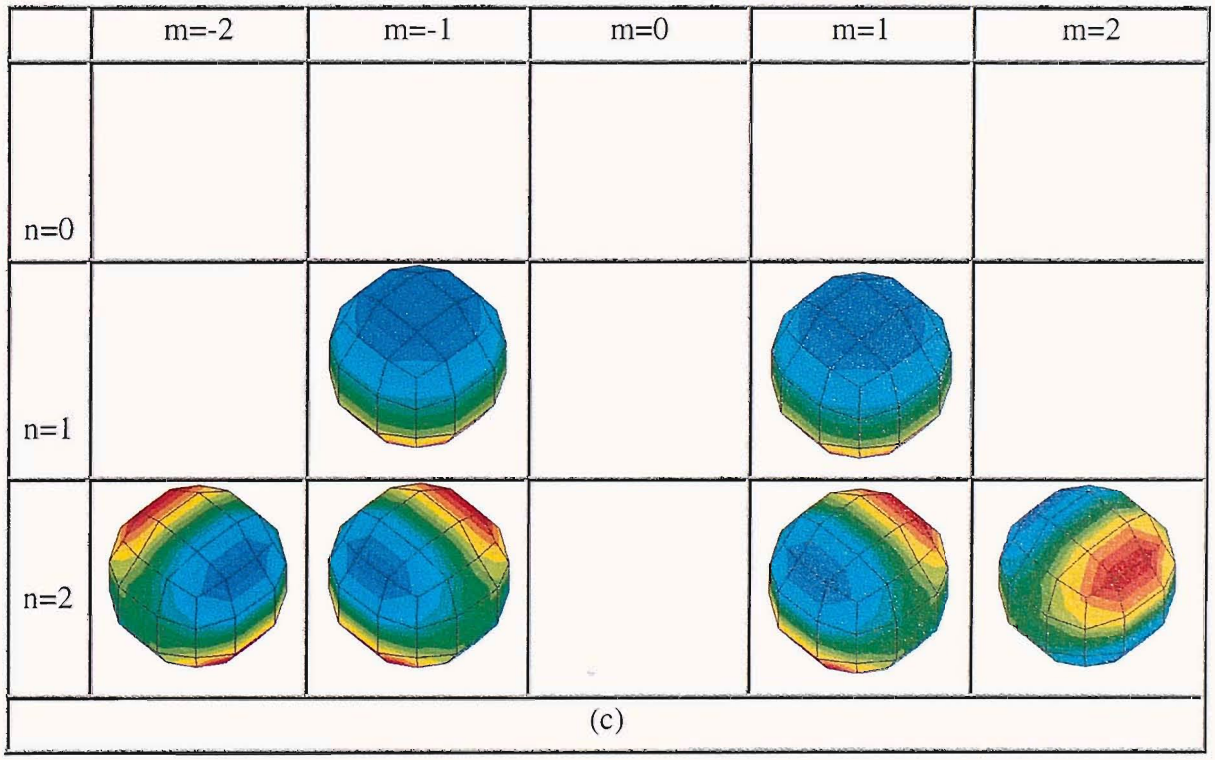


Figure 7-3 (cont.): Three dimensional colour maps illustrating (b) the normalised real part of the spherical harmonics evaluated over the same spherical surface, (c) the corresponding normalised imaginary parts of the spherical harmonics. Values in red correspond to +1, and values in blue correspond to -1.

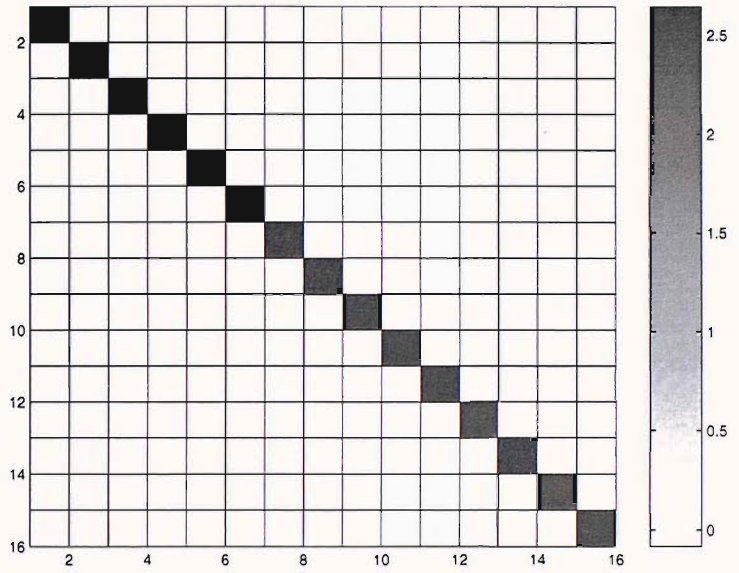
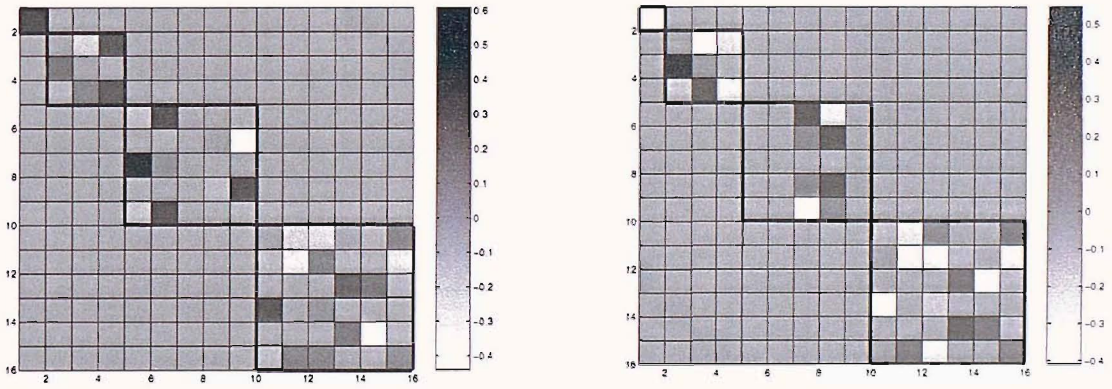
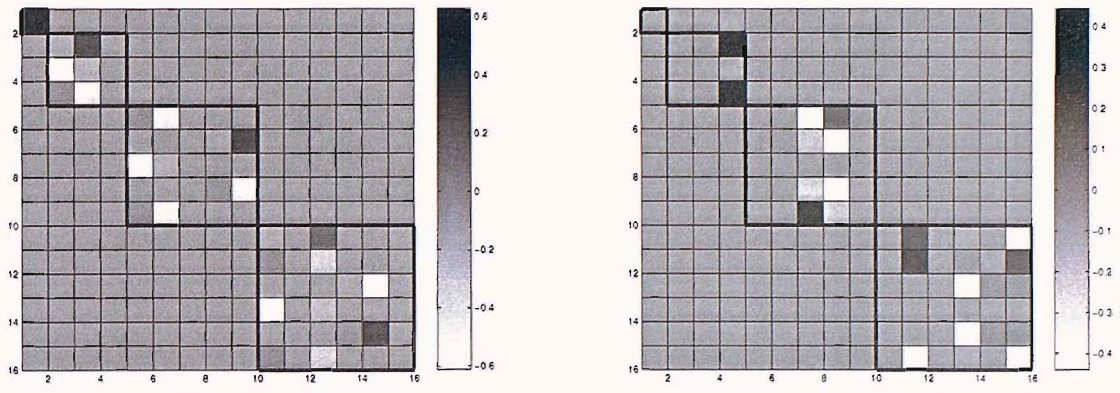


Figure 7-4: A grey scale plot of the values of the real parts of the elements of the matrix $\mathbf{Y}(\mathbf{r}_k)^{\mathrm{H}}\mathbf{Y}(\mathbf{r}_k)$ where the values of $\mathbf{r}_k$ chosen correspond to 32 uniformly distributed points on the sphere.



(a) (b)

Figure 7-5: A grey scale plot of the (a) real and (b) imaginary parts of the elements of the matrix $\mathbf{T}(\mathbf{r}_k)$ associated with the analytically generated 32 × 32 Green function matrix.



(a) (b)

Figure 7-6: A grey scale plot of the (a) real and (b) imaginary parts of the elements of the matrix $\mathbf{T}(\mathbf{r}_l)$ associated with the analytically generated 32 × 32 Green function matrix.

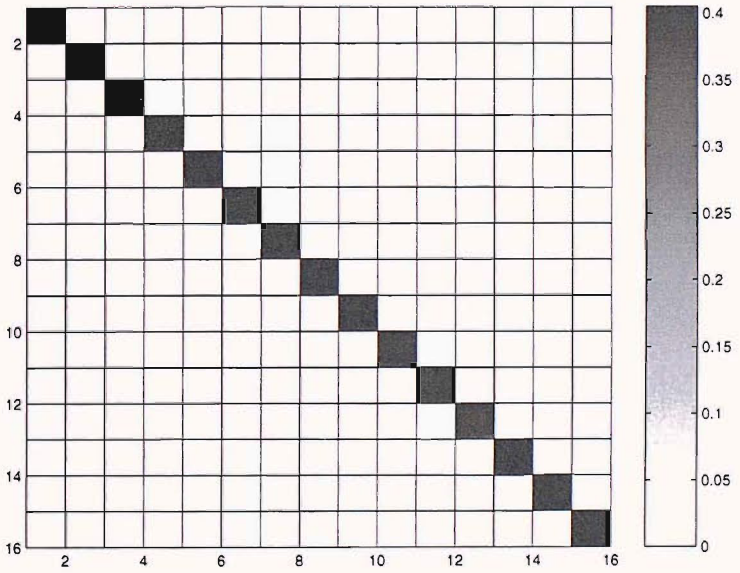


Figure 7-7: A grey scale plot of the real parts of the elements of the matrix $\mathbf{T}(\mathbf{r}_k)^{\mathrm{H}}\mathbf{T}(\mathbf{r}_k)$ associated with the analytically generated 32 × 32 Green function matrix.

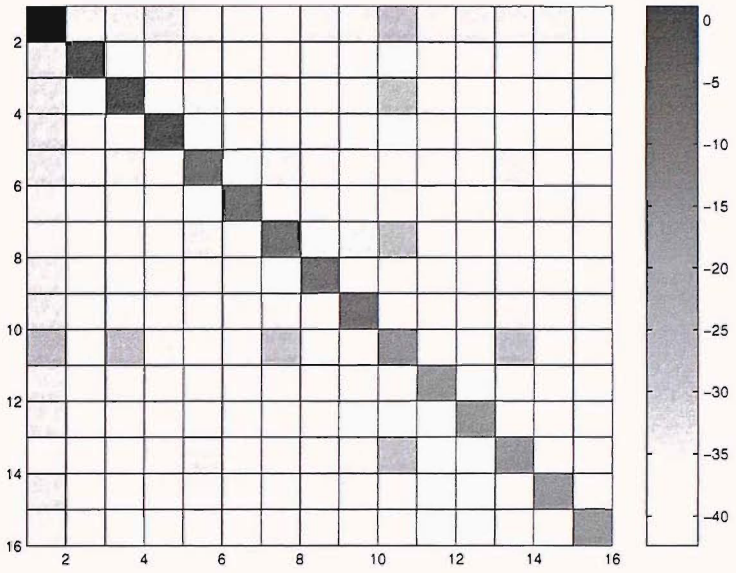


Figure 7-8: A grey scale plot (on a logarithmic scale) of the real parts of the elements of the matrix $\mathbf{T}(\mathbf{r}_k)\boldsymbol{\Sigma}_N\,\mathbf{T}(\hat{\mathbf{r}}_l)^{\mathrm{H}}$ associated with the 32 × 32 analytically generated Green function matrix.

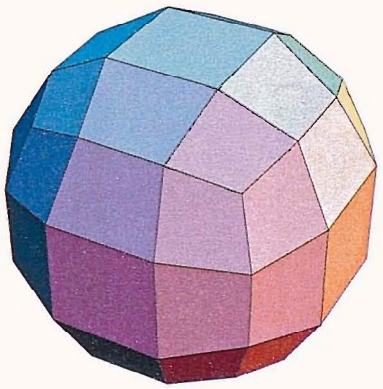


Figure 7-9: A spherical surface sampled at 56 locations with an approximately uniform distribution of points. The points used are at the vertices of the quadrilateral mesh elements.

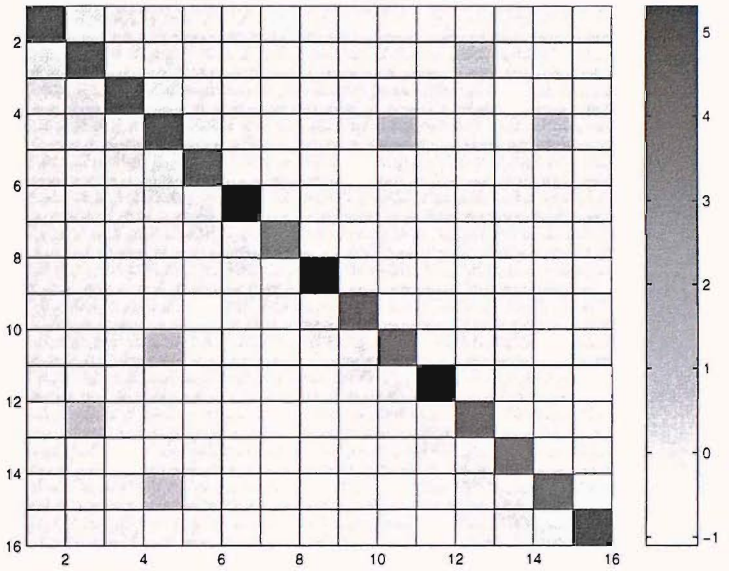
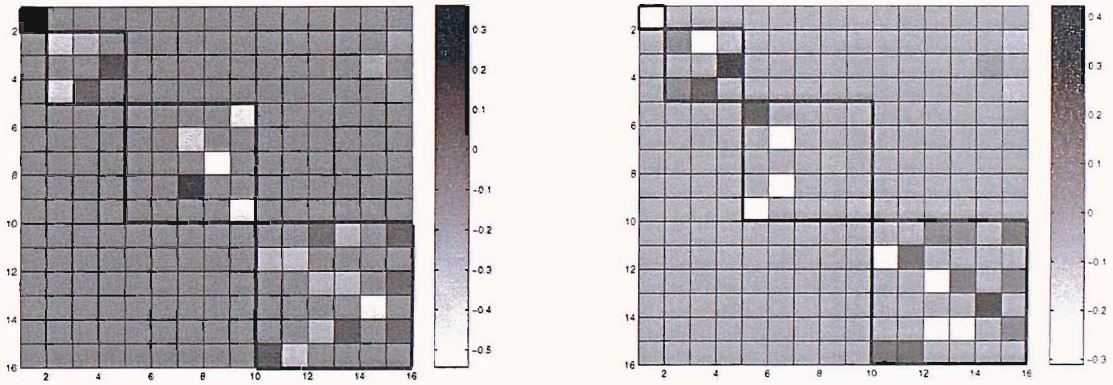


Figure 7-10: A grey scale plot of the values of the real parts of the elements of the matrix $\mathbf{Y}(\mathbf{r}_k)^H \mathbf{Y}(\mathbf{r}_k)$ where the values of $\mathbf{r}_k$ chosen correspond to 56 points distributed on the far field of a sphere in the manner illustrated in Figure 7-9.



(a) (b)

Figure 7-11: A grey scale plot of the (a) real and (b) imaginary parts of the elements of the matrix $\mathbf{T}(\mathbf{r}_k)$ associated with the analytically generated 56 × 56 Green function matrix.

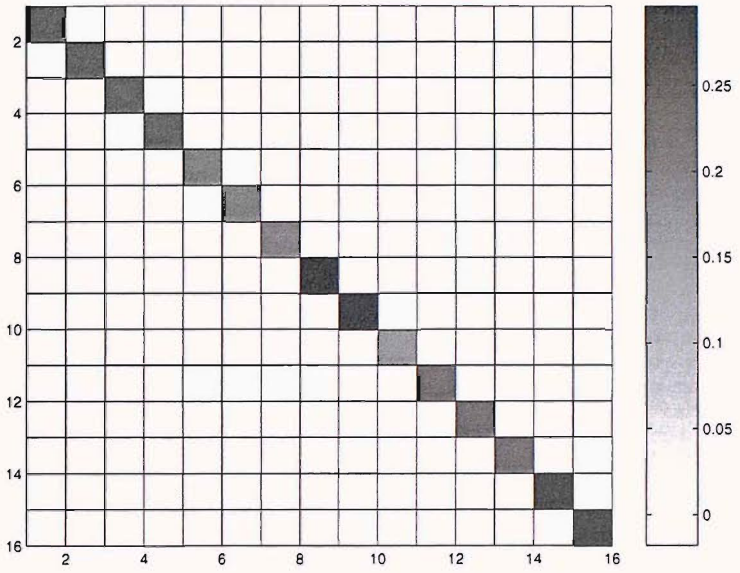


Figure 7-12: A grey scale plot of the real parts of the elements of the matrix $\mathbf{T}(\mathbf{r}_k)^{\mathrm{H}}\mathbf{T}(\mathbf{r}_k)$ associated with the analytically generated 56 × 56 Green function matrix.

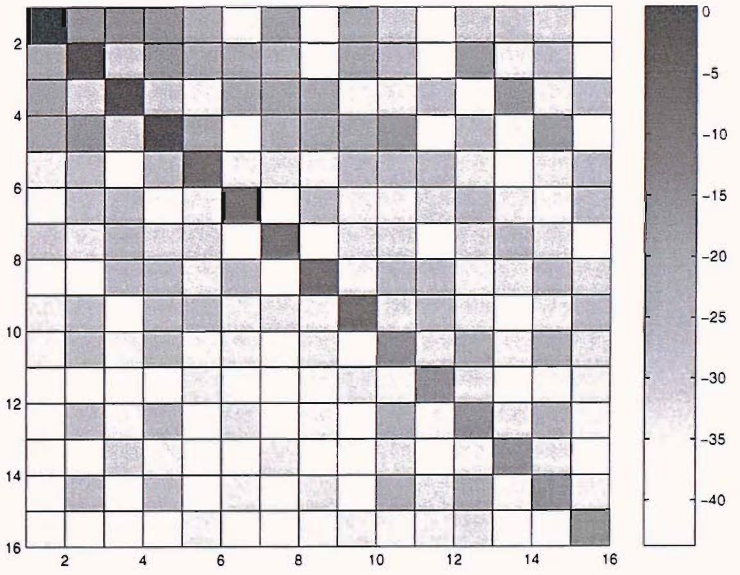


Figure 7-13: A grey scale plot (on a logarithmic scale) of the real parts of the elements of the matrix $\mathbf{T}(\mathbf{r}_k)\sum_N \mathbf{T}(\hat{\mathbf{r}}_l)^{\mathrm{H}}$ associated with the analytically generated 56 × 56 Green function matrix.

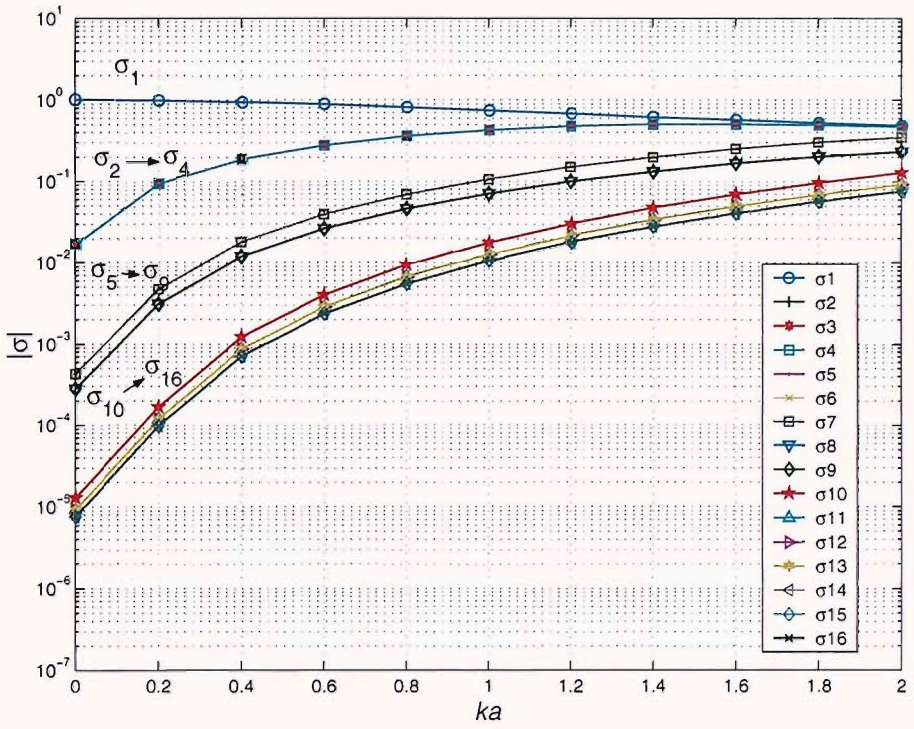


Figure 7-14: The singular values of the analytically generated 56 × 56 Green function matrix relating points on the surface of a rigid sphere of radius *a* (0.1 m) to points on the surface of a far field surface of radius 100 *a* (10 m). Both spherical surfaces were sampled using the mesh illustrated in Figure 7-9.

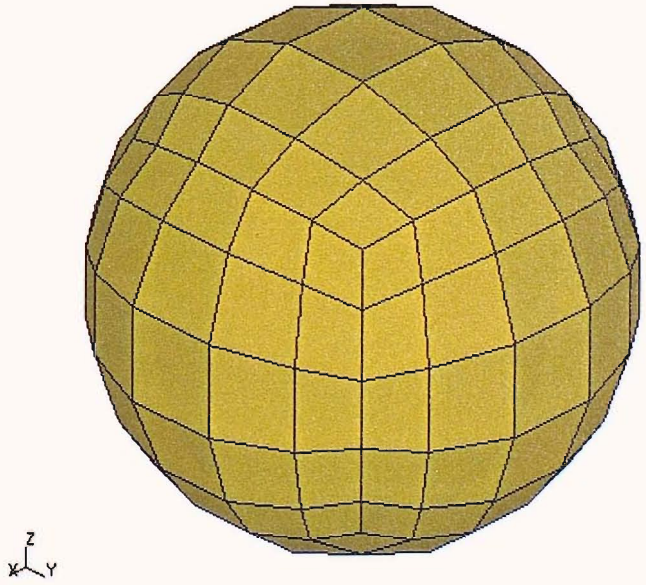


Figure 7-15: A spherical surface sampled at 152 approximately uniformly distributed points. The points used are at the vertices of the quadrilateral elements.

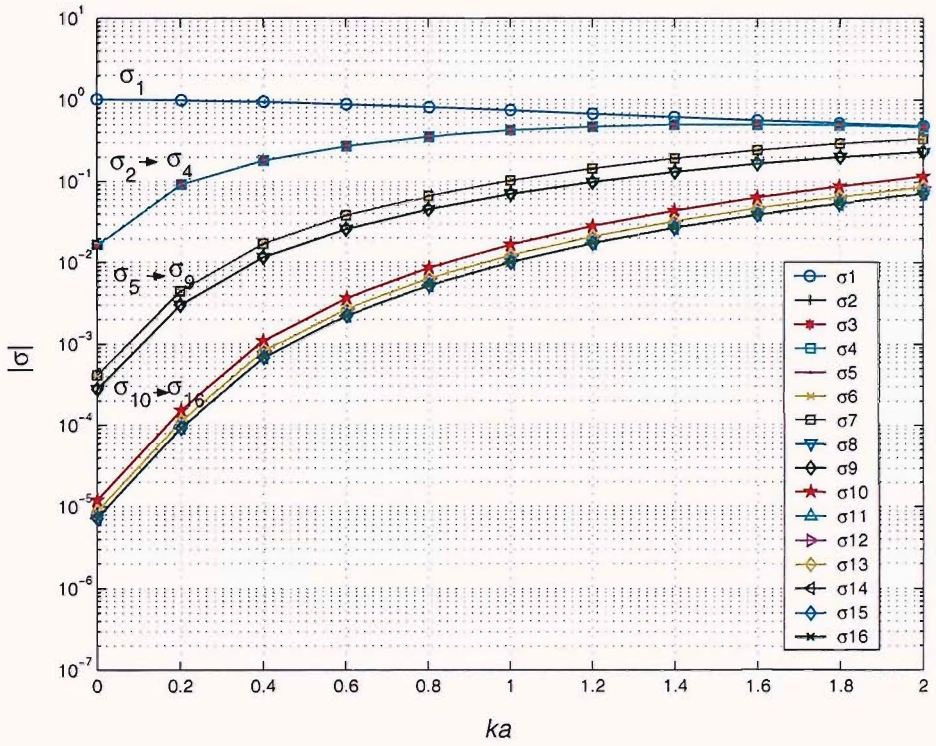


Figure 7-16: The singular values of the numerically generated 152 × 152 Green function matrix relating points on the surface of a rigid sphere of radius *a* (0.1 m) to points on a far field spherical surface of radius 100 *a* (10 m). Both spherical surfaces were sampled using the mesh illustrated in Figure 7-15.

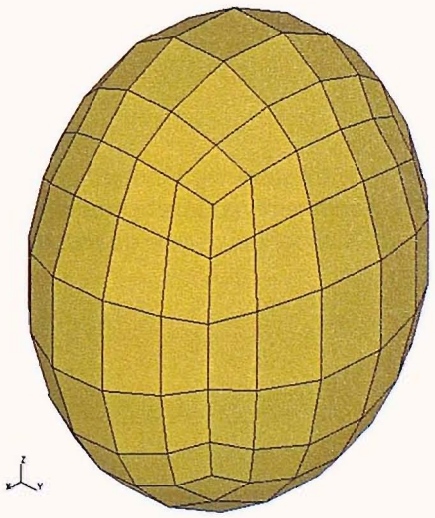


Figure 7-17: An ellipsoidal surface sampled at 152 approximately uniformly distributed points. The points used are at the vertices of the quadrilateral elements. The ellipsoid has semi-axes of dimensions a_x = 9.6 cm, a_y = 7.9 cm, a_z = 11.6 cm.

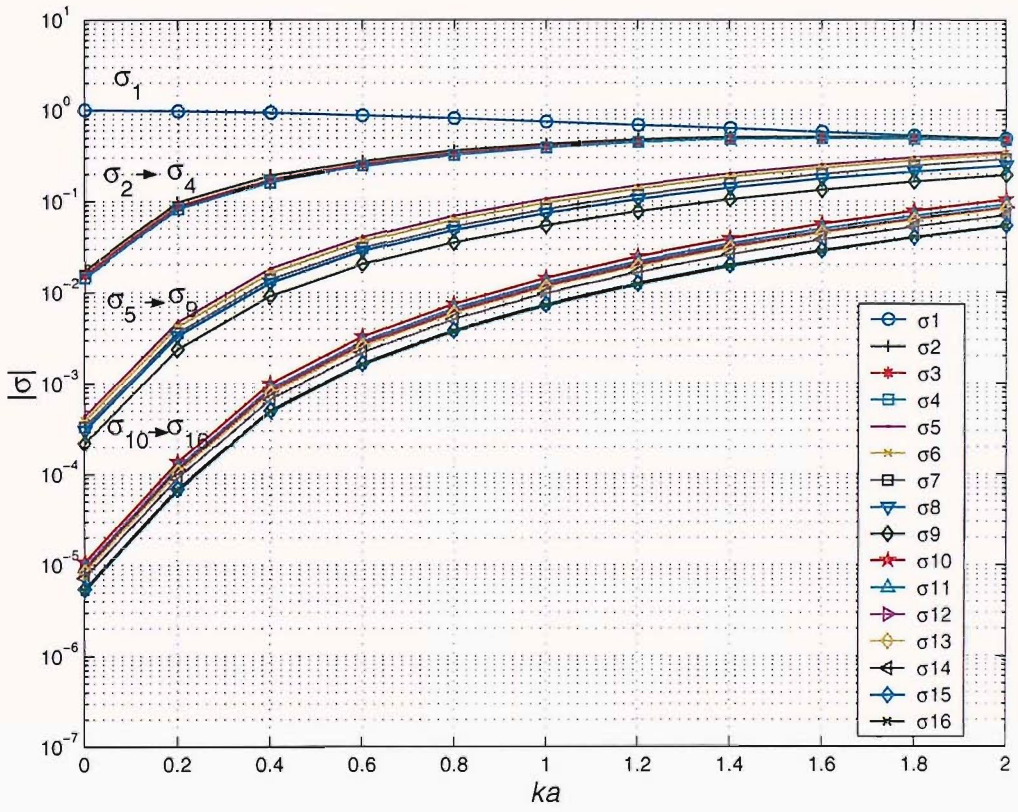


Figure 7-18: The singular values of numerically generated 152 x 152 Green function matrix relating points on the surface of a rigid ellipsoid (distributed as shown in Figure 7-17) to points on the surface of a far field sphere (distributed as shown in Figure 7-15) with a radius of 104 a_x (10 m).

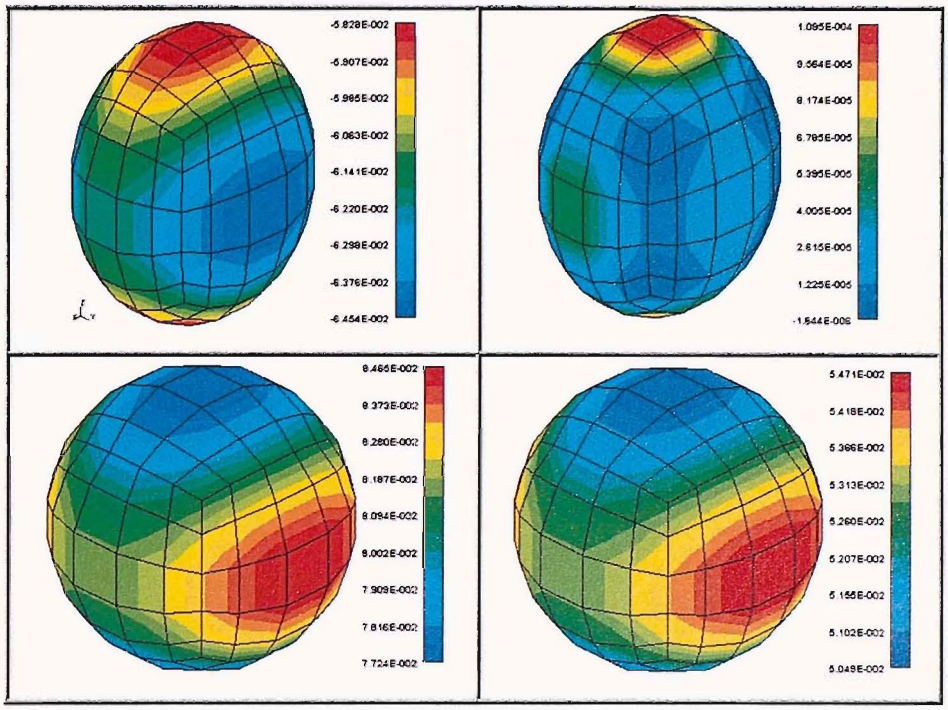


Figure 7-19: The real part of the singular vectors of the 'source points' on the ellipsoid, and the 'field points' on the sphere (with a radius of 3 m). The results shown are associated with the first singular value - σ_1 at 500 Hz.

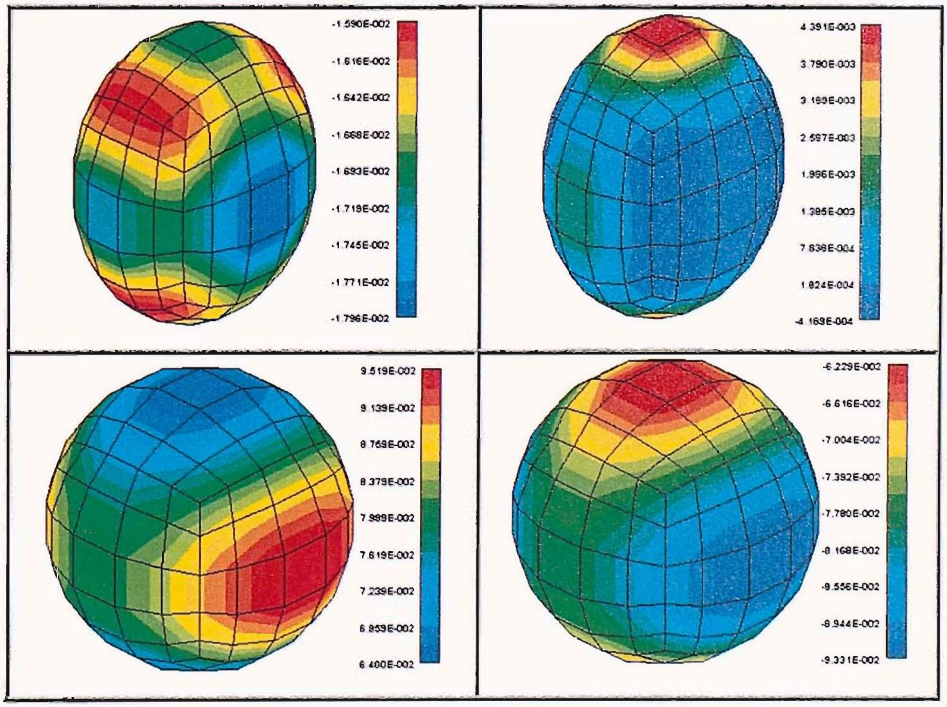


Figure 7-20: As in Figure 7-20 but the singular vectors are associated with σ_1 at 1 kHz.

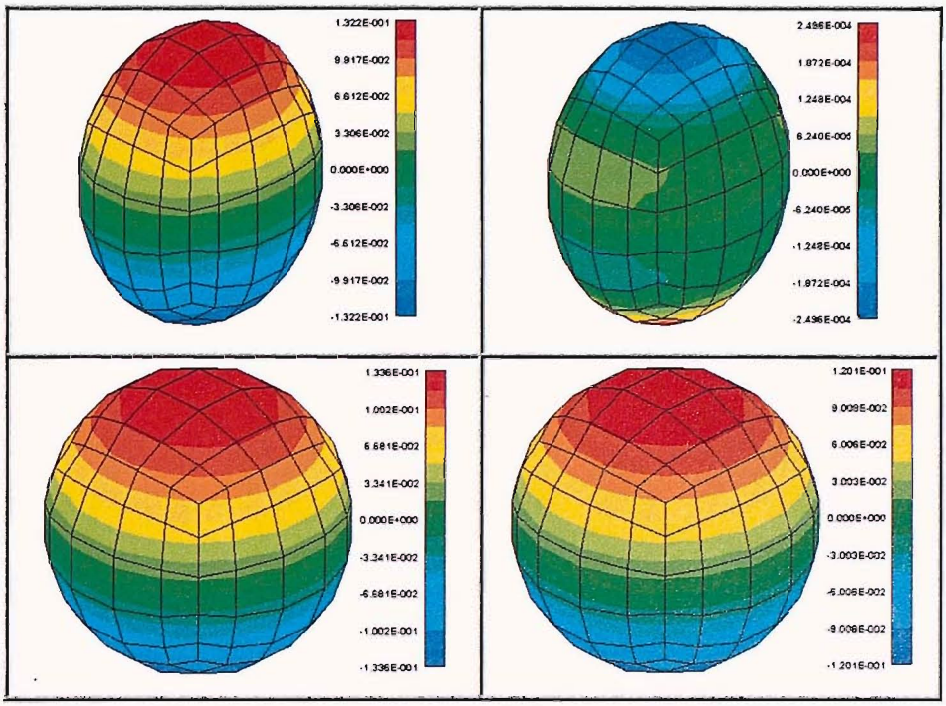


Figure 7-21: As in Figure 7-20 but the singular vectors are associated with σ_2 at 500 Hz.

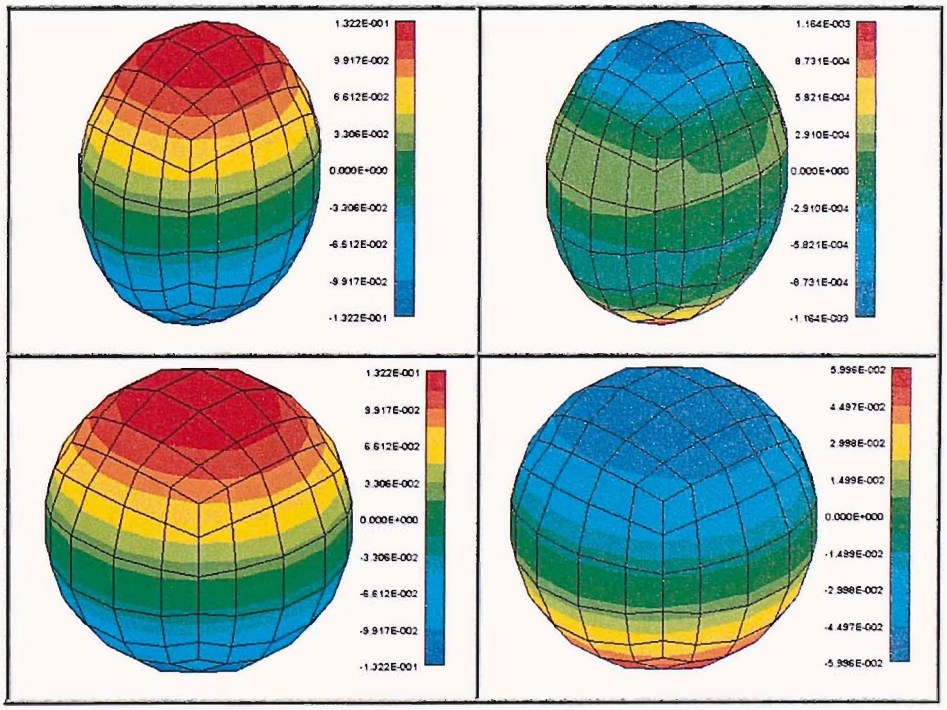


Figure 7-22: As in Figure 7-20 but the singular vectors are associated with σ_2 at 1 kHz.

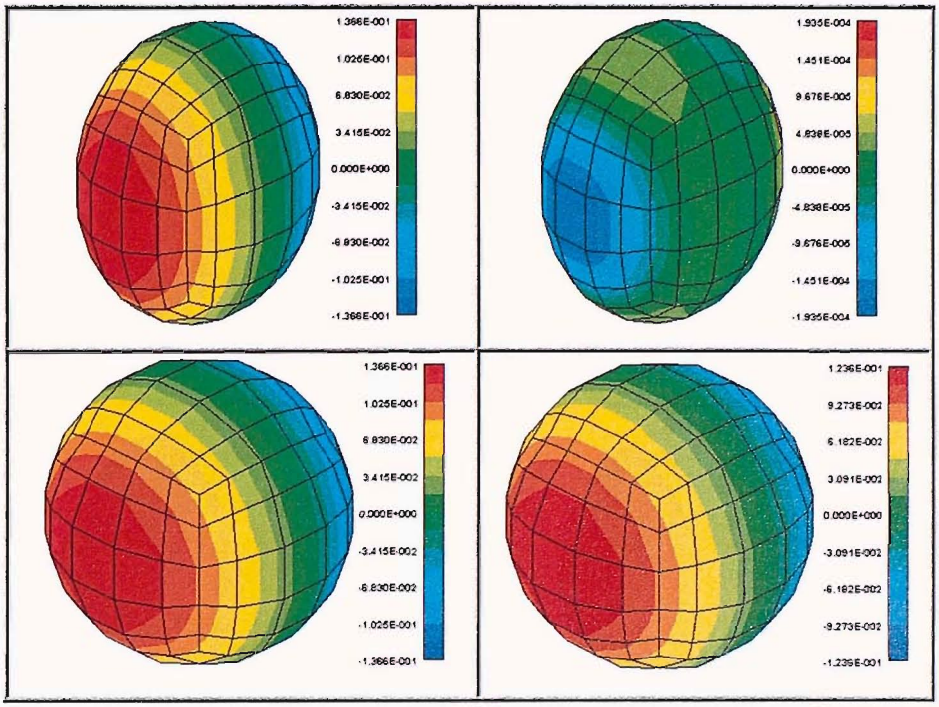


Figure 7-23: As in Figure 7-20 but the singular vectors are associated with σ_3 at 500 Hz.

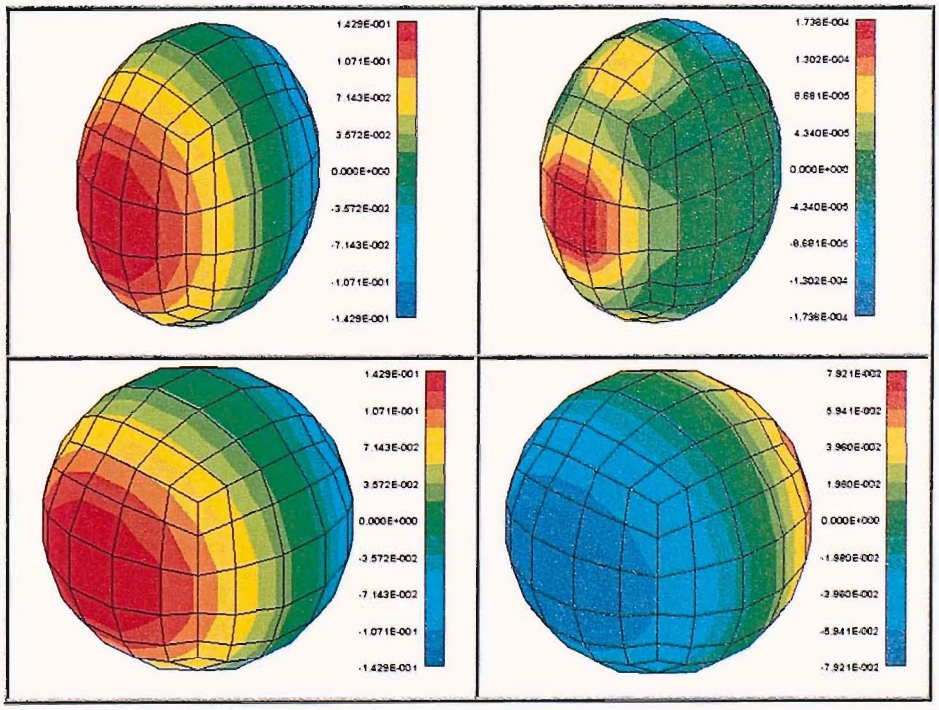


Figure 7-24: As in Figure 7-20 but the singular vectors are associated with σ_3 at 1 kHz.

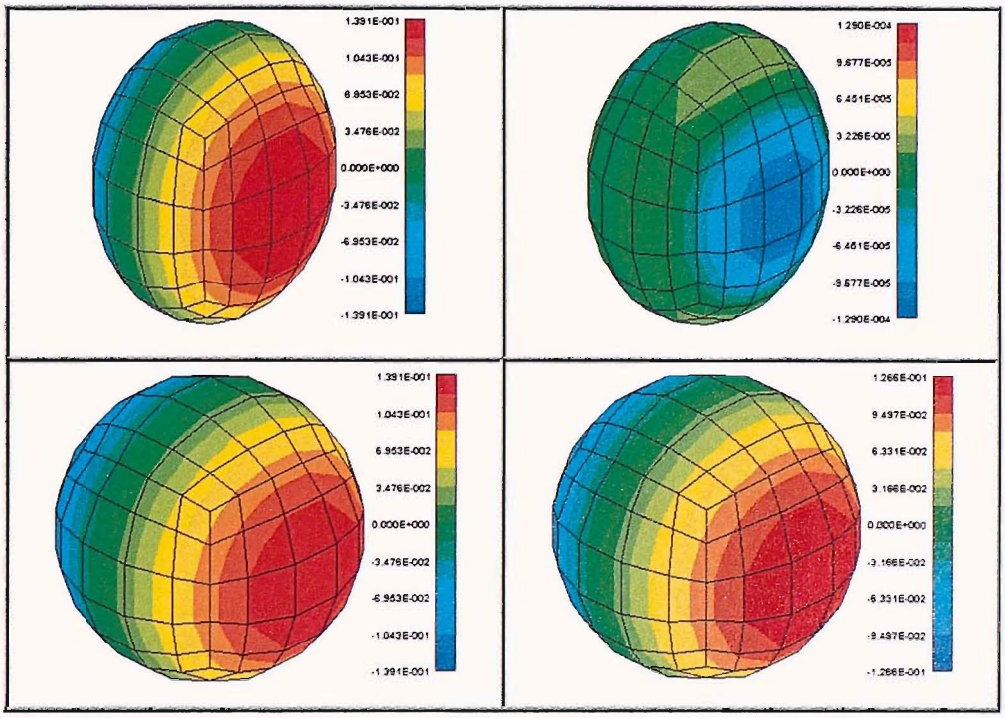


Figure 7-25: As in Figure 7-20 but the singular vectors are associated with σ_4 at 500 Hz.

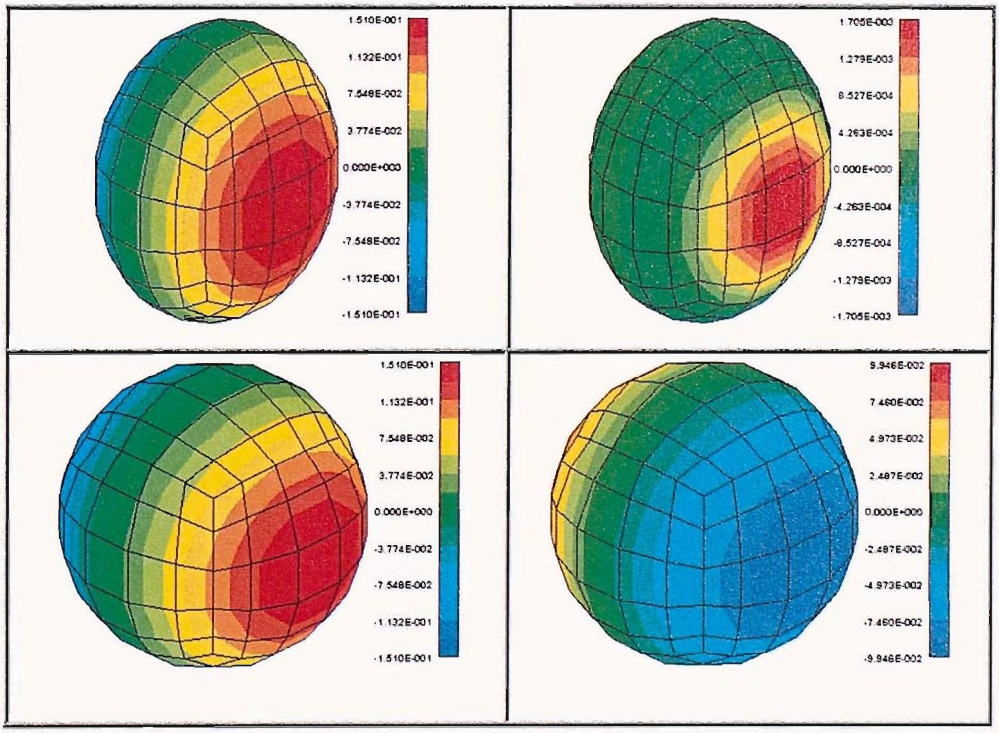


Figure 7-26: As in Figure 7-20 but the singular vectors are associated with σ_4 at 1 kHz.

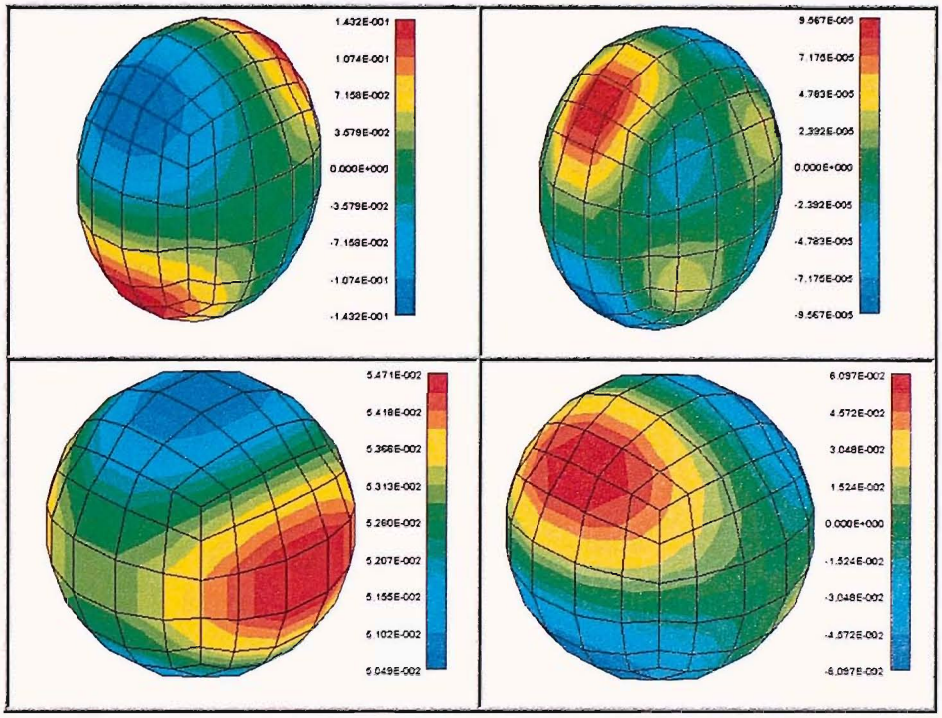


Figure 7-27: As in Figure 7-20 but the singular vectors are associated with σ_5 at 500 Hz

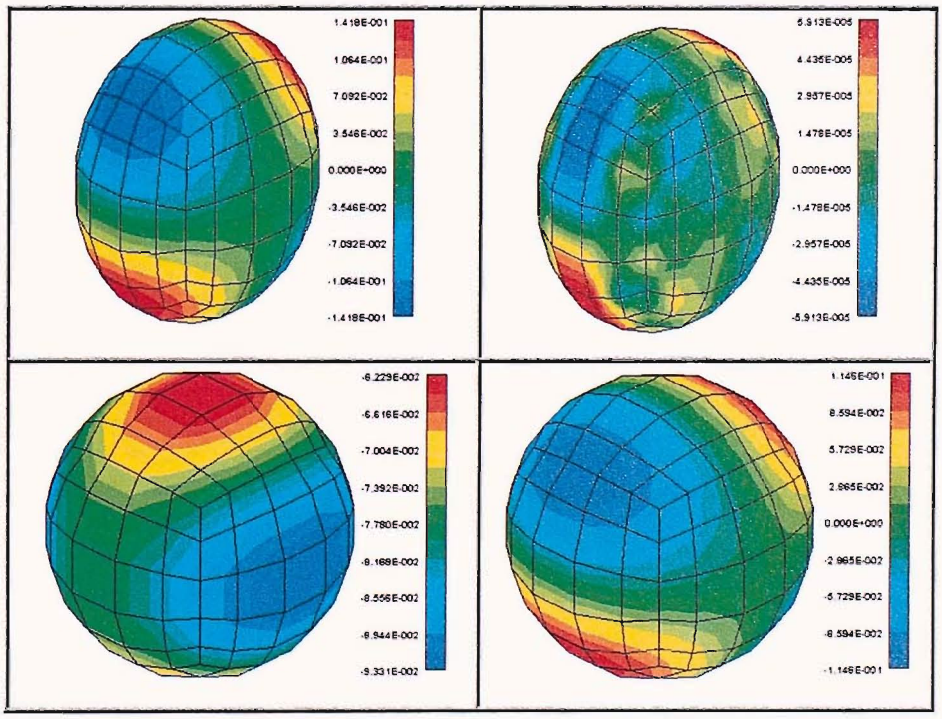
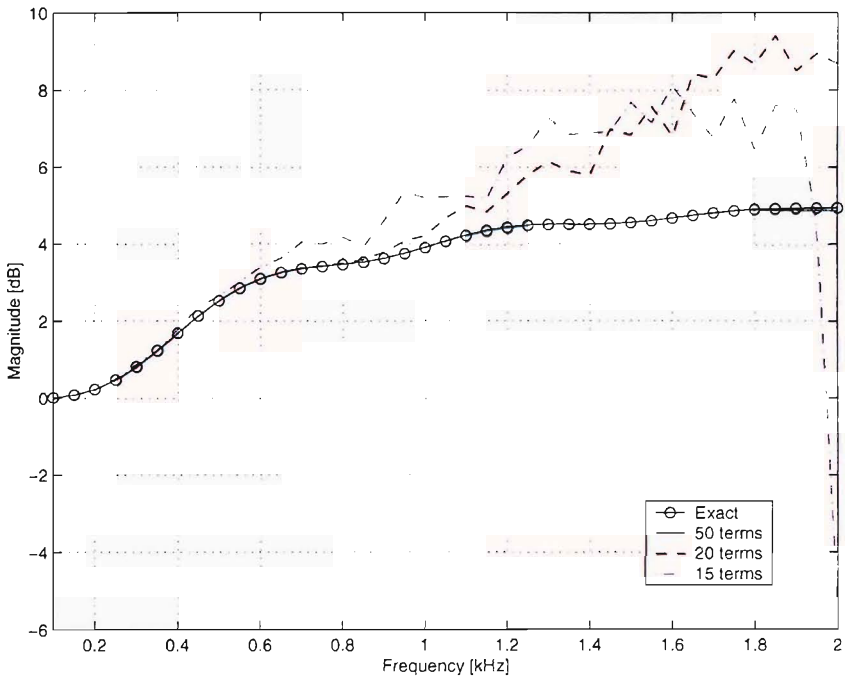
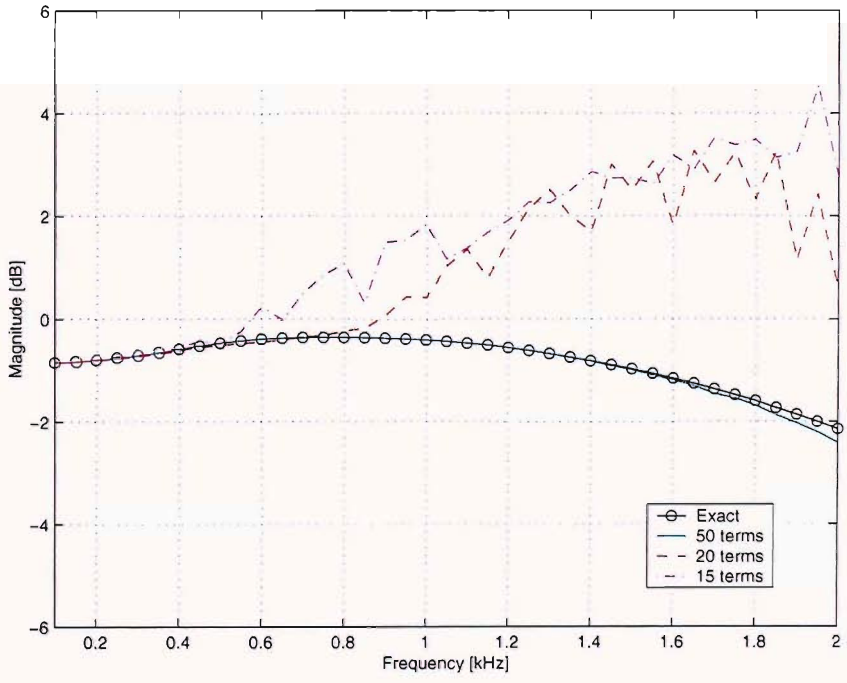


Figure 7-28: As in Figure 7-20 but the singular vectors are associated with σ_5 at 1 kHz.



(a)



(b)

Figure 7-29: Frequency response decomposition of two points on the surface of the sphere (a) in the front, $\phi = 0°, \theta = 0°$ (b) at the rear, $\phi = 180°, \theta = 0°$. In both cases a monopole source is positioned at $\phi = 0°, \theta = 0°$ at a distance of 3 m. The response shown with exact calculation (using the DBEM) and reconstruction with 15 to 50 terms.

CHAPTER 8

MODELLING THE SPATIAL BASIS FUNCTIONS OF THE PINNA USING THE SINGULAR VALUE DECOMPOSITION

8.1 INTRODUCTION

In the previous chapter, the spatial basis functions ('mode shapes') of the sphere have been presented using the SVD formulation and a Green function matrix relating a number of points on the surface of the sphere to a number of points on a sphere in the far field at discrete frequencies. It was also concluded that whenever the space or the sphere is not sampled uniformly, or the shape of the sphere is modified, for example, to an ellipsoid, the mode shapes are not directly related to the spherical harmonic basis functions and become frequency dependent.

In this chapter, we investigate further the properties of the SVD formulation with the aid of numerical modelling of more complex geometries, with an emphasis on the properties of the human pinna. We hypothesise that at a specific 'resonance' frequency, the first singular value σ_1 will have significantly higher amplitude than the others, and the complex singular vectors associated with this singular value will have distinctive patterns. Although found by using a different approach, Shaw (1974, 1997) identified three types of 'mode shapes': a 'monopole like' pattern, and horizontal and vertical 'dipole-like' patterns. We demonstrate here that these common modes of the pinna, or similar patterns can be extracted by defining the appropriate Green function.

This chapter is organised as follows: the SVD formulation together with the generation of a Green function matrix is investigated for a few cases: first, the characteristics of a simple baffled Helmholtz resonator are shown, followed by a shallow baffled cylinder that models the concha. Accurate pinna geometries are investigated with two cases of source distribution: when the sources are uniformly distributed to give grazing incidence, and also, when they are approximately uniformly distributed in the upper hemisphere above the infinite baffle. Finally, the mode shapes found by Shaw (1974, 1997) are validated using a similar procedure to that used in his measurements. This is, however, undertaken numerically here. The relationship between these mode shapes and the basis functions found with the SVD is discussed.

8.2 The Helmholtz resonator

A Helmholtz resonator is a simple example of a device, which behaves as a harmonic oscillator with one degree of freedom (Kinsler *et al*, 1982). It consists of a rigid-walled cavity of volume V with a neck of area S and length L. The main characteristics of the resonator are the strong amplification of pressure at the resonance frequency, and that generally the resonance occurs only in a narrow frequency bandwidth.

The main motivation of analysing the Helmholtz resonator with the SVD formulation is to investigate the behaviour of the singular values. This model can be regarded as a very simple model of the concha, like the baffled cylinder that was investigated initially by Teranishi and Shaw (1968), but when limiting the area size of the opening, only a single narrow resonance frequency is expected to appear.

In our numerical example we model a baffled cylinder with similar dimensions to those defined by Teranishi and Shaw (1968): the cylinder depth is 10 mm, and its diameter is 22 mm. Its top (which coincides with the plane of the infinite baffle) is blocked, except for

a small opening, through which sound waves can propagate in both directions. The modelling technique is based on the 'IBEM transparency' formulation.

The mesh includes 190 nodes and 360 elements. (The 'IBEM transparency' formulation requires coupling of the acoustic mesh with a structural mesh. Therefore a 'dummy' mesh with four elements was added far away from the acoustic model in order to minimise errors). Only six elements at the centre, at the top plane of the cylinder have been defined as 'transparent' (see Figure 8-1). The total area of these was $S = 27.9$ mm^2, the volume of the cavity was $V = 3800$ mm^2.

In the case of an opening consisting of a hole in a thin wall of a baffled resonator ($L = 0$ the effective length of the neck $L' = 1.7a$ (where a is the radius of a circular opening) is affected by its radiation-mass loading. The resonance frequency can be predicted from the analytical expression given by Kinsler *et al*, 1982, pp. 225-227)

$$f_0 = \frac{c_0}{2\pi}\sqrt{\frac{S}{L'V}} = \frac{340}{2\cdot\pi}\sqrt{\frac{27.9\cdot 10^6}{1.7\cdot 2.98\cdot 3800}} = 2060 \text{ Hz} \qquad (8.1)$$

Since our opening is not exactly circular we can expect this equation to predict only approximately the resonance frequency.

A Green function matrix was generated relating 336 field points at the bottom (z =-10 mm) of the resonator and 121 sources distributed approximately uniformly on a hemisphere in the far field, at a radius of 1 m on the upper hemisphere. The calculation is undertaken at 301 frequencies (from 1 kHz to 4 kHz in steps of 10 Hz). The normalised singular values are presented in Figure 8-2a and 8-2b. When the first 10 singular values are plotted on a linear scale, only the curve of the first singular value is noticed with a narrow frequency bandwidth resonance. The logarithmic scale reveals the separation between the first singular values and the remaining nine with a factor of >10^3.

The frequency response at an arbitrary position at the bottom of the resonator due to a source positioned arbitrarily in space is calculated using Equation 7.57 and presented in Figure 8-3. A 'reduced order' calculation is undertaken with the first three terms and also with only the first term in the summation. As expected from Figure 8-2b, due to the high significance of the first singular value, the frequency response is reconstructed with high precision when only the first term is taken into account. The physical interpretation is that the radiation from a Helmholtz resonator can be described by a single monopole 'radiation' operator (as described by Borgiotti, 1990).

8.3 THE CYLINDER IN AN INFINITE BAFFLE

8.3.1 The singular values

The cylinder, as a basic model of the concha was investigated by Shaw and Teranishi (1968) in their early work. Since the first resonance of a blocked meatus pinna corresponds to a quarter wavelength 'depth' mode, it is possible to match the volume of a cylinder to produce the desired resonance frequency as well as controlling the amplification in its base. In this case we use the same cylinder model used in the previous section but now the entire top section of the cylinder is defined as transparent (see Figure 8-4a). A model with similar dimensions was investigated by Teranishi and Shaw (1968) who showed that the first resonance frequency around 4.5 kHz is determined by the volume of the cylinder. The response is substantially independent of the angle of incidence up to about 7 kHz but becomes strongly dependent on the angle of excitation at the first transverse mode at approximately 11 kHz (see Teranishi and Shaw, Figure 2b, 1968, and also in Chapter 4, Figure 4-18). As before the calculation was undertaken using the 'IBEM transparency' formulation.

Clearly the characteristics of the singular values depend on the geometry of the radiation body as well as the source positions defined in the Green function matrix. By defining the field points not only at the bottom of the cylinder (as in the case of the Helmholtz resonator) but also on the cylinder walls and on a small area on the baffle, the properties of the singular values can be investigated where some parts of the 'radiating' body have different resonance frequencies. Therefore the field point mesh (see Figure 8-4b) was used to calculate the pressure values and includes 336 nodes and 457 elements. As before, the sources were positioned for each run at one of the 121 positions defined on a mesh grid with 108 elements*.

Figure 8-5 presents the variation of the normalised singular values with frequency. The main difference between this graph and the previous cases of the sphere (Figure 7-2, 7-14 and 7-16) and of the ellipsoid (Figure 7-18), is that now two resonance frequencies are characterised with strong peaks. The 'resonance' frequencies (the frequencies at which the curves of the singular values reach a peak) do not need to coincide with resonance frequencies for a particular source position and a particular field point. Nevertheless, due to the relatively large bandwidth of these 'resonance' frequencies, at each specific case these will not shift significantly from the centre frequency (see also Section 8.3.3).

The first mode, at 4.2 kHz is known to be associated with a quarter wavelength depth resonance, with a peak in σ_1 only. The second mode appearing at 10.8 kHz is a transverse mode (see also Figure 4-19), with peaks in both σ_1 and σ_2. It is interesting to observe the connectivity and intersection between the curves, as the SVD automatically positions σ_1 before σ_2 (see Equation (7.31). It was found that when too low a resolution sampled mesh of source points was used, the same resonance frequency appeared, but the curves were not

* It is estimated that much lower density meshes could have been sufficient, and the high resolution was used to investigate the effect on the intersection of the curves of the singular values as a function of frequency (see below).

intersecting each other, as we shall see in the case of modelling of accurate representation of pinnae (Section 8.4). Following the curve of σ_1 (a 'monopole' mode), its radiation efficiency is predominant up to 4.2 kHz, and gradually decreasing with increasing frequencies. The curves of σ_2 and σ_3 ('dipole' modes) start with a similar singular value at 2 kHz and increase with frequency up to 10.8 kHz, and then have decreasing amplitude. As frequency increases the contribution of lower order singular values is increasing, such that at 12 kHz σ_{10} is almost 50% of σ_1. Another important conclusion that arises from this graph is that certain modes can be associated with more than a single σ value.

8.3.2 The patterns of the singular vectors

In Figure 8-6a the 'mode shapes' associated with σ_1 at 4.2 kHz are presented†. The non-dimensional basis function has a uniform distribution in the base of the cylinder (where the scale has only positive values), which is related to a uniform distribution in the far field (and the scale has only negative values). Note that the sign of the values depends on the relative distance between these surfaces. In Figure 8-6b the 'mode shapes' associated with σ_2 at 4.2 kHz are presented. Both 'mode shapes' on the cylinder and on the hemisphere show a 'dipole-like' variation. Since at this particular frequency σ_2 does not contribute significantly in the Green function matrix (as illustrated in Figure 8-5), no particular variation is noticed in the cylinder base. It is hypothesised that a resonance mode will produce varying values of basis functions on the surface mesh with little variation at other positions on the baffle. Figure 8-6c and Figure 8-6d illustrate the 'mode shapes' associated with σ_1 and σ_2, respectively at 10.8 kHz. This transverse mode results in two basis functions, which are out of phase. Note that for both singular values the variation of the singular vectors on the baffle are small compared with the variation on the base of the

cylinder, and also that the scales of the singular vectors associated with σ_1 and σ_2 are different.

8.3.3 Extraction of the frequency response

Based on Equation 7.57 we can reconstruct a single frequency response from an arbitrary position in the far field (in this case at a distance of 1 m, at x=-0.53 m, y=-0.26 m, and z=-0.8 m). We can first obtain directly an exact response by positioning a monopole source at the above position and solve for the pressure at the desired position at the base of the cylinder (in this case at x=0.05 m, y=0.003 m, and z=-0.01 m) with the 'IBEM transparency'. As a second stage, we can verify that the same response is obtained, this time with the SVD method and the singular vectors and singular values are used with all of their terms, and finally, the number of terms in the summation (Equation 7.57) is reduced to only the first few terms. Figure 8-7 presents the frequency response between these points. It is concluded that if the first five basis functions are used, the error is less than 0.5 dB up to 12 kHz, and if only the first three basis functions are used, the same accuracy is obtained, but the maximum frequency is reduced to 8 kHz. Note that a significant error occurs from around 8.5 kHz if only three terms are used. It can be seen in Figure 8-5 that at this frequency σ_4 intersects with σ_1 and its value increases, so although σ_4 does not have as strong a peak as the first three highest singular values, it is still important in the reconstruction of the frequency response (its value is around 50% of σ_2 and σ_3 which is clearly not negligible).

Nevertheless, this solution demonstrates that it might be possible to describe a frequency response at any given position in space, based on only few basis functions if these have strong peaks at certain frequencies which are significantly higher than the remaining

[†] Although the basis functions are complex, only the real values are presented here to demonstrate the principle.

singular values. This will ensure at least an approximate reconstruction of the peaks but not the notches, as will be discussed in Section 8.4.2.3.

8.4 AN ACCURATE MODEL OF A PINNA IN AN INFINITE BAFFLE

The principles used in the previous sections are implemented now with a baffled accurate pinna model (DB60). The model includes 6887 nodes and 13488 elements, and has a blocked ear canal. For this size of problem, the exterior DBEM formulation was the most efficient and 30 over-determination points were used to remove 'irregular frequencies'.

8.4.1 Source excitation at grazing incidence

8.4.1.1 The singular values of the DB60

A Green function matrix was generated that related the 6887 points on the surface of the pinna and 36 source positions. These were uniformly distributed on a circle with a radius of 1 m in grazing incidence, at a resolution of 10°. The purpose is to investigate the 'mode shapes' of a similar source and microphone arrangement to that used by Shaw (1974, 1997).

Figure 8-8 presents the singular values for frequencies between 2 kHz and 20 kHz (the response was calculated at 91 frequencies in steps of 200 Hz).

This plot highlights the behaviour of the first dominant singular value and also shows the frequency dependence of the next nine largest singular values. The most significant feature of these results is that the dominant singular value has peaks appearing at certain frequencies. It has been found that the corresponding left and right singular vectors also reveal distinctive spatial patterns at these frequencies.

In addition, note that the curves do not intersect (as in the case of the baffled cylinder, Figure 8-5), although the trend of 'flipping' modes can be seen. For example, the curve of σ_1 reaches its maximum at 4.5 kHz and its value is declining up to 6 kHz, and then

continues (albeit with a gap) with the second curve. σ_2 changes from the second curve to the first curve at the same frequency (6 kHz). It was found that the curves do not intersect due to a low sampling of the source positions in space (this was investigated with the small model of the cylinder).

The main 'resonance' frequencies associated with σ_1 are (4.5, 7.8, 10.3, 13.8 and 17.3) kHz. It is difficult though to specify if peaks occur at 'resonance' frequencies that are associated with σ_2, σ_3, etc. due to acoustical behaviour or due to a low sampling resolution of the hemisphere.

8.4.1.2 The singular vectors of the DB60

In order to investigate the characteristics of the singular vectors in three dimensions in space and on the surface of the pinna, a simple example is presented in Figure 8-9 where the excitation was at DC (1 Hz). Only the real parts of the left and right singular vectors are presented. These vectors show the variation of the basis functions. These are equivalent to free-field since the pinna does not contribute to these patterns at this frequency. The singular vectors associated with the first singular value are omnidirectional (the x-axis on the right figures is the angular position from $\theta = 0°$ to $\theta = 350°$). For higher modes the patterns on the circle and on the surface of the pinna have systematic 2-D and 3-D variations of the spherical harmonics, respectively. Note that both real and imaginary patterns are required to describe the 'mode shapes' at each frequency.

The first peak in the variation of the singular values with frequency (Figure 8-8) appears at 4.5 kHz. The first three real parts of the left and right singular vectors are presented in Figure 8-10. The real part of the singular vector on the surface of the pinna is characterised by an omnidirectional variation in the concha. It should be noted that the sign of the values is a function of the phase and it is determined by the distance between the sources and the

pinna. The real part of the first right singular value at the position of the sources still shows only positive values, as shown in Figure 8-9a, but the deviation from a perfect 'monopole' directivity pattern is noticed. The following patterns associated with the second and third singular values are presented in Figure 8-10b and in Figure 8-10c. These show some distortion to the 'dipole-like' patterns shown at DC. The variations also noticed outside the pinna that means that these 'modes' are not tuned for pinna resonance. This is also supported by the fact that the singular values at this frequency have no certain peaks (in Figure 8-8). Therefore, for higher 'resonance' frequencies we will concentrate only on the first dominant singular values since these are probably the most dominant in identifying 'resonance' frequencies, and also since the reliability of lower order singular values is poorer due to low resolution of the sampled source positions.

Figures 8-11a to 8-11d show the real part of the singular vectors associated with the first singular values at (7.8, 10.3, 13.8 and 17.3) kHz, respectively. It is clear that there is some resemblance between the 'modes' on the surface of the pinna, and the 'modes' appearing at source positions. However the interpretation of the shapes, especially at source positions become more complex as frequency increases. In addition, the interpretation requires the visualisation of the real and imaginary values of both singular vectors.

It is evident that at frequencies of which the singular values reach a peak, a tuned mode appears on the surface of the pinna. The variation of the non-dimensional basis function in the surrounding of the pinna is close to zero, and the maximum and minimum values appear mainly in the cavum concha and cymba concha, and also in the fossa of helix and antihelix.

Only five peaks are detected in Figure 8-8. Based on the simulation of the frequency response of the DB60 in Chapter 5 (Figure 5-16 to Figure 5-18), measurements of the response of the DB60 at grazing incidence in the anechoic chamber, and the investigation of the normal mode shapes in Chapter 8, Section 8.5, it is evident that the resonance

frequencies are shifted upwards due to the relatively small dimensions of the pinna. As Shaw found, the first 'mode' has a 'monopole-like' pattern, the next two modes have 'vertical dipole' modes with variation between the positive values in the cavum concha and negative values in the cymba concha and the fossa of helix in the second mode, and positive values in the cavum concha and fossa of helix, nodal line on the *crus helias,* and negative value in the cymba concha in the third mode. In the next two patterns the variation seems to be in the horizontal direction, with a single nodal line in the cavum concha in Figure 8-11c, and two nodal lines in Figure 8-11d. This is in general agreement with the results of Shaw.

8.4.2 Source excitation on the upper hemisphere

8.4.2.1 The singular values of the DB60

The SVD method is used in this case in a very similar procedure to that discussed in the previous section. The only difference in this case is the distribution of source positions on the upper hemisphere. We investigate again the characteristics of the singular values and the singular vectors and compare the results with the excitation at grazing incidence. Since the computational cost of operating the SVD procedure at many frequencies is expensive (see Chapter 10 for a summary), a smaller mesh is used (with 2825 nodes compared to 6887 nodes in the previous case). In this case the pinna is not smoothed gradually to the baffle plane but is used with its rectangular frame such that its bottom is snapped to the baffle plane. It was checked that the fact that the pinna is now physically 7 mm above the plane did not change significantly the response extracted with the SVD. The dimensions of the matrix of Green functions relating a number of points on the surface of the DB-60 pinna mounted on an infinite baffle and a number of points on a far field hemispherical surface are now 2825 × 209. The procedure is repeated 91 times for frequencies between 2 kHz and

20 kHz in steps of 200 Hz, so overall more than 50 million complex pressure values have been calculated.

The frequency dependence of the resulting singular values is illustrated in Figure 8-12. Note the similarities found in the resonance frequencies appearing in this graph and those found in Figure 8-8. The only shift in 'resonance' frequency is noticed for the second peak. Its relatively low amplitude and broad resonance shifts from 7.8 kHz in the case of excitation at grazing incidence to 8.8 kHz in the current case. The other four peaks have similar centre frequencies, and similar amplitudes in both cases. In addition, the characteristics of the non-intersecting curves appear again, and the trend of 'flipping' σ at certain frequencies is evident, as before, at lower values of σ. For example, it is not clear whether σ_2 at 10 kHz has a peak due to similar physical behaviour of the baffled cylinder in this frequency range (see the curve of σ_2 at 10.8 kHz in Figure 8-5), or because of a break that appears in the 'flipping' curves of σ_1 and σ_2.

8.4.2.2 The singular vectors of the DB60

Figure 8-13 shows the real parts of the left and right singular vectors associated with the first dominant singular value (σ_1) appearing in Figure 8-12 at 4.6 kHz. Figure 8-13a illustrates the first quarter wavelength depth resonance. The concha has a 'monopole' mode, with a similar pattern appearing in the far field. Figure 8-13b and Figure 8-13c show the real parts of the singular vectors associated with σ_2 and σ_3. Similar patterns appear on the pinna surface when those are compared to Figure 8-10b and Figure 8-10c, and also slight distortions of the 'dipole' patterns appear on the hemisphere in a similar way to those that appeared when the sources were distributed on a circle (the right column in Figure 8-10b and Figure 8-10c). Figure 8-14 shows the imaginary part of the singular vectors with similar patterns to those that appeared in Figure 8-13 albeit with changes on the phase.

Figure 8-15a and Figure 8-15b show the real and imaginary singular vectors at 8.8 kHz associated with σ_1. The 'vertical' dipole patterns identified in all four plots support the similar patterns obtained in the case of excitation at grazing incidence (Figure 8-11a). These involve oscillatory flow between the cavum concha, cymba concha and the antihelix. In addition, the patterns on the pinna are comparable with the first 'vertical' mode identified by Shaw at 7.1 kHz. The third pair of real and imaginary singular vectors associated with the first singular value that has a peak at 10.8 kHz (in Figure 8-12) is presented in Figure 8-16. In a similar behaviour to that noticed with excitation at grazing incidence (Figure 8-11c), another 'vertical' mode is noticed on the surface of the pinna and involves oscillatory flow between the cavum concha, cymba concha and the antihelix, as before but with different signs in the cymba concha and the antihelix. The variation in the far field is more difficult to analyse visually and different patterns appear for the real and imaginary singular vectors.

Figure 8-17 presents the only singular vectors associated with σ_2. It is probable that these vectors are related to the first transverse mode appearing in the case of the baffled cylinder (see Figure 8-6c and Figure 8-6d). Although the four plots have a clear 'vertical dipole' in the far field and 'horizontal dipole' patterns on the surface of the pinna with oscillatory flow between the fossa of helix and the antihelix, the cavum and cymba concha patterns remain unchanged.

The patterns become more complex in Figure 8-18 and Figure 8-19. The patterns on the surface of the pinna are in general agreement with the results of Shaw although in this case the pattern at 13.8 kHz match the 12.1 kHz average frequency given by Shaw, and the final mode here, at 17 kHz is comparable with the 14.4 kHz identified by Shaw. In both figures the interpretation of the modes in the far field is not possible, and it is not clear if increasing

the resolution of source positions in space will change the patterns or if these are a result of the high variation of efficient excitation of the pinna at high frequencies.

The resolution of the sources was investigated for the 'resonance' frequency of the second peak. Figure 8-20 presents four cases where the real parts of the left singular values are presented on the surface of the DB60. The source distributions varied from an approximately uniform distribution on the upper hemisphere with 17, 57 and 209 sources, in Figure 8-20a, Figure 8-20b and Figure 8-20c, respectively, and also 36 sources distributed uniformly in grazing incidence (Figure 8-20d). As clearly seen, the oscillatory flow between the cavum concha and the cymba concha and antihelix, producing a 'vertical dipole', appears in all cases. This suggests that the 'mode shapes' of the pinna do not need a high density. However, using more than 209 sources is beyond the capability of the available computational resources at this stage (SGI, Origin 2000 with parallel processing).

8.4.2.3 Frequency response reconstruction of the DB60

The procedure, based on Equation (7-57) was investigated for the case of the Helmholtz resonator (Figure 8-3) and the baffle cylinder (Figure 8-7). In this case the blocked entrance to the ear canal is used, and four source positions are investigated. Figure 8-21 presents the efficiency and accuracies of 'reduced order' frequency responses using only three to ten terms. It is concluded that the reconstruction produces different accuracies depending on the source positions. In all cases the reconstruction of peaks is better than notches mainly since the peaks in the singular values are dominant. Whenever the first singular value reaches a minimum value, the relative contribution of the low order singular values increases and more terms are required in the summation. Nevertheless, for source excitation at a grazing angle in the front (at $\phi = 0°, \theta = 0°$) presented in Figure 8-21a, three terms are sufficient to reconstruct the response up to 14 kHz with errors less than 2 dB (except in the minimum around 8, and 12 kHz where the error is slightly higher). The results obtained

with three to ten terms are compared to the calculation made with the DBEM. When the source is at $\phi = 0°, \theta = 90°$ (on the baffle, above the pinna, Figure 8-21b) the curve is more complex and three terms are sufficient for accurate reconstruction only up to 7 kHz. Above this frequency the number of terms should be between five and ten. Similar responses are obtained when the source is below ($\phi = 0°, \theta = 270°$, Figure 8-21c). For excitation of the source at the rear ($\phi = 180°, \theta = 0°$), presented in Figure 8-21d, more errors occur (more than 10 dB) if only three terms are used, and high accuracy is achieved if ten terms are used in the summation.

8.4.3 The singular values of additional pinnae

In the following section the B&K, DB65, CORTEX and YK pinnae are investigated, and the Green function matrix relating the points on each pinna and 209 source positions on the upper hemisphere. The results are presented in Figure 8-22 to Figure 8-25. As in the previous case, the 'resonance' frequencies are found by the peaks of the first singular value. The amplitudes vary among pinnae for each centre of 'resonance' frequency.

Table 8-1 summarises the centre 'resonance frequencies' of all of the pinnae investigated. For the DB60 column (a) is related to excitation with 209 sources and (b) for excitation of 36 sources in grazing incidence. In general, similar frequencies appear when compared with the averages given by Shaw, especially for the first three modes. As frequency increases, variations in pinnae shapes and size as well as low density of source positions in the Green function matrix make the comparison less clear.

B&K	DB60		DB65	CORTEX	YK pinna
	a	b			
4.0	4.6	4.5	4.0	4.0	3.9
7.4	8.8	7.8	6.8	6.8	7.6
10.3	10.3	10.3	9.5	9.8	10.8
12.2	-	-	12.5	12.7	-
15.4	13.8	13.8	16.2	14.6	14.0
18.9	17.5	17.3	18.2	-	18.6

Table 8-1: 'Resonance' frequencies (in kHz) of pinnae obtained with the SVD. In each case the Green function matrix is based on 209 sources approximately distributed in the upper hemisphere, except the case of DB60(b) where 36 sources were distributed uniformly in grazing incidence (every 10°). Note the similarities of the resonance frequencies and the average values given by Shaw (1997) and summarised in Table 5-1.

8.5 CONVENTIONAL MODE EXTRACTION OF BAFFLED PINNAE

8.5.1 Method

In this section we extract with numerical modelling the normal modes of the pinna, following the same experimental procedure as that used by Shaw (1974, 1982). He used the following procedure. A plate was attached to the head, so that diffraction around the head was minimised. A progressive wave source was moved in the grazing incidence plane and the pressure was measured with a probe microphone at the blocked entrance to the ear canal. The variation of pressure under these conditions resembles the response under free-field conditions for median plane sources. A mode was found when maxima appeared as the frequency and source positions were moved. Then the pressure amplitude and phase were recorded at the base of the concha and the fossa of helix as well as the angle at which the mode was most excited.

In numerical modelling of this experiment, the small baffle Shaw attached to his subjects around their pinnae is replaced here by an infinite baffle. His 'wave progressive source' is

replaced here first[‡] by an ideal monopole source that is placed 1 mm away from the blocked entrance to the ear canal. Using the principle of reciprocity, this is used to calculate the pressure variations 1 m away (Shaw was limited to the near field only with a distance of 8 cm from the microphone, due to the size of the baffle and calibration of the transducers). With the reciprocity simulation technique, the calculation of the pressure at field points in the post-process stage is straightforward and fast. However, we limited these points to be on a circle with a radius of 1 m at grazing incidence. The calculation is undertaken at a resolution of 1°, and the frequency step was 200 Hz from 1 kHz to 20 kHz. Figure 5-16(DB60) shows the variations of peaks and notches detected at the blocked entrance to the ear canal with the above parameters. In fact, for all pinnae presented in Figure 5-16, peaks appear at certain frequencies, regardless of the angle (i.e. the amplitude of these will change, but not the frequency). It can be seen that five or six resonance frequencies appear which might be explained by the different sizes of the pinnae, where smaller pinnae will shift the resonance to a higher frequency. By analysing Shaw's measurement for individuals (Shaw, 1997, pp. 34-35), it can be seen, for example, that the averaged 12.1 kHz resonance frequency has distinct behaviour with subjects J and H, and a very small change in the curvature of the curve around this frequency with subjects A, B and G. Also the averaged 14.4 kHz resonance frequency is absent in subject H.

8.5.2 Results

The resonance frequencies found for DB60 in Figure 5-16 are: (4.5, 7.8, 10.3, 14 and 17) kHz. The angle, at which the excitation is maximal was found by plotting a directivity/vector plot, at each resonance frequency. In the left plots in Figure 8-26, the length of the arrows facing the centre of the pinna is corresponding to the magnitude of the

[‡] This to find the maxima of pressure variation in the far field. Later a plane wave source was used to excite the pinna, to eliminate spherical attenuation.

pressure at the blocked entrance to the ear canal due to a source at that location. Also the colour of the arrow changes with the magnitude: red for maximum response and blue for minimum response. The colours assigned to each vertex on the pinna mesh (right hand pictures) correspond to the absolute pressure, and negative values indicate negative phase at the particular vertex (compared to the phase at the blocked entrance to the ear canal). Note that a more accurate way of presenting these results would be with two plots; either of the real and imaginary values of the pressure, or the magnitude and phase, but we use this format for consistency with Shaw's plots.

The first resonance, appearing at 4.5 kHz was excited the most from a source positioned at 21° in elevation (at grazing incidence). The directivity plot is almost omnidirectional, i.e. the pinna is excited almost with the same efficiency from every direction in grazing incidence. The pressure variation at the base of the concha is also almost uniform with amplification factor of 5 compared with the response detected at the centre of the baffle (which doubles the pressure, compared to free-field equalisation).

The next two 'vertical' modes, in Figure 8-26b and in Figure 8-26c are excited the most effectively from $\theta = 41°$ and $\theta = 76°$, respectively. The patterns in the concha are characterised with a single nodal line in the vicinity of the crus helias. The difference between these modes is the negative phase in the fossa of helix at 7.8 kHz and positive phase at the same position at 10.3 kHz. Note also the similar variation of the pressure amplification at 10.3 kHz compared with the results of Shaw, with average factors in the range of 8-10).

The next two modes have more complex patterns. The 'horizontal' mode in Figure 8-26d at 14 kHz has a very similar pattern to the 12.1 kHz mode identified by Shaw, and is also excited from the front. It is not clear, at this point, if this mode is perfectly tuned, and the shift in resonance frequency is due to the small dimensions of the pinna. The final mode is

believed not to be tuned as maximum efficiency of excitation appears from two directions, one from the front as found by Shaw, and one from above. There is clearly a relationship between the patterns in the concha and the angle of excitation: source from above will create 'vertical' modes, and from the front, 'horizontal' modes. As a result it was decided to investigate further, another pinna, the larger DB65. As can be seen in Figure 5-16 and in Table 5-1, the resonance frequencies of this 'average male' pinna (Maxwell and Burkhard, 1979) are similar to the average responses of Shaw.

Figure 8-27 presents the six modes found for this pinna and the best angle of excitation in a similar format used in Figure 8-26. The first resonance, at 4.2 kHz have a similar pattern to the 4.6 kHz resonance found in DB60 with almost a uniform distribution of pressure amplification in the concha. The next two modes are characterised with efficient excitation in elevation (these, at $\theta = 60°$ and $\theta = 94°$ are within the range measured by Shaw) and also the nodal line appears in the vicinity of the crus helias.

The next three modes are clearly 'horizontal' with all excitation have maxima at the front (again, the angle variations are in agreement with Shaw's results). Also the nodal lines and phase changes have similar patterns to his measurements.

8.6 CONCLUSIONS

In this chapter the mathematical formulation of the SVD was used to investigate the 'mode shapes' of arbitrary geometries. It is shown that the SVD formulation with the aid of exact numerical modelling techniques (the DBEM and the IBEM) extend the interpretation and use of orthogonal basis functions by relating spatially the 'mode shapes' on the body and the mode shapes on the hemisphere or on a particular plane on which the radiated field is defined.

The numerical results of a baffled Helmholtz resonator, a baffled cylinder and an accurately modelled pinna demonstrate that the singular vectors are frequency dependent and at particular resonance frequencies these 'modes' radiate more efficiently than at other frequencies. Distinctive spatial patterns that are associated with these singular values appear both on the surface of the body as well as in the far field.

The mode shapes of two pinnae (the DB60 and the DB65) were also investigated by the excitation by sources positioned at grazing incidence. When compared with the 'modes' obtained with the SVD, the resonance frequencies are nearly identical. A high level of agreement was also obtained when the mode shapes and the angle of excitation were compared with the work of Shaw, especially for the case of the DB65.

Due to a limited number of source positions used in the SVD, it is difficult to analyse the behaviour of lower order singular values and their singular vectors. This was due to computing and hardware limitations.

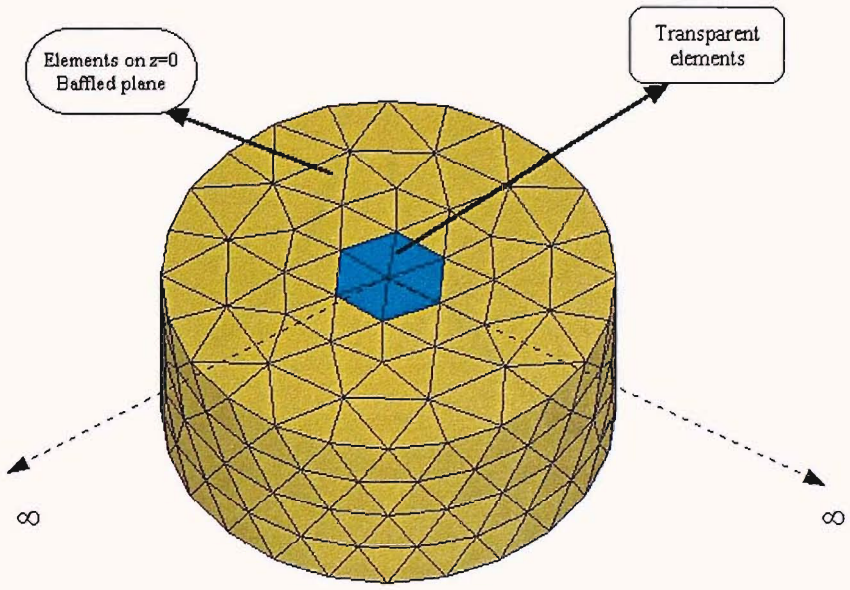
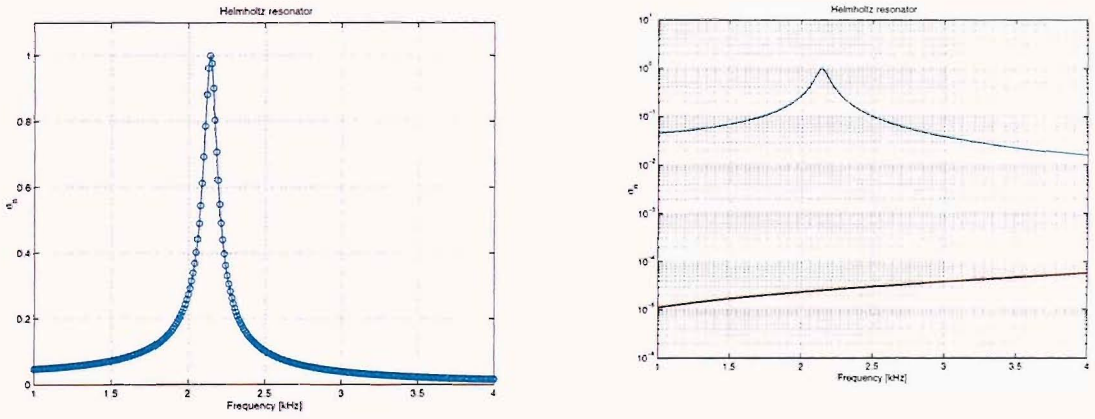


Figure 8-1: Baffled Helmholtz resonator. The mesh includes 190 nodes and 360 elements. Six elements at the top are 'transparent' so that waves can propagate through this opening.



(a) (b)

Figure 8-2: The singular values of the 336 x 121 Green function matrix of the Helmholtz resonator (a) linear scale (b) logarithm scale. The simulation is undertaken at 301 frequencies in steps of 10 Hz.

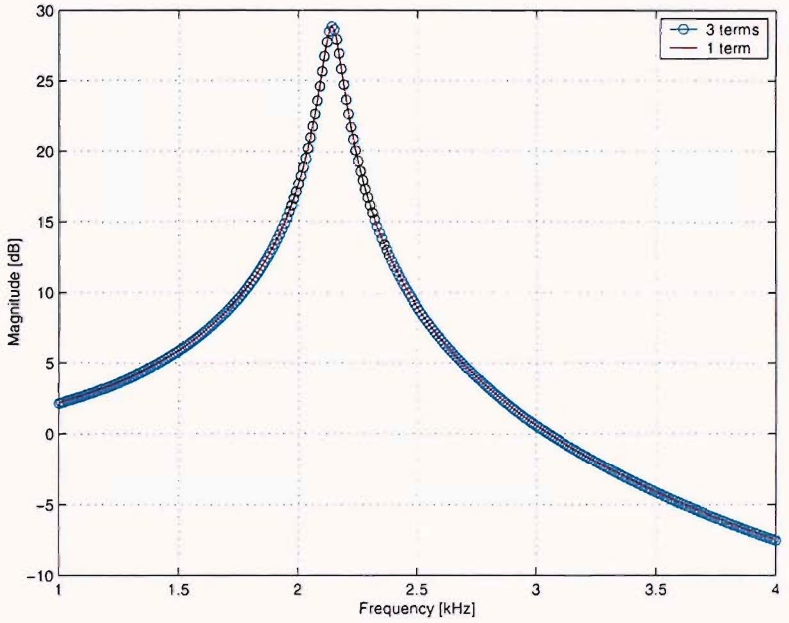
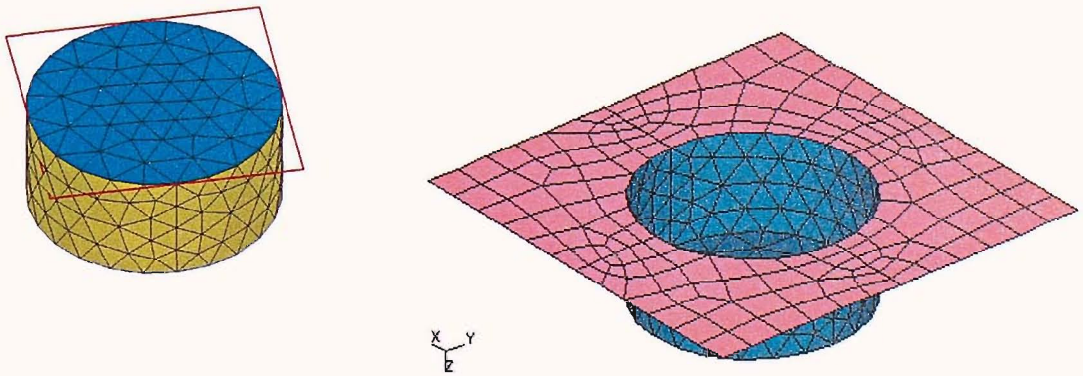


Figure 8-3: The frequency response at the bottom of the Helmholtz resonator reconstructed using the first three singular values and also with only the first singular value.



(a) (b)

Figure 8-4: The IBEM transparency mesh model (a) A mesh cylinder with a volume of 3.8 cm^3 to represent the concha is composed of 376 linear elements and 190 vertices. Its top is aligned with the infinite baffle ($z = 0$), where all top elements are transparent so waves can propagate through both sides of the baffle. (b) 457 elements and 336 vertices of field points are located both on the walls of the cylinder and the baffle.

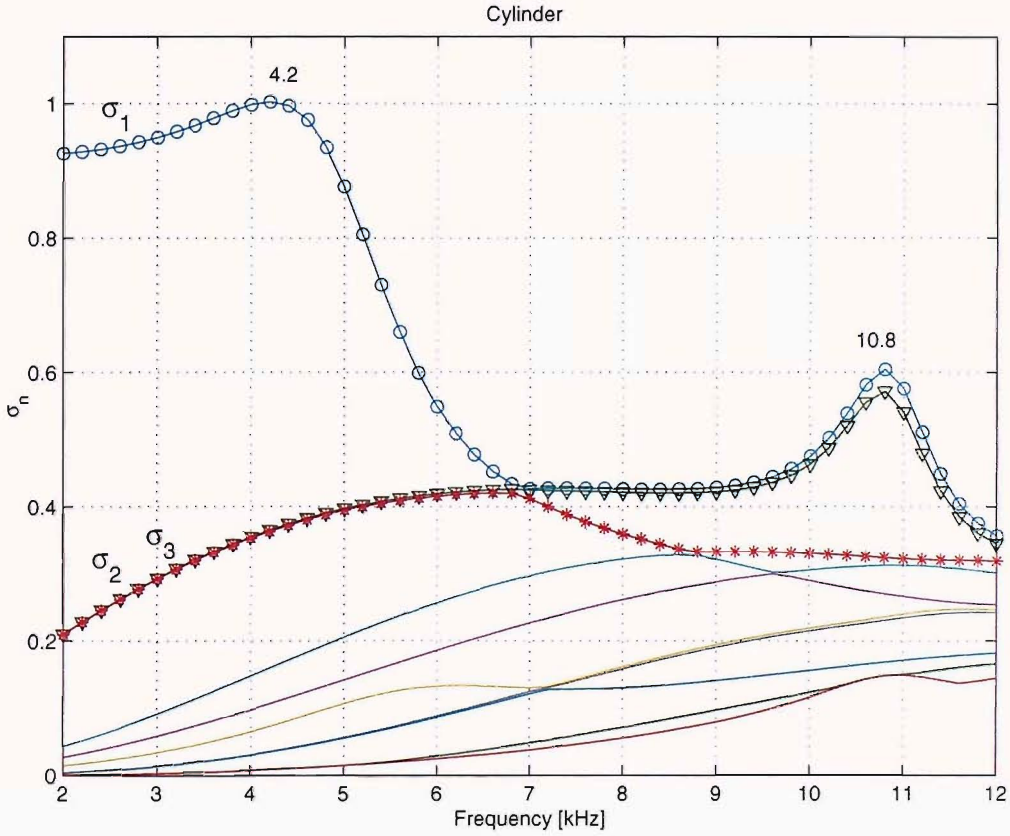


Figure 8-5: The singular values of the numerically generated 336 × 121 Green function matrix relating points on the cylinder and its surrounding to points on the hemisphere with radius of 1 m. The simulation is undertaken at 51 frequencies.

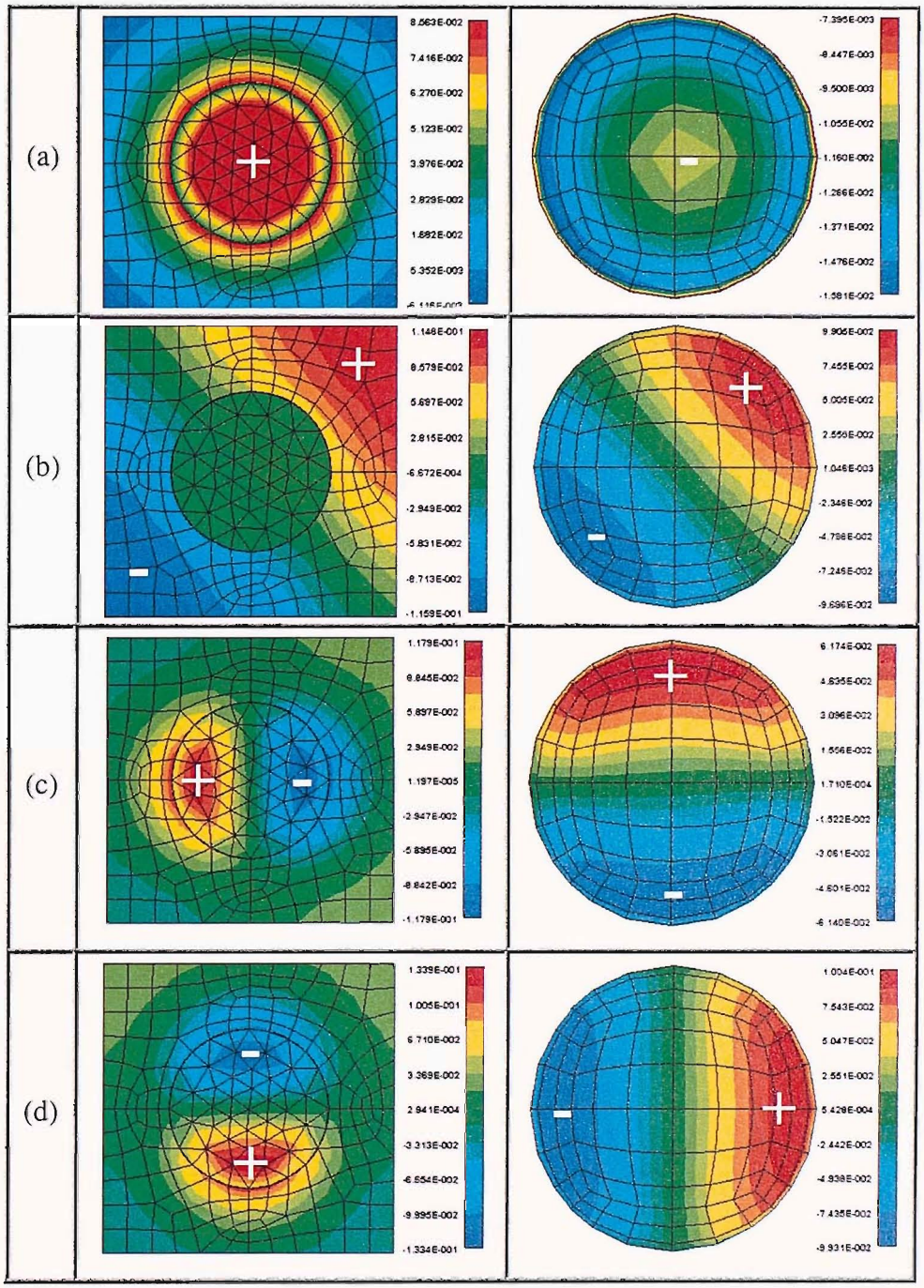


Figure 8-6: The figures on the left show colour maps of the real part of the left singular vectors of the numerically generated 336 × 121 Green function matrix. The figures on the right show the real part of the right singular vectors in the far field (a) 4.2 kHz - σ_1 (b) 4.2 kHz - σ_2 (c) 10.8 kHz - σ_1 (d) 10.8 kHz - σ_2.

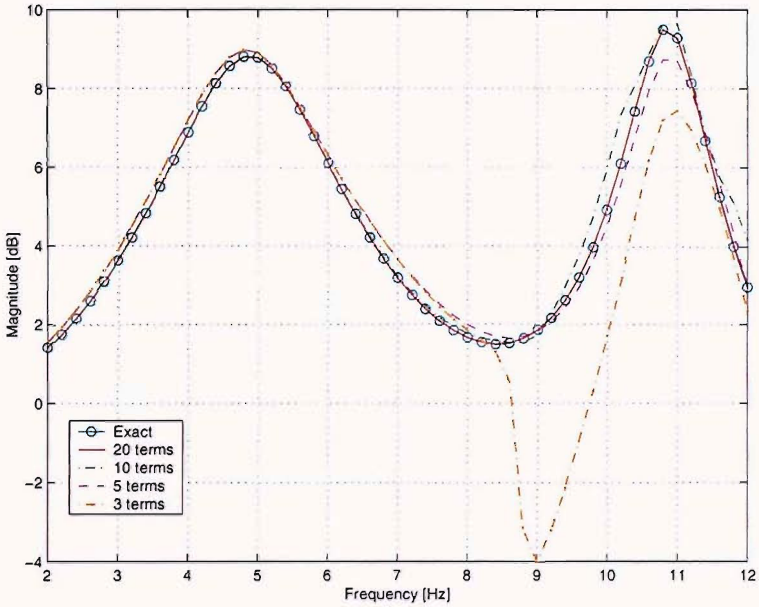


Figure 8-7: Reconstruction of the frequency response detected at the bottom of the cylinder (x=0.005 m, y=0.003 m, and z=-0.01 m) due to a source, arbitrarily chosen at x=-0.53 m, y=-0.26 m, and z=-0.8 m, based on the basis functions found on the surface of the cylinder mesh and in the far field.

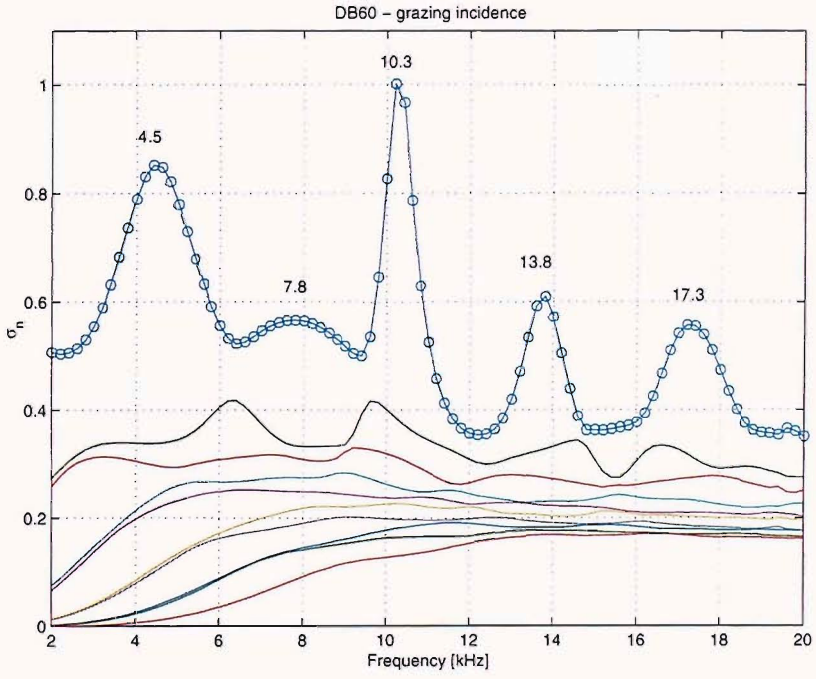


Figure 8-8: The singular values of the numerically generated 6887 × 36 Green function matrix relating 6887 points on the surface of the blocked meatus of the DB60 KEMAR pinna mounted on a rigid baffle to 36 points distributed uniformly, every 10° on a circle at grazing incidence. The calculation is undertaken at 91 frequencies.

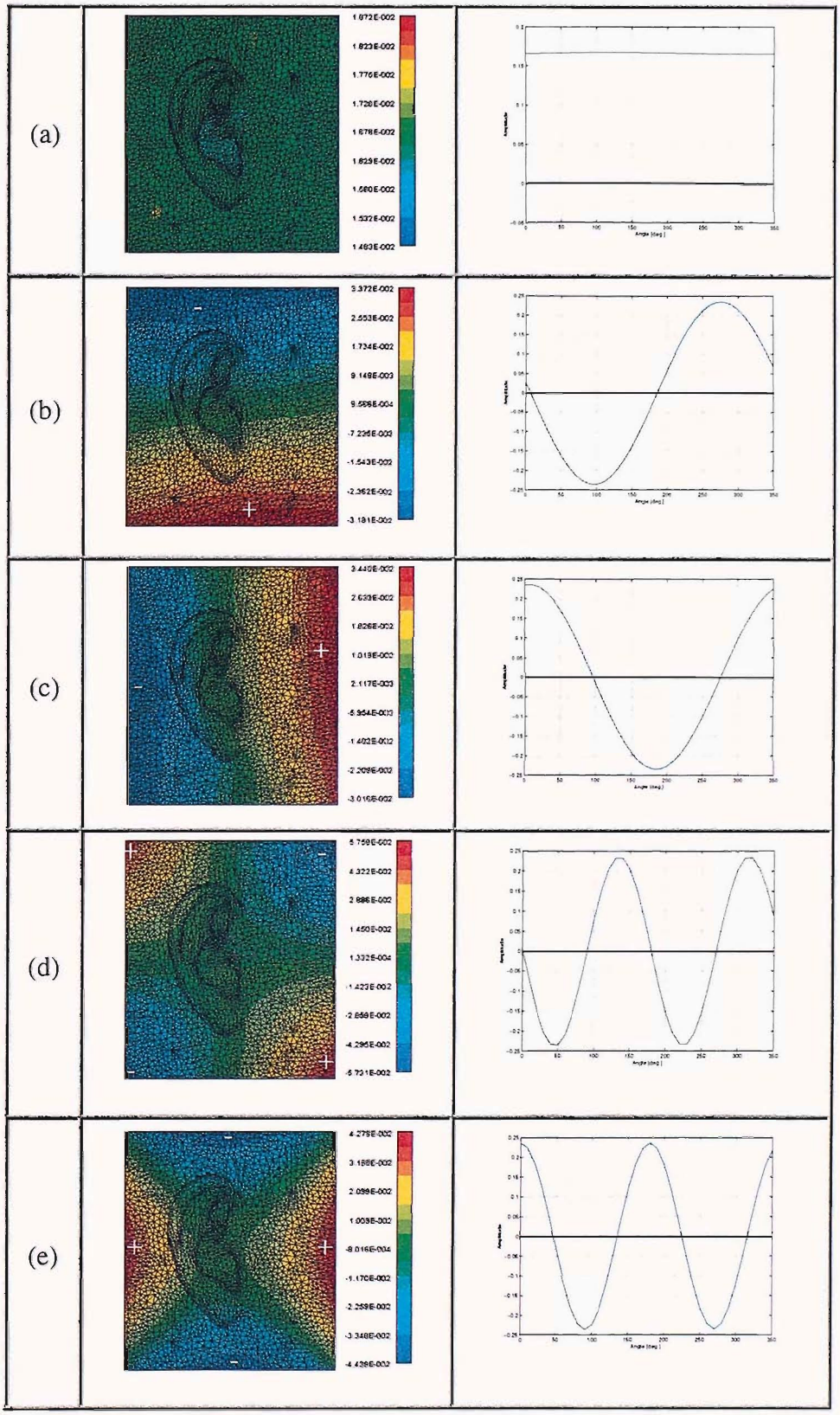
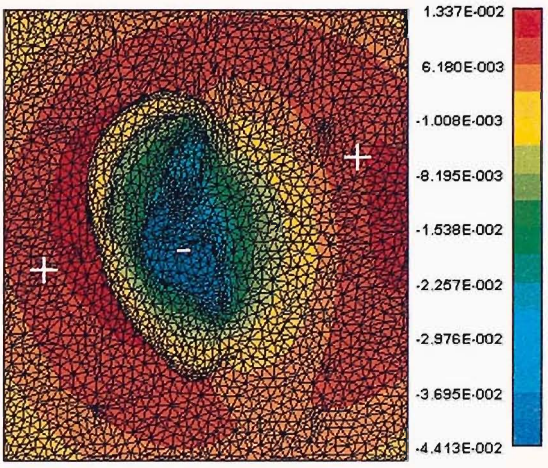
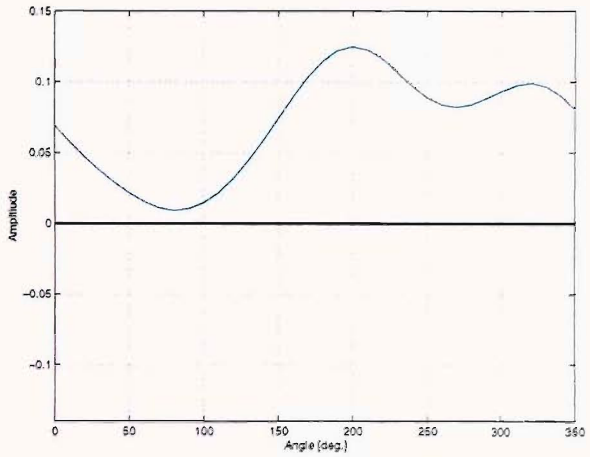
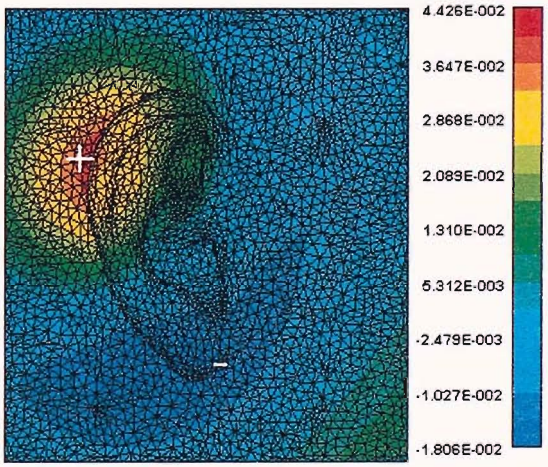


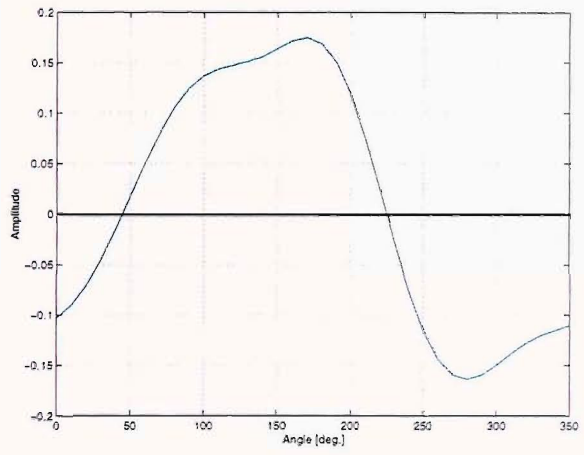
Figure 8-9: The real part of the left and right singular vectors at DC (1 Hz) associated with the dominant singular values of the numerically modelled 6887 × 36 Green function for the DB60 pinna.



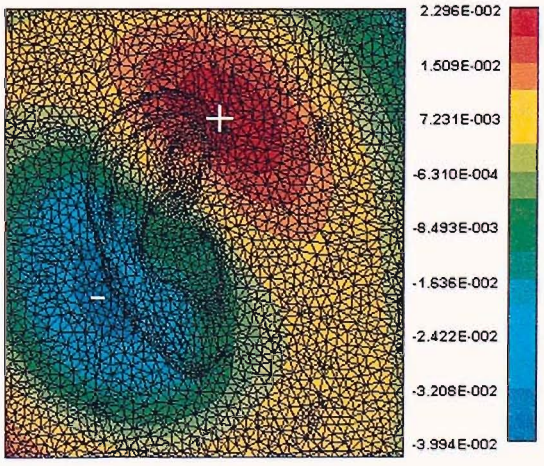


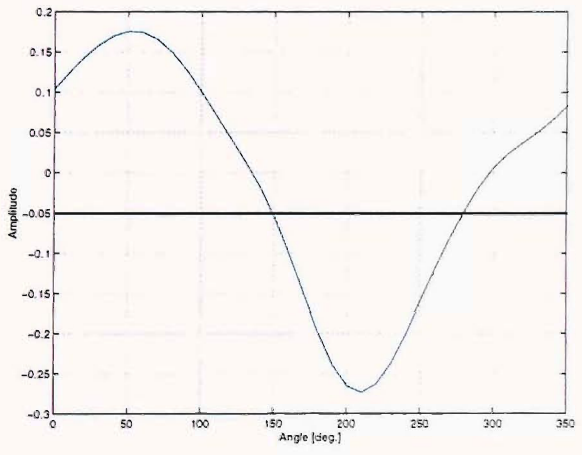
(a)





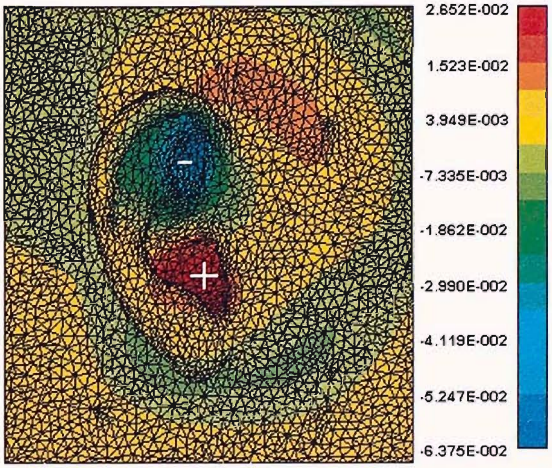
(b)

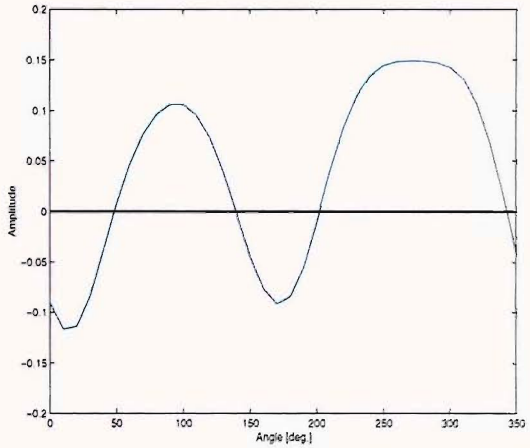




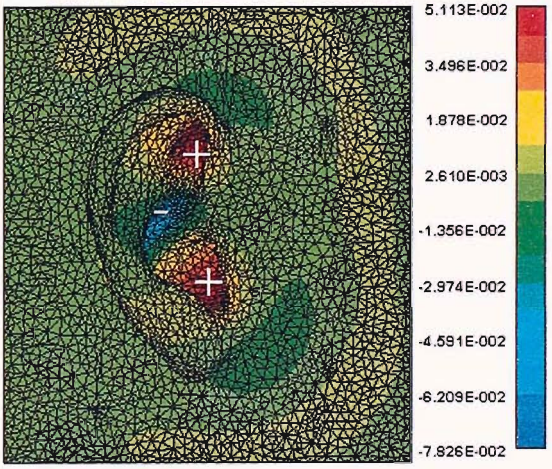
(c)

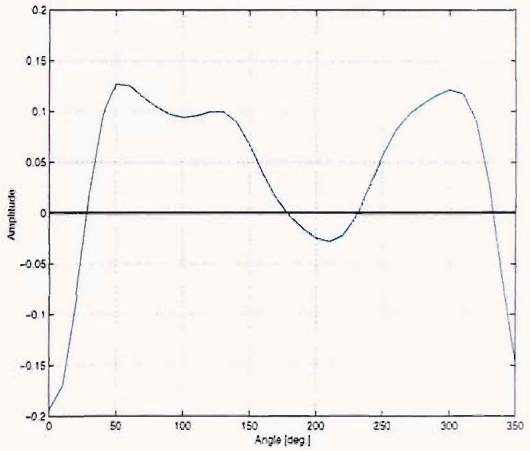
Figure 8-10: The first three real parts of the left and right singular vectors of the DB60 at 4.5 kHz (a) σ_1 (b) σ_2 (c) σ_3. The Green function matrix is as defined in Figure 8-9.


2.652E-002
1.523E-002
3.949E-003
-7.335E-003
-1.862E-002
-2.990E-002
-4.119E-002
-5.247E-002
-6.375E-002

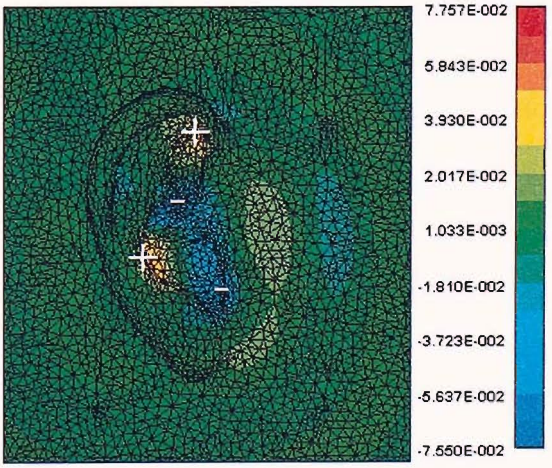

Amplitude
Angle [deg.]

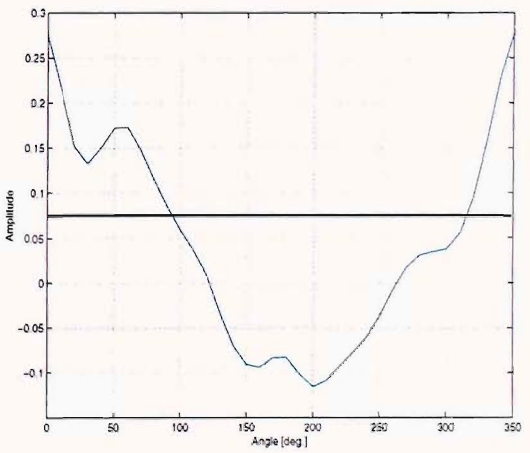
(a)


5.113E-002
3.496E-002
1.878E-002
2.610E-003
-1.356E-002
-2.974E-002
-4.591E-002
-6.209E-002
-7.826E-002

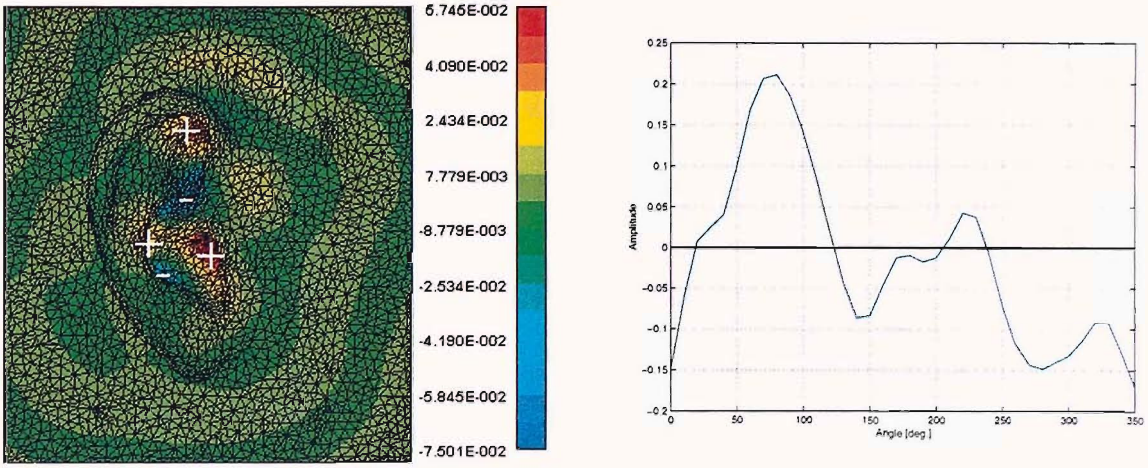

Amplitude
Angle [deg.]

(b)


7.757E-002
5.843E-002
3.930E-002
2.017E-002
1.033E-003
-1.810E-002
-3.723E-002
-5.637E-002
-7.550E-002


Amplitude
Angle [deg.]

(c)



(d)

Figure 8-11: The real parts of the left and right singular vectors of the DB60 associated with the first singular value (σ_1) at (a) 7.8 kHz (b) 10.3 kHz (c) 13.8 kHz (d) 17.3 kHz. The Green function matrix is as defined in Figure 8-9.

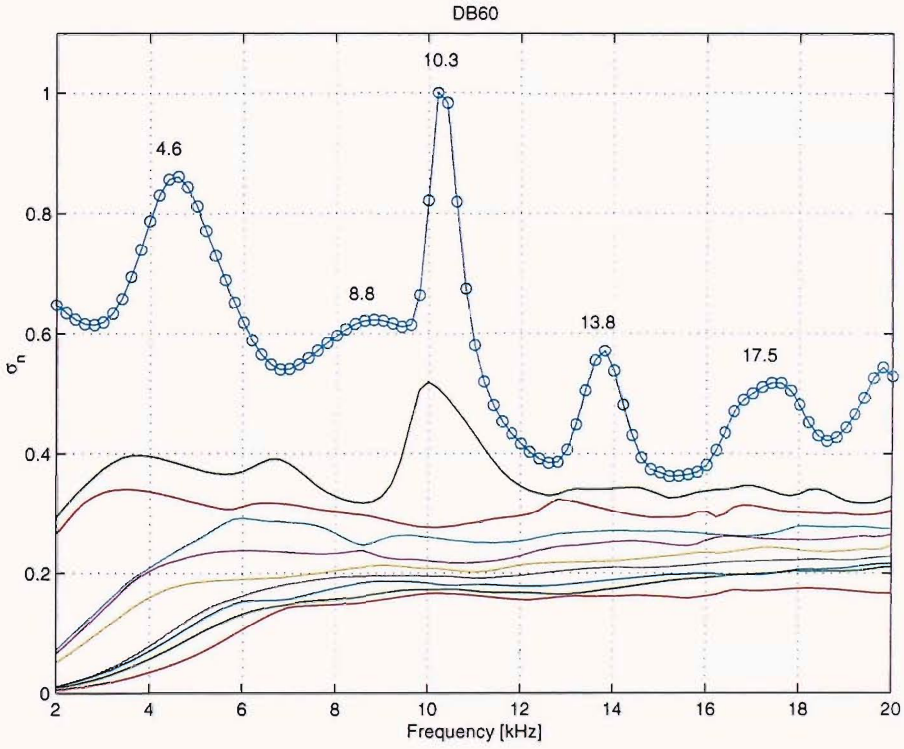
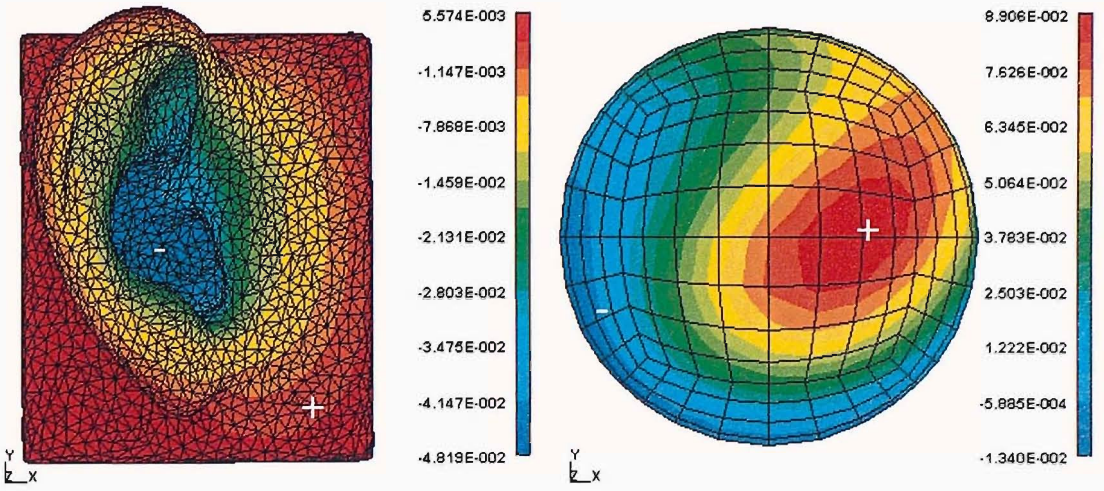
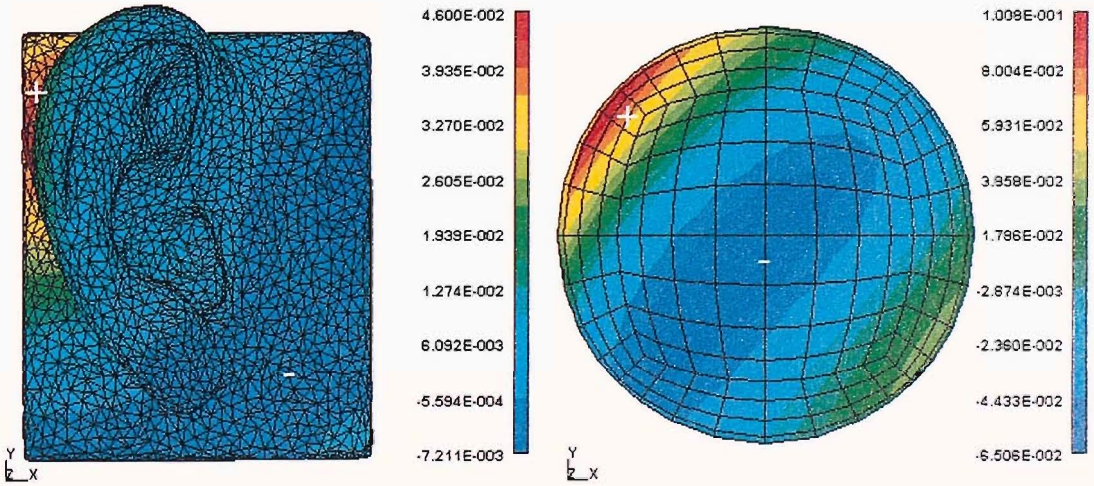


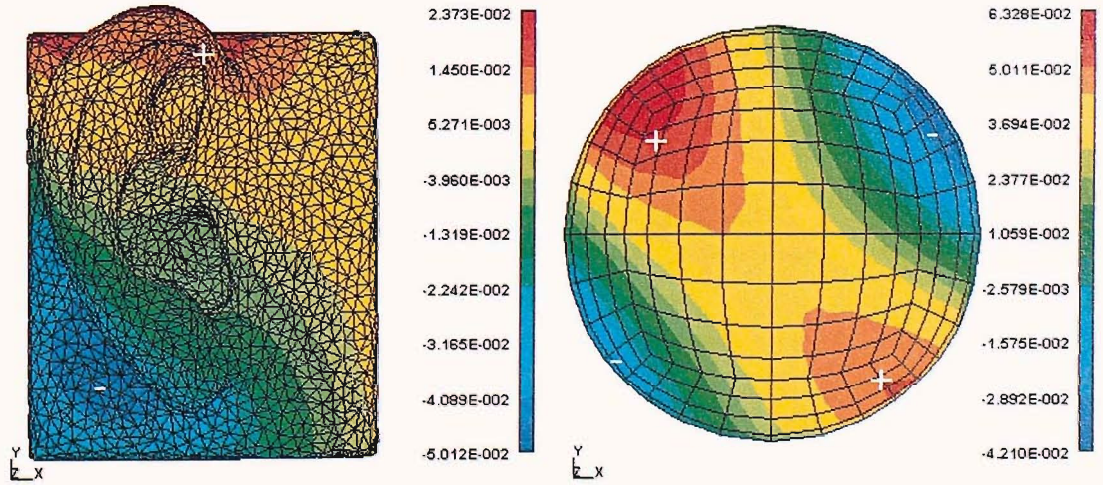
Figure 8-12: The singular values of the numerically generated 2825 × 209 Green function matrix relating 2825 points on the surface of the blocked meatus of the DB60 KEMAR pinna mounted on a rigid baffle to 209 points distributed approximately uniformly on the upper hemisphere. The calculation is undertaken at 91 frequencies.



(a)

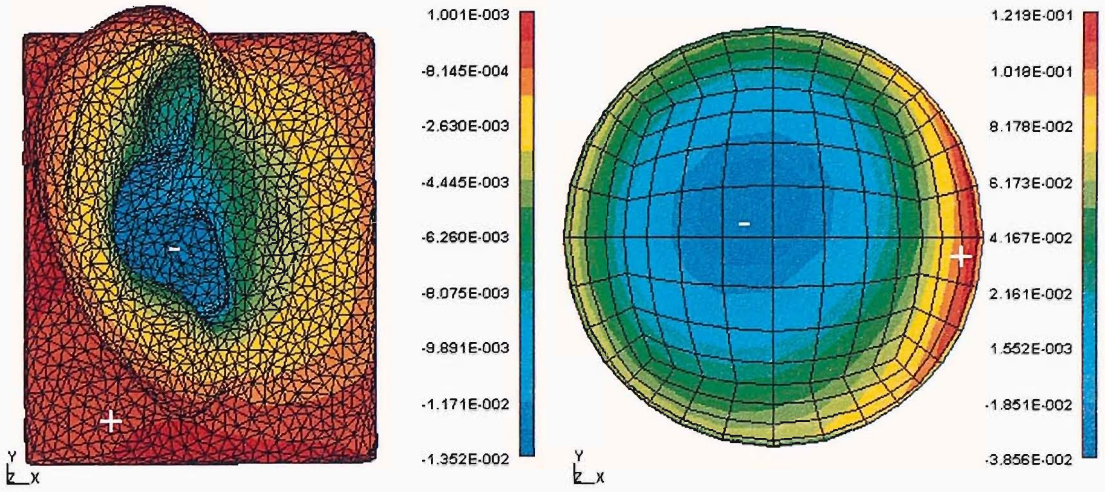


(b)

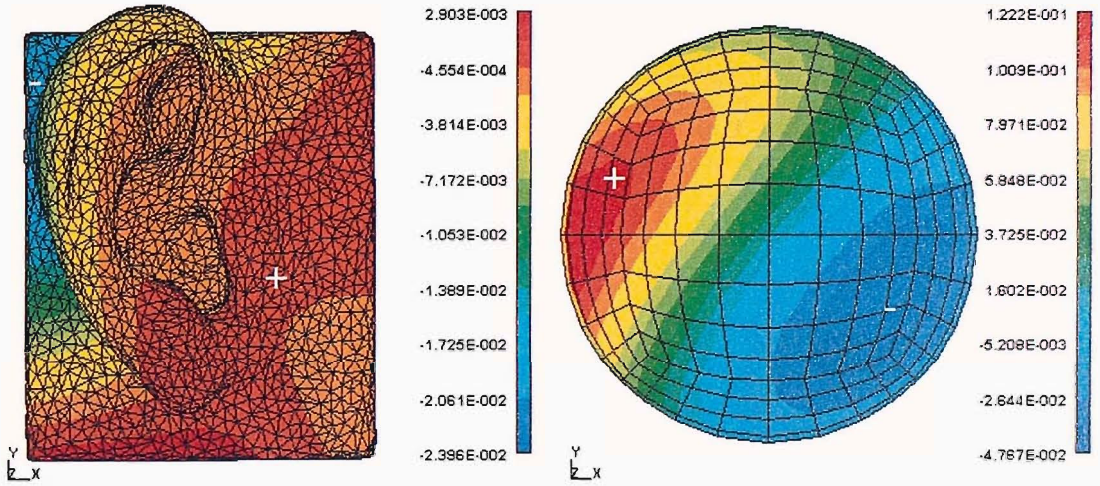


(c)

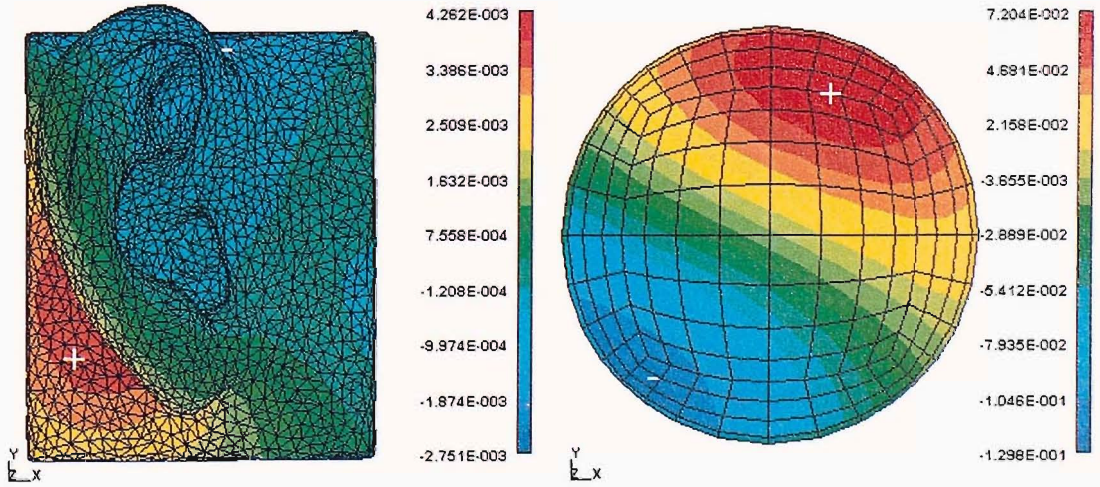
Figure 8-13: The real parts of the left and right singular vectors associated with the first three dominant singular values of the numerically generated 2825 × 209 Green function matrix for the DB60 pinna at 4.6 kHz (a) σ_1 (b) σ_2 (c) σ_3.



(a)

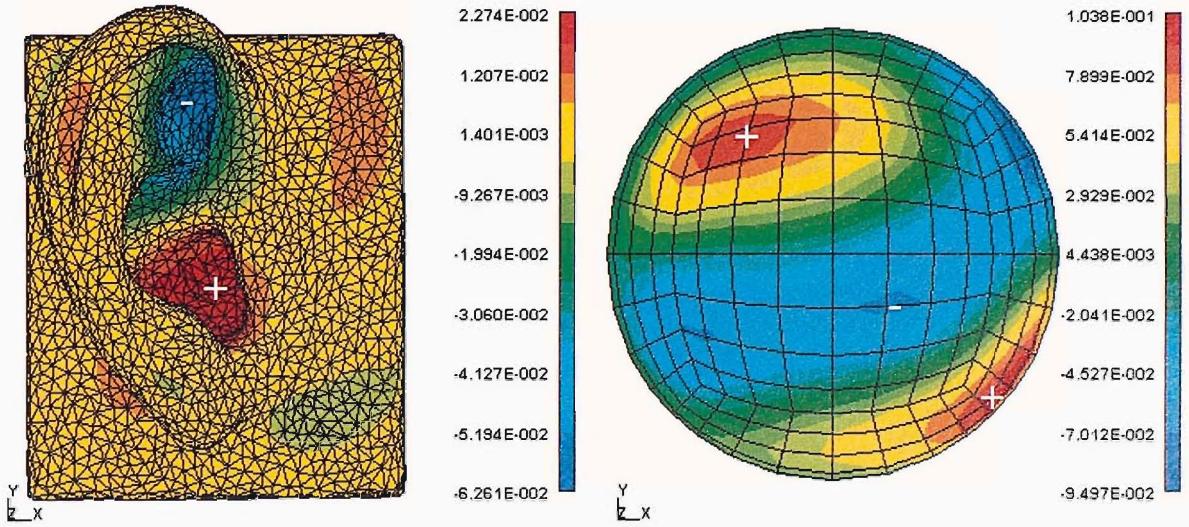


(b)

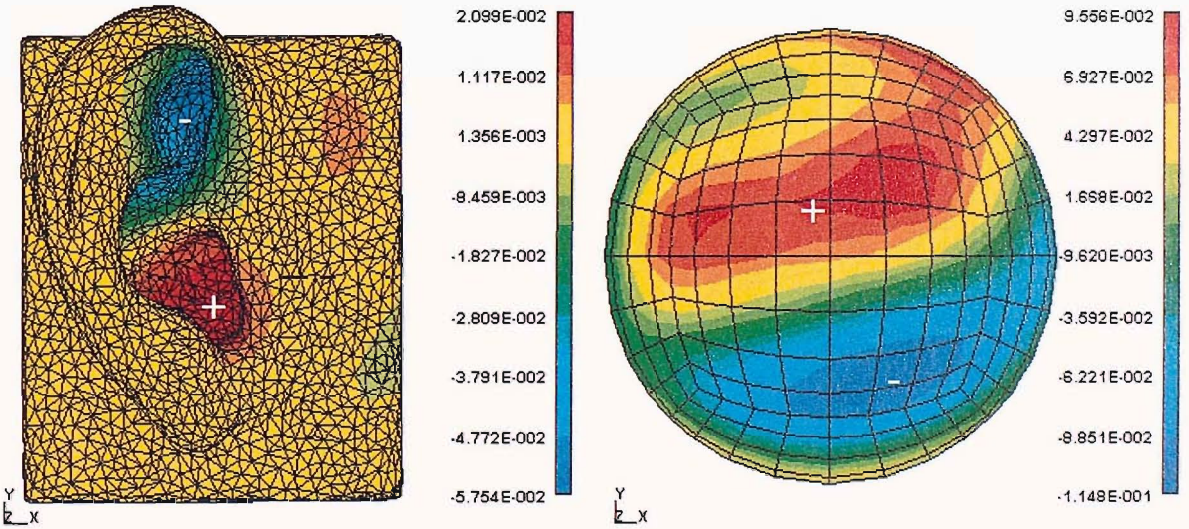


(c)

Figure 8-14: The imaginary parts of the left and right singular vectors associated with the first three dominant singular values of the numerically generated 2825 × 209 Green function matrix for the DB60 pinna at 4.6 kHz (a) σ_1 (b) σ_2 (c) σ_3.

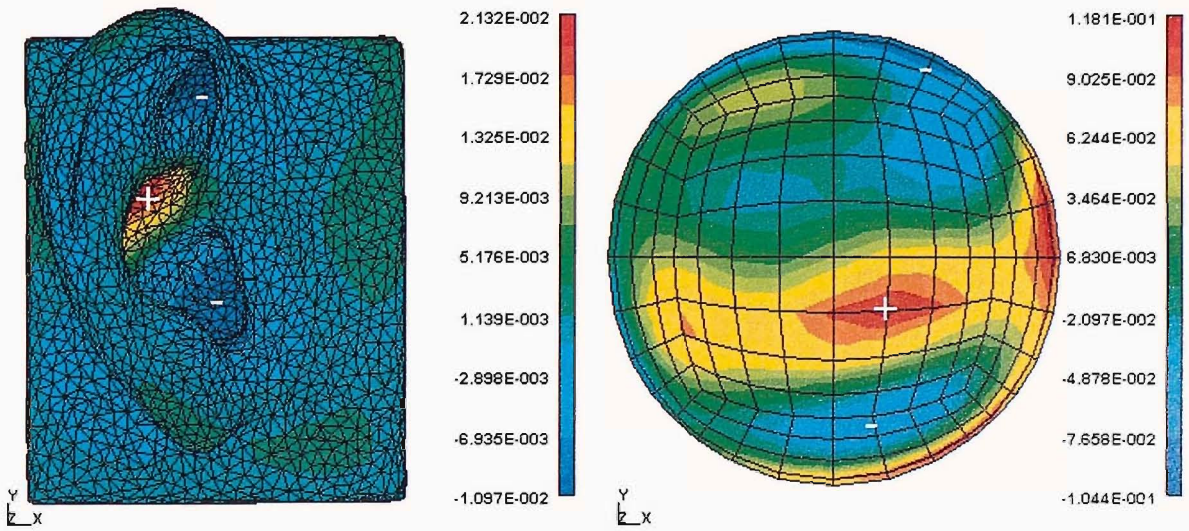


(a)

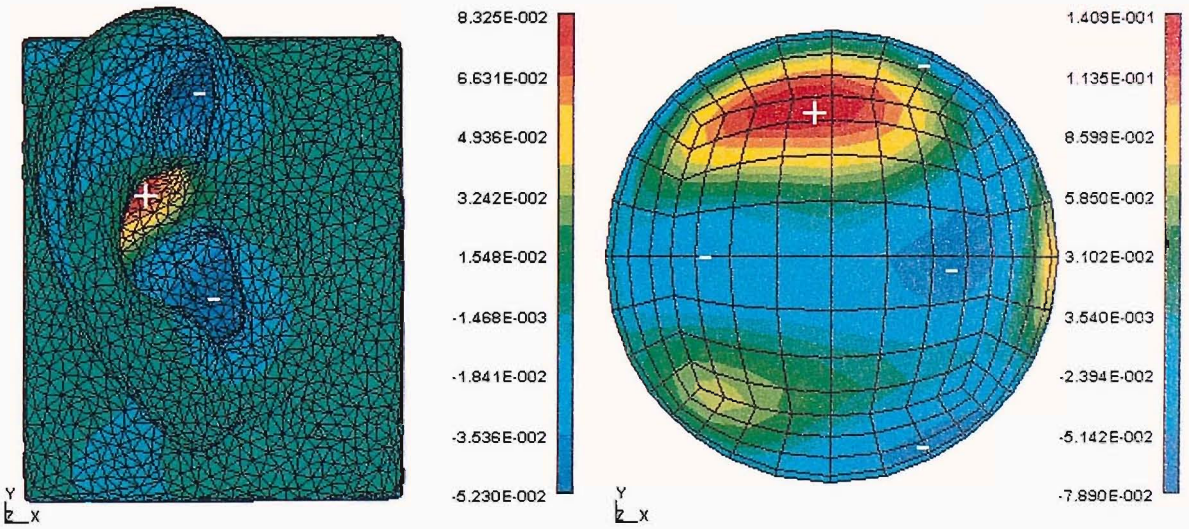


(b)

Figure 8-15: The left and right singular vectors associated with the dominant singular value (σ_1) of the numerically generated 2825 × 209 Green function matrix for the DB60 pinna at 8.8 kHz (a) real values and (b) imaginary values.

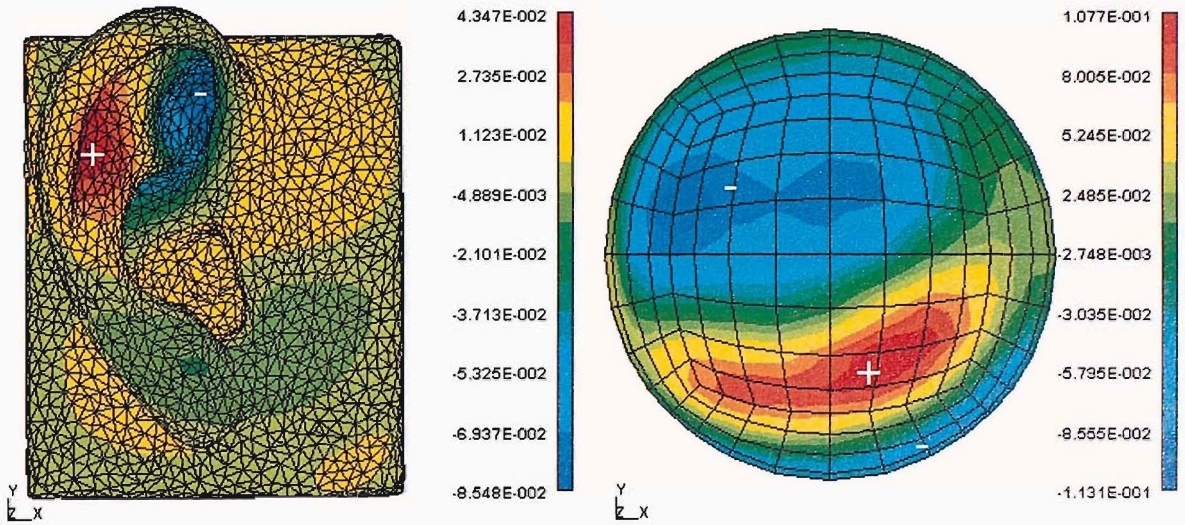


(a)

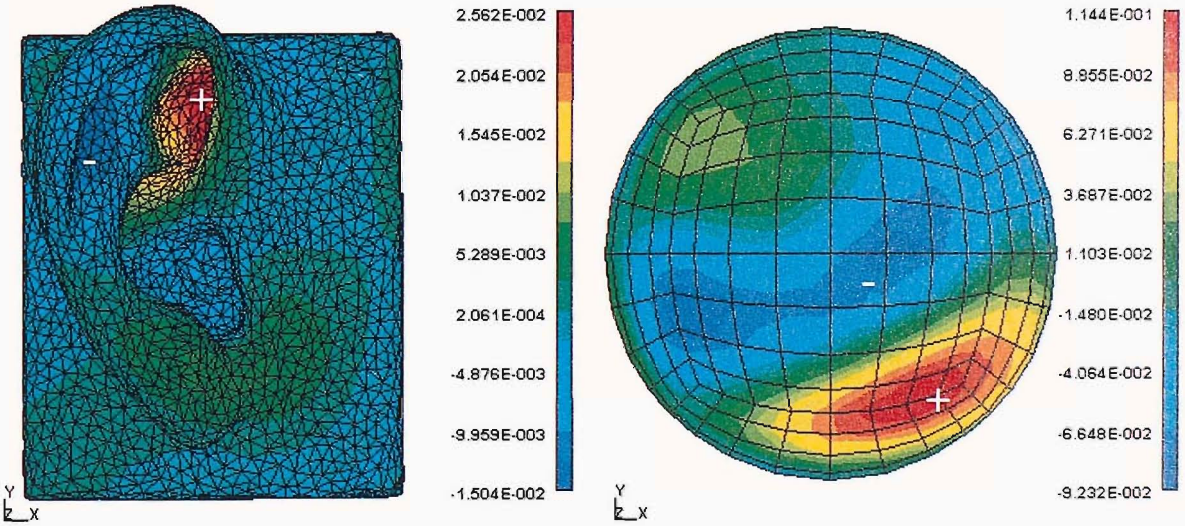


(b)

Figure 8-16: The left and right singular vectors associated with the dominant singular value (σ_1) of the numerically generated 2825 × 209 Green function matrix for the DB60 pinna at 10.3 kHz (a) real values and (b) imaginary values.

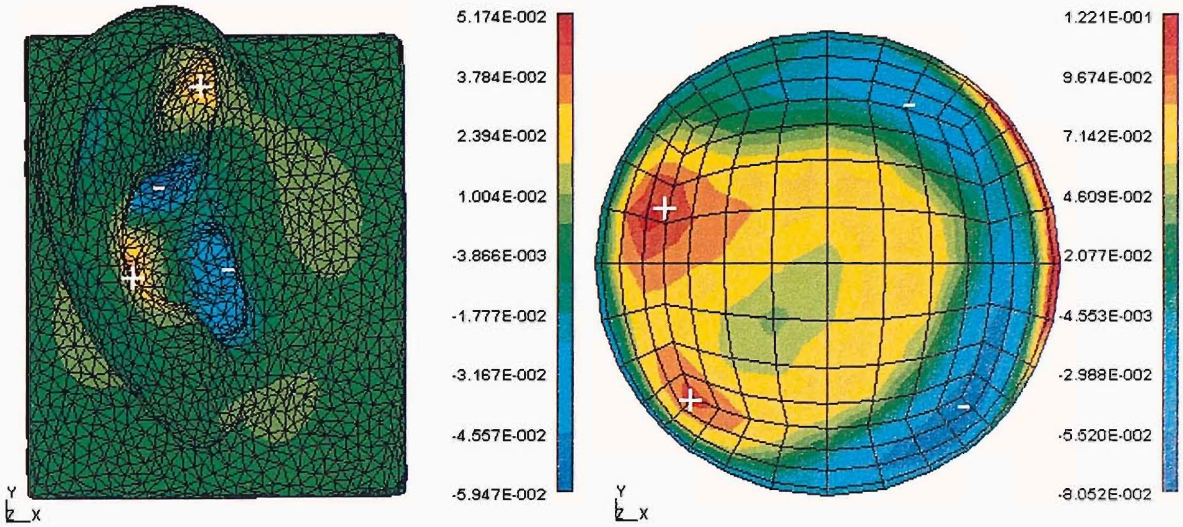


(a)

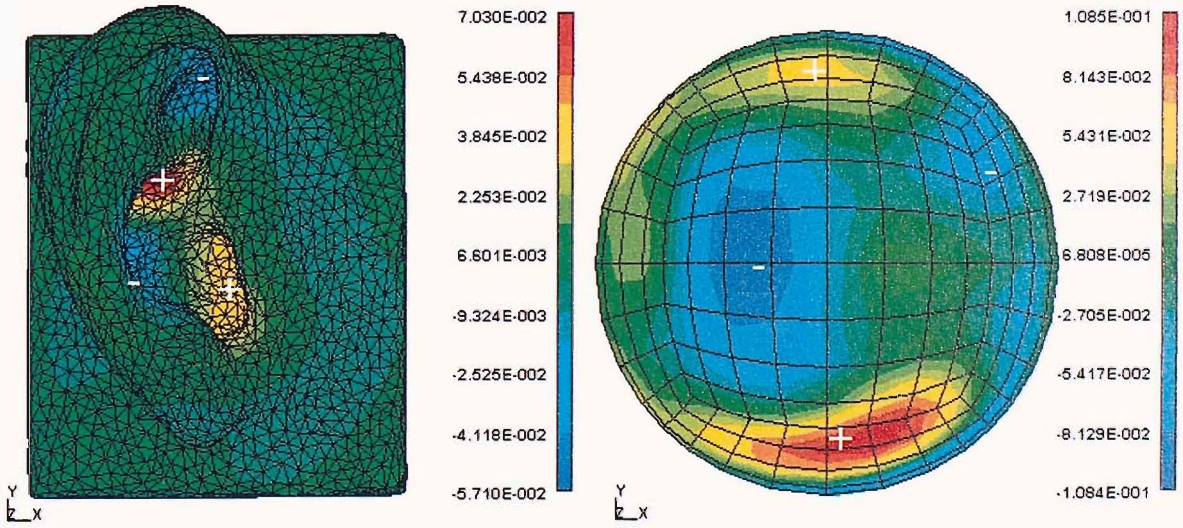


(b)

Figure 8-17: The left and right singular vectors associated with the second dominant singular value (σ_2) of the numerically generated 2825 × 209 Green function matrix for the DB60 pinna at 10.3 kHz (a) real values and (b) imaginary values.

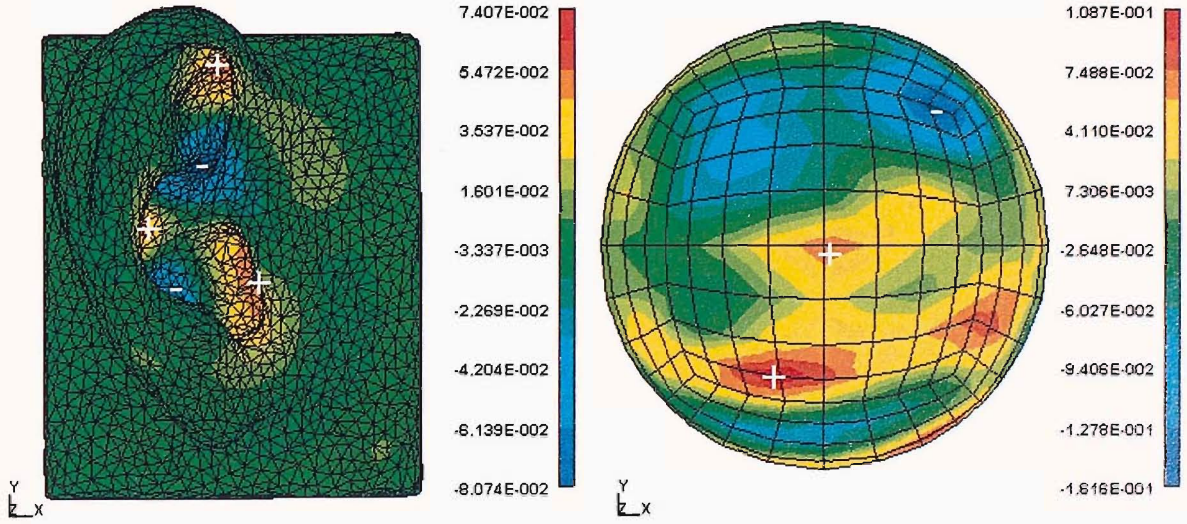


(a)

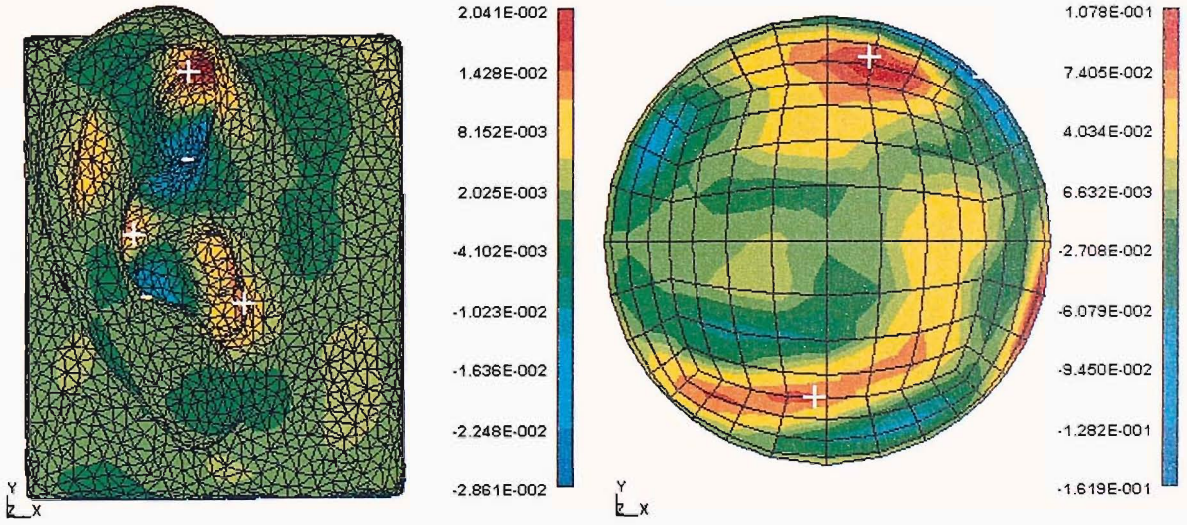


(b)

Figure 8-18: The left and right singular vectors associated with the dominant singular value (σ_1) of the numerically generated 2825 × 209 Green function matrix for the DB60 pinna at 13.8 kHz (a) real values and (b) imaginary values.



(a)



(b)

Figure 8-19: The left and right singular vectors associated with the dominant singular value (σ_1) of the numerically generated 2825 x 209 Green function matrix for the DB60 pinna at 17.5 kHz (a) real values and (b) imaginary values.

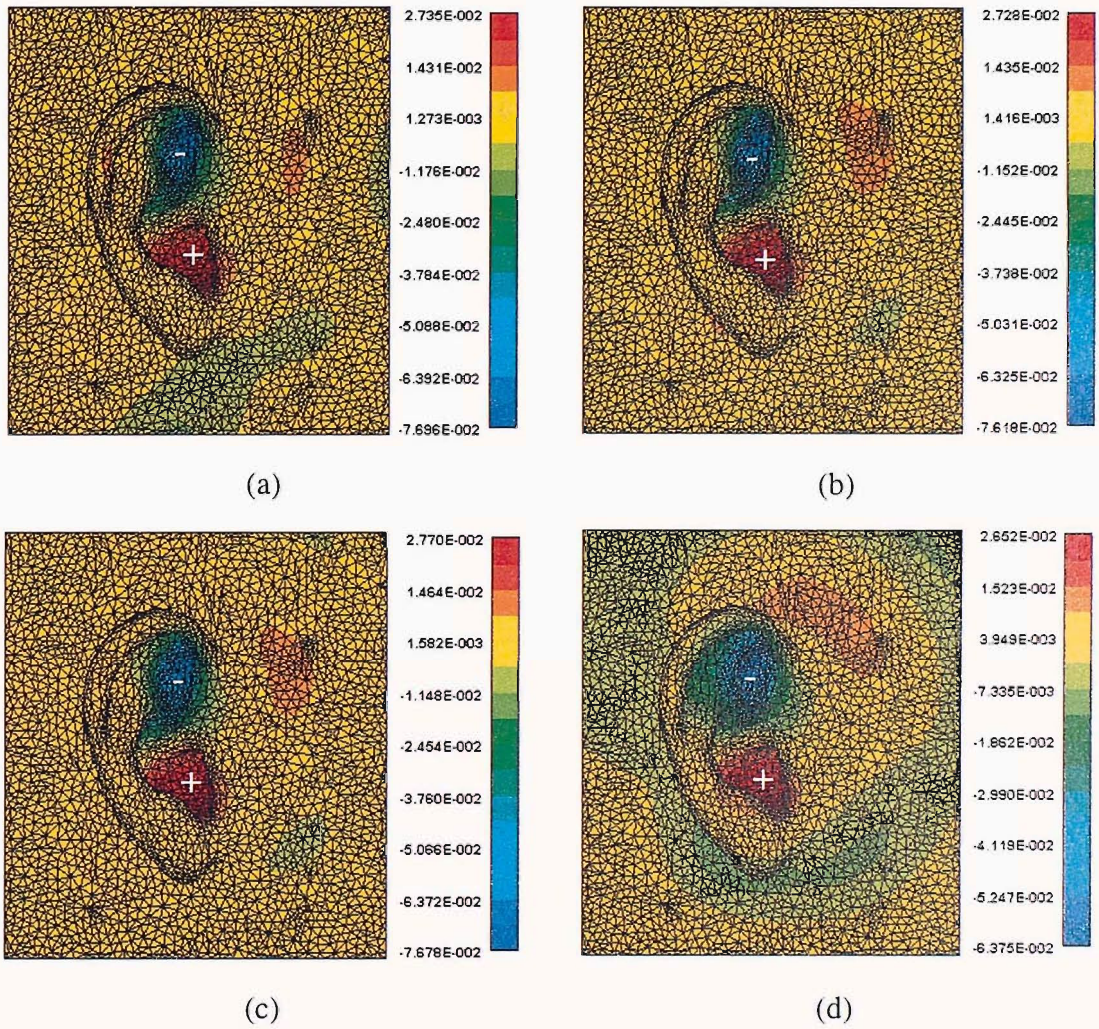
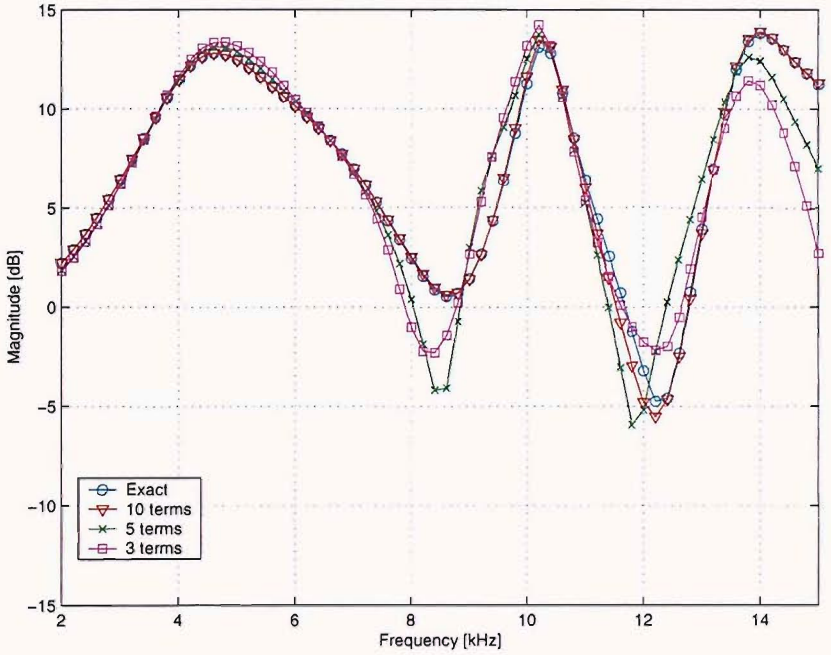
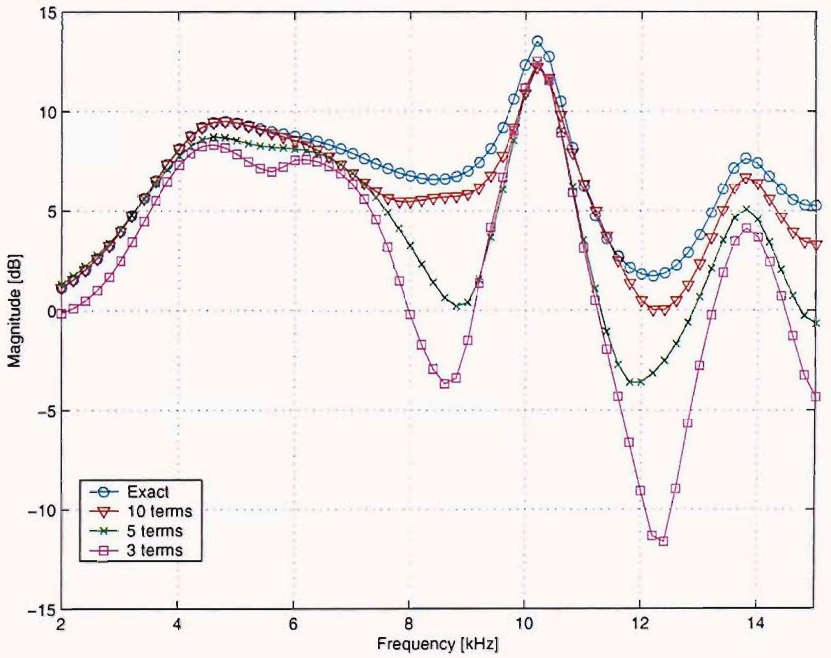


Figure 8-20: Comparison of spatial basis function (the real part of the left singular vectors) associated with the first dominant singular value (σ_1) of the second peak (see Figure 8-8 and Figure 8-12). Different number of sources and positions are used (a) 17 sources approximately uniformly distributed on the upper hemisphere (at 7.8 kHz) (b) 57 sources approximately uniformly distributed on the upper hemisphere (at 7.8 kHz) (c) 209 sources approximately uniformly distributed on the upper hemisphere (at 7.8 kHz) (d) 36 sources uniformly distributed at grazing incidence (at 8.8 kHz).

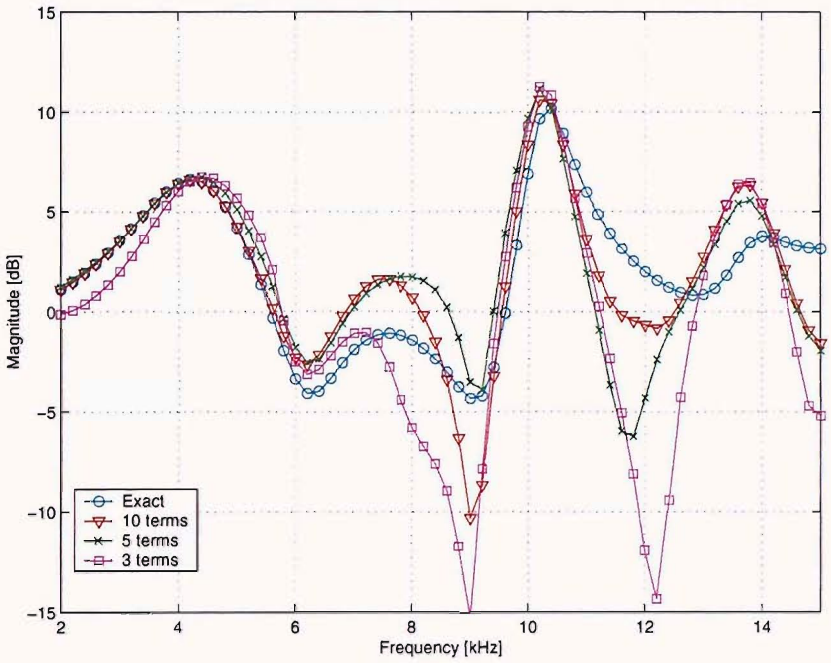


(a)

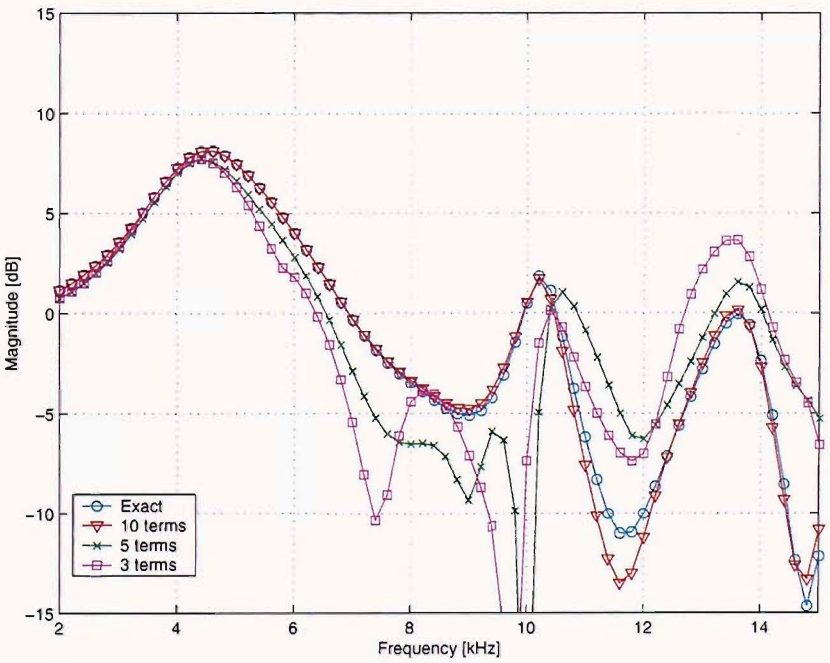


(b)

Figure 8-21: Reconstruction of the frequency response detected at the blocked ear canal of baffled DB60. The curves show comparisons of the response obtained with direct calculation (using the DBEM) and with the SVD with limited number of terms in the series. The source is positioned at grazing incidence (a) front ($\phi = 0°, \theta = 0°$) (b) above ($\phi = 0°, \theta = 90°$).



(c)



(d)

Figure 8-21 (cont.): (c) below ($\phi = 0°, \theta = 270°$) (d) rear ($\phi = 0°, \theta = 180°$).

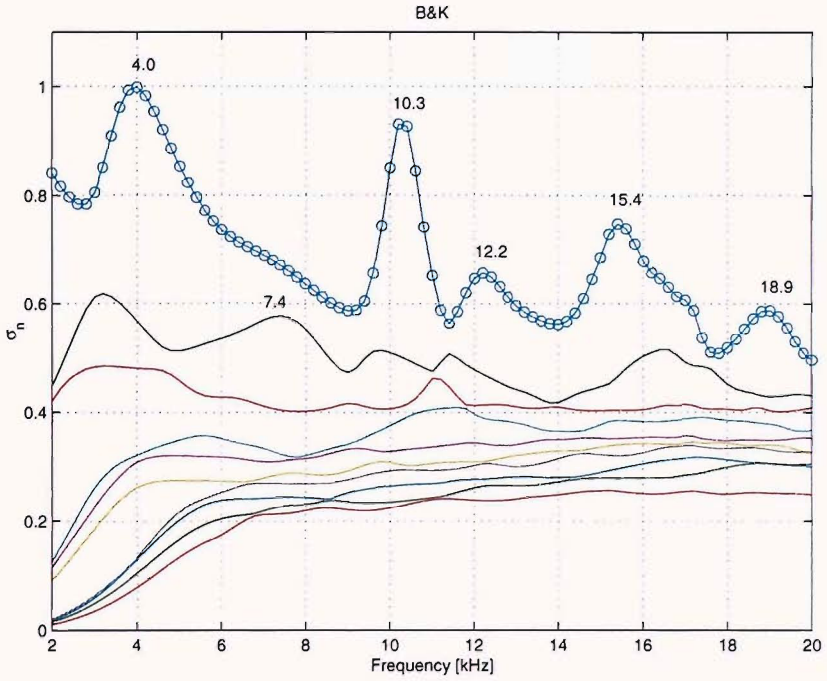


Figure 8-22: The singular values of the numerically generated 3906 × 209 Green function matrix relating 3906 points on the surface of the blocked meatus of B&K pinna mounted on a rigid baffle to 209 points distributed approximately uniformly on the upper hemisphere. The calculation is undertaken at 91 frequencies.

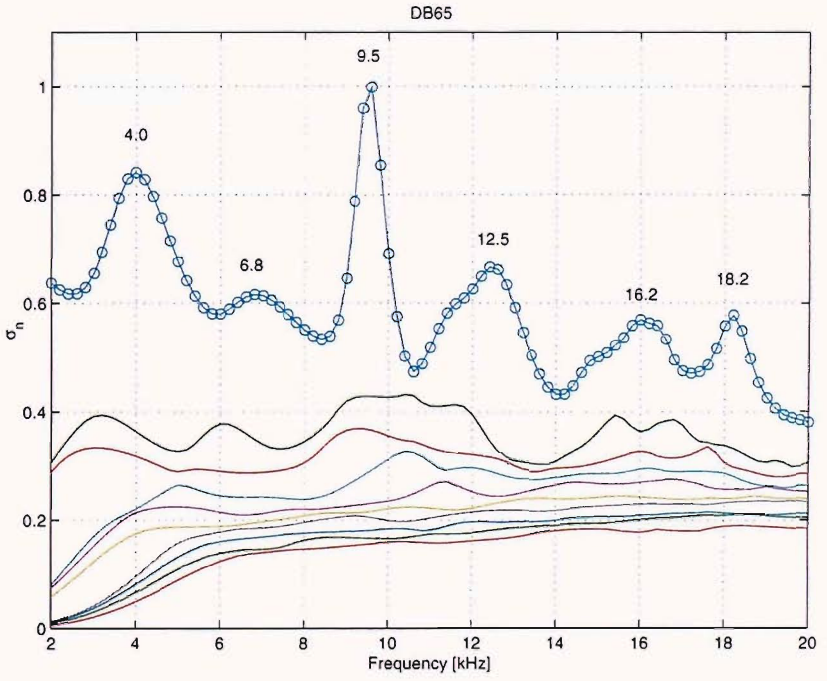


Figure 8-23: The singular values of the numerically generated 3389 × 209 Green function matrix relating 3389 points on the surface of the blocked meatus of DB65 pinna mounted on a rigid baffle to 209 points distributed approximately uniformly on the upper hemisphere. The calculation is undertaken at 91 frequencies.

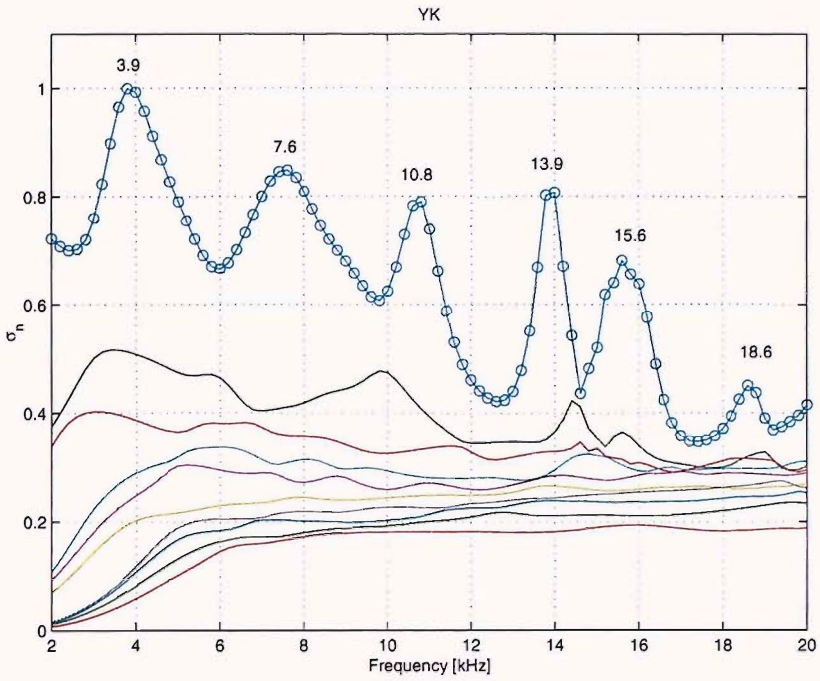


Figure 8-24: The singular values of the numerically generated 3392 × 209 Green function matrix relating 3392 points on the surface of the blocked meatus of YK pinna mounted on a rigid baffle to 209 points distributed approximately uniformly on the upper hemisphere. The calculation is undertaken at 91 frequencies.

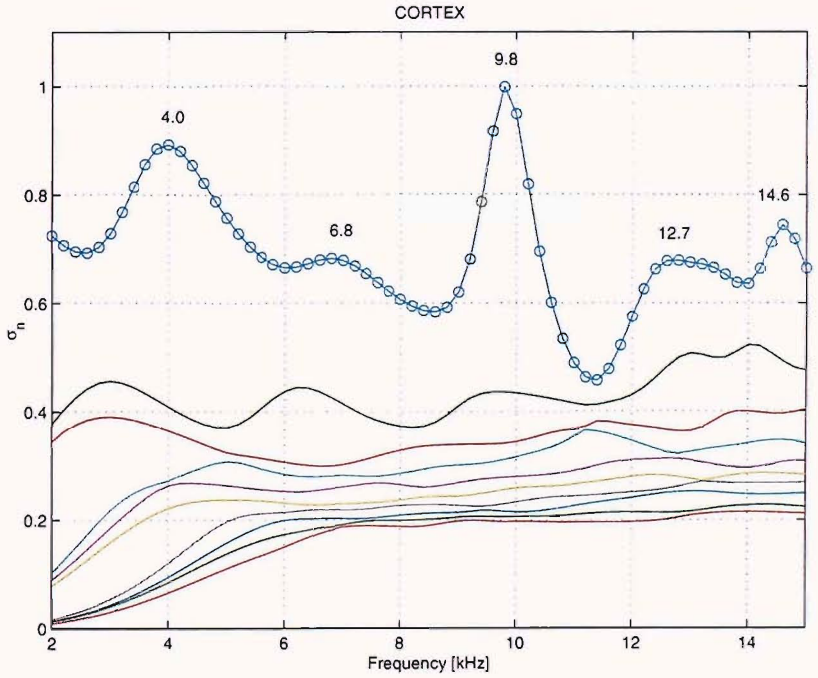


Figure 8-25: The singular values of the numerically generated 3390 × 209 Green function matrix relating 3390 points on the surface of the CORTEX pinna mounted on a rigid baffle to 209 points distributed approximately uniformly on a far field hemisphere of radius 3m. The calculation is undertaken at 66 frequencies.

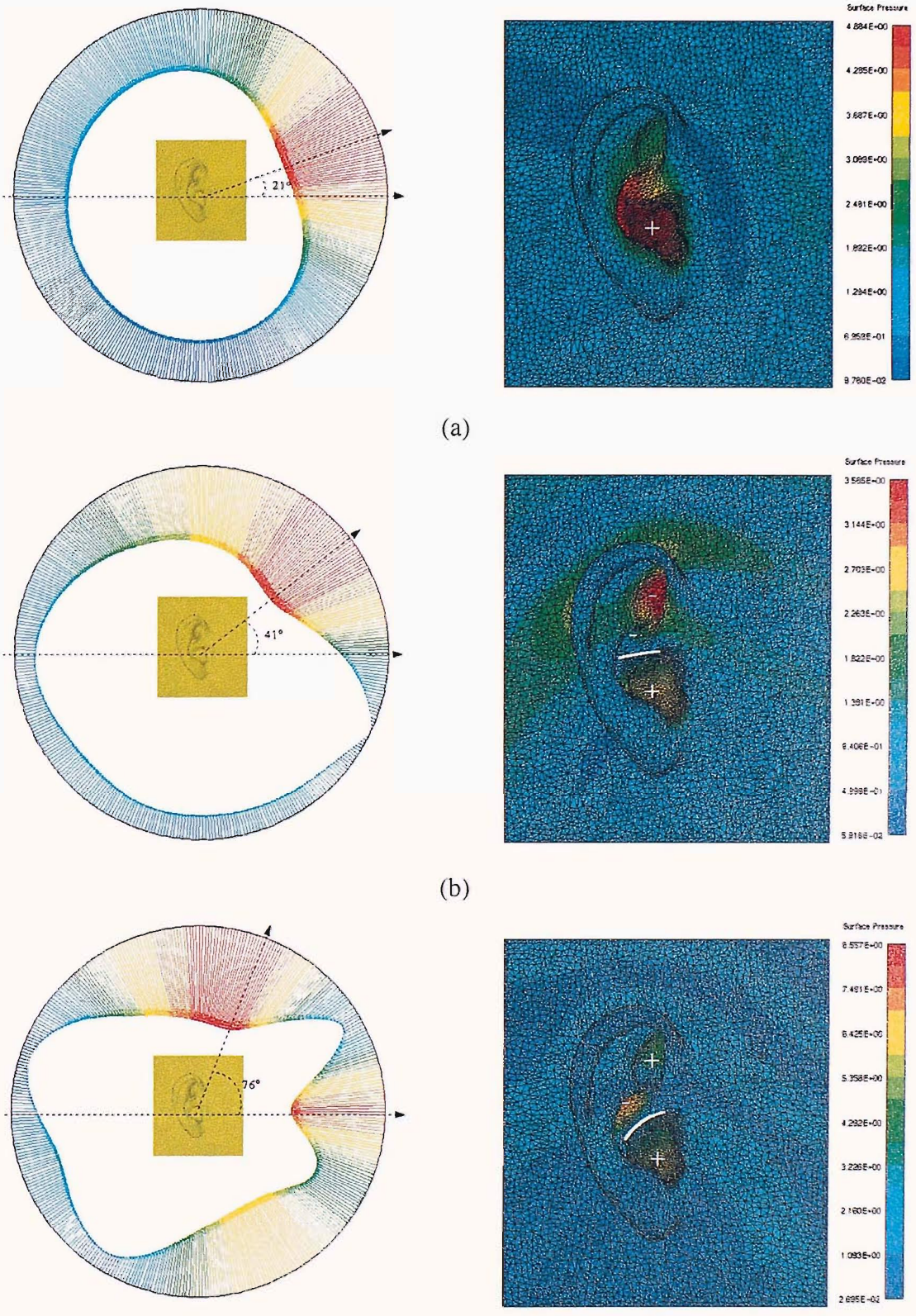
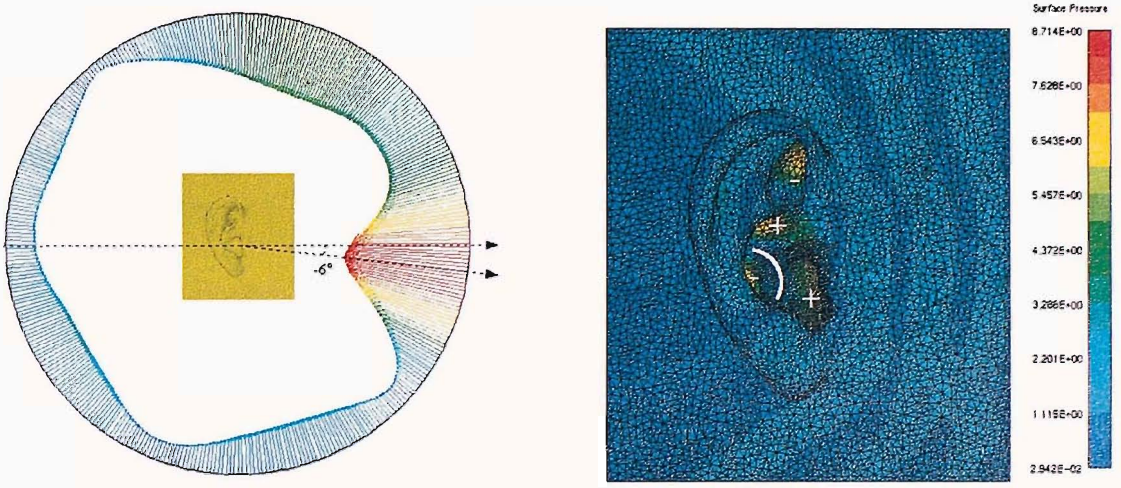
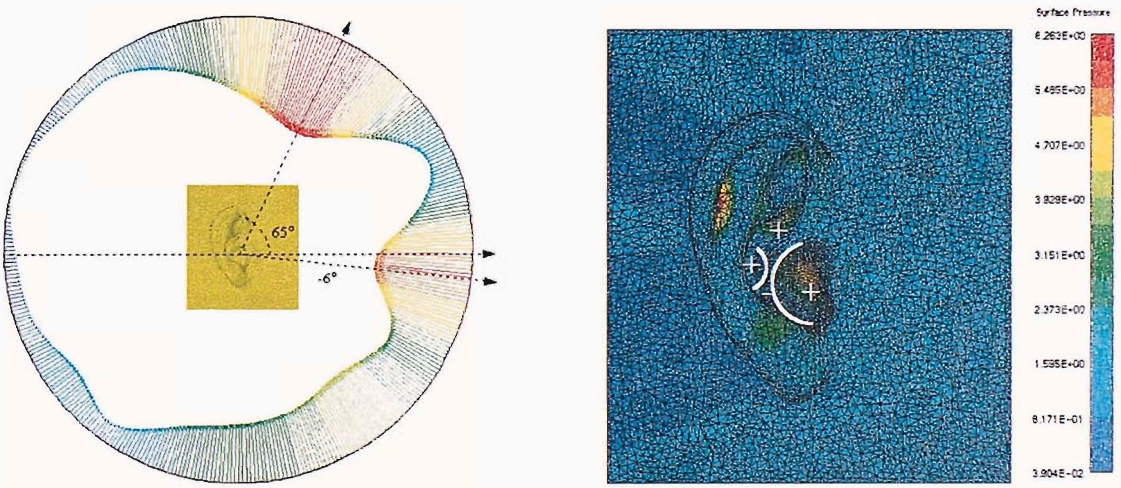


Figure 8-26: The figures on the left show the level of the response at the blocked meatus of DB-60 at a specific frequency in a vector/directivity format as a function of the location of the source on the circle. The higher the response, the longer the arrow facing the centre of the pinna. On the right, the absolute pressure is shown when the pinna is excited with a plane wave source at the angle specified on the left. Negative/positive values indicate negative/positive phase, respectively. The figures correspond to the following frequencies (a) 4.5 kHz (b) 7.8 kHz (c) 10.3 kHz.

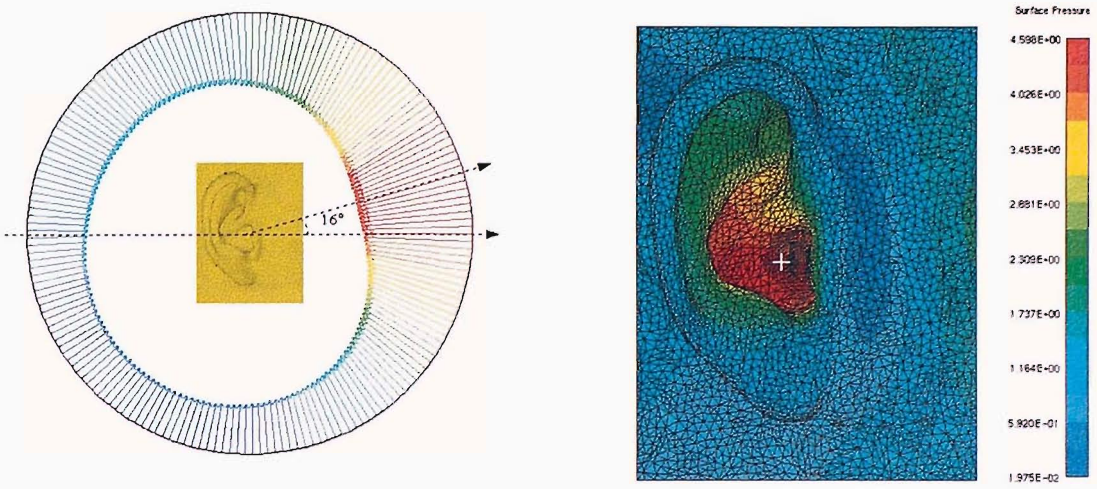


(d)

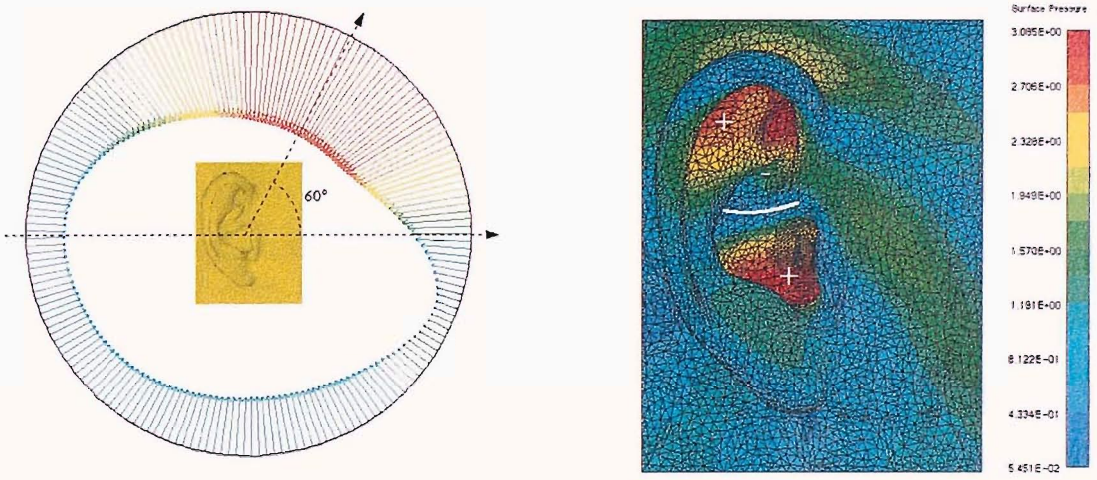


(e)

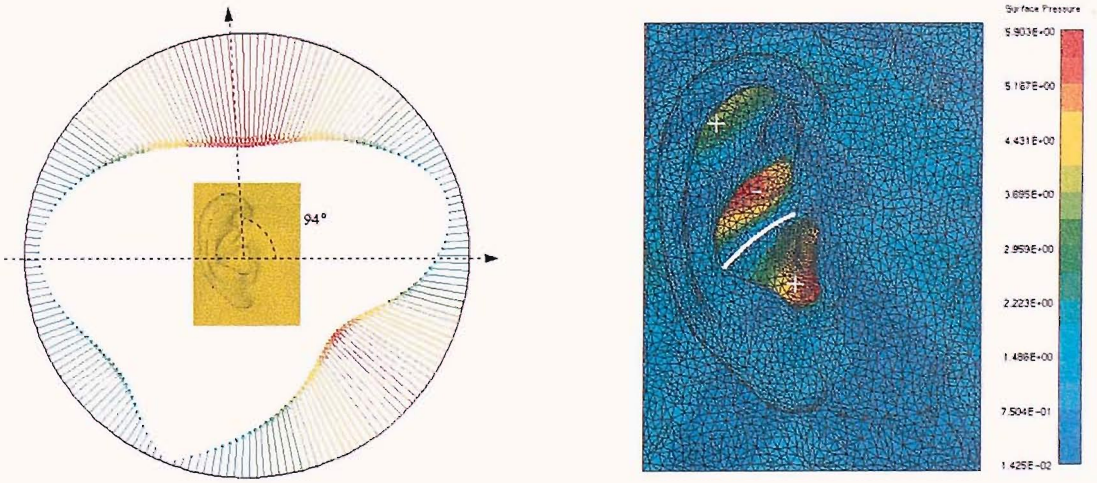
Figure 8-26 (cont.): (d) 13.8 kHz (e) 17.0 kHz.



(a)

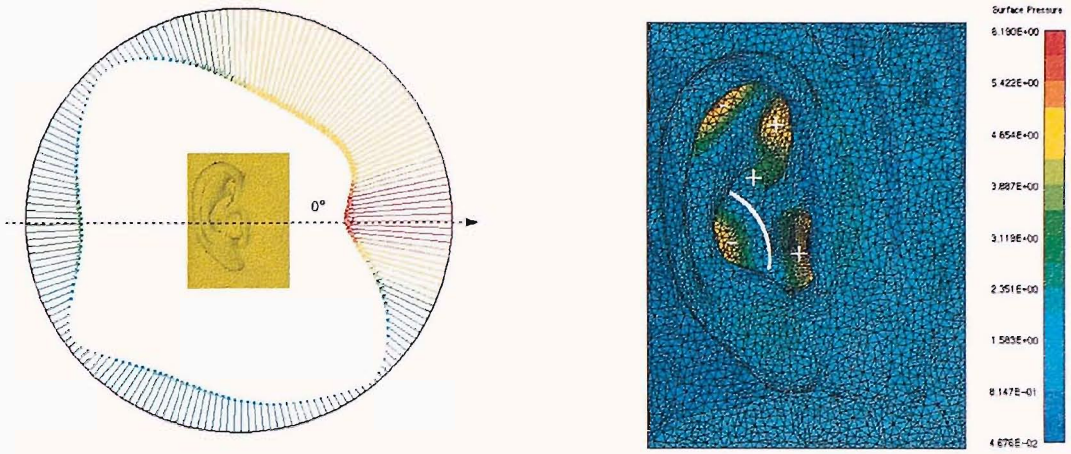


(b)

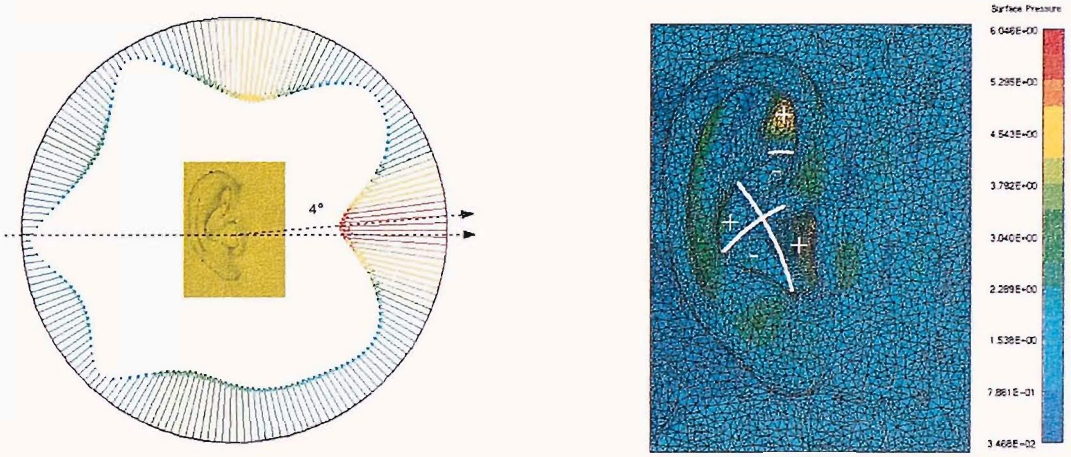


(c)

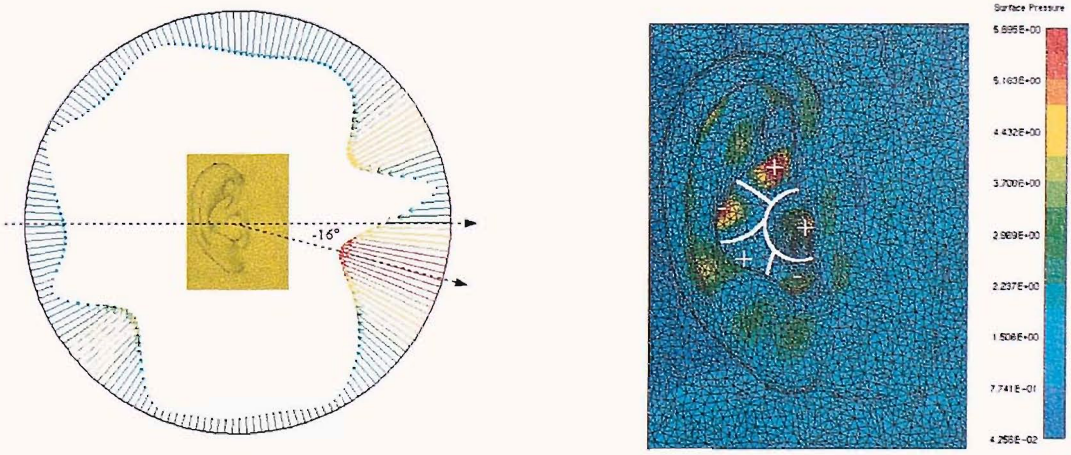
Figure 8-27: The modes of the DB65 modes with a similar format presented in Figure 8-26. The figures correspond to the following frequencies (a) 4.2 kHz (b) 7.2 kHz (c) 9.6 kHz.



(d)



(e)



(f)

Figure 8-27 (cont.): (d) 11.6 kHz (e) 14.8 kHz (f) 17.8 kHz.

CHAPTER 9

EXAMPLES OF THE SOUND FIELDS OF VIRTUAL ACOUSTIC IMAGING SYSTEMS

9.1 INTRODUCTION

Currently, one of the main limitations of the production of virtual acoustic images with loudspeakers is the equalization zone, also known as the 'sweet spot'. In principle, by using appropriate inverse filtering, which eliminates the cross-talk paths (Atal and Schroeder, 1962), it is possible to transform the 'binaural synthesis' or the 'binaural recording' techniques from headphones to loudspeakers. The common feature in both cases is that the goal is to reproduce the sound pressure at two points, the eardrums. In the case of headphones, the acoustical characteristics of the cavity and the headphones can be compensated to a certain degree. When loudspeakers are used, still, the pressure at only *two* points is reconstructed (it is possible however to eliminate this problem and reconstruct the wave field with a large number of loudspeakers (Berkhout, 1993)). Therefore the designer of any virtual acoustic system using loudspeakers does not know or cannot predict the variation of the sound field in the vicinity of the head.

The positions of the loudspeakers seem to play a significant part in the size of the 'sweet-spot'. The recursive behaviour of the cross talk signals shown in the patent of Atal and Schroeder (1962), and by Nelson *et al* (1997) and Kirkeby *et al* (1998) causes a 'ringing' behaviour (in the time domain) of the interacting waves. This results in 'colouration'

whenever the listener is not positioned in the 'sweet-spot' or even when non-individualised HRTFs are used.

In the following sections we demonstrate a very useful feature of the BEM: the inputs to the loudspeakers are filtered and the sound field is solved to produce the desired signals at the ears of the head model. The visualisation of the total sound pressure can be useful for example in the design of virtual acoustic systems using individual HRTFs with adaptive systems (either with laser trackers, video cameras, or electromagnetic trackers) where the filters used on the input to the loudspeakers are updated in real time in order to compensate for head movements.

An outline of this chapter is as follows: first, the sound field around KEMAR due to a single monopole source, both in the frequency and time domains, is calculated with the BEM and IFEM. Subsequently, cross-talk cancellation is implemented with a general multi-channel formulation. The HRTFs and sound field around KEMAR for the 2×2 Stereo Dipole is calculated and visualised, and finally the response of the head and the sound fields of 4×4 systems (four loudspeakers and four microphones) for a single, or two listeners are presented.

9.2 HEAD SCATTERED ACOUSTIC FIELD DUE TO A MONOPOLE SOURCE

9.2.1 Frequency domain response with the DBEM

When the human head is simplified to the classical rigid sphere, analytical expressions exist for both the frequency response on the sphere surface (Equation (4.1)) and also the total sound field (Equation (4.2)) due to an excitation of a plane wave or a monopole source (Equation (7.29)). However, it was demonstrated in Chapters 4, 5 and 6 that the pinna can produce a much more complex structure of frequency response at high frequencies. An

example of the sound field around KEMAR (without a torso) calculated with the DBEM is shown in Figure 9-1. In this case the mesh model of KEMAR included 10283 nodes and 20562 elements (of a full head), and the mesh of the field points around the head comprises of 200 × 200 points. The total (incident + scattered) sound pressure in an area of 0.5 m × 0.5 m is shown on a linear scale. The source is positioned in front of the head at a distance of 0.5 m from the centre of the head. In Figure 9-1a and Figure 9-1b the characteristics of the sound fields are very similar to the case of a sphere where at 200 Hz the head is almost transparent to the propagating waves. At 1 kHz the shadow zones appear at the rear, and an increase of pressure occurs in front of the head due to the superposition of the incident and reflected waves. As frequency increases (at 2 kHz and 5 kHz, presented in Figures 9-1c and 9-1d) the interference between these waves causes more complex patterns in front of the head and higher attenuation is noticed at the rear. In addition, the scattered sound field is slightly uneven on the left and right sides due probably to slight misalignment and due to the asymmetrical shape of KEMAR mesh model with DB60 and DB61 pinnae.

9.2.2 Frequency domain response with the IFEM

As indicated in Chapter 4, the IFEM method was implemented successfully for the case of a rigid sphere and a feasibility study was undertaken in order to investigate the success of the method when the KEMAR head is modelled.

Currently, one of the main disadvantages of using the IFEM in practice for 'real' problems is the generation of the IFEM mesh. An IFEM mesh was constructed with PRE-SYSNOISE. Due to the dense resolution of the vertices in the vicinity of the pinna the procedure of filling tetrahedral elements to the smallest circumscribing ellipsoid results in a very large number of elements. A relatively small mesh was generated which includes 10800 elements of the original BEM mesh as the inner layer, 41100 tetrahedral elements of

the conventional FEM, and 2400 elements of the infinite layer (see Figure 9-2). This mesh with its maximum edge length corresponds to a maximum frequency of 2.5 kHz. Two cases were compared with the BEM in Figure 9-3: first order approximation (only the first term in the $1/r$ expansion, see Equation (3.39)), and with order 4. For the latter, the average error of the total pressure due to a source at position ($r = 1$ m, at $\phi = 45°, \theta = 45°$) was 1.3%. The highest sensitivity appears at the rear due to large attenuation and limited mesh resolution.

No improvement in CPU time was achieved in this case but it is claimed that the efficiency increases for larger models. This issue requires more advanced mesh generation tools.

9.2.3 Time domain response with the DBEM

The direct simulation of impulse responses using the time domain formulation of the BEM is still not recommended due to stability problems in the solutions. The method undertaken and described below is based on using the Inverse Fourier transforms of the computed responses in the frequency domain.

It is possible to visualise the propagation of the spherical wave, impinging on the surface of the head, by calculating the inverse Fourier transform of the frequency response at each field point. To ensure causality of the impulse response of each point, and also for the clarity of animation, the response was multiplied with a half Hanning window with the Nyquist frequency being at 6.4 kHz. The following steps were followed:

- The frequency response of each field point around KEMAR is calculated with the DBEM around the head up to 6.4 kHz.
- The frequency is 'mirrored' at the Nyquist frequency by appending the conjugate values.

- The new frequency response is multiplied by half of the required window in the frequency domain.
- An inverse FFT is applied followed by a cyclic shift to produce a delay at half the number of the sampling points.

Examples of the effect of applying various windows for the generation of the impulse response of field points in the case of scattering from a sphere are given in Kahana (1997a).

Figure 9-4 includes 12 snapshots of the animated sound pressure. When the spherical wave impinges on the head, a secondary wave is seen to propagate backwards. When the main wave leaves the head, another spherical wave appears. The amplitude of this wave is small, compared with the incident wave, and its effect was emphasised in this simulation by positioning the source in the near field at 0.5 m away. The further away the source, the less significant this effect.

9.3 MULTI-CHANNEL CROSS-TALK CANCELLATION

The calculation of the sound field for virtual acoustic imaging systems is given below for a general number and arrangement of loudspeakers and microphones. A diagram describing the signal processing problem of the sound reproduction system is shown in Figure 9-5. By assuming that the system is working in discrete time, we can adopt the definition of the signals in the z-domain. It is assumed that the T microphones which record the sound field are restricted to the horizontal plane. The T observed signals are described by the vector $\mathbf{u}(z)$. The objective is to reproduce these signals as closely as possible at the equivalent locations in the listener space. S sources are used to reproduce the field and their input signals are described by the vector $\mathbf{v}(z)$. These sources produce signals at R locations in the listening space, these signals comprising the vector $\mathbf{w}(z)$. For our purposes, we assume that the microphones are positioned at R locations in the listening space, which are equivalent to

the T locations of the microphones in the recording space such that T=R. The R desired signals are described by the vector $\mathbf{d}(z)$, and the R performance error signals are described by the vector $\mathbf{e}(z)$.

The goal is to design the matrix $\mathbf{H}_{m,\mathbf{A}}(z)$ which minimises the magnitude of the vector of error signals, where $\mathbf{A}(z)$ and $\mathbf{C}(z)$ are given. From the block diagram shown in Figure 2 and Equations 1-3, we can derive the following relationships:

$$\mathbf{v}(z) = \mathbf{H}_{m,\mathbf{A}}(z)\mathbf{u}(z) \tag{9.1}$$

$$\mathbf{d}(z) = z^{-m}\mathbf{A}(z)\mathbf{u}(z) \tag{9.2}$$

$$\mathbf{w}(z) = \mathbf{C}(z)\mathbf{v}(z) \tag{9.3}$$

and

$$\mathbf{e}(z)=\mathbf{d}(z)-\mathbf{w}(z) \tag{9.4}$$

The calculation of the optimal filters in the least squares sense exposes some of the fundamental problems encountered with deconvolution of a multi-channel system; these filters are constrained to be stable but not constrained to be either causal or finite duration. The goal is to find the signal vector $\mathbf{v}(e^{j\omega\Delta})$ which operates in discrete time on sampled input signals, and minimises the sum of squared errors between the desired and reproduced signals.

The quadratic cost function that is to be minimised is given by

$$J(e^{j\omega\Delta}) = \mathbf{e}^{\mathrm{H}}(e^{j\omega\Delta})\mathbf{e}(e^{j\omega\Delta}) + \beta\mathbf{v}^{\mathrm{H}}(e^{j\omega\Delta})\mathbf{v}(e^{j\omega\Delta}) \tag{9.5}$$

The cost function thus consists of the sum of the squared errors $\mathbf{e}^{\mathrm{H}}(e^{j\omega\Delta})\mathbf{e}(e^{j\omega\Delta})$ plus the sum of squared source input voltages $\mathbf{v}^{\mathrm{H}}(e^{j\omega\Delta})\mathbf{v}(e^{j\omega\Delta})$ multiplied by a positive real number β, where $\mathbf{e}(e^{j\omega\Delta})$ and $\mathbf{v}(e^{j\omega\Delta})$ are vectors containing the Fourier transforms of the

error signals and source input signals respectively, and β is a regularisation parameter. The latter quantifies the relative weighting in the cost function given to the 'effort' used in minimising the sum of squared errors. By varying β from zero to infinity, the solution changes gradually from minimising only the performance error to minimising only the effort cost . The optimal vector of source input signals $\mathbf{v}_{\text{opt}}(e^{j\omega\Delta})$ is given by

$$\mathbf{v}_{\text{opt}}(e^{j\omega\Delta}) = [\mathbf{C}^{\text{H}}(e^{j\omega\Delta})\mathbf{C}(e^{j\omega\Delta}) + \beta\mathbf{I}]^{-1}\mathbf{C}^{\text{H}}(e^{j\omega\Delta})\mathbf{A}(e^{j\omega\Delta})\mathbf{u}(e^{j\omega\Delta}) \tag{9.6}$$

With no modelling delay (m=0) and according to Equation (9.1), the optimal filter matrix becomes in the z-domain,

$$\mathbf{H}_{0,\mathbf{A}}(z) = [\mathbf{C}^{\text{T}}(z^{-1})\mathbf{C}(z) + \beta\mathbf{I}]^{-1}\mathbf{C}^{\text{T}}(z^{-1})\mathbf{A}(z) \tag{9.7}$$

The matrix of optimal inverse filters can be represented by $\mathbf{H}_{0,\mathbf{I}}$ (again, in a special case where m=0) which we define as the generalised cross-talk cancellation matrix, given by

$$\mathbf{H}_{0,\mathbf{I}}(z) = [\mathbf{C}^{\text{T}}(z^{-1})\mathbf{C}(z) + \beta\mathbf{I}]^{-1}\mathbf{C}^{\text{T}}(z^{-1}) \tag{9.8}$$

This term is used whenever the target matrix is an identity matrix regardless of the value of m. As mentioned above, it is crucial to include a modeling delay with a sufficient length in order to ensure the best performance under the constraint of causality of all the filters.

The method of calculating the sound field with a numerical solution includes the following stages:

- The pressure at each of the blocked ear-canals is calculated due to a single source at a time. The position of the source is equivalent to the position of a *real* source.
- A matrix of the 'electro-acoustic' response is obtained (2×2 for two speakers and two ears, and 4×4 for four speakers and four ears). e.g. C_{12} is the response at ear no. 1 due to source no. 2.

- A desired signal is determined. Generally "0" pressure is assigned to one of the ears, and "1" to the other ear, i.e. $\mathbf{d}=[1\ 0]^T$ for a 2×2 system, where $\mathbf{d}$ is the vector of the desired signals (this is equivalent to attenuation of ∞ dB at one ear compared with the other).

- Each source is filtered (via appropriately designed inverse filters) to produce the desired signals with a table of complex pressure values at each frequency. The regularisation value was chosen for each head and loudspeaker arrangement as described below. In all cases the regularisation parameter was chosen to be 0.001 to remove ill conditioning due to symmetric acoustic paths.

9.4 TWO CLOSELY SPACED LOUDSPEAKERS: THE STEREO DIPOLE SYSTEM

A closely two spaced loudspeaker arrangement, the Stereo Dipole, was developed at the ISVR (Nelson *et al*, 1997, Kirkeby *et al*, 1998) to produce virtual acoustic images with a larger 'sweet spot'. Not only do the filtering of the signals result in quicker fade of the recursive nature of the cross-talk cancellation (see Kirkeby *et al*, 1998), and a higher frequency of the 'ringing' of the filters, but also the resulting sound field is assumed to have a smoother form with minimal interaction between the waves propagating from the two loudspeakers.

In the current investigation, we visualise the total sound pressure in the frequency domain on the surface of the head, around it, and also the time domain animated sound field with a few snapshots, by using a similar technique to that used in the previous section.

Figure 9-6 shows the sound pressure variation on the head. Our goal is to control locally the pressure at the entrance to the ear canal only. Our cross-talk cancellation matrix is designed to produce zero pressure at the right ear of KEMAR and pressure of unity in its left ear. It is

seen that the proximity of the speaker divides the pressure zones to the right and left sides of the head. As frequency increases the equalisation zone on the head is reduced, and only accurate attenuation is obtained at the entrance to the ear canal and within the concha. The scale is linear where red is limited to unity and blue to zero pressure. Note that this simulation is ideal, and it does not highlight the problem of loudspeaker behaviour due to the ill-conditioning at low frequencies.

Figure 9-7 shows the resulting sound field around the head in the frequency domain for the same filtered signals as used in the previous figure. The characteristic common to all frequencies is the symmetry with respect to the centre axis between the sources, i.e. the separation between the right and left channels is preserved even if the head is not positioned in the exact 'sweet-spot' but along this axis.

Figure 9-8 presents 12 snapshots of the animated sound field around the head with the Stereo Dipole loudspeaker arrangement. It is seen that the loudspeakers produce waves which are almost out of phase and the interaction between them is minimal. Note that the Hanning window attenuates the high frequencies in the signals and therefore pinnae resonances do not affect these results.

9.5 FOUR-CHANNEL CROSS-TALK CANCELLATION

The signal processing is only briefly introduced here, and more examples of using this scheme with objective and subjective studies are given in Kahana (1997a, b, c, and 1999b).

9.5.1 A "four-ear" listener system

In this system we do not attempt to reconstruct the sound pressure at the eardrum exactly. It has been shown that the main advantage of such an arrangement is that the rate of change of the interaural cross-correlation (IACC) function with respect to head rotations in a four channel system, with loudspeaker also placed behind the listener, can be reproduced. Our

goal in this case is to reconstruct the accurate pair of HRTFs at the blocked ear canal when the head is rotated at angles of $\pm 5^0$. In Figure 9-9 we demonstrate the operation of the cross talk cancellation by reproducing "1" at ear no. 1 and "0" at all the others (this is achieved by multiplying the vector $[1\ 0\ 0\ 0]^T$ with the cross-talk cancellation matrix **H**). We can notice the effect of the cross talk cancellation on the robustness of head rotations: the zero pressure in the right side is maintained when the head is rotated, and on the left side a rapid change between "0" and "1" is achieved. Two conclusions can be reached from this example: the use of four ears for calculation of filters for a single listener proves very robust to head rotations and can provide important cues for front and back discrimination. Secondly, the 'sweet spot' is reduced in comparison to the 2×2 system presented in the previous example. If the head is moved forward or backward the cross-talk cancellation effect disappears.

In Figure 9-10 the sound field in the vicinity of KEMAR is presented with the same filtered signals that were used to calculate the desired signals and the response on the surface of KEMAR in Figure 9-9. The plots show the amplitude of the resultant pressure at discrete frequencies. Since the desired signals include two points to each ear: one is moved slightly towards the frontal hemisphere, and the other towards the rear, the sound pressure varies rapidly in the left ear of the dummy-head when it is rotated slightly, whereas the sound pressure in the vicinity of its right ear is maintained under the conditions of the cross-talk cancellation with zero pressure. The complexity of the sound field and the effects of the interference between the waves are clearly shown at high frequencies.

9.5.2 Virtual acoustic images for two listeners

The analysis of this virtual acoustic imaging system for two listeners has been investigated in Kahana (1998, 1999b). In Figure 9-11 we present the sound field (with the dimensions of

1 m $\times$ 0.5 m) around two head models of KEMAR. The desired signals in this case were $\mathbf{d}=[1\ 0\ 1\ 0]^T$ so that each head can 'listen' to filtered binaural recordings (the left channel of an artificial head recording is calculated to be heard at ears 1 and 3, with the right channel at ears 2 and 4). It can be noticed that the 'sweet spot' which appears now is very small (a few centimetres in an optimal calculation at 1 kHz, which would be even smaller at higher frequencies), and any movement can put the listeners out of the 'sweet spot' zone. It can also be noted that even with a symmetric arrangement of loudspeakers and listeners, the 'sweet spot' for one listener will always be narrower and also the sound quality will be better for one listener than for the other. The calculation was undertaken using the property of symmetry with the IBEM, as described in Section 4.1.2.3 since both loudspeakers and sources are arranged symmetrically for the left and right sides.

9.6 CONCLUSIONS

A few examples of the visualisation of the total pressure on the head and around it have been presented. At low frequencies (up to 1 kHz) the sound field is not very different for a head model of KEMAR and a rigid sphere. As frequency increases, the sound field in free-field becomes more complex and can be simulated based on the modelled individualised HRTF. Time domain representations have been presented for scattered sound field due to a monopole, and for the case of the Stereo Dipole system. In this case, the impulse response included mainly frequencies below the first resonance of the concha, but in principle high frequencies can be visualised albeit with larger models and consequently longer calculation times.

The visualisation of the sound field for multi-channel systems reveals the resultant sound field due to the interaction between the waves and the heads(s). These systems create very complex sound fields, and their performance can be of good quality only if low frequencies

are present. This information can be used in the design of an interactive virtual system, when the filters are updated with the movement of the listener.

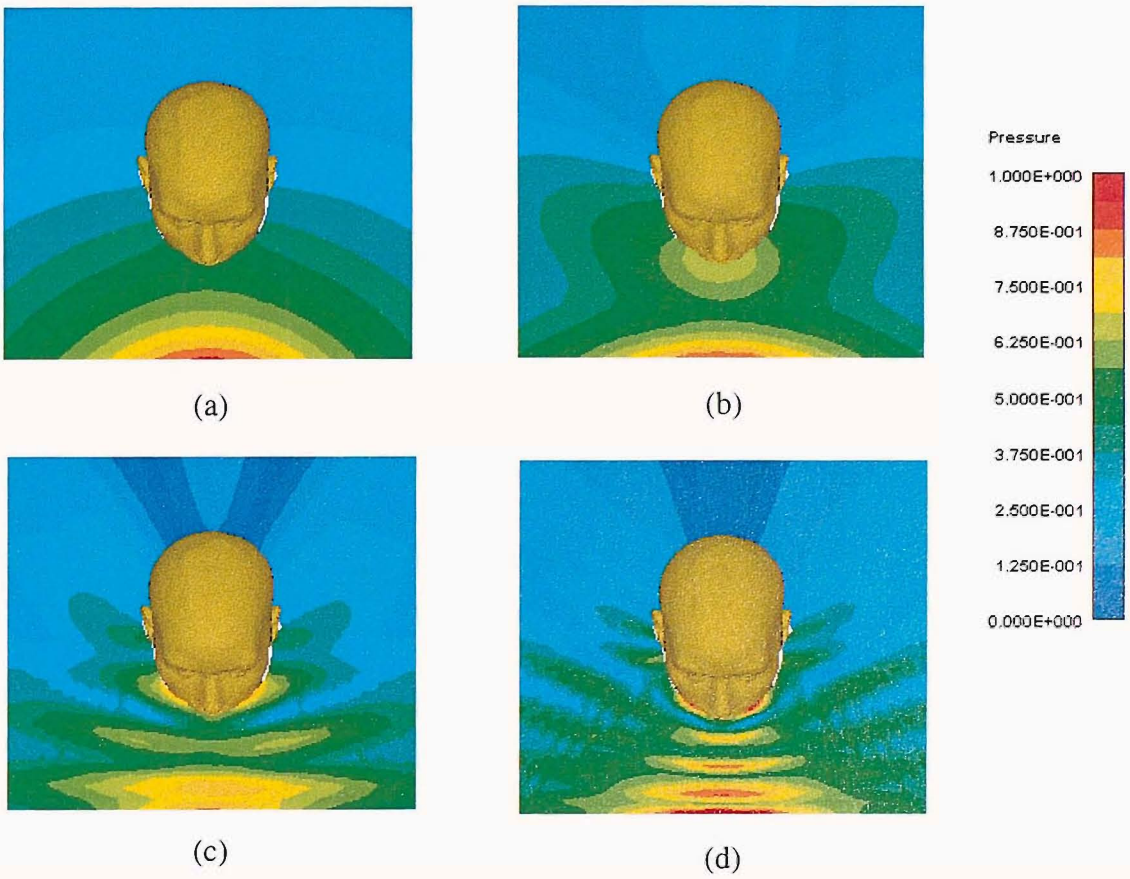


Figure 9-1: Scattered sound field due to a monopole around KEMAR at (a) 200 Hz (b) 1 kHz (c) 3 kHz (d) 5 kHz. The mesh includes 10283 nodes and 20562 elements. The mesh of the field points consists of 40000 nodes.

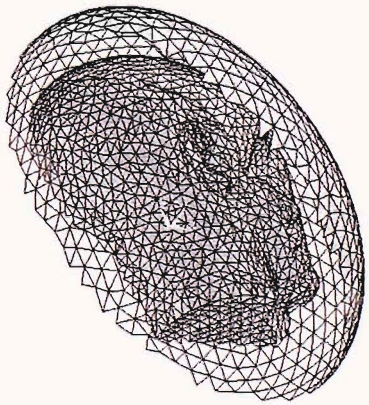
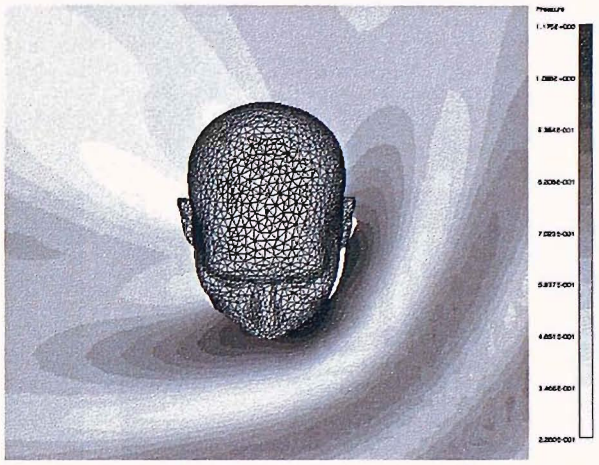
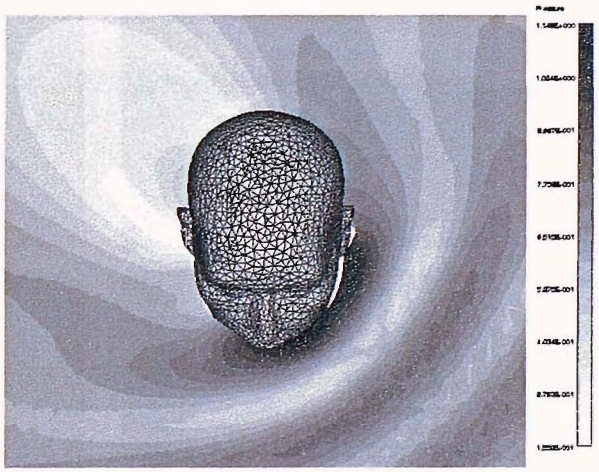


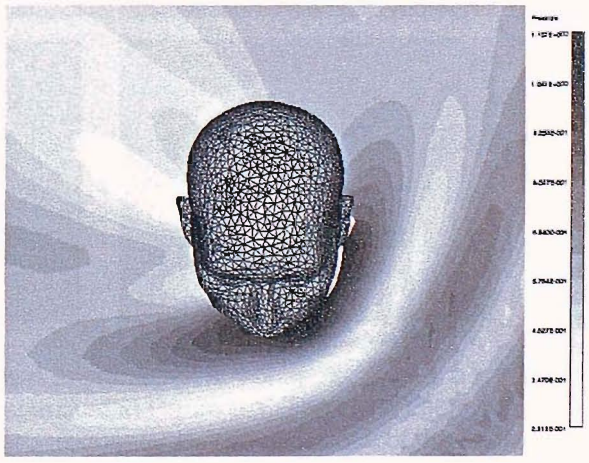
Figure 9-2: An IFEM mesh of KEMAR. The figure shows a cross-section of the original mesh (10800 elements), and the outside layer with 2400 infinite elements. The 41100 tetra elements which connect the two layers are not shown. The infinite layer was optimised with minimal ellipsoidal dimensions.



(a)



(b)



(c)

Figure 9-3: The amplitude of the total sound pressure around KEMAR at 2 kHz. The source is positioned at 1;1;1 ($\phi = 45°, \theta = 45°$) (a) BEM (b) IFEM - order 1 (c) IFEM - order 4.

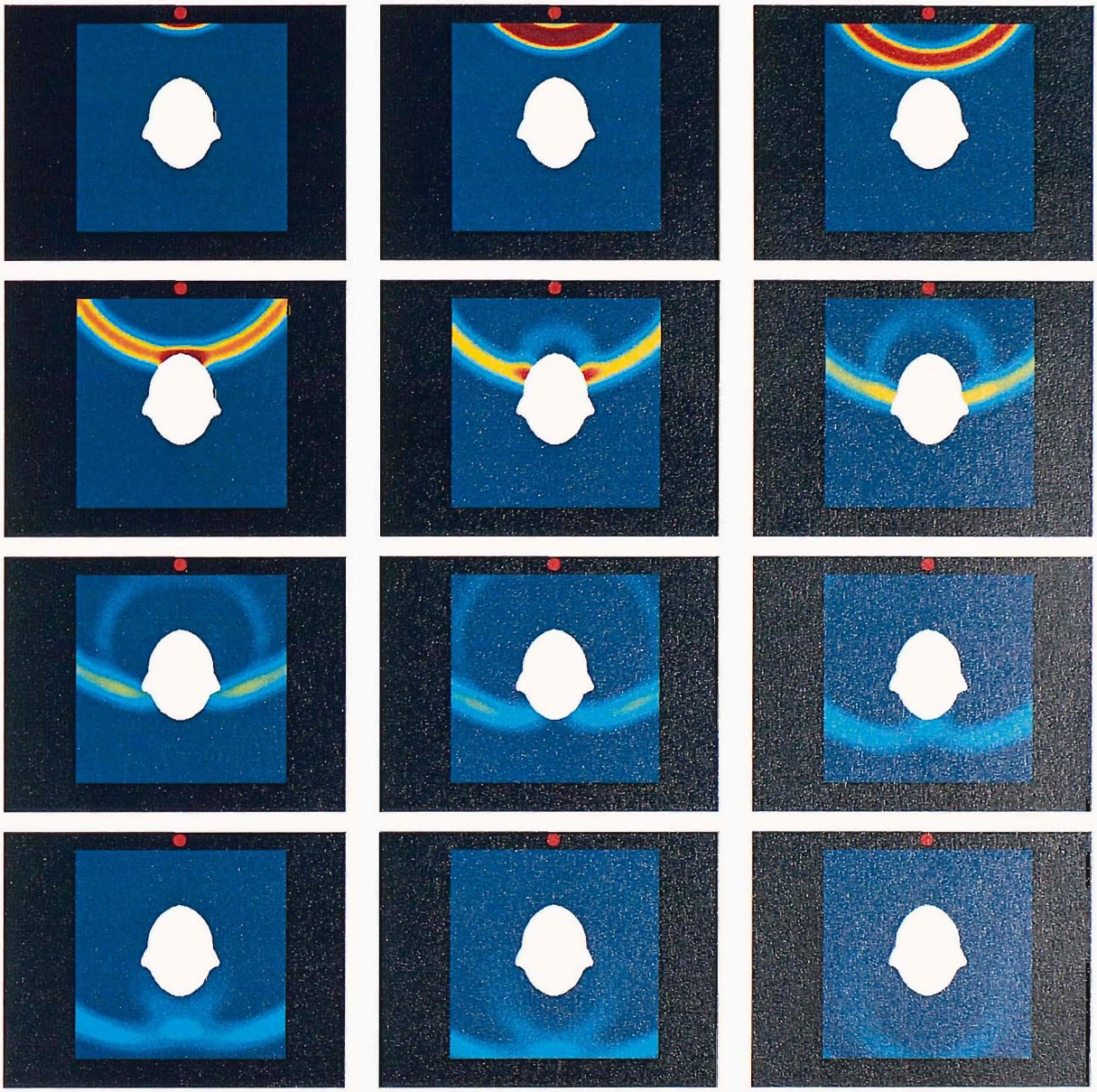


Figure 9-4: Time domain animation of a monopole wave scattered by KEMAR. The mesh model and field points are as specified in Figure 9-1. The calculation was undertaken using the DBEM up to 6.4 kHz. The modelled frequency response of each of the field points was multiplied with a half Hanning window before operating the Inverse Fourier Transform (IFFT). The sequence is presented with snapshots arranged from to upper left to bottom right. The head is not shown for clarity of the visualisation. The source is positioned 0.5 m away from the centre of the head to emphasise the secondary wave propagation in the shadow zone.

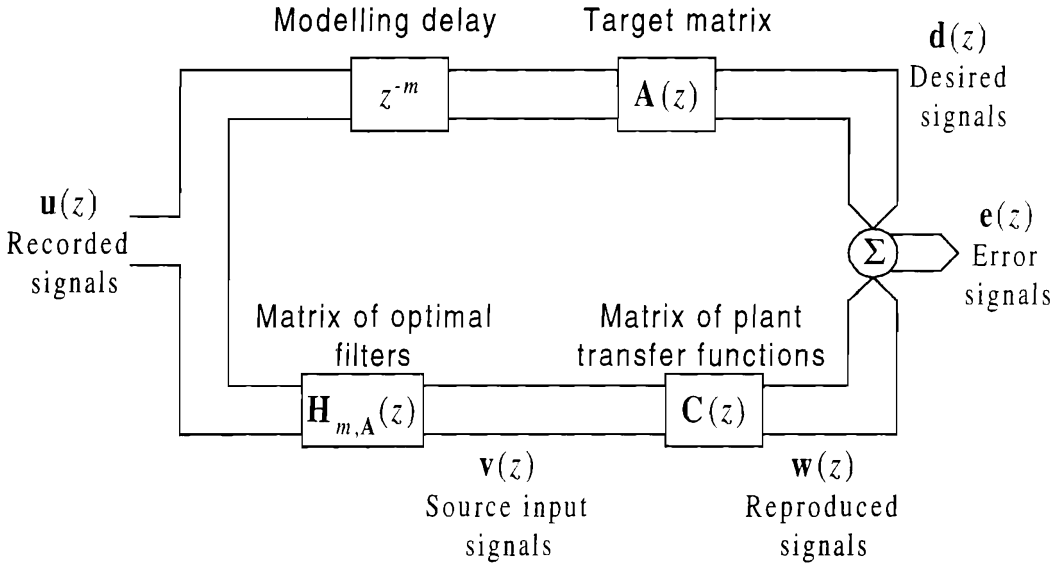


Figure 9-5: A block diagram of the multi-channel sound reproduction system.

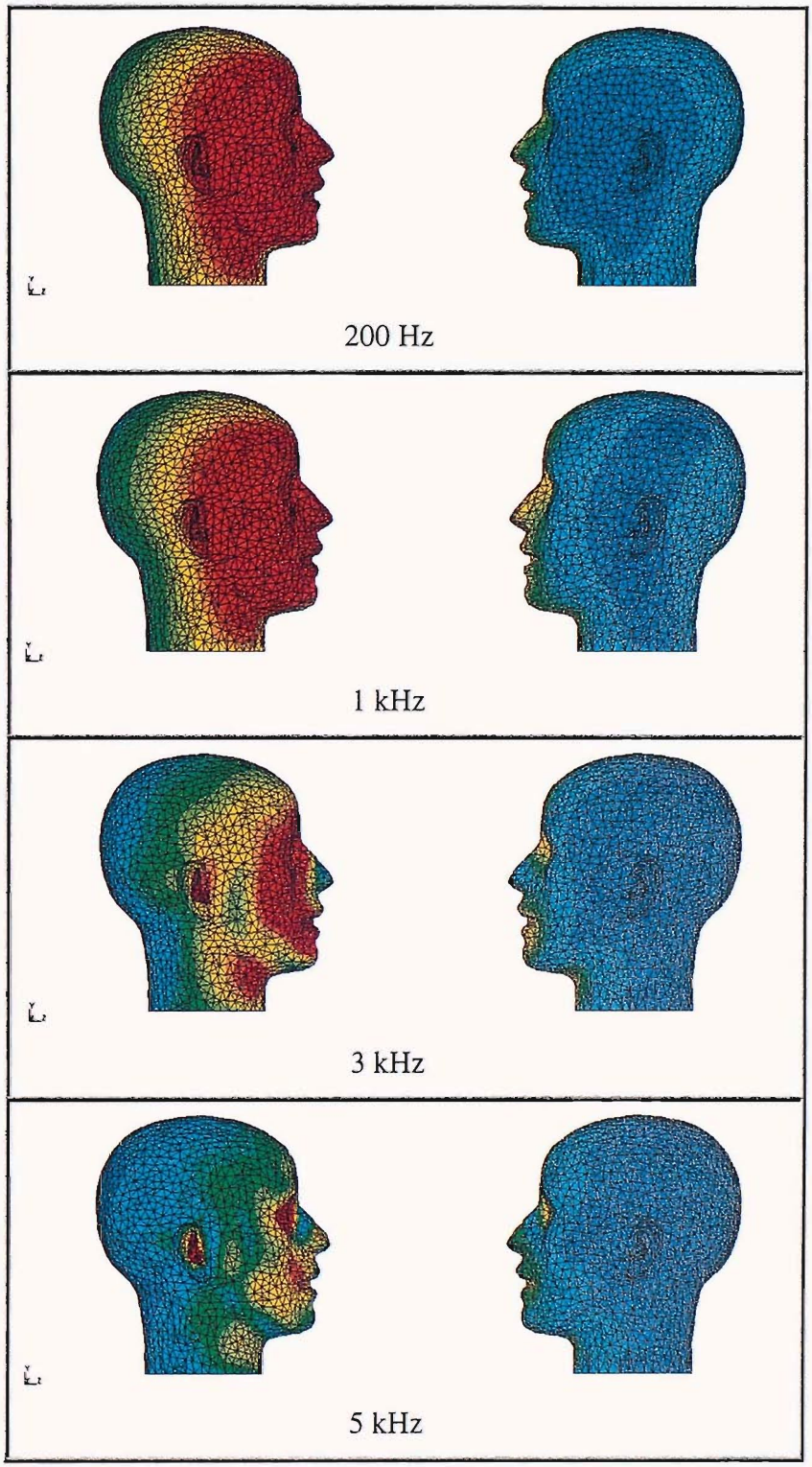


Figure 9-6: Linear sound pressure values on the surface of KEMAR artificial head with the Stereo Dipole system. The desired signals are '1' at the right ear of KEMAR (red) and '0' at his left ear (blue). Due to the proximity of the sources to each other efficient cross-talk cancellation is obtained for the two sides of the head.

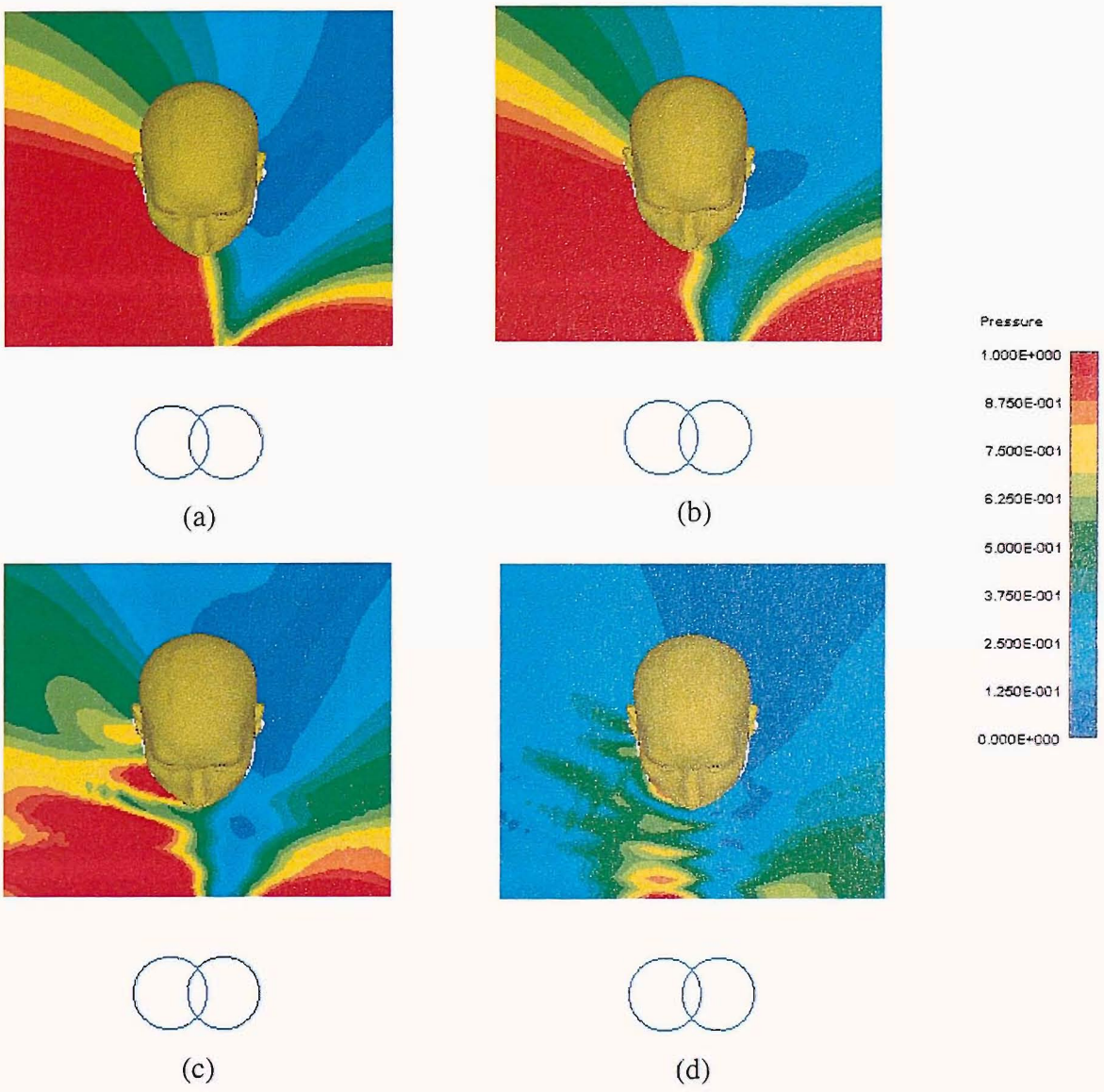


Figure 9-7: The total sound pressure of the Stereo Dipole system on a linear scale in the vicinity of the head (0.5 m × 0.5 m), presented in discrete frequencies. The source inputs were filtered to produce the desired signals as described in Figure 9-6. The figures correspond to the following frequencies (a) 200 Hz (b) 1 kHz (c) 3 kHz (d) 5 kHz.

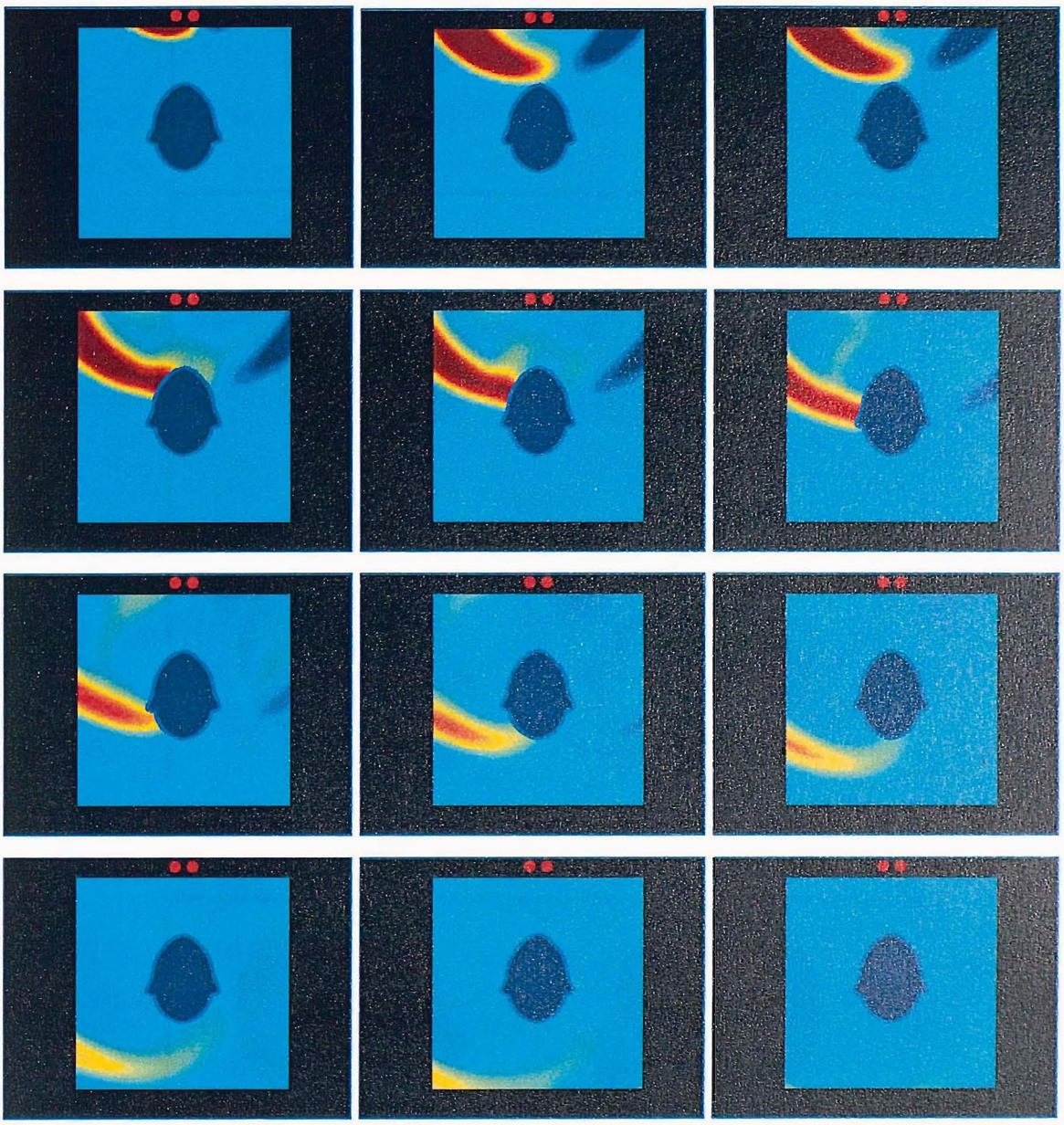


Figure 9-8: Time domain sound field of the Stereo Dipole system. The order of the snapshots is similar to those given in Figure 9-4. The procedure for the conversion of the frequency response of the field points presented in Figure 9-7 is similar, and described in Figure 9-4.

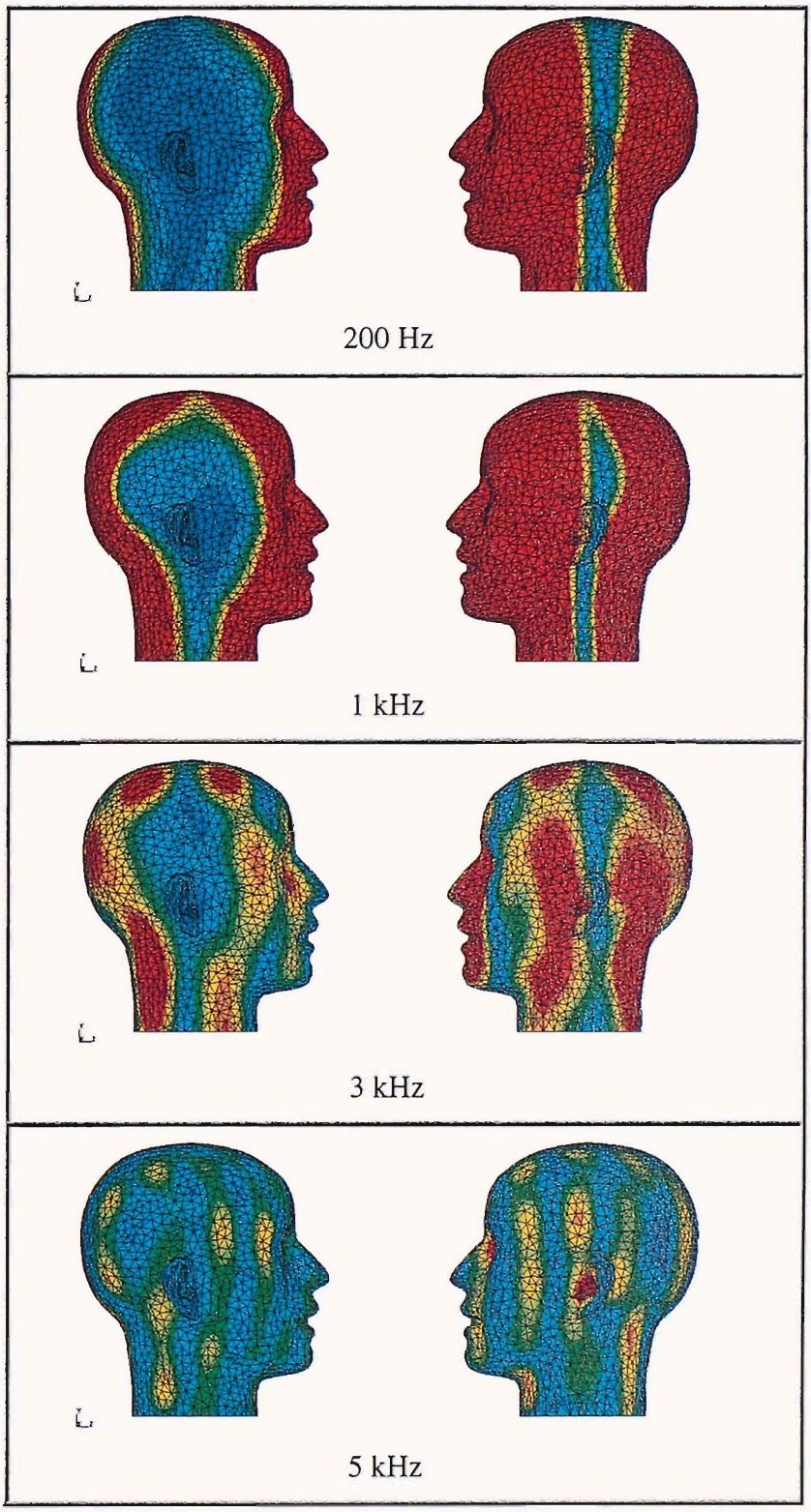


Figure 9-9: Linear sound pressure values on the surface of KEMAR artificial head with the 4×4 system. Four loudspeakers are used with their inputs are filtered to produce the following desired signals: '1' at the left ear of KEMAR and '0' at his right ear when the head is rotated 5° to the right. When the head is rotated 5° to the left cancellation is obtained at both ears.

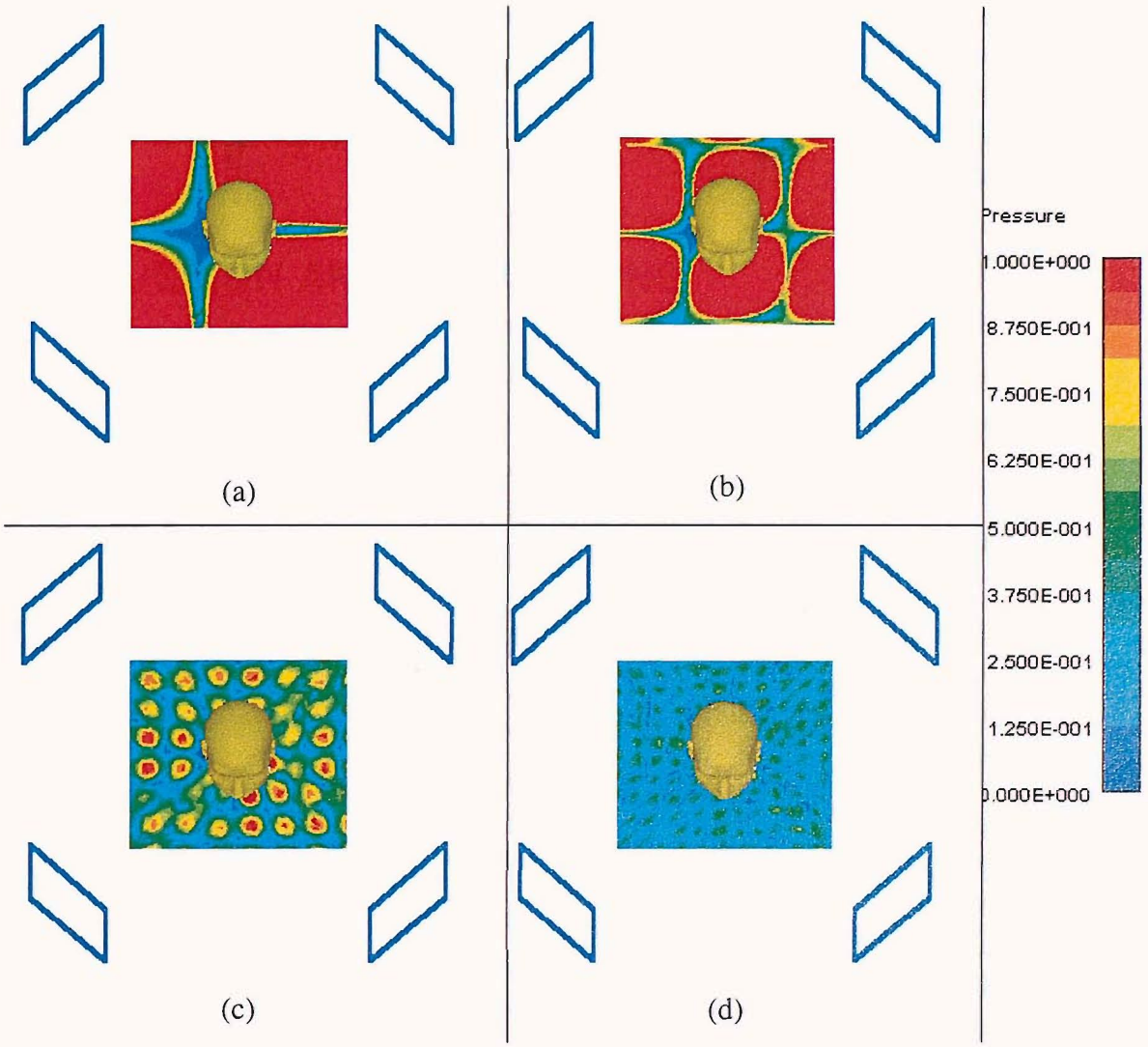


Figure 9-10: 'The sound field around KEMAR for the 4×4 system. The source inputs are filtered to produce the desired signals that are described in Figure 9-9. The figures correspond to the following frequencies (a) 200 Hz (b) 1 kHz (c) 3 kHz (d) 5 kHz.

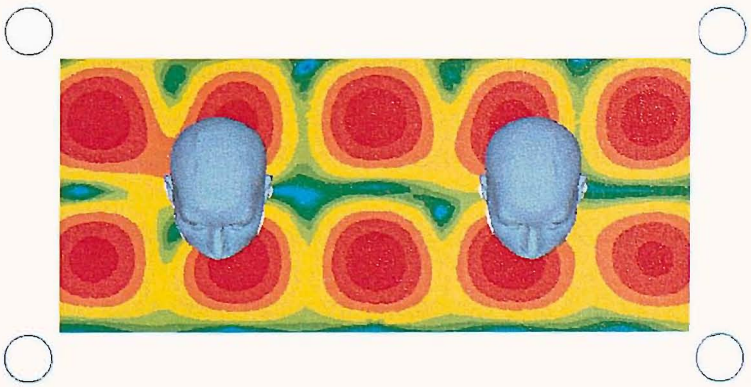


Figure 9-11: A 4x4 multicahnnel virtual acoustic imaging system for two listeners using cross-talk cancellation. The desired signals are **d**=[1 0 1 0] (the left channel of a dummy head recording is calculated to be heard at ears 1 and 3, with the right channel at ears 2 and 4, with an ideal cross-talk cancellation). The total sound field is calculated at 1 kHz in a grid with dimensions of 1×0.5 m^2 around two head models of KEMAR.

CHAPTER 10

DISCUSSION AND CONCUSIONS

10.1 SUMMARY

In this thesis, the feasibility of modelling the HRTF with numerical techniques has been investigated. It has been recognised for a while that in principle, the solution of the wave equation could provide the ultimate solution when modelling the frequency response of the external ear, since this method relies on the *exact* solution for arbitrary, complex bodies. However, only very recently has the combination of improved integral equation techniques, advanced computing hardware, advanced visual capturing devices, and an understanding of the significance of HRTFs in psychoacoustic studies enabled us to carry out such research.

The solution of exterior problems in acoustics can be derived using two approaches: the BEM, and recently also the IFEM. Although the latter approach proved to be much more efficient for large problems than the BEM (Burnett, 1994), the problem of mesh generation is still intricate for complex geometries. Therefore, our study has been focused mainly on the use of BEM techniques: the DBEM, the IBEM and their special formulations.

HRTFs include the transformation of sound due to the contribution of the different parts of the human body: the head, torso and shoulders, pinna, ear canal and eardrum. The contributions of these have been analysed. Although it is still not possible to include all of the above in a single BEM model, the investigation has included the analysis of the contribution of separate models.

Accurate simulation and measurement of the frequency response at high frequencies (above 5 kHz) is a difficult task. Therefore, an emphasis was put on minimising errors as much as possible for the comparison between the physical world (measurements apparatus, head and pinnae), and the virtual world (computer models of the same head and pinnae mentioned above, geometric arrangement, etc). The computer-controlled rotating arc was designed to help measure the response of baffled pinnae, as well as HRTFs in the anechoic chamber, with a high resolution of spatial sampling. Consequently, it was possible to track the variations of the frequency and time responses in various planes with both measurement and simulation procedures.

The spatial mode shapes of the human pinna have been investigated. The results have been compared with the literature, and the inter-variability among six pinnae has been studied. The formulation for obtaining numerically the spatial basis functions of the pinna, based on the SVD technique, extended the interpretation of the 'mode shapes' obtained by Shaw (1974) with the excitation of the pinna at grazing incidence. These spatial basis functions can now be used to extract a 'reduced order' frequency response of the pinna, by selecting only the first few singular values, which contribute the most to the total sound pressure.

A few examples of sound fields for virtual acoustic imaging systems have been presented. These demonstrate the possibilities of using this tool to predict the sound field in a free-field environment, with individualised HRTFs and various loudspeaker arrangements.

The main results and conclusions of the research presented in this thesis are presented below.

10.2 DISCUSSION AND CONCLUSIONS

10.2.1 HRTFs of simple geometrical models

The simulation of the response of simple geometrical models has served two purposes in this study. Firstly, to investigate if numerical techniques are accurate when compared to analytical solutions and to examine what attention the numerical implementation requires in order to ensure that the results are not degraded. Secondly, to investigate numerically models which cannot be modelled analytically and can be used in the study of the basic features of the HRTF.

Good agreement was achieved when the sphere was modelled and compared with analytical solutions. Removing 'irregular frequencies' of the sphere is not a trivial task, probably due to its symmetry. The effectiveness of treating these singularities is different in the DBEM and the IBEM, and there are no exact rules governing the methods to correct these errors without a trace, especially at high frequencies. In the DBEM, over-determination points should be added but kept to a minimum, especially if the number of nodes in the model is not large. They should be added only after the problem was run first, and then the problem should be run again only in the 'corrupted' frequency range. In the IBEM, either absorbent 'singular elements' should be added in the cavity, or the inner side of the elements should be assigned with positive admittance values. This does not, however, guarantee to remove the severe peaks and notches completely.

Four different methods were used to investigate the sound field around the sphere (the DBEM, the IBEM, the axisymmetric IBEM, and the IFEM). Although different parameters could control the accuracy of the modelling, overall, with approximately the same accuracy, the high efficiency of the IFEM and the axisymmetric IBEM compared to 3-D

BEM was demonstrated. All of these methods can be used very efficiently to model the response of any spheroid, just by scaling the model with different factors along its axes.

The response of an ellipsoid with dimensions matching approximately the shape of an artificial head was investigated using the principle of reciprocity. The monaural and interaural responses were solved up to 5 kHz with a high spatial resolution. The ITD was extracted using the differences in the interaural group delays.

The implementation of the response of a large volume mesh such as the torso and shoulders imposes great difficulties due to many 'irregular frequencies'. Dividing the large model into half, using the symmetry property, thus reducing the volume, results in a reduction of the number of severe peaks and notches that need to be removed. The response of the torso is direction dependent. The modulation of the amplitude, when compared to a sphere or an ellipsoid, is a result of the reflection that causes an echo in the time domain, with varying delays as a function of the angle of the source.

10.2.2 Mesh resolution

The number of elements per acoustic wavelength is a crucial factor in the preparation of the mesh, and the reliability and predictability of the results. The general guidelines are that the mesh should not include edges in any element that are larger than 1/6th of the wavelength. It was found in our cases of scattering problems, that when the pressure on the body is required in a position that is on the same side of the source (e.g. the ipsilateral ear) four elements per wavelength could be used as a the minimum mesh resolution. Although the results with such a low requirement do not 'break up' with unpredicted severe peaks or notches, the accuracy is slightly reduced. Six elements per wavelength are sufficient and a higher resolution is not required. However, for points in the shadow zone (e.g. the contralateral ear) a higher resolution is required. In this case, six elements per wavelength

are essentially the minimum, and the accuracy can increase with a higher resolution. The exact resolution could not be analysed for the complex shape of the pinna with the large model of the head since the response on the contralateral side is very sensitive to the geometry of the pinna, and in order to increase the mesh resolution only very large models should be used to eliminate geometry changes. This was beyond our computing capabilities.

10.2.3 Mesh accuracy

In this study, an emphasis was put on the minimal simplification possible in every part of the modelling. Is it crucial to use a high accuracy laser scanner? How accurately can the acoustic response be modelled, when the various parts of the pinna, and especially the concha are roughly approximated? The response of the DB60 pinna was modelled (under baffled conditions, to eliminate irrelevant contributions due to diffraction around the head) with two mesh models obtained with different type of scanners. The acoustical response was significantly degraded with a 'low-resolution scanner', and the resonance peaks and notches lost their exact structure with a shift to higher frequencies due to shallower volume of the concha and lack of details in the pinna. It is concluded that the use of simpler approaches to obtain a mesh (such as extraction of 3-D information from 2-D still pictures) must first be compared with the performance achieved with the 'high-resolution' scanner.

The accurate mesh models of the KEMAR and YK heads were obtained by integrating and manipulating the 3-D models, obtained with both types of laser scanners. To the best of the author's knowledge a single visual capturing device which can produce such high accuracy does not exist.

10.2.4 Acoustical characteristics of the external ear

The pinna has a complex shape, and therefore high directivity and variation of the frequency response at high frequencies. The response of the external ear was calculated when attached to an infinite baffle and to the head. The similarities in the resonance frequencies between the two cases enable great improvement in calculation efficiency, the ignoring of the variability between head shapes and dimensions, and the orientation of the pinnae. Many of the spectral changes reported in the literature have been observed, and the variability between the ears has been shown to be high at frequencies above 8-10 kHz.

The HRTFs of YK and KEMAR heads (both without torso) were simulated. The results of the latter compared with measurements with high precision. Since the ipsilateral ear produces higher accuracy with only four elements per wavelength than the contralateral ear with at least six elements per wavelength, ILD values in the ITF can be used reliably only up to 10 kHz.

10.2.5 Spatial mode shapes of the pinna

The modes of the human pinna have been simulated using two approaches: by simulating the procedure used with the classical measurements carried out by Shaw, and also by using the SVD formulation relating the Green function matrix of field and source points. With the first case, we observed the following: Shaw identified six resonance frequencies but, by looking at his ten pinnae responses, it is clear that some frequencies have a very weak excitation. We repeated his approach of finding the resonance frequencies that are excited the most. Then by observing the variation of pressure amplitude and phase, we found similarities but also deviations from his results. First the DB60 was investigated. In this pinna in both simulation and measurements the 12 kHz mode reported by Shaw (and mentioned also by Middlebrooks *et al*, 1989) could not be observed, and only five

resonance frequencies were found. When the larger pinna, the DB65 was investigated, all six modes were found with similar centre frequencies to the averages given by Shaw. Exact variations of amplitude and phase on the surface of the pinnae were given. Variations appeared in the amplitude levels in the concha, fossa of helix and antihelix when compared with the results of Shaw, but he presented only average data. In addition, slightly different conditions were used in our case (such as far field excitation and using an ideal plane wave source). Nevertheless the patterns have a very similar structure.

These pressure variations reveal the transverse modes of the pinna at high frequencies, but cannot be used directly in the construction of the frequency response at the entrance to the ear canal. Since the BEM can provide the accurate pressure values at any points on an arbitrary body, we hypothesised that by using the SVD method we can observe certain common characteristics that appear at the resonance frequencies. The mathematical formulation showed that the left and right singular vectors of the SVD yield two sets of basis functions which respectively represent a series of mutually orthogonal complex 'field mode shapes' and mutually orthogonal complex 'source mode shapes'. The connection shown with the spherical harmonics demonstrates that under certain specific conditions, such as using only uniformly sampled spheres, the frequency response and the Green function matrix can be composed of 'spatial patterns' which are frequency independent, and a frequency dependent operator which is spatially independent.

When the formulation was implemented numerically on the pinnae, spatial patterns of both 'source mode shapes' and 'field mode shapes' revealed the resonance frequencies that take into account excitation from anywhere on the hemisphere. Although clear 'monopole-, dipole-, and quadrupole-like' modes appear on the surface of the pinna, and similar patterns in the source points, larger variations appear among pinnae for the latter. It is not clear if

the relatively low sampling of the space contributes to the variation, especially at high frequencies, or the high variation is derived purely from physical reasons.

The frequency response at the entrance to the ear canal reconstructed with only a few terms show promising results for the generation of a reduced order technique. The performance is at its best in the reconstruction of peaks rather than notches, since the first singular value of the resonance frequency contributes the most. At minima, many low order singular values make a contribution of the same order.

10.2.6 HRTF measurement

The measurement of the response of baffled pinnae and HRTFs has been carried out in the anechoic chamber using a specially designed and built apparatus. Its main objective was measuring the response of baffled pinnae. Since it is possible to control the position of the arc in a resolution that is much smaller than 1°, it was possible to equalise each measurement with the equivalent measurement position, when the head or the pinna was absent. The measured dynamic range of KEMAR was more than 50 dB over the entire sphere, and 40 dB for the baffled pinnae.

10.2.7 Sound fields for virtual acoustic imaging systems

Most psychoacoustical studies are carried out with headphones. This is because there is a greater control of the acoustic environment. Recently, the addition of head trackers into these studies helped in the study of front-back reversals as well as externalisation. These studies are not used with loudspeakers, mainly since the wavelength becomes so short, that it is impossible to control high frequencies (at 10 kHz, the wavelength is in the order of 3 cm). If a head tracker is used in virtual loudspeaker reproduction, it might be beneficial to study the characteristics of the sound field for a specific loudspeaker arrangement.

A few examples have been presented and have demonstrated the simple sound field achieved with two closely spaced loudspeakers in both frequency and time domains. Multi-channel cross-talk cancellation suffers from a narrow sweet spot, mainly because of the sensitivity of the sound field at high frequencies, as well as discrepancies in individualised HRTFs, the imperfection of the transducers etc.

10.2.8 Computational cost

The main hurdle of modelling HRTFs with the BEM is the very high computational cost. Although it is anticipated that the use of the IFEM will alleviate this limitation in the future, currently due to the difficulties in mesh generation we are confined to using the BEM .

At high frequencies, the size of the elements comprising the mesh model needs to be refined. As the dimensions of the model become larger and larger, and the requirements for accurate modelling at high frequency remains fixed, the efficiency of the BEM is dramatically reduced. This is the main reason why the BEM is mainly associated with the 'low' frequency range.

The main difficulties in using large models are summarised as follows:

- Processing speed. Currently the trend of CPU performance of the PC is said to double every 18 months. HRTFs can already be modelled on home PCs. The improvement of parallel computers and their interface made it possible to multiply performance depending on the number of processors available. More cost-effective parallel PCs have just started to be used. We therefore predict that the improvement in computational speed will make simulating HRTFs more accessible in the near future.

- Physical memory size. The performance of the BEM is optimised for the in-core solver. A few Gb are required for each processor (if a parallel computer is used). As before, this limitation is already diminishing.

- Singularities. Inherently part of the formulation, singularities are very difficult to remove. This is especially prominent when the cavity is large and many resonance exist (such as in the case of a head with torso).

- Geometric accuracy. Rapid improvement in accuracy and cost effectiveness of accurate scanning devices together with advances in computer graphics, make it possible to obtain excellent geometric description of the models. It is predicted that even simpler capturing techniques such as digital cameras could be used in the future.

- Mesh manipulation. Currently this is a complex and iterative process. It might be possible to develop an automatic procedure to convert digital images or scanned models directly to the BEM or the IFEM models.

- Boundary conditions. Approximate boundary conditions of the eardrum impedance, hair and clothing could be implemented, based on published average data or empirical data. Individualised data of this kind cannot be achieved without acoustical measurements.

Our research started in 1997 with using SYSNOISE 5.2, on a 486-PC with 16 Mb of RAM, disk space of 200 Mb, and the Windows 3.1 operating system. This was used to analyse the response and sound fields scattered from a rigid sphere at low-medium frequencies only (ka=4). The final results have been obtained in 2000 using SYSNOISE 5.4, on a parallel SGI/Origin 2000 with 4.5 Gb of RAM and a hard disk of 100 Gb. With this combination of hardware and software, it has become possible to investigate our problem with high

precision at high frequencies. In addition, the modelling of baffled pinnae was undertaken on a Pentium II-400 PC with 256 Mb of RAM using SYSNOISE 5.4 and the Windows NT operating system.

The performance of the DBEM and the IBEM with various configurations of hardware and software is summarised and presented in Table 10-1 to 10-4 and in Figures 10.1 to 10.9. The bench test results summarise the CPU time and space required to solve our HRTF problems with a PC with Windows 95 and Windows NT operating systems (which produced large differences in performance), SGI Origin 2000 and HP exampler parallel computers.

Time estimates		Memory requirements	
BEM collocation (direct method)	BEM variational (indirect method)	BEM collocation uncoupled (direct method)	BEM variational uncoupled (indirect method)
Assembling time $= \frac{(\alpha+\beta.n)n^{3/2}}{m_a}$	Assembling time $= \frac{(\alpha)e^{3/2}}{m_a}$	In-core solver $M = 16((n+o)^2+(n+o))$	In-core solver $M = 16\left(\frac{n(n+1)}{2}+n\right)$
Solution time $= \frac{\gamma.n^3}{m_s}$	Solution time $= \frac{\beta.n^3}{m_s}$	Out-of-core solver $M = 16\sqrt{b}(n+o+1)$	Out-of-core solver $M = 16(2b+n)$
With $\alpha = 2.3\ 10^{-2}, \beta = 8.0\ 10^{-6}, \gamma = 2.2\ 10^{-6}$	With $\alpha = 5\ 10^{-2}, \beta = 1.15\ 10^{-6}$	With b block size, M memory in bytes, n number of nodes, o number of over-determination points	

Table 10-1 Theoretical predictions of time and memory requirements for solving the BEM with SYSNOISE. The values of α, β, γ were found in a bench test. The equations and numbers are given as approximations only.

Whenever SYSNOISE is run with the out-of-core solver, the performance is reduced significantly. This can be explained by the following: the solver finds the solution to the system of equations [**A**]{**x**} = {**b**} by decomposing [**A**] in a block-wise manner. The [**A**] matrix is split into blocks, and each block is decomposed (factored) separately. The size of these blocks may be user-defined. Varying block size will vary the performance (execution time) of the solver. Certain block size values will provide optimal performance. This

optimal block size may vary from machine to machine. Being an out-of-core solver, a temporary (scratch) file is written by the solver. This file can be quite large.

Tables 10-2 to 10-3 summarise the relations between the size of the mesh models (for both heads and pinnae), the memory required and the total CPU time. Since the calculation of the 'mode shapes' using the SVD is very computationally expensive, Table 10-4 summarises the CPU time for each pinna model.

No. of nodes	No. of elements	Max. edge length [mm]	Approx. Max. frequency [Hz] (6 e/w)	Memory (Gb)		CPU time- HP Exampler [min] SYSNOISE 5.3		CPU time- SGI Origin [min] SYSNOISE 5.4	
				DBEM	IBEM	DBEM	IBEM	DBEM	IBEM
202000	400000	3.1	18000	640	320	-	-	-	-
27000	50000	6.0	15000	12	6	-	-	-	-
14000	30000	7.2	10000	4	2	2800	4600	-	830
9900	20000	9.0	6400	1.8	0.9	540	900	-	281
5060	10000	11.5	5000	0.4	0.2	60	85	45	46
2600	5000	13.0	4500	0.12	0.06	20	40	18	10

Table 10-2: The relation between the size of the BEM head models and the maximum frequency, memory and CPU time requirements (the running time might be different for different platforms, and the maximum frequency assumes six elements per wavelength). All models are the left half of the KEMAR head and modelled using the symmetry property in the BEM, and the time presented is the overall time calculation. Fields are blank whenever it is not possible to solve the problem with the in-core solver.

pinna		No. of nodes	Number of elements	Max. edge length [mm]	Max. freq. [Hz]		CPU time-DBEM/SGI [min]
					6 e/w	4 e/w	
DB60	a	6887	13488	3.64	15560	23340	50
	b	2825	5421	3.75	15100	22700	5
DB65	a	5199	10216	3.67	15420	23100	24
	b	3389	6656	3.78	14980	22470	8
YK	a	6361	12523	2.96	19120	28600	41
	b	3392	6579	4.31	13160	19740	8
CORTEX	a	5923	11632	2.99	18900	28360	34
	b	3390	6658	3.82	14840	22250	8
B&K	a	5442	10898	3.42	16540	24810	28
	b	3906	7671	3.88	14590	21890	11
DB90		4799	9409	3.12	18130	27200	20
DB95		4728	9267	3.20	17680	26520	19

Table 10-3: The relation between the size of the BEM pinnae models and the maximum frequency, memory and CPU time requirements (the running time might be different for different platforms, and the maximum frequency assumes six elements per wavelength).

pinna	Max. freq. [Hz] 4 e/w	Green function matrix dimensions	No. of frequencies	CPU time-DBEM/SGI [hours]
DB60	22700	2825 × 209	91	21
DB65	22470	3389 × 209	91	43
YK	19740	3392 × 209	91	43
B&K	21890	3906 × 209	91	60
CORTEX	22250	3390 × 209	66	31

Table 10-4: The CPU time of the SVD calculation for optimised baffled pinnae.

10.3 FUTURE WORK

In the course of the study, various questions have been raised. As we tried to investigate various aspects of both the numerical techniques and the properties of HRTFs, a few questions have remained unanswered, and in the following section we summarise the possibilities for future research into this topic:

- **Modelling with the IFEM**

The advantages of the IFEM are very important for exterior modelling of large problems. The computational speed should be reduced significantly and the problems of singularities in the integrals do not exist. In addition, a time domain formulation can be used to obtain HRIR directly. Specially designed mesh generators such as PRE-SYSNOISE might be improved and developed to produce automatically an optimised IFEM mesh.

- **Reduced order HRTFs**

Any reduced order modelling technique can be investigated and compared to an originally modelled HRTF with the BEM, since the simulation results are accurate anywhere in space. The spatial patterns obtained with the SVD formulation require further investigation in the reconstruction of HRTF, both objectively and subjectively with various pinnae models.

- **Geometric manipulation**

By using a 'reverse engineering' approach, is it possible to work the problem backwards so that by stretching and altering the pinna shape, a close match to a desired set of frequency response curves could be achieved, i.e. to a 'good' listener (Wenzel *et al*, 1993, Morimoto and Ando, 1982) or to a 'super normal' listener (as described by Shinn-Cunningham in Carlile, 1996, pp. 220-222). Using existing 'sculpting' mesh tools, it is possible to control

the size, shape and orientation of specific parts of the pinna. Will it be possible to solve this 3-D problem using genetic algorithm approaches?

Currently, the manipulation process is difficult, interactive and tedious. The development of an automatic mesh processing toolbox is essential if an individualised HRTFs is required to be modelled.

- **Structural modelling of HRTF**

Generating individualised HRTFs could be obtained by hybrid approaches, such as matching the size, shape, orientation of an individualised pinna, based on visual recognition by a digital camera. The ITD could be matched by using a look-up table with matched size heads and ellipsoids, scaling the first resonance of the concha by calculating the volume of the concha, etc.

In the analysis of interpolation techniques, various strategies have been developed. The proposed numerical technique is an ideal tool, to develop further the objective and subjective effects of interpolation approaches, especially at high frequencies.

- **Individualised HpTF**

It is demonstrated in this research that Individualised HRTFs can be modelled. However, when listening to headphones, it is necessary to obtain the individualised Headphones Transfer Functions (HpTF). In this case, the baffled pinna approach can be used, and it will be required to include the cavity of the headphones with the appropriate boundary conditions, which can be supported by empirical data.

- **Digital filter design for virtual acoustic systems**

In the applications of head tracked systems using loudspeakers, it can be beneficial to predict the equalization zones in the vicinity of the listener. The sound pressure, phase,

intensity and other computed values such as IACC, can be taken into account in the design of the filters.

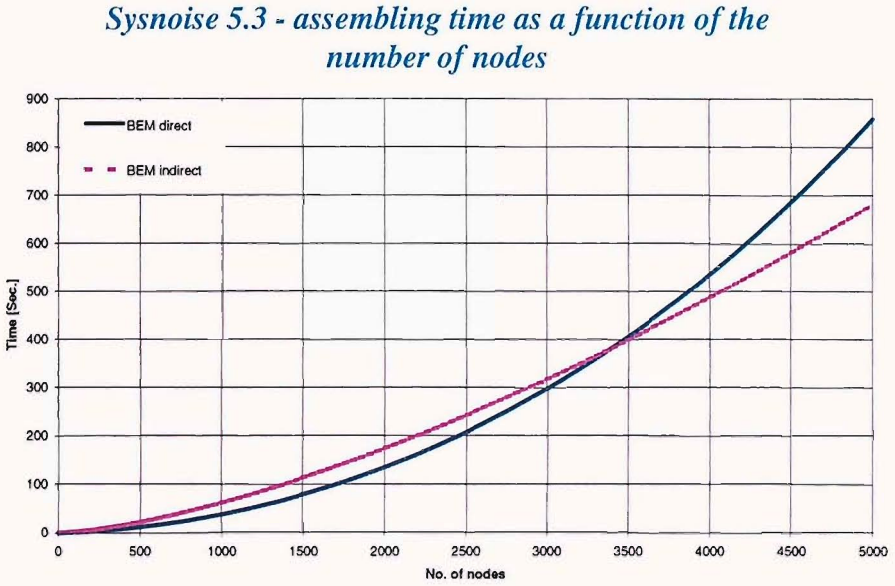


Figure 10-1: Assembling time of the matrices as a function of the number of nodes with SYSNOISE 5.3. When closed boundary surface meshes are used, the problem can be solved with either the IBEM or the DBEM. For 'small' problems (up to 3500 nodes) the DBEM is more efficient than the IBEM (in this research all baffled pinnae were modelled using the DBEM).

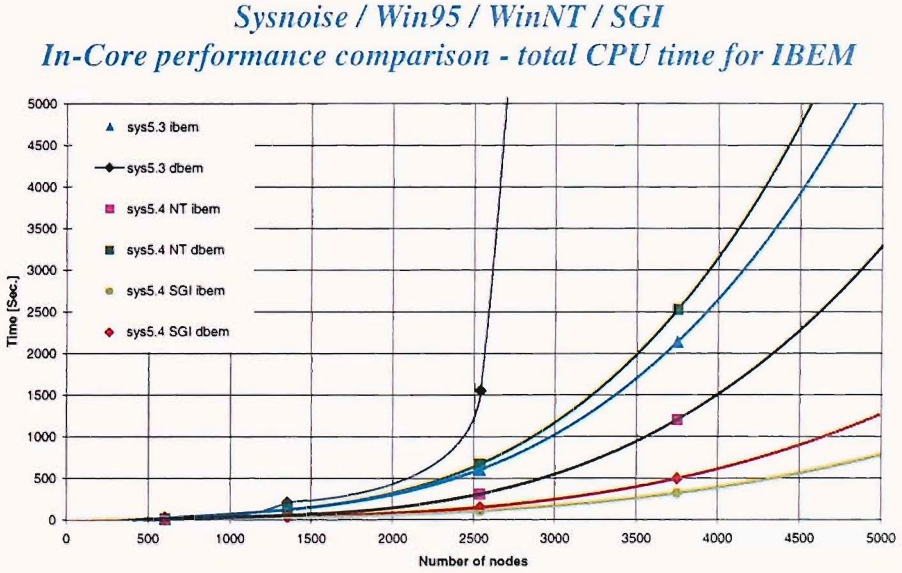


Figure 10-2: Total CPU time comparison with different platforms and versions of SYSNOISE.

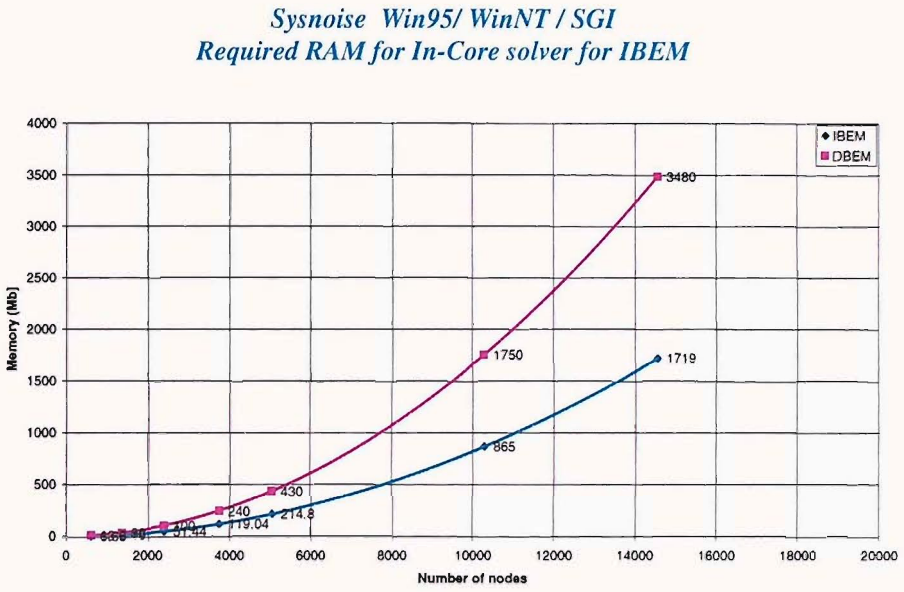


Figure 10-3: The required RAM for efficient modelling with the IBEM. The maximum frequency is limited by the physical RAM available.

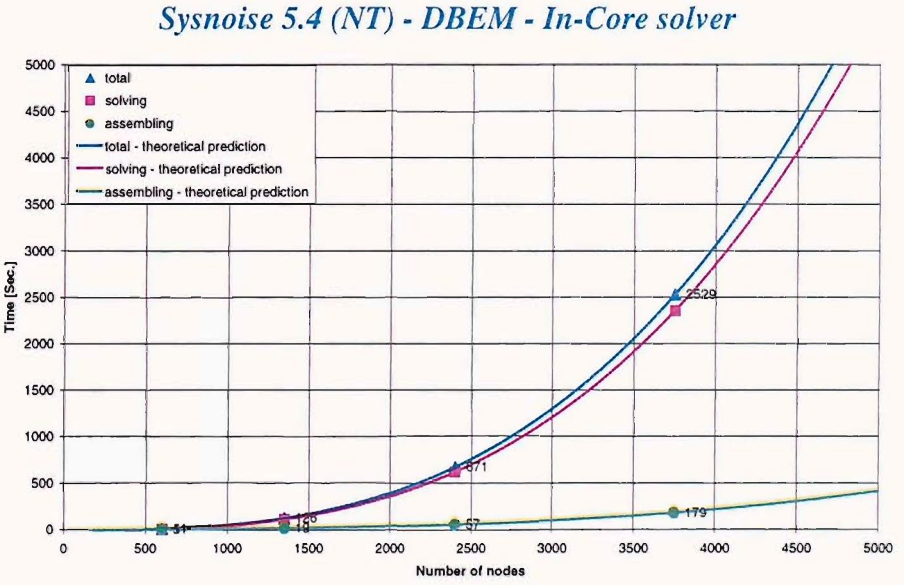


Figure 10-4: The CPU time for SYSNOISE 5.4 with Windows NT using the DBEM and the in-core solver. As the model gets larger, the total CPU time is dominated by the solving time.

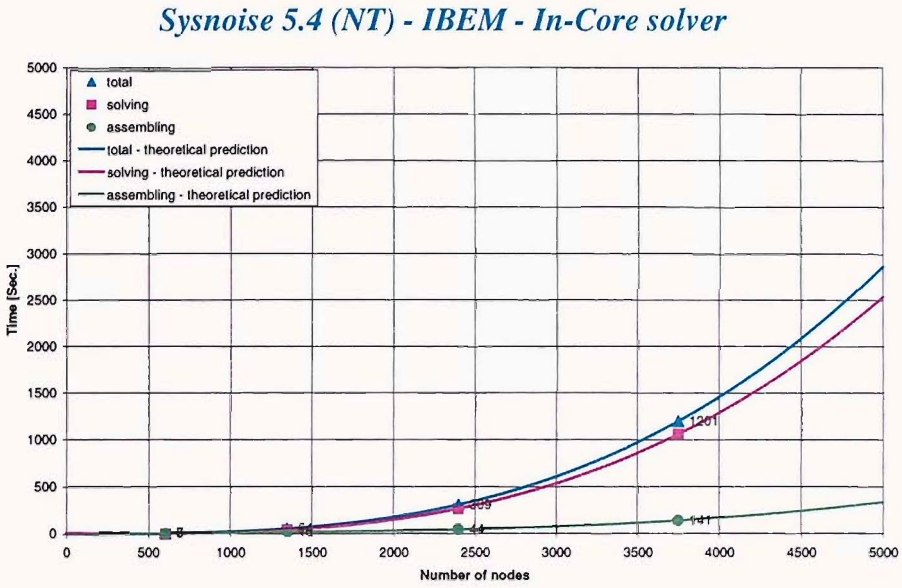


Figure 10-5: The CPU time for SYSNOISE 5.4 with Windows NT using the DBEM and the in-core solver. The relation between the number of nodes, the maximum frequency and the total CPU time can be found also in Tables 10-1 to 10-4.

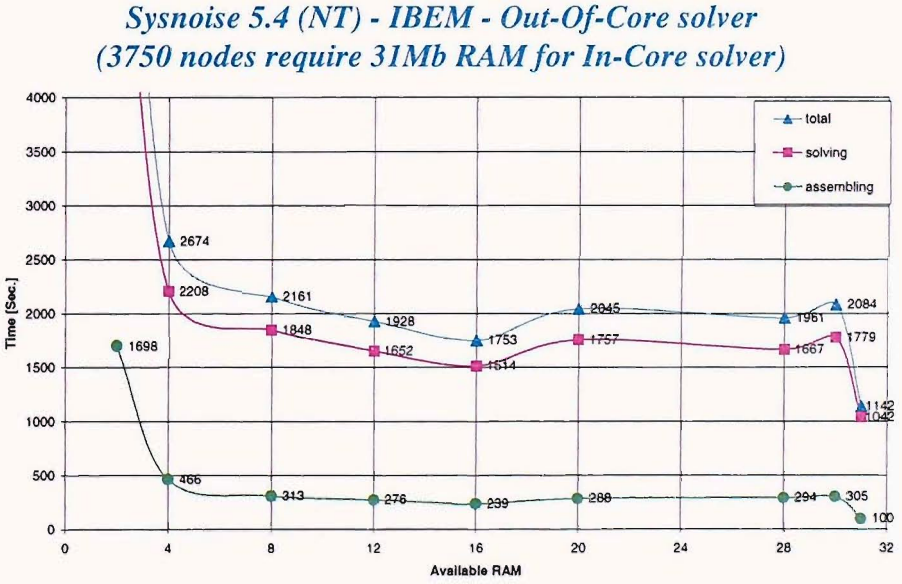


Figure 10-6: The total CPU time as a function of available RAM. Out-of-core solver with almost the same amount of RAM was found to be as twice as slow when compared with the in-core solver. It is recommended that large problems should be solved with the IBEM or the DBEM with only the in-core solver for reasonable calculation times.

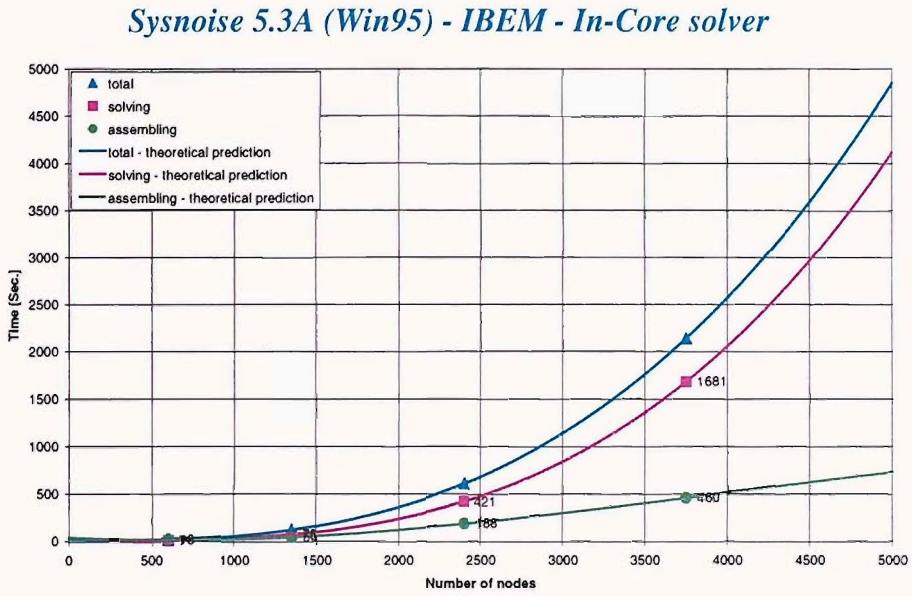


Figure 10-7: The CPU time for SYSNOISE 5.3 with Windows 95 using the IBEM with the in-core solver.

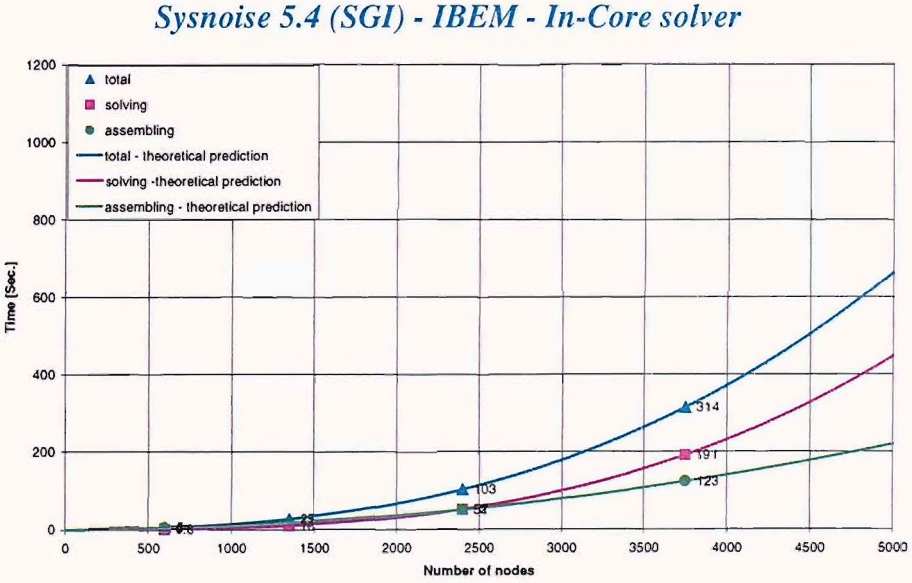


Figure 10-8: The CPU time for SYSNOISE 5.4 with SGI (Irix 6.2) using the IBEM with the in-core solver.

REFERENCES

Alexander, M. and Laubach, L. (1968). Anthropometry of the human ear. Aerospace medical research laboratories, Wright-Patterson air force base, Ohio.

Arfken, G. (1970). Mathematical Methods for Physicists (2nd Edition) New York: Academic Press.

Astley, R.J., Macaulay, G.J. and Coyette, J.P. (1994). Mapped wave envelope elements for acoustical radiation and scattering. *Journal of Sound and Vibration* **170**(1), 97-118.

Astley, R.J., Macaulay, G.J., Coyette, J.P. and Cremers, L. (1998). Three-dimensional wave-envelope elements of variable order for acoustic radiation and scattering. Part I. Formulation in the frequency domain, *Journal of the Acoustical Society of America* **103**, 49-63.

Atal, B.S. and Schroeder, M.R. (1962). Apparent sound source translator. U.S. patent 3236949.

AUDIS (1998). The AUDIS Catalog of Human HRTFs. http://www.ika.ruhr-uni-bochum.de/EAA/

Avendano, C., Algazi, V.R. and Duda, R.O. (1999). A head-and-torso model for low-frequency binaural elevation effects, Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New York.

Batteau, D.W. and Plante, R.L. (1962). The mechanisms of human localization of sounds with application in remote environments. China Lake, Calif., United Research, Inc. under U.S. Navy Contract N123-(60530) 27872A.

Batteau, D.W. (1967). The role of the pinna in human localization. Proc. Royal Society London, series B 168: 158-180.

Begault, D.R. (1994). 3-D sound for virtual reality and multimedia, AP Professional, Cambridge MA.

Begault, D.R., Wenzel, E.M., Lee, A.S. and Anderson, M.R. (2000). Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source, *Audio Engineering Society*, 108th convention, preprint 5134, Paris.

Berkhout, A.J., de Vries, D. and Vogel, P. (1993). Acoustic control by wave field synthesis, *Journal of the Acoustical Society of America* **93**, 2764-2778.

Blauert, J. (1997). Spatial hearing: the psychophysics of human sound localisation, MIT Press, Cambridge, MA.

Bloom, P.J. (1977a). Creating Source Elevation Illusions by Spectral Manipulation, *Journal of the Audio Engineering Society*, **25**, 560-565.

Bloom, P.J. (1977b). Determination of Monaural Sensitivity Changes Due to the Pinna by Use of Minimum Audible-Field Measurements in the Lateral Vertical Plane. *Journal of the Acoustical Society of America* **61**, 820-828.

Borgiotti, G.V. (1990). The power radiated by vibrating body in an acoustic fluid and its determination from boundary measurements. *Journal of the Acoustical Society of America* **88**(4), 1884-1893.

Brown, C.P. and Duda, R.O. (1998). A structural model for binaural sound synthesis, IEEE transactions on speech and audio processing, Vol. 6, No. 5.

Brüel and Kjær, HATS artificial head, Brüel and Kjær, Nærum, Denmark

Burkhard, M.D. and Sachs, R.M. (1975). Anthropometric Manikin for Acoustic Research. *Journal of the Acoustical Society of America*. **58**, 214-222.

Burnett, D.S. and Holford, R.L. (1998a). An ellipsoidal acoustic infinite element, Computer Methods in Applied Mechanics and Engineering, 164, 49-76.

Burnett, D.S. and Holford, R.L. (1998b). Prolate and oblate spheroidal acoustic infinite elements, Computer Methods in Applied Mechanics and Engineering, 158, 117-141.

Burnett, D.S. (1994). A three-dimensional acoustic infinite element based on a prolate spheroidal multipole expansion. *Journal of the Acoustical Society of America* **96** (5), Pt.1, 2798-2816.

Butler, R.A. and Belendiuk, K. (1977). Spectral cues utilized in the localization of sound in the median sagittal plane. *Journal of the Acoustical Society of America* **61,** 1264-1269.

Butler, R.A. (1997). Spatial referents of stimulus frequency: Their role in sound localization. In Gilkey, R.H. and Anderson, T.R. Binaural and spatial hearing in real and virtual environments

Carlile, S. and Pralong, D. (1994). The localization-dependent nature of perceptually salient features of the human head-related transfer functions. *Journal of the Acoustical Society of America* **95**, 3445-3459.

Carlile, S. (1996). Virtual auditory space: Generation and applications, Springer-Verlag.

Chen, J., Van Veen, B.D. and Hecox, K.E. (1992). External ear transfer function modelling: a beamforming approach. *Journal of the Acoustical Society of America* **91**, 1333-1344.

Chen, J., Van Veen, B.D. and Hecox, K.E. (1995). A spatial feature extraction and regularization model for the head-related transfer function. *Journal of the Acoustical Society of America*. **97**(1), 439-452.

Christensen, F., Møller, H., Minnaar, P.J., Plogsties, J. and Olesen, S.K. (1999). Interpolation between Head-related transfer functions measured with low directional resolution, *Audio Engineering Society*, 107th Convention, preprint 5047, New York.

Ciscowski, R.D. and Brebbia, C.A. (1991). Boundary Element methods in acoustics. Computational Mechanics Publications & Elsevier Applied Science, Southampton.

CORTEX (1996). Binaural recording head, MK1, Neutrik Cortex Instruments (NCI)

Colton, D. and Kress, R. (1998). Inverse Acoustic and Electromagnetic Scattering Theory (2nd Edition) Berlin: Springer-Verlag.

Coyette, J.P and Lecomte, C. (1997). A generalized axisymmetric boundary element model for complex acoustic problems. LMS, internal report.

Coyette, J.P. Migeot, J.L., Lecomte, C. and McCulloch, C.F. (1999). Numerical techniques for solving acoustic transparency problems. LMS, internal report.

Cremers, L., Fyfe, K.R. and Coyette, J.P. (1994). A variable order infinite acoustic wave envelope element. *Journal of Sound and Vibration* **171**(4), 483-508.

Cremers, L. and Fyfe, K.R. (1995). On the use of variable order infinite wave envelope elements for acoustic radiation and scattering. *Journal of the Acoustical Society of America* **97**(4), 2028-2040.

Currey, M.N. and Cunefare, K.A. (1995). The radiation modes of baffled finite plates. *Journal of the Acoustical Society of America* **98**(3), 1570-1580.

Desmet, W. (1997). Introduction to numerical acoustics, International seminar on applied acoustics (ISAAC), University of Leuven, Belgium.

Dryfus, H. (1967). The measure of man. Whitney library of design, New-York.

Duda, R.O. (1997). Elevation dependence of the interaural transfer function. In Gilkey, R.H. and Anderson, T.R. Binaural and spatial hearing in real and virtual environments, Lawrence Erlbaum Associates, Mahwah, NJ.

Duda, R.O., Avendano, C. and Algazi, V.R. (1999). An adaptable ellipsoidal head model for the interaural time difference. IEEE International Conference on Acoustics, Speech, and Signal Processing, Proceedings Vol. 2, 965-968.

Elliott, S.J. and Johnson, M.E. (1993). Radiation modes and active control of sound power. *Journal of the Acoustical Society of America* **94**(4), 2194-2204.

Evans, J.E., Angus, J.A.S. and Tew, A.I. (1998). Analysing head-related transfer function measurements using surface spherical harmonics, *Journal of the Acoustical Society of America* **104**(4), 2400-2411.

Fillipi, P.J.T. (1977). Layer potentials and acoustic diffraction. *Journal of Sound and Vibration* **54**, 473-500.

Fillipi, P.J.T., Habault, D. and Piraux, J. (1998). Noise source modelling and identification procedure. *Journal of Sound and Vibration* **124**, 285-296.

Flemming, C., Møller, H., Minnaar, P., Plogsties, J. and Olesen, S.K. (1999). Interpolating between Head-Related Transfer Functions Measured with Low Directional Resolution, *Audio Engineering Society,* 107th convention, Preprint 5047, New-York.

Foley, J., Dam, A.V., Feiner, S and Hughes, J. (1990). Computer Graphics: Principles and Practice. Addison-Wesley, New York, (1990).

Gardner, M.B. and Gardner, R.S. (1973). Problem of Localization in the Median Plane: Effect of Pinna Cavity Occlusion. *Journal of the Acoustical Society of America* **53**, 400-408.

Gardner, W.G., Martin, K (1994). HRTF measurements of KEMAR dummy head microphone. Technical report 280, Massachusetts Institute of Technology - Media Lab Perceptual Computing Group.

Gardner, W.G. and Martin, K. (1995). KEMAR HRTF measurements, MIT's Media Lab through http://sound.media.mit.edu/KEMAR.html.

Gardner, W.G. (1999). Reduced-rank modeling of head-related impulse responses using subset selection. In Proceedings of the IEEE Workshop on Application of Signal Processing to Audio and Acoustics.

Genuit, K. (1986). A description of the human outer ear transfer function by elements of communication theory, Proceeding of the 12th International Acoustics, Toronto, Canada. ADSTR. B6-8

Genuit, K. (1987). Method and apparatus for simulating outer ear free-field transfer function, U.S. Patent 4672569.

Gilkey, R.H. and Anderson, T.R. (1997). Binaural and spatial hearing in real and virtual environments, Lawrence Erlbaum Associates, Mahwah, NJ.

Givoli, D. (1993). Numerical methods for problems in infinite domains. Elsevier Science, Amsterdam.

Grace, S.P., Atassi, H.M. and Blake, W.K. (1996). AIAA Journal 34, 2233-2240. Inverse aeroacoustic problem for a streamlined body, Part 1: Basic formulation.

Grace, S.P., Atassi, H.M. and Blake, W.K. (1996). AIAA Journal 34, 2241-2246. Inverse aeroacoustic problem for a streamlined body, Part 2: Accuracy of solutions.

Hammershøi, D. and Møller, H. (1996). Sound transmission to and within the human ear canal. *Journal of the Acoustical Society of America* **100**(1), 408-427.

Han, H.L (1994). Measuring a dummy head in search of pinna cues, *Journal of Audio Engineering Society* **42**(1/2), 15-37.

Hebrank, J. and Wright, D. (1974). Spectral cues used in the localization of sound sources on the median plane, *Journal of the Acoustical Society of America* **56**(6), 1829-1834.

Heckbert, P. and Garland, M. (1995). Survey of polygonal surface simplification algorithms, Carnegie Mellon University, School of Computer Science, Tech. Report (CMU-CS-97-194)

Hiranaka, Y. and Yamasaki, H. (1983). Envelope representations of pinna Impulse responses relating to three-dimensional localization of Sound Sources. *Journal of the Acoustical Society of America* **73**, pp. 291-296.

Huopaniemi, J. and Karjalainen, M. (1997). Review of digital filter design and implementation methods for 3-D sound, *Audio Engineering Society*, 102nd convention, preprint 4461, Munich.

Huopaniemi, J., Zacharov, N. and Karjalainen, M. (1999). Objective and subjective evaluation of Head-related transfer function filter design, *Journal of the Audio Engineering Society*, **47**(4), 218-239.

IEC 959 (1990). Provisional head and torso simulator for acoustic measurements on air conduction hearing aids. International Electrotechnical Commission (IEC), international standard IEC/TR0 60959 (1990-05). ED 1.0. ICS code 17.140.50 Electroacoustics

Johnson, A.E. and Hebert, M. (1997). Control of Polygonal Mesh Resolution for 3-D Computer Vision, Carnegie Mellon University, School of Computer Science, Tech. Report (CMU-RI-TR-96-20).

Jot, J.M., Larcher, V. and Pernaux, J.M (1999). A comparative study of 3-D audio encoding and rendering techniques. *Audio Engineering Society*, 16th international conference, 281-300, Rovaneimi.

Kahana, Y. (1997a). Multi-channel sound reproduction with a four-ear dummy-head. *M.Sc. thesis*. Institute of Sound and Vibration Research, University of Southampton, England.

Kahana, Y., Nelson, P.A., Kirkeby, O. and Hamada, H. (1997b). Multi-channel sound reproduction using a four-ear dummy-head, *Audio Engineering Society*, 102nd Convention, preprint 4465, Munich.

Kahana, Y., Nelson, P.A., Kirkeby, O. and Hamada, H. (1997c). Objective and subjective assessment of systems for the production of virtual acoustic images for multiple listeners, *Audio Engineering Society,* 103rd Convention, preprint 4573, New-York.

Kahana, Y., Nelson, P.A., Petyt, M. and Choi, S. (1998). Boundary Element simulation of HRTFs and sound fields produced by virtual acoustic imaging systems, *Audio Engineering Society*, 105th convention, preprint 4817, San Francisco.

Kahana, Y., Nelson, P.A., Kirkeby, O. and Hamada, H. (1999a). Multiple microphone technique for the reproduction of virtual acoustic images. *Journal of the Acoustical Society of America* **105**(3), 1503-1516.

Kahana, Y., Nelson, P.A. and Yoon, S. (1999b). Experiments on the synthesis of virtual acoustic sources in automotive interiors, *Audio Engineering Society,* 16th international conference, 218-232, Rovaneimi.

Kahana, Y., Nelson, P.A., Petyt, M. and Choi, S. (1999c). Numerical modeling of the transfer functions of a dummy-head and of the external ear, *Audio Engineering Society*, 16th international conference, 330-345, Rovaneimi.

Kahana, Y. and Nelson, P.A. (2000). Spatial acoustic mode shapes of the human pinna, *Audio Engineering Society,* 109th convention, Los-Angeles.

Katz, B.F.G. (1998). Measurement and calculation of individual head-related transfer functions using a boundary element model including the measurement and effect of skin and hair impedance. *Ph.D Thesis*. Pennsylvania State University.

KEMAR (1978). Manikin measurement, Proceedings of a conference organised by M.D. Burkhard, Industrial Research Products, Inc, A Knowles Company

Kim, G.T. and Lee, B.H. (1990). 3-D sound reconstruction and field projection using the Helmholtz integral equation. *Journal of Sound and Vibration* **136**, 245-26.

Kinsler, L.E. ,Frey, A.R., Coppens, A.B and Sanders, J.V. (1982). Fundamentals of acoustics, John Wiley & Sons, (1982).

Kirkeby, O., Nelson, P.A. and Hamada, H. (1998a). Local sound field reproduction using two closely spaced loudspeakers. *Journal of the Acoustical Society of Americ* **104**(4), 1973-1981.

Kirkeby, O., Nelson, P.A. and Hamada, H. (1998b). The "stereo dipole" - A virtual source imaging system using two closely spaced loudspeakers. *Journal of the Audio Engineering Society*, **46**(5), 387-395.

Kistler, D.J. and Wightman, F.L. (1992). A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. *Journal of the Acoustical Society of Americ* **91**(3), 1637-1647.

Kuhn, F. (1977). Model for the interaural time difference in the azimuthal plane. *Journal of the Acoustical Society of America* **62**(1), 157-167.

Kuhn, G.F. (1982). Towards a Model for Sound Localization, in Localization of Sound: Theory and Applications, R. W. Gatehouse, Ed. (Amphora Pres Groton, CT), pp. 51-64.

Kuhn, G.F. and Guernsey R.M. (1983). Sound pressure distribution about the human head and torso. *Journal of the Acoustical Society of America* **73**(1), 95-105.

Kuhn, G.F. (1987). Physical acoustics and measurements pertaining to directional hearing. Directional Hearing. W. A. Yost and G. Gourevitch. New York, Springer-Verlag: Chap. 1. 3-25.

Kulkarni, A., Isabelle, S.K. and Colburn, H.S. (1999). Sensitivity of human subjects to Head-related transfer-function phase spectra. *Journal of the Acoustical Society of America* **105**(5), 2821-2840.

Kyriakis, K. and Holman, T. (1997). Video-based head tracking for improvements in multichannel loudspeaker audio. *Audio Engineering Society*, 105th convention, Preprint 4845, San Francisco.

Larcher, V., Jot, J.M., Guyard, J. and Warusfel, O. (2000), Study and comparison of efficient methods for 3-D audio spatialisation based on linear decomposition of HRTF data, *Audio Engineering Society*, 108th convention, preprint 5097, Paris.

Mackensen, P., Fruhmann, M., Thanner, M., Theile, G., Horbach, U. and Karamustafaiyglu, A. (2000). Head-Tracker based auralization systems: additional consideration of vertical head movements, *Audio Engineering Society*, 108th convention, preprint 5135, Paris.

Martens, W.L. (1987). Principal components analysis and resynthesis of spectral cues to perceived direction, The International Computer Conference, edited by J. Beauchamp (International Computer Music Association, San Francisco), 274-281.

MATHEMATICA 3.0 Users Manual, Wolfram Research Inc.

MATLAB 5.3 Users Manual, MathWorks. Inc.

Maxwell, R.J. and Burkhard, M.D. (1979). Larger Ear Replica for KEMAR Manikin. *Journal of the Acoustical Society of America* **65**, 1055-1058.

Mehrgardt, S. and Mellert, V. (1977). Transformation characteristics of the external human ear. *Journal of the Acoustical Society of America* **61**(6), 1567-1576.

Middlebrooks, J.C., Makous, J.C and Green, D.M. (1989). Directional sensitivity of sound pressure levels in the human ear canal. *Journal of the Acoustical Society of America* **86**, 89-108.

Middlebrooks, J.C. and Green, D.M. (1990). Directional dependence of interaural envelope delays. *Journal of the Acoustical Society of America* **87**(5), 2149-2162.

Middlebrooks, J.C. (1997). Spectral shape cues for sound localiztion, in Gilkey, R.H. and Anderson, T.R. (1997). Binaural and spatial hearing in real and virtual environments, Lawrence Erlbaum Associates, Mahwah, NJ.

Minnaar, P.J., Plogeties, J., ølesen, S.K., Christensen, F. and Møller, H. (2000). The interaural time difference in binaural synthesis. *Audio Engineering Society,* 108th convention, Preprint 5133, Paris.

MLSSA, (1998). Maximum-length sequence analyzer, Reference manual, version 10W, DRA laborastories, D. Rife

Møller, H. (1992). Fundamentals of binaural technology. Applied acoustics, vol. 36, 171-218.

Møller, H., Sørensen, M.F., Hammershøi, D. and Jensen C.B. (1995). Head-related transfer function of human subjects, *Journal of the Audio Engineering Society,* **43**(5), 300-321.

Møller, H., Sørensen, M., Hammershøi, D. and Jensen, C.B. (1996). Binaural technique: Do we need individual recordings? *Journal of the Audio Engineering Society,* **44** (6), 451-469.

Møller, H., Jensen, C.B., Hammershøi, D. and Sørensen, M (1999). Evaluation of artificial heads in listening tests, *Journal of the Audio Engineering Society,* **47** (3), 83-100.

Morimoto, M. and Ando, Y. (1982). Simulation of sound localisation. In R.W. Gatehouse (Ed.), Localization of sound: Theory and application, The Amphora Press, Groton, CT. pp. 85-89.

Morse, P.M. and Feshbach, H. (1953). Methods of Theoretical Physics New York: McGraw Hill.

Morse, P.M. and Ingard, K.U. (1968). Theoretical Acoustics Princeton New Jersey: Princeton University Press.

Musicant, A.D. and Butler, R.A. (1984). The influence of pinna-based spectral cues on sound localization. *Journal of the Acoustical Society of America* **75**, 195-1200.

Musicant, A.D. and Butler, R.A. (1985). Influence of monaural spectral cues on binaural localization. *Journal of the Acoustical Society of America* **77**, 202-208.

Musicant, A.D. (1995). The relationship between tone frequency and perceived elevation under headphone listening conditions. *Journal of the Acoustical Society of America* **97**(5), 3279.

Nelson, P.A., Kirkeby, O., Takeuchi, T. and Hamada, H. (1997). Sound fields for the production of virtual acoustic images. *Journal of Sound and Vibration* **2**, 386-396.

Nelson, P.A. (1999). Some inverse problems in acoustics. Proceedings of the Seventh International Congress on Sound and Vibration, Copenhagen 1, 7-32.

Nelson, P. A. and Yoon, S.H. (2000). Estimation of acoustic source strength by inverse methods: Part I, Conditioning of the inverse problem. *Journal of Sound and Vibration* (in press).

Nelson, P.A and Kahana, Y. (2000). Spherical harmonics, singular value decomposition and the head related transfer function. Special edition of the *Journal of Sound and Vibration* (in press).

Novy R.W. (1998). Characerizing elevation effects of a prolate spheroidal HRTF model, San Jose State University, Department of Electrical Engineering, Tech. Report No. 20.

Ono, K., Komiyama, S. and Nakabayashi, K. (1998). A Method of Reproducing Concert Hall Sounds by "Loudspeaker Walls". *Journal of the Audio Engineering Society*, **46**(11), 988-995.

Oldfield, S.R. and Parker, P.A. (1984). Acuity of sound localization: a topography of auditory space II, pinna cues absent. *Perception* **13**, 601-617.

PARKER automation (1999). Mini Stepping drive, User guide, PDFX Series, Electromechanical Division of Parker Hannifin, plc.

Pierce, A.D. and Wu, X.F. (1983). Variational method for prediction of acoustic radiation from vibrating bodies. *Journal of the Acoustical Society of America* **74**, Suppl. 1, S107.

Photiadis, D.M. (1990). The relationship of singular value decomposition to wave-vector filtering in sound radiation problems. *Journal of the Acoustical Society of America* **88**(2), 1152-1159.

Plogesties, J., Minnaar, P.J., Olesen, S.K., Christensen, F. and Møller, H. (2000). Audibility of all-pass components in head-related transfer functions. *Audio Engineering Society*, 108th Convention, Paris. (Preprint)

Pralong, D. and Carlile, S. (1994). Measuring the human head-related transfer functions: A novel method for the construction and calibration of a miniature "in ear" recording system. *Journal of the Acoustical Society of America* **95**, 3435-3444.

Rayleigh, Lord. (1907). (J. W. Strutt, 3rd Baron of Rayleigh), On our perception of sound direction, Philos. Mag. 13, 214-232.

Rife, D. and Vanderkooy, J. (1989). Transfer-function measurement with maximum-length sequences, *Journal of the Audio Engineering Society,* **37** (6), 419-444.

Sandvad, J. (1996). Dynamic aspects of auditory virtual environments. *Audio Engineering Society,* 100th convention, preprint 4226, Copenhagen.

Schenck, H.A. (1968). Improved integral formulation for acoustic radiation problems. *Journal of the Acoustical Society of America* **44**, 41-58.

Searle, C.L., Braida, L.D., Cuddy, D.R. and Davis, M.F. (1975). Model for auditory localization. *Journal of the Acoustical Society of America* **60**, 1164-1175.

Seybert, A.F., Soenarko, B., Rizzo, F.J. and Shippy, D.J. (1985). An advanced computational method for radiation and scattering of acoustic waves in three dimensions. *Journal of the Acoustical Society of America* **77**, 362-368.

Shaw, E.A.G. and Teranishi, R. (1968). Sound pressure generated in an external-ear replica and real human ears by a nearby point source, *Journal of the Acoustical Society of America* **44**(1), 240-249.

Shaw, E.A.G. (1974). Transformation of sound pressure level from the free-field to the eardrum in the horizontal plane. *Journal of the Acoustical Society of America* **56**(6), 1848-1861.

Shaw, E.A.G. (1975). The external ear: New Knowledge in Earmolds and Associated Problems, Proceedings of the Seventh Danavox Symposium edited by S.C. Dalsgaard, Scand. Audiol. (supp. 5) pp. 24-50.

Shaw, E.A.G. (1979). Rayleigh Medal Lecture: The elusive connection in Localisation of Sound: Theory and Applications, edited by R.W. Gatehouse (Amphora Press, Groton, CT), pp. 18-29.

Shaw, E.A.G. (1982). External Ear Response and Sound Localization, in Localization of Sound: Theory and Applications, R. W. Gatehouse, Ed. (Amphora Press, Groton, CT), pp. 30-41.

Shaw, E.A.G. and Vaillancourt M. M. (1985). Transformation of sound-pressure level from the free-field to the eardrum presented in numerical form. *Journal of the Acoustical Society of America* **78**(3), 1120-1123.

Shaw, E.A.G. (1997). Acoustical features of the human external ear; in Binaural and Spatial Hearing in Real and Virtual Environments, R.H. Gilkey and T.R. Anderson (Eds), Mahwah, New Jersey; Lawrence Erlbaum Associates.

Shinn-Cunningham, B. and Kulkarni A, (1996). Recent developments in virtual auditory space, in Virtual auditory space: generation and application, by Simon Carlile, Springer-Verlag.

Spence, R.D. and Granger, S. (1951). The scattering of sound from prolate spheroid. *Journal of the Acoustical Society of America* **23**, 701-706.

Speyer, G. (1999). A boundary element model for predicting the Head related transfer function. *Ms.c thesis,*Tel-Aviv University, Israel.

Stinson, M.R. (1985). The spatial distribution of sound pressure within scaled replicas of the human ear canal. *Journal of the Acoustical Society of America* **78**, 1596-1602.

Stinson, M.R. and Lawton, B.W. (1989). Specification of the geometry of the human ear canal for the prediction of sound-pressure level distribution. *Journal of the Acoustical Society of America* **85**, 2492-2503.

Sugiyama, K. and Irii, H. (1991). Prolate spheroid Comparison of the sound pressure radiation from a prolate spheroid and the human mouth. *Acustica* **73**(5), 271-276.

SYSNOISE 5.4 (1993-1999). Users Manual, LMS International, Leuven, Belgium. LMS Numerical Integration Technologies, SYSNOISE theoretical manual, ver. 5.0-5.4, Leuven, Belgium.

Temkin, S. (1981). Elements of Acoustics New York: John Wiley.

Teranishi, R. and Shaw, E.A.G. (1968). External-ear acoustic models with simple geometry. *Journal of the Acoustical Society of America* **44**(1), 257-263.

Wallach, H. (1940). The role of head movements and vestibular and visual cues in sound localization. *J. Exp. Psych.* **27**. 339-368.

Watkins, A.J. (1978). Psychoacoustical Aspects of Synthesized Vertical Locale Cues. *Journal of the Acoustical Society of America* **63**, 1152-1165.

Weinrich, S. (1984). Sound field calculations around the human head, *Technical report* 37, The acoustics laboratory, Technical university of Denmark.

Wenzel, E.M. Arruda, M., Kistlerb D.J. and Wightman, F.L. (1993). Localization using non-individualised head-related transfer functions. *Journal of the Acoustical Society of America* **94**(1), 111-123.

Wenzel, E.M., Miller, J.D. and Abel, J.S. (2000). Sound Lab: A real-Time, Software-Based system for the study of spatial hearing, *Audio Engineering Society,* 108th convention, Preprint 5140, Paris.

Wightman, F.L. and Kistler, D.J. (1989). Headphone simulation of free-field I and II: Psychophysical validation. *Journal of the Acoustical Society of America* **85**, 858-878.

Wightman, F.L. and Kistler, D.J. (1992). The Dominant Role of Low-Frequency Interaural Time Differences in Sound Localization. *Journal of the Acoustical Society of America* **91**(3), 1648-1661.

Wightman, F.L. and Kistler, D.J. (1997). Factors affecting the relative slence of sound localization. In Gilkey, R.H. and Anderson, T.R. Binaural and spatial hearing in real and virtual environments, Lawrence Erlbaum Associates, Mahwah, NJ.

Wright, D., Hebrank J.H. and B. Wilson, B. (1974). Pinna Reflections as Cues for Localization. *Journal of the Acoustical Society of America* **56**, 957-962.

Yoon, S.H. and Nelson, P.A. (2000). Estimation of acoustic source strength by inverse methods: Part II, Experimental investigation of methods for choosing regularisation parameters. *Journal of Sound and Vibration* (in press).

Zhou, B, Green, D.M. and Middlebrooks, J.C (1992). Characterization of external ear impulse responses using Golay codes. *Journal of the Acoustical Society of America* **92**(2), 169-1171.

Zwislocki, J. (1970). An acoustic coupler for earphone calibration. Special report LCS-S-7, Syracuse University, New-York.

APPENDIX 1 MESH MANIPULATION

The mesh models that can be handled by the BEM require that acoustic elements are interconnected by their edges, which are connected through the vertices. Fortunately a similar format of geometry is provided by the laser scanner. However, the geometrical model cannot be used directly in the BEM for the following reasons:

- The integration procedures in the BEM requires that certain properties of the mesh must be met, such as a minimum edge length, no holes should appear in the mesh (for the DBEM), consistency of element normals, additional elements for special formulations such as the 'IBEM transparency', baffled and coupled structures, etc.
- The original mesh resolution is much higher than required by the BEM, and in any case cannot be handled with the current computing hardware. Since the BEM is very inefficient for solving large problems, the size of the mesh must be reduced.

A few tools have been developed, but the process is still interactive and difficult. It is doubtful, if with the current scanning procedures, an automated manipulation tool could be devised. The goal of this Appendix is to review the main tools used, and to summarise the geometrical properties of the head and pinnae models used in this research.

A1-1 SCANNER PERFORMANCE

- **'Low-resolution' scanner - Motion platform 3030**

In the case of scanning the head with the 'low-resolution' scanner, the scanner produces an incomplete mesh: this is since the geometry is captured 'in one go' and, if some parts cannot be viewed, the mesh will include holes. In fact, many holes appear, with some of

them easy to fill and some not. The top of the head is a more difficult problem, since extrapolation of the curvature is required in three views. Fortunately, the scanner software has this capability, although a few iterations are required in order to obtain a smooth filling.

- **'High-resolution' scanner – HIREZ /mini motion platform**

This scanner accumulates the geometry by tens of scans, and it was used to scan pinnae only. The accuracy is much higher, but also in this case a large number of holes exist at the end of many scans. The operator of the scanner is required to decide when a sufficient level of accuracy is obtained, so that his interactive filling of the holes (especially in the folds of the concha and fossa of helix) will not produce too excessive a distortion of the geometry. In particular, it was found that although the entrance to the ear canal in all rubber pinnae models was parallel to the base of the model, and was circular, in practice this was not possible to achieve even with this high resolution of scanner. The final model includes details all around the pinna (front, sides and back), although we are interested only in the frontal shell.

The motion platforms are presented in Figure A1-1. Table A1-1 summarises the typical accuracies of the Cyberware scanners, although it should be noted that the accuracies in practice are determined by the curvature and complexity of the object, where with the 'mini high-resolution' scanner, this problem is somewhat alleviated.

	Head and face 3030RGB SCANHEAD	High resolution 3030RGB/HIREZ
X	typically 250 μm - 1.0 mm	typically 150 μm - 1.0 mm
Y	700 μm	313 μm
Z	minimum 100 μm	average of 50-200 μm

Table A1-1: The accuracies of the scanner and motion platforms. Note that the accuracies also depend on the complexity of the object. The high-resolution scanner can accumulate data through repeated scans at different angles.

A1-2 STAGE 1 - MESH MANIPULATION/ PRE-PROCESSING

- **Data format**

Surface meshes are described by a list of 3-D points and triangular faces between the points that create a piecewise linear representation of a surface. All mesh manipulation tools used in this research, imported and exported data files using the VRML 1.0 format. Later, each file was converted to other formats for further manipulation, and the final format was the SYSNOISE 'user-defined', or 'free format'.

The 3-D position of the vertices (as expressed in the `'Coordinate3 field'`), the connectivity of the mesh (as expressed by the `'IndexedFaceSet field'`), and the colour of the mesh (as expressed in the `'diffuse color field'`) can change from mesh to mesh. The rest of the mesh and the order of the fields must remain as presented below. The VRML file below represents a red mesh with two triangular faces and four vertices. For more information about the VRML format, the 'Open Inventor Mentor' in the online section of most SGI's can be consulted. This convention for meshes does not assume that the faces in the mesh are oriented; the `'ShapeHints field'` insures that both sides of the mesh are displayed.

```
#VRML V1.0 ascii
Material {
DiffuseColor [1 0 0]
}
ShapeHints {
vertexOrdering COUNTERCLOCKWISE
shapeType UNKNOWN_SHAPE_TYPE
}
Coordinate3 { point [
0.0 0.0 0.0,
2.0 0.0 0.0,
2.0 2.0 1.0,
0.0 2.0 1.0
]
}
IndexedFaceSet {coordIndex [
0,1,2,-1,
1,2,3,-1
]
}
```

- **Reorientation of the head**

The final mesh of the head is an open volume. The head is not aligned by any means to any specific co-ordinate system. The mesh should be reoriented in space and transformed in 6 DOF such that the following criteria are met:

- The two entrances to the ear canals should be positioned at constant height at $y=0$.
- The centre of the line that connects these two points (the 'interaural axis') is the origin of the co-ordinate system ($x=y=z=0$).
- When the head is viewed from the top, the entrance to the ear canal of the left ear should be at $+\Delta x/2$ and the other ear at $-\Delta x/2$.
- The plane that cuts the head vertically to two identical parts through the centre of the nose should be at $\phi = 90^\circ$ with the horizontal plane that coincides with the interaural axis.
- With KEMAR, the bottom of the head is marked through the neck extensions (see Figure 5-2). These should be approximately parallel to the y (height) axis.

The three rotation angles and three translation variables are found in a program where the input includes five co-ordinates: of the right blocked entrance to the ear canal, the left, the tip of the nose, and front and rear points at the level of the neck extension (the bottom of the head).

- **Closing the neck**

A large hole is present for any scan of the head, as a result of the open neck. Closing this area using conventional techniques results in triangles with high aspect ratio. Since we need to add vertices on the 2-D plane of the closed area and to connect them with triangular elements, the Delaunay triangulation method can be used. The following stages are followed:

- The mesh resolution of the original model is calculated (see details below).
- Vertices are added with the distance between them approximately the mesh resolution, starting from the boundaries, inwards.
- Elements are added using Delaunay triangulation.

- **Mesh check procedure**

Using conventional software tools (ANSYS, HYPERMESH, SYSNOISE), the mesh is checked for the following:

- No holes exist, and a volume can be calculated.
- All normals are facing the same direction (outside is defined as positive normal).
- All the elements are legitimate (i.e. no elements with area size of zero, each edge shares only two elements, no elements with low aspect ratio exist, no isolated nodes, etc.).

A1-3 STAGE 2 - MESH MANIPULATION/ CONTROL OF MESH RESOLUTION

In the literature, curves and surface mesh simplification algorithms are investigated from different perspectives in the following fields: cartography, geographic information systems (GIS), virtual reality, computer vision, computer graphics, scientific visualisation, computer-aided geometric design, approximation theory, computational geometry and finite and boundary element methods. For a detailed review of more than 100 simplifications for the above areas see Heckbert and Garland (1997). In most cases, however, commercial packages for mesh decimation are designed to preserve the accuracy of the rendered model and do not necessarily operate to the same restrictions that the BEM models would require.

The algorithm used in this research has been developed and described in detail by Johnson and Hebert (1997). Its main advantage is in successfully handling the two forces in mesh decimation: preserving the shape by limiting a defined maximum 'global shape error' and distributing the vertices homogeneously by local operators.

At this stage, the head and pinnae models are ready to be converted to the BEM model. The size of these models is huge (approximately 150000 elements for the pinnae and 400000 elements for the head, see below) and is required to be reduced. Since the CPU time of the BEM increases drastically with the number of nodes (see Chapter 10), it is crucial to optimise the size of the mesh. It is well known that the maximum frequency for which accurate results may be obtained in the BEM/FEM/IFEM corresponds to the longest edge in the mesh. Any alteration to this global limit will distort the overall results. Therefore, a homogeneous distribution of the nodes and elements is required. This is achieved by changing the positions of the nodes on the surface while still adequately describing the shape of the object. We first define the following terms:

- *Mesh resolution* - the median of the lengths of the edges in the mesh.
- *Edge length spread* - the upper quartile of edge lengths minus the lower quartile lengths (half widths) in the mesh.
- *Length normalisation* - the process of which the resolution of the original mesh is adjusted to a desired resolution while minimising the edge length spread.
- *Edge length weight* - this is derived from a Gaussian of edge length: $W = \exp\left[(l - L_0)/(L_d)^2\right]$ where l is the length of the edge is, L_0 is the desired resolution and L_d is the acceptable edge length spread.
- *Shape change measure* - each time an edge is collapsed (see below) the shape of the mesh changes slightly. The shape change measure is defined as the maximum distance between the mesh before and after the edge was collapsed.

The principle of the *length normalisation* algorithm used here is similar to other mesh simplification algorithms in that it iteratively changes the mesh by applying local mesh operators. The input data required for the *length normalisation* algorithm includes the upper and lower bounds on edge lengths: $L_{\min} = L_0 - L_d/2$ and $L_{\max} = L_0 + L_d/2$, respectively, and also the maximum global *shape change measure*.

The general flow of iterative simplification algorithms is as follows: first, a dynamically ordered queue is created from all edges in the mesh. The position of an edge in the priority queue is determined by the product of an *edge length weight* and the *shape change measure* of the edge. Next, the first edge in the priority queue is operated on. Two operations are iteratively applied to the edges to achieve the desired resolution: 'edge-split' is used to remove long edges (if the edge of the length is greater than $L_{\max}$, the edge is split at its mid point), and 'edge-collapse' is used to remove short edges (if the edge length is less than

L_{min}, the edge is collapsed. See Figure A1.2). During 'edge-split' an edge is divided in the middle and produces a new vertex, two new edges and two elements, without changing the accuracy of the model. However, the operation of 'edge-collapse' does change the accuracy. Since an edge is reduced to a point (an edge and two elements are eliminated, and two vertices are replaced by a new vertex) expansion and shrinkage can be minimised if the new point is positioned not in the middle of the eliminated edge, but at the projection of the midpoint of the edge on to the planes of the surrounding faces (Figure A1-2b). When an edge is collapsed its *shape change measure* is added to the accumulated shape change of the edges in the new neighbourhood of the edge. After that the priority queue is updated according to the accumulated shape change for each edge.

In addition to these local operations, a maximum allowable shape change for the mesh is defined to prevent the mesh from changing too much. As a result, the low priority edges are treated first, and the program is stopped only if the global shape change criterion is met.

It was found that the decimation algorithm produced minimal errors noticed visually (with errors up to 0.3 mm, see Figure A1-3) as well as acoustically: the deviation of the frequency response calculated with the original raw data of the pinna on a baffle when compared to a decimated pinna was less than ±1 dB.

A1-4 STAGE 3 - INTEGRATING THE PINNA TO THE HEAD AND SLICING THE MODEL

The integration of the pinna required the 'cut' of the 'low-resolution' pinna from the head and identifying key features and record the positions in global co-ordinates. The accurate pinna was transformed to these co-ordinates . The process is still iterative, since the shape of the low-resolution pinna is not always comparable with the high-resolution, accurate

pinna. The mesh resolution was adjusted around the pinna to remove the long edges used to connect the new pinna and the surrounding frame.

Slicing the models was required in three cases:

- At an arbitrary position (parallel to one of the three axes). A contour is then provided, that connects all the x and y values for a given z plane of intersection. Then different decimated mesh models can be compared at positions in which complex changes occur in the geometry, such as the face and the pinna (as in Figure A1-3).
- Slicing the bottom of the head, at a constant y value. The 2-D area around the neck is closed using a Delaunay triangulation.
- Dividing the head into two identical parts, when the symmetric formulation is used with the DBEM or the IBEM.

A1-5 GEOMETRICAL PROPERTIES OF THE MODELS

A summary of the properties of all the original models and the decimated models used in BEM modelling are presented in Tables A1-1 and A1-2. The variables are defined. The bounding box is aligned with the XYZ axes of world co-ordinates. When the spread of length is small (in the order of 25%) then the surface described is more or less uniformly sampled.

		YK_head Original data	KEMAR Original data	YK_pinna Original data	DB60 Original data
Number of mesh points		209718	202338	112170	72849
Number of mesh edges		627740	606449	335762	217839
Number of mesh faces		418023	404112	223592	144988
Mesh resolution		0.959	0.978	0.360	0.366
Average mesh edge length		0.982	0.993	0.359	0.359
StDev mesh edge length		0.254	0.261	0.114	0.111
Spread of length		25.92%	26.28%	31.69%	30.85%
Max mesh edge length		4.076	3.191	2.932	2.994
Min meshedge length		0.066	0.073	0.009	0.014
Bounding Box minimum	X	-49.708	-56.191	-30.817	-25.151
	Y	-121.429	-122.091	-30.265	-30.761
	Z	-104.617	-107.633	-12.856	-5.876
Bounding Box Maximum	X	32.197	31.688	24.683	24.535
	Y	130.981	123.122	42.219	33.914
	Z	127.124	111.203	21.682	19.079

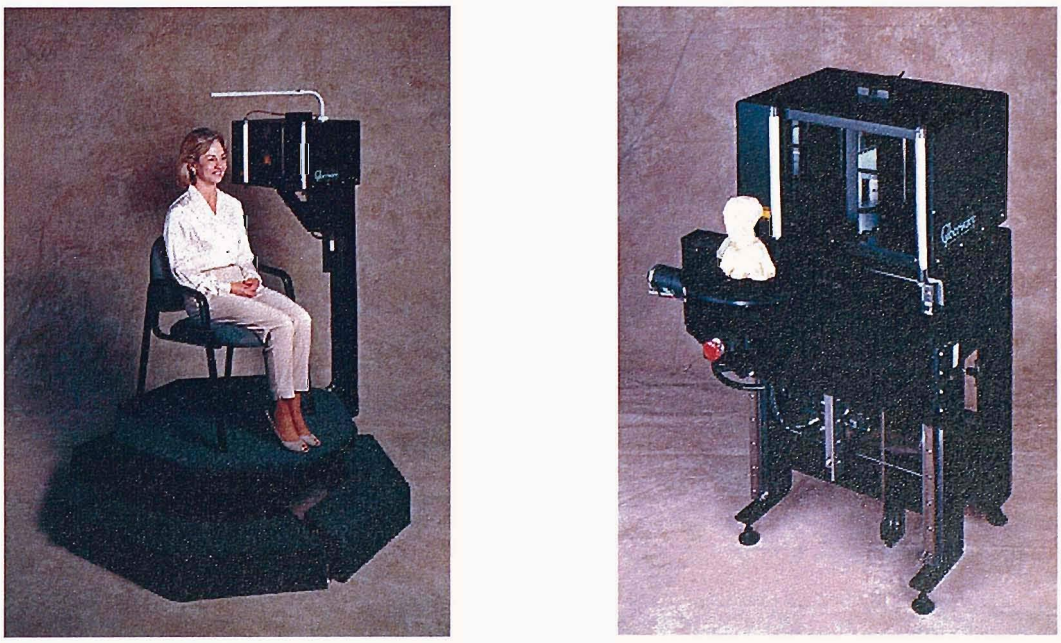
DB65	DB90	DB95	BK	CORTEX
117721	115946	120751	130423	97623
353163	347838	362069	391203	290529
235444	231892	241319	260781	192802
0.355	0.355	0.355	0.357	0.371
0.349	0.351	0.349	0.348	0.385
0.085	0.088	0.087	0.088	-0.097
24.36%	25.25%	24.974%	25.4929%	25.25%
1.084	1.038	1.460	0.953	1.648
0.018	0.020	0.021	0.021	0.017
-27.612	-28.580	-28.330	-33.044	-30.063
-30.711	-29.129	-29.978	-29.754	-30.971
-6.965	-7.418	-12.186	-10.369	-28.793
24.095	23.660	23.799	25.148	30.029
34.728	35.678	35.489	33.351	33.158
23.915	22.903	23.679	20.971	27.258

Table A1-2: Original mesh models. Statistical analysis of mesh properties. All lengths are in millimetres.

	YK_head Decimated data	KEMAR Decimated data	YK_pinna Decimated data	DB60 Decimated data
Number of mesh points	13518	11634	8417	6505
Number of mesh edges	40008	34590	24784	19245
Number of mesh faces	26491	22957	16367	12740
Mesh resolution	2.510	2.774	1.058	1.058
Average mesh edge length	2.542	2.826	1.082	1.087
StDev mesh edge length	0.712	0.634	0.312	0.306
Spread of length	26.9%	22.43%	28.85%	28.14%
Max meshedge length	6.125	5.634	2.79884	2.94121
Min meshedge length	0.236	0.303	0.089	0.047

DB65	DB90	DB95	BK	CORTEX
7564	7505	7782	7885	8023
22317	22151	22784	23197	23586
14753	14647	15214	15312	15562
1.038	1.035	1.024	1.048	0.988
1.067	1.067	1.043	1.073	0.996
0.290	0.322	0.302	0.299	0.293
27.21%	26.9%	28.3%	27.9%	29.9%
2.685	2.690	2.813	2.708	2.736
0.084	0.124	0.073	0.074	0.083

Table A1-3: Decimated BEM models. Statistical analysis of mesh properties. All lengths are in millimetres.



(a) (b)

Figure A1-1: The scanners used in this research (a) the 'low-resolution' and (b) the high resolution 3030RGB/HIREZ/MM is ideal for scanning.

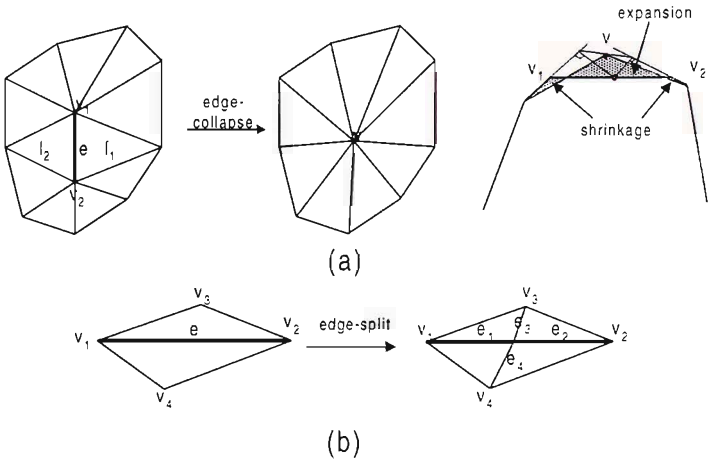


Figure A1-2: The homogeneous mesh decimation algorithm is based on two local operations: edge-split, and edge-collapse. (a) By merging vertices v_1 and v_2 into v, two faces f_1, f_2 and edge e are eliminated, and the original topology is slightly changed. A minimal local error is obtained by placing the new vertex v off the edge. (b) The operation of edge-split creates a new vertex in the middle of edge e, adding two edges and two faces. This operation does not change the overall accuracy of the mesh. (After Johnson and Hebert, 1997).

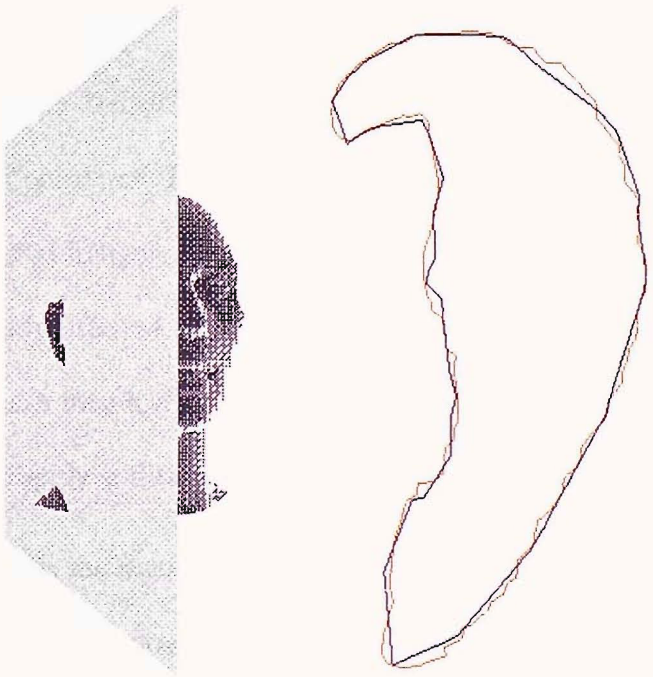


Figure A1-3: The effect of the algorithm on the accuracy of the decimated model. A cutting plane is positioned at an arbitrary part of the pinna, and is slicing a cross-section. Two models are overlapped on the left: the original model with 400,000 elements and the minimum model with 4,000 elements. The two slices are compared on the right thus resulting in a minimal error margin.

APPENDIX 2 SINGULAR VECTORS OF PINNAE

In this Appendix the real and imaginary singular vectors of various pinnae are presented. These are calculated at frequencies at which a maximum peak is obtained in the singular values of the Green function matrix relating the points on the pinna and the points in the far field.

For the first 'resonance' frequency, i.e. the quarter wavelength resonance, the first three singular vectors are presented. In this case, the 'mode shapes' change only with respect to the phase. The real and imaginary values of the left singular vectors seem to have a similar pattern (although with different magnitude levels). The real and imaginary parts of the right singular values are not identical and clearly more sensitive, i.e. the 'monopole' and 'dipole' modes are excited slightly at different locations for each pinna.

As frequency increases, the second mode shown still has many common features among the pinnae investigated. The patterns of the real and imaginary left singular vectors are similar to before, but in this case some phase shifts occur although in all case the 'vertical dipole' pattern is clearly seen. Also 'dipole' patterns appear in the real and imaginary right singular vectors, but with slight distortions.

Lower order singular values at this frequency and higher frequencies are not presented because these do not show significant radiation patterns. For the next, third peak, three pinnae are shown with similar patterns on their surface, but with large variations on the hemisphere. Similar behaviour is noticed also for the fourth peak. It is believed the relatively low resolution of the hemisphere is not sufficient to extract correctly all 'mode shapes' as frequency increases.

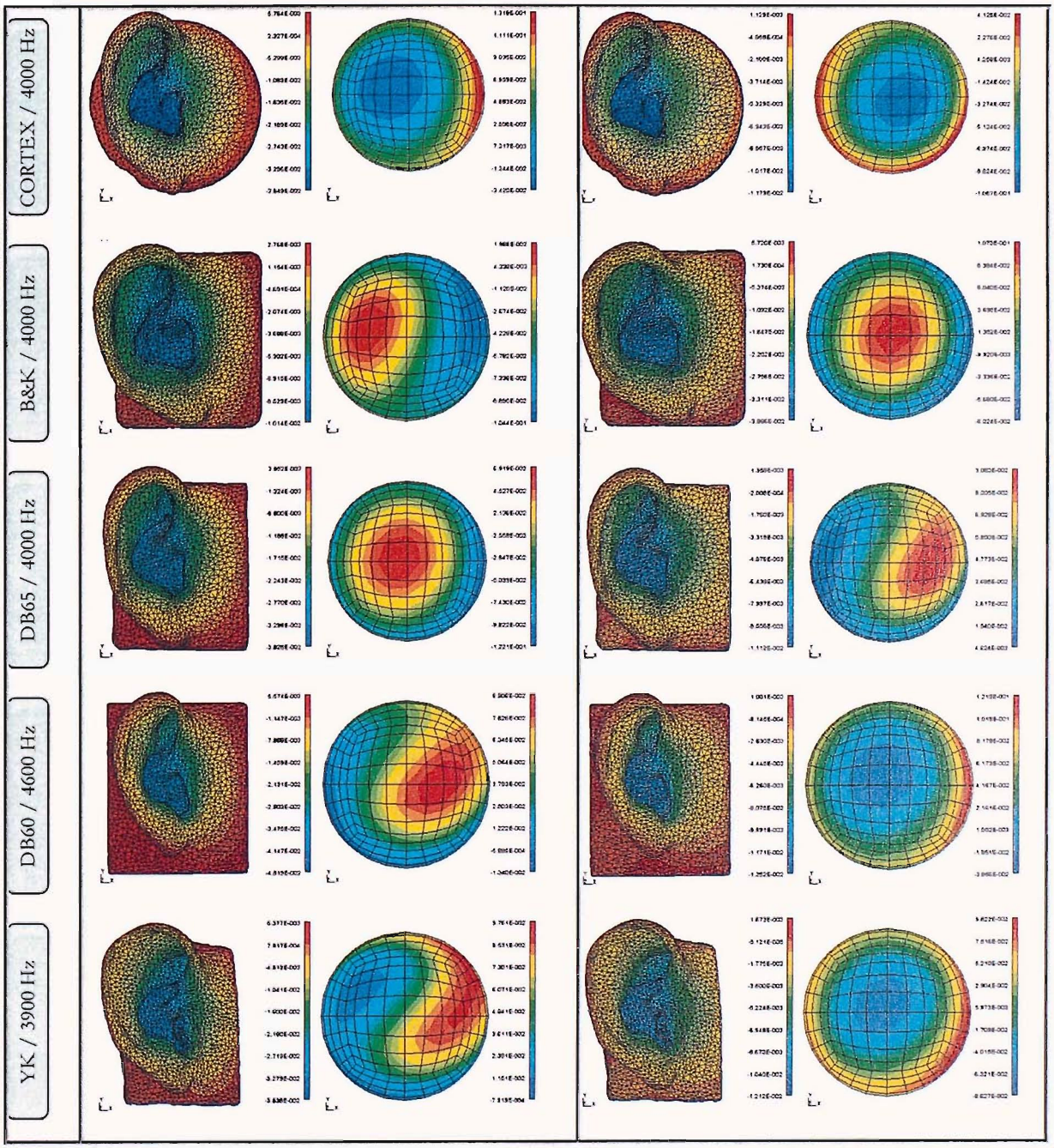


Figure A2-1: The real and imaginary left and right singular vectors associated the first singular value, σ_1 at the first peak, around 4 kHz.

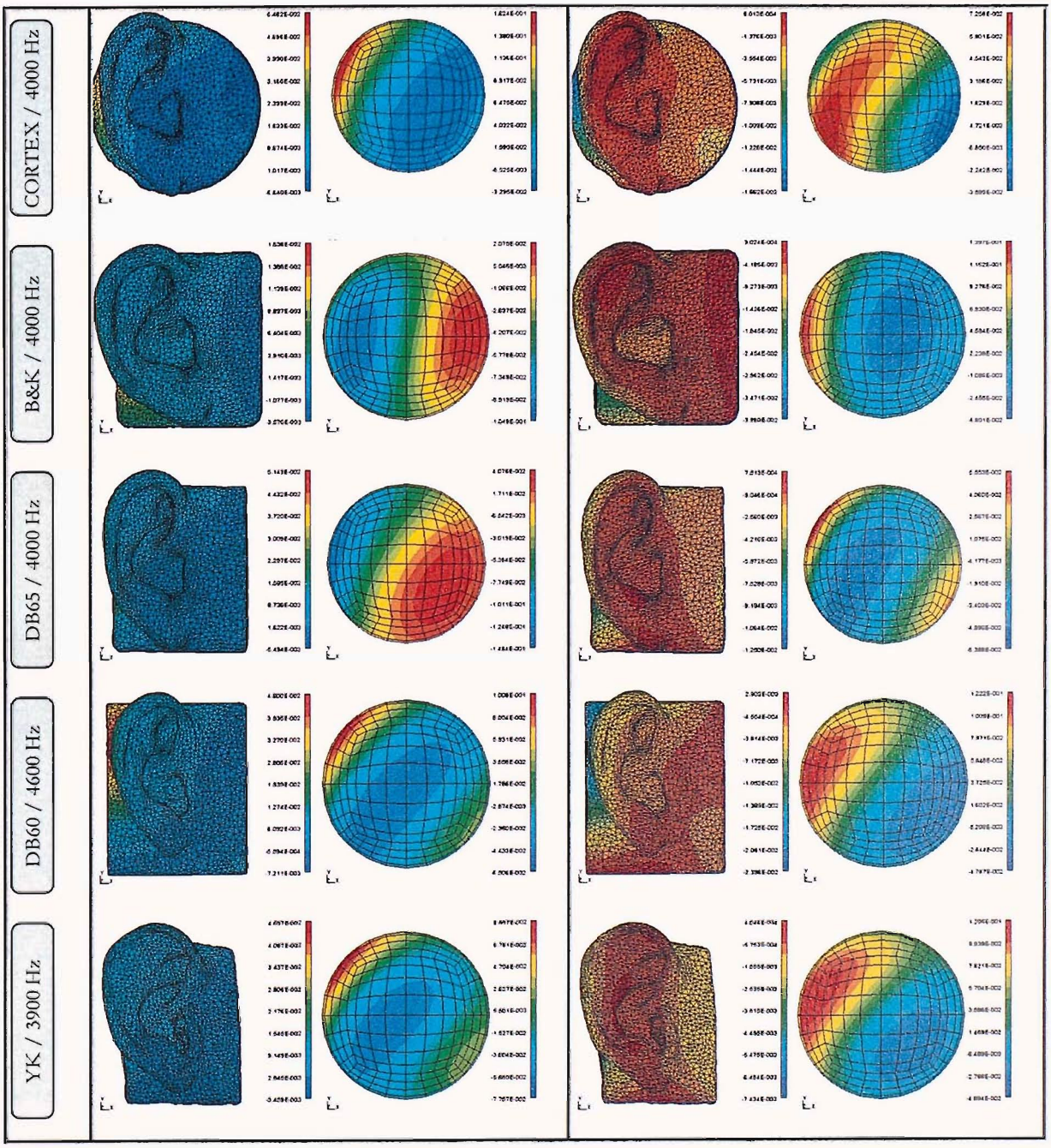


Figure A2-2: The real and imaginary left and right singular vectors associated the second singular value, σ_2 at the first peak, around 4 kHz.

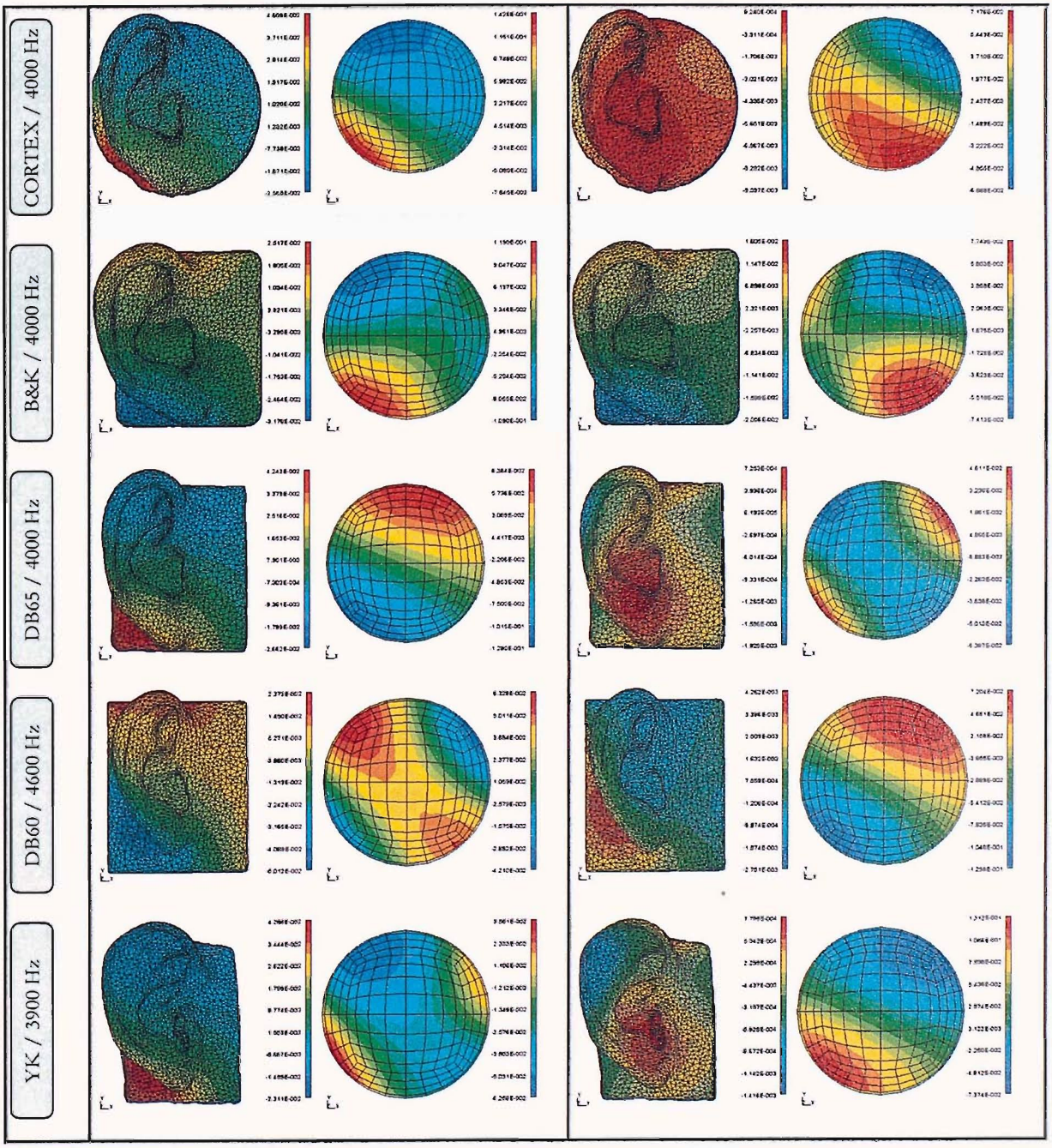


Figure A2-3: The real and imaginary left and right singular vectors associated the second singular value, σ_3 at the first peak, around 4 kHz.

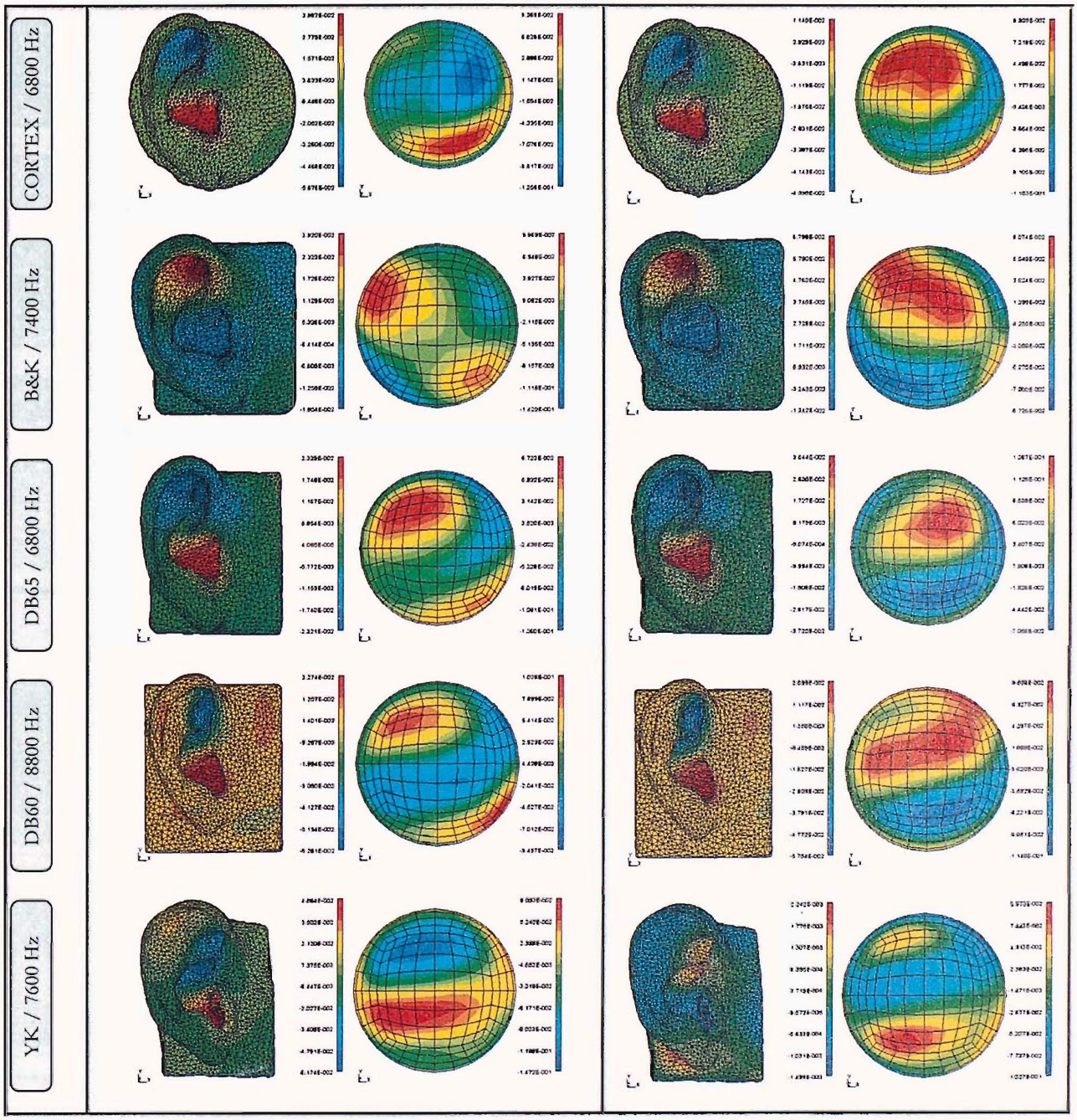


Figure A2-4: The real and imaginary left and right singular vectors associated the first singular value, σ_1 at the second peak, around 7 kHz.

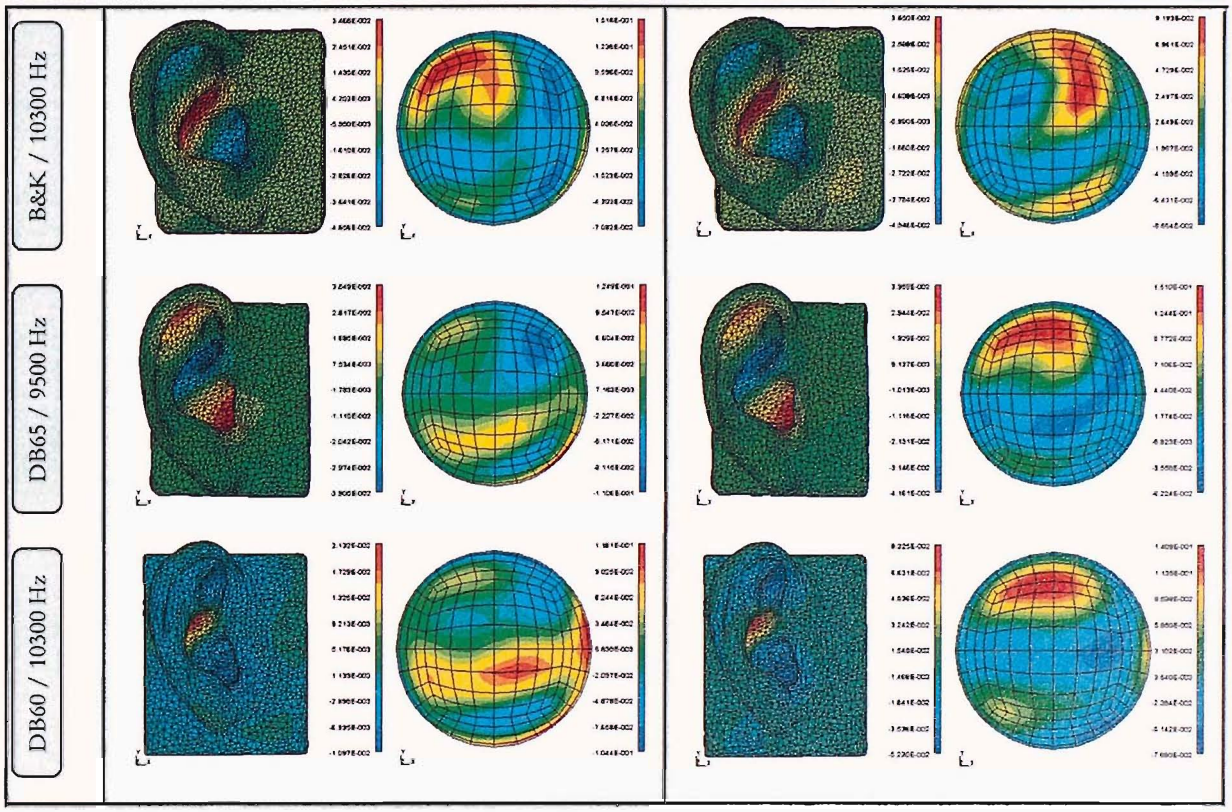


Figure A2-5: Caption as before, with σ_1, at the thrid peak around 10 kHz.

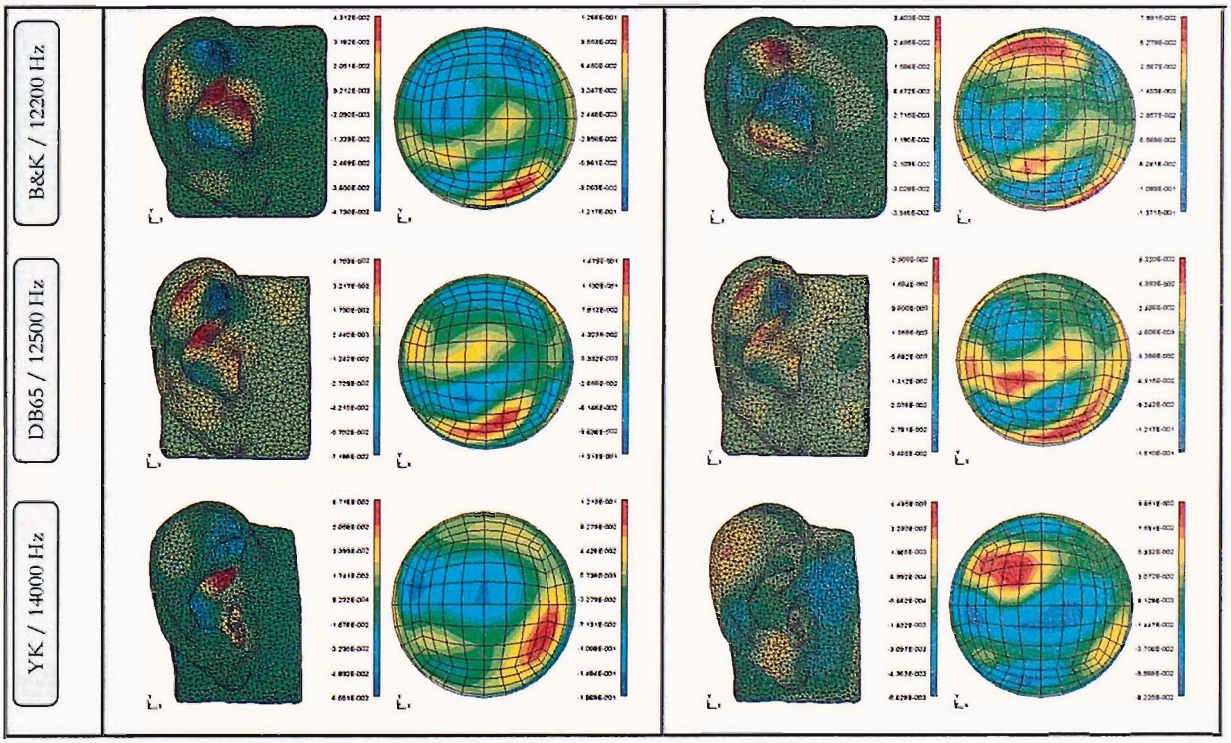


Figure A2-6: Caption as before, with σ_1, at the forth peak around 12-14 kHz.