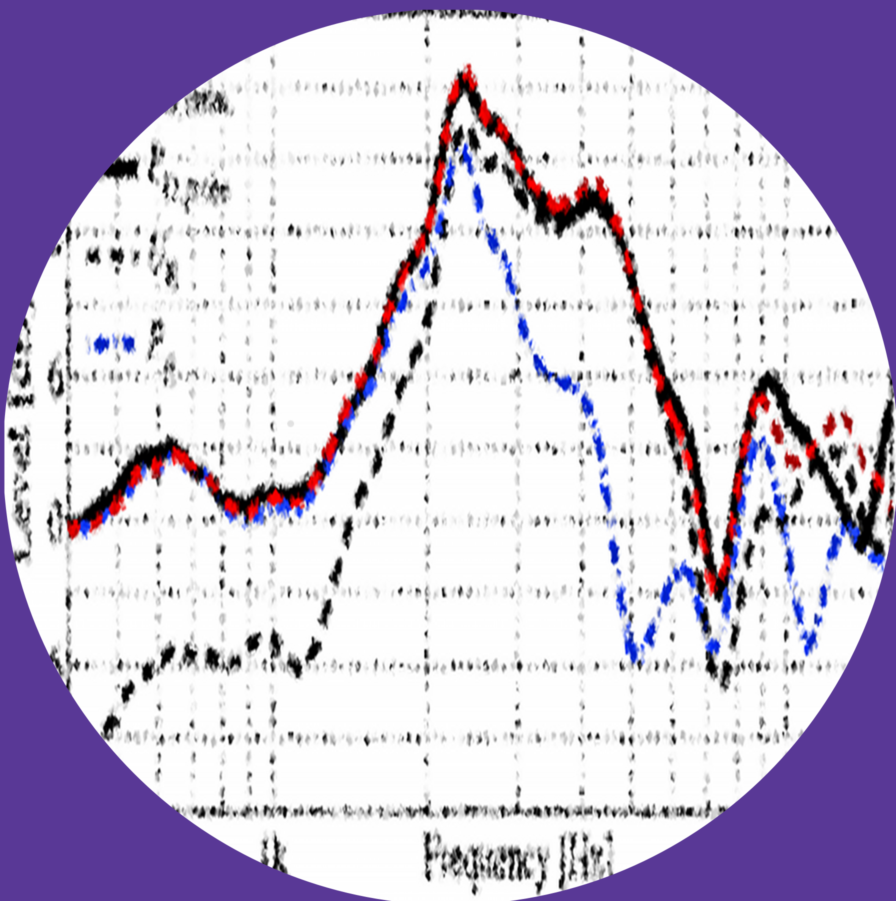
Department of Signal Processing and Acoustics

Estimating pressure at the eardrum for binaural reproduction

Marko Hiipakka



Aalto University

DOCTORAL DISSERTATIONS

Aalto University publication series
DOCTORAL DISSERTATIONS 149/2012

Estimating pressure at the eardrum for binaural reproduction

Marko Hiipakka

Doctoral dissertation for the degree of Doctor of Science in Technology to be presented with due permission of the School of Electrical Engineering for public examination and debate in Auditorium S1 at the Aalto University School of Electrical Engineering (Espoo, Finland) on the 7th of December 2012 at noon (at 12 o'clock).

Aalto University
School of Electrical Engineering
Department of Signal Processing and Acoustics
Spatial sound

Supervising professor
Professor Ville Pulkki

Thesis advisor
Professor Ville Pulkki

Preliminary examiners
Professor Dorte Hammershøi, Aalborg University, Denmark
Professor Matthias Blau, Jade Hochschule, Germany

Opponent
Professor Sascha Spors, University of Rostock, Germany

Aalto University publication series
DOCTORAL DISSERTATIONS 149/2012

ISBN 978-952-60-4864-2 (printed)
ISBN 978-952-60-4865-9 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)
http://urn.fi/URN:ISBN:978-952-60-4865-9

Unigrafia Oy
Helsinki 2012

Finland

441 697
Printed matter

Aalto University

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Abstract

Author
Marko Hiipakka

Name of the doctoral dissertation
Estimating pressure at the eardrum for binaural reproduction

Publisher School of Electrical Engineering

Unit Department of Signal Processing and Acoustics

Series Aalto University publication series DOCTORAL DISSERTATIONS 149/2012

Field of research Acoustics and Audio Signal Processing

Manuscript submitted 5 March 2012 **Date of the defence** 7 December 2012

Permission to publish granted (date) 8 October 2012 **Language** English

☐ **Monograph** ☒ **Article dissertation (summary + original articles)**

Abstract

The head-related transfer function (HRTF) characterizes the spectral transformation of sound on the path from a point sound source in free field to the eardrum. When using HRTFs in binaural reproduction, individual measurements should be favored, since non-individual HRTFs tend to cause coloration and errors in localization of virtual sound sources. HRTF measurements are commonly made at the entrances of the ear canals, which are blocked with earplugs. However, depending on the type of headphone used, this method may cause timbral coloration in reproduction, which is caused by, e.g., the difference in the acoustic coupling of the ear canal to the headphone in comparison to free air. An easy method of measuring HRTFs and headphone transfer functions (HpTFs), i.e., the pressure at the eardrum rather than at the blocked ear canal entrance would solve the aforementioned problem. In addition, HRTFs with the eardrum as point of reference can be used in binaural reproduction with insert headphones without the need to do additional estimations of the transfer function from the blocked ear canal entrance to the eardrum. Probe microphone measurements of the pressure at the eardrum are generally not considered reliable above 4 kHz, and they also involve the risk of damaging the eardrums. Hence, a non-invasive method of measuring or estimating the pressure at the eardrum is needed.

In this work, a miniature-sized acoustic pressure-velocity sensor is used to measure energy density at the ear canal entrance, which in turn is used to estimate the sound pressure at the eardrum. The reliability of the estimation method is verified through measurements with simulators and human subjects. In addition, HRTF filters are designed using the estimated HRTFs and HpTFs, and are used in a listening test. The result of the listening test shows that the estimation method presented can reduce coloration in binaural reproduction. In conclusion, the method presented facilitates the obtaining of individual HRTFs and HpTFs at the eardrum using non-invasive measurements.

An important procedure in binaural reproduction is the equalization of headphones using measured HpTFs. For insert headphones, however, neither the blocked ear canal nor the probe microphone measurement methods are applicable, since the inserts themselves block the ear canal entrance and since the transducer ports of the inserts are inside the ear canals. This study develops an alternative method of obtaining HpTFs of inserts using measurements with in-ear microphones, computational modeling, and electro-acoustic source models of the inserts. The success of this approach is verified through measurements at the eardrums of simulators and human subjects as well as through a listening test. The method facilitates individual in-situ equalization of inserts for binaural reproduction.

Keywords HRTF, pressure-velocity measurement, modeling, insert headphones.

ISBN (printed) 978-952-60-4864-2 **ISBN (pdf)** 978-952-60-4865-9

ISSN-L 1799-4934 **ISSN (printed)** 1799-4934 **ISSN (pdf)** 1799-4942

Location of publisher Espoo **Location of printing** Helsinki **Year** 2012

Pages 149 **urn** http://urn.fi/URN:ISBN:978-952-60-4865-9

Aalto-yliopisto

Aalto-yliopisto, PL 11000, 00076 Aalto www.aalto.fi

Tiivistelmä

Tekijä
Marko Hiipakka

Väitöskirjan nimi
Tärykalvoon kohdistuvan äänenpaineen estimoiminen binauraalista äänentoistoa varten

Julkaisija Sähkötekniikan korkeakoulu

Yksikkö Signaalinkäsittelyn ja akustiikan laitos

Sarja Aalto University publication series DOCTORAL DISSERTATIONS 149/2012

Tutkimusala Akustiikka ja äänenkäsittelytekniikka

Käsikirjoituksen pvm 05.03.2012 **Väitöspäivä** 07.12.2012

Julkaisuluvan myöntämispäivä 08.10.2012 **Kieli** Englanti

☐ **Monografia** ☒ **Yhdistelmäväitöskirja (yhteenveto-osa + erillisartikkelit)**

Tiivistelmä

Akustisen paineen siirtofunktio vapaassa kentässä sijaitsevasta pistemäisestä äänilähteestä tärykalvolle (engl. Head-related transfer function, HRTF) kuvaa äänen muutosta matkalla lähteestä tärykalvolle.

Parhaan laadun saavuttamiseksi binauraalisessa toistossa olisi suosittava yksilöllisesti mitattuja HRTF:iä, sillä ei-yksilöllisillä HRTF:llä on taipumus aiheuttaa värittymää äänentoistossa ja virheitä virtuaalisten äänilähteiden paikallistumisessa. Yksilölliset HRTF:t mitataan useimmiten tärykalvon sijaan suljetun korvakäytävän suulta, jolloin mikrofoni sijoitetaan esim. korvatulppaan. Tällainen mittausmenetelmä saattaa kuitenkin myös aiheuttaa värittymää äänentoistossa käytetyn kuulokkeen ominaisuuksista riippuen. Helposti toteutettavissa oleva menetelmä mitata HRTF:t ja kuulokkeiden siirtofunktiot (engl. Headphone transfer function, HpTF) tärykalvolta suljetun korvakäytävän edustan sijaan ratkaisisi edellä mainitun ongelman. Lisäksi tärykalvolta mitattuja HRTF:iä voidaan käyttää binauraalisessa äänentoistossa tulppakuulokkeilla välttäen tarpeen estimoida erikseen siirtofunktioita korvakäytävän suulta tärykalvolle. Valitettavasti mikrofonilla tehtyjä mittauksia tärykalvolta ei voi pitää luotettavina 4 kHz suuremmilla taajuuksilla ja niihin liittyy myös tärykalvojen vaurioitumisriski. Näin ollen ei-invasiivinen menetelmä mitata tai estimoida HRTF ja HpTF tärykalvolla olisi erittäin tarpeellinen.

Tässä työssä käytetään erittäin pienikokoista akustista sensoria, jolla mitataan akustista painetta ja hiukkasnopeutta avoimen korvakäytävän suulla. Mitattua painetta ja hiukkasnopeutta käytetään puolestaan tärykalvon akustisen paineen estimoimiseen. Estimointimenetelmän toimivuus todennetaan mittauksilla simulaattoreiden ja koehenkilöiden tärykalvoilta. Lisäksi estimoiduista HRTF:stä ja HpTF:stä luodaan suotimet kuuntelukoetta vasten. Kokeen tulokset osoittavat, että estimointimenetelmän avulla voidaan vähentää värittymiä binauraalisessa äänentoistossa.

Tärkeä operaatio binauraalisessa äänentoistossa on kuulokkeiden ekvalisointi, eli kalibrointi, jota varten on mitattava kuulokkeiden HpTF:t. Tulppakuulokkeiden tapauksessa HpTF:iä ei voi mitata suljettujen korvakäytävien menetelmällä eikä luotainmikrofonilla, sillä kuulokkeet itsessään sulkevat korvakäytävän. Tässä työssä on kehitetty vaihtoehtoinen menetelmä estimoida tulppakuulokkeiden HpTF:t käyttäen mittauksia korvakäytävämikrofonilla, laskennallista mallintamista ja kuulokkeiden sähköakustisia lähdemalleja. Menetelmän toimivuus on todennettu mittauksilla simulaattoreiden ja koehenkilöiden tärykalvoilta sekä kuuntelukokeen avulla. Menetelmä mahdollistaa tulppakuulokkeiden yksilöllisen kalibroinnin binauraalista äänentoistoa varten.

Avainsanat HRTF, tilavuus- ja hiukkasnopeus, mallinnus, tulppakuuloke.

ISBN (painettu) 978-952-60-4864-2 **ISBN (pdf)** 978-952-60-4865-9

ISSN-L 1799-4934 **ISSN (painettu)** 1799-4934 **ISSN (pdf)** 1799-4942

Julkaisupaikka Espoo **Painopaikka** Helsinki **Vuosi** 2012

Sivumäärä 149 **urn** http://urn.fi/URN:ISBN:978-952-60-4865-9

Preface

This thesis is the result of research carried out at the Laboratory of Acoustics and Audio Signal Processing of the Department of Signal Processing and Acoustics, School of Electrical Engineering, Aalto University (until 31.12.2009 Helsinki University of Technology) in Espoo, Finland.

I am honored and grateful to have had the opportunity to work together with my first supervisor, Professor Matti Karjalainen, who gave me the opportunity to do some interesting research at the Laboratory of Acoustics. His passing was a great loss.

I am most grateful to my instructor and supervisor Professor Ville Pulkki for his continuous interest in my work and for the positive attitude and fruitful discussions during the course of the research. Professor Pulkki's inspiring and insightful support has been of tremendous help in the work with the publications. Also, his expert advice was invaluable in the completion of this thesis.

I wish to thank Professor Paavo Alku for his support in the completion of this thesis. My thanks too to pre-examiners Professor Dorte Hammershøi and Professor Matthias Blau for offering their expertise and valuable feedback and for their encouraging and positive evaluation of this thesis.

I would also like to express my gratitude to The Academy of Finland, Nokia Foundation, Emil Aaltonen Foundation, and the Research Foundation of Helsinki University of Technology who have supported this work. The research leading to these results received funding from the European Research Council under the European Community Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no [240453].

I am grateful to my co-authors Miikka Tikander, Teemu Kinnari, Marko Takanen, Symeon Delikaris-Manias, and Archontis Politis. You did a wonderful job. The many discussions we have had with Seppo Uosukainen have definitely been of incalculable help in the development of the methods used in this thesis. Special thanks goes to all my test subjects who boldly volunteered for the sometimes daring measurements. David Mitchell, too, deserves my deepest thanks for all his efforts in proofreading. Thank you Antti Aarnisalo at the Department of Otolaryngology at Helsinki University Central Hospital for fruitful co-operation.

Huge thanks to co-workers Sami Oksanen, Ville Sivonen, Simone Spagnol, Javier Gómez Bolaños, Mikko-Ville Laitinen, Olli Santala, Jussi Rämö, Jussi Pekonen, Tapani Pihlajamäki, Olli Rummukainen, Jukka Ahonen, Teemu Koski, and Juha Vilkamo. I highly appreciate Jussi Hynninen, Lea Söderman, Heidi Koponen, Mirja Lemetyinen, Martti Rahkila, Henri Penttinen, Okko Räsänen, Tuomo Raitio, as well as Professors Vesa Välimäki, Unto Laine, and Jorma Skyttä and all the people working at the Department of Signal Processing and Acoustics, not least for providing an exhilarating and pleasant working atmosphere. I am grateful to the Doctoral Programme Committee and the personnel at the student services office for their contribution. I also acknowledge the Acoustical Society of Finland and its president, Professor Tapio Lokki and the researchers at the Department of Media Technology. I am very thankful to all the researchers at The Laboratory of Visualization and Perception at Minho University, Portugal, for the many new things I learned during my research exchange there.

I would also like to thank the president of the Driving School and his students as well as friends in Bratislawa Youghurt and the Polytech Choir for many intriguing conversations and for providing motivation.

I hereby express my most sincere thanks to Henrik and Riitta Fröjdman, and all of my friends, who have been of great support during the course of my academic career.

My gratitude goes especially to Sakari Hiipakka and Hilja Latva-aho, Tarja Hiipakka and Peter Hoffstadt, Linn and Mira Klingberg as well as Erkki Hiipakka and all the big family for being there for me all these years. And most importantly, thank you Catarina.

Guimarães, Portugal, October 8, 2012

Marko Hiipakka

Contents

List of Publications

This thesis consists of an overview and of the following peer-reviewed publications, which are referred to in the text by their Roman numerals.

I Marko Hiipakka, Miikka Tikander, and Matti Karjalainen. Modeling of external ear acoustics for insert headphone usage. *Journal of the Audio Engineering Society*, Vol. 58, No. 4, pp. 269–281, Apr 2010.

II Marko Hiipakka, Matti Karjalainen, and Ville Pulkki. Estimating pressure at eardrum with pressure-velocity measurement from ear canal entrance. In *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, (New Paltz, NY, USA), pp. 289–292, Oct 2009.

III Marko Hiipakka. Estimating pressure and volume velocity in the ear canal for insert headphones. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, (Prague, Czech Republic), pp. 289–292, May 2011.

IV Marko Hiipakka, Teemu Kinnari, and Ville Pulkki. Estimating head-related transfer functions of human subjects from pressure-velocity measurements. *Journal of the Acoustical Society of America*, Vol. 131, No. 5, pp. 4051–4061, May 2012.

V Marko Takanen, Marko Hiipakka, and Ville Pulkki. Audibility of coloration artifacts in HRTF filter designs. In *Proc. AES 45th International Conference*, (Helsinki, Finland), Mar 2012.

VI Marko Hiipakka, Marko Takanen, Symeon Delikaris-Manias, Archontis Politis, and Ville Pulkki. Localization in binaural reproduction with insert headphones. In *Proc. AES 132nd Convention*, (Budapest, Hungary), Apr 2012.

Author's Contribution

Publication I: "Modeling of external ear acoustics for insert headphone usage"

The present author assembled the measurement equipment and performed all the measurements for the study. The computational models were created with the assistance of the third author of the paper, but the final versions were perfected by the present author. Hence, all the main results were obtained and analyzed by the present author. The article was written by the present author with the assistance of the third author of the paper, especially in Section 2.2 and from the second author in Section 0.3.

Publication II: "Estimating pressure at eardrum with pressure-velocity measurement from ear canal entrance"

The present author performed the measurements and constructed the models used for estimating the pressure at the eardrum, including the new power-based estimation method presented in the paper. The present author analyzed the results as well as wrote the article with the guidance of the second and third authors of the paper.

Publication III: "Estimating pressure and volume velocity in the ear canal for insert headphones"

The present author is solely responsible for the study.

Publication IV: "Estimating head-related transfer functions of human subjects from pressure-velocity measurements"

The present author designed the new measurement equipment for the study in cooperation with Microflown Technologies. All the measurements, the computational modeling, estimations, analysis of the results, as well as the writing of the article were the responsibilities of the present author.

Publication V: "Audibility of coloration artifacts in HRTF filter designs"

The present author computed the individual frequency responses for the HRTF filter design in the study and collaborated with the first author of the article in the design of the listening test. The present author produced Figures 1–6 for the article and wrote a significant part of the text, especially in Section 2.

Publication VI: "Localization in binaural reproduction with insert headphones"

The present author collaborated with the third and fourth authors of the paper in performing the measurements, the recording of the binaural signals, and the planning of the listening test for the study. The present author also constructed the Norton models of the insert headphones and produced the individually filtered recordings for the listening test. The present author collaborated with the second author of the paper in the analysis of the listening test results. The present author produced Figures 1–6 and carried the main responsibility of the writing of the article.

List of Abbreviations

AES	Audio Engineering Society
BR	Binaural reproduction
CAD	Computer aided design
DirAC	Directional Audio Coding
FEC	Free-air equivalent coupling
FFT	Fast Fourier transform
HpTF	Headphone transfer function
HRTF	Head-related transfer function
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
ILD	Interaural level difference
IPD	Interaural phase difference
ITD	Interaural time difference
ITU	International Telecommunication Union
PDR	Pressure division ratio
PU	Pressure-velocity (pressure and velocity)
SPL	Sound pressure level

List of symbols

A	Cross-sectional area of a duct
c	The speed of sound in air
E	Sound energy density
L_{c}	Characteristic length of an electrical circuit
p	Pressure
P_{D}	HRTF and/or HpTF
P_{freq}	Pressure frequency response at the eardrum regardless of source
q	Volume velocity
u	Particle velocity
Z_0	Characteristic impedance of air
Z_{W}	Wave impedance of a wave propagating in a duct
λ	Operating wavelength of an electrical circuit
ρ	Density of air

1. Introduction

The first major break-through in audio reproduction took place in 1877 when Edison presented the phonograph, which could both record and reproduce sounds. Shortly after, in Paris in 1881, the first binaural audio system was presented by Clément Ader – using telephonic transmission the Théâtrophone enabled stereophonic perception of an opera performance for an audience at a distant location outside the opera house. Headphones and portable audio players came to general awareness in the early 1980's with the introduction of the Sony Walkman cassette player. In the 1990's, digital music reproduction developed quickly. One example is the mp3-player that replaced the short-lived portable CD-player as the most popular portable mobile audio player at around the turn of the millennium. In three-dimensional (3-D) audio reproduction an important innovation was the use of head-related transfer functions (see Section 2.4) as filters in binaural reproduction, i.e., 3-D audio reproduction over headphones (see Section 4). Today, the popularity of portable audio players is steadily growing, which has given rise to a discussion on music-induced hearing disorders. While 3-D audio reproduction is studied extensively and has gained in popularity over stereophonic listening, the vast majority of all audio recordings produced are still designed for stereophonic listening.

The goal of the research presented in this thesis is to develop new techniques that allow easy implementation of individually customized high-quality binaural reproduction. The research can be divided into two sub-classes, the first being related to measuring and estimating individually the acoustical characteristics of the external ear, which are needed to process audio signals for binaural reproduction. The commonly used methods for measuring these individual parameters are not suitable for reproduction with all kinds of headphones, which may lead to undesired coloration and localization errors in the reproduction phase. The second sub-class focuses on the reproduction over headphones and, more precisely, on fast, easy, and individual in-situ equalization of insert headphones. Such equalization – or calibration – is not feasible with known techniques, since commonly they require probe microphones or individual ear canal models to be used. With the method presented, the required measurements can be made using the headphone itself and no additional equipment is needed.

Individual measurements of acoustic pressure and particle velocity at the ear canal entrance are used to develop optimal filters for reproduction. The measurements are carried out with miniature-sized pressure-velocity (PU) sensors designed for the purposes of this study. New approaches to the equalization of headphones are applied in order to calibrate individually the headphones used in binaural reproduction. In addition to the various measurement methods and computational models used in the verification of the results, perceptual evaluation is finally used to test the applicability of the methods proposed.

One of the implications of the results of this thesis is that individually optimized 3-D audio reproduction becomes easier to implement. In addition, high-fidelity 3-D audio reproduction becomes available to a larger audience due to the possibility of using low-cost headphones in the reproduction. These kinds of new techniques could generate a strong growth in the popularity of 3-D audio amongst ordinary consumers.

This thesis consists of an introductory part and six peer-reviewed publications. Two of these publications have been published in international peer reviewed journals and four in international peer reviewed scientific conferences. The structure of the introductory part is as follows. The basic concepts related to sound and binaural hearing are presented in Section 2. Section 3 deals with measuring, modeling, and estimating sound pressure at the eardrum. Here the emphasis is to give a compact overview, whereas the new results of the topics can be found in the publications. How the pressure signals obtained at the eardrum are utilized in binaural reproduction is discussed at a general level in Section 4. Brief summaries and the most important results of the publications are presented separately in Section 5. Finally, cohesive summaries of the two main research themes and the main methods of this thesis, future research directions, and conclusions are presented in Section 6.

2. Acoustics of hearing

In this section, a brief introduction to some of the basic concepts of the physics of sound is given, the focus being on subjects relevant to this thesis. One of the main themes of this thesis is the acoustics of the external ear, which will be discussed on several occasions later on. Nonetheless, the different parts of the external ear and some of their characteristics are presented already in this section. Two essential subjects or terms the reader should be acquainted with are binaural hearing and the head-related transfer function, short explanations of which are also given here.

Understanding the basics of binaural hearing and the behavior of sound in and around the ear helps the reader to appreciate the results of this thesis. Due to the concise format used, many important details are not discussed here, but complementary reading can be found, e.g., in [1, 2, 3, 4, 5].

2.1 Sound and hearing

Sound can be defined as wave motion in a medium that causes an auditory sensation in the brain. An event that evokes audible sound pressure waves that propagate in air is determined as a sound event. The frequency range of sounds audible to humans is approximately 20 Hz – 20 kHz.

When a sound pressure wave reaches the external ears and propagates through the ear canal it undergoes spectral transformations, which are discussed in Section 2.2. The eardrum and the ossicles convert the small variation in air pressure in the ear canal to vibrations in the fluid inside the cochlea, where different frequencies create resonances at different traveling wave excitation positions on the basilar membrane [6, 7]. Hence, the cochlea operates as a time-frequency analyzer that can be modeled as a bandpass filter bank, that divides the audible frequency range into critical bands. [8, 9].

The hair cells transform the vibrations of the basilar membrane to nerve impulses in the fibers of the cochlear nerve. An auditory event emerges in the brain when these nerve impulses reach the central auditory system and the primary auditory cortex. Attributes describing the perceived auditory event are, e.g., direction and distance, which will be discussed further in Section 2.3. Based on the frequency contents of the sound entering the ear canals, the auditory system translates the frequency to a subjective measure called pitch, which is related to how high or low perceptually a sound is on a musical scale. Another quality of an auditory event is loudness, which is related to the amplitude of the sound wave entering the ears of the listener.

While the mean atmospheric pressure is approximately 105 Pa, sound pressure level (SPL) is defined as the magnitude of the pressure relative to 2×10^{-5} Pa, which is approximately only 10^{-10} standard atmosphere. Hence, audible SPLs (approximately 0 dB and louder at 1 kHz) represent extremely small variations of the static pressure. The ratio between the SPL of the sounds that bring pain to the ears and the weakest audible sounds is more than 10^{12}.

Next, some essential physical quantities related to sound are presented due to their importance in this thesis. Sound pressure p and particle velocity u are the two physical quantity measures of sound. Particle velocity is the velocity of the molecules in air, which is caused by sound wave motion. Particle velocity should not be confused with the speed of sound, i.e., the speed of propagation of a sound wave. The specific acoustic impedance of an acoustic component, such as the eardrum, is $Z = p/u$, where p and u are pressure and particle velocity at the connection point, respectively. The characteristic impedance of air is defined as $Z_0 = \rho c$, where c is the speed of sound and ρ is the density of air. Volume velocity q is related to particle velocity through the cross-sectional area A of an acoustical duct in which a sound wave propagates: $q = uA$. The wave impedance of a wave propagating in a duct is defined as $Z_\mathrm{W} = p/q$. Finally, acoustic energy density in an acoustical waveguide

$$E = \frac{1}{2}\rho|\vec{u}|^2 + \frac{|p|^2}{2\rho c^2}. \tag{2.1}$$

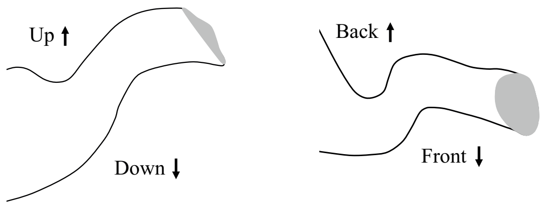


Figure 2.1. Outline of a typical (right) ear canal as seen from the front and from above. The shaded area represents the approximate location of the eardrum.

2.2 Acoustics of the external ear

From an acoustical perspective the external ear comprises the torso, the head, the pinna, the concha, and the ear canal up to the eardrum, since all these parts of the human body have at least some effect on hearing. For some directions of sound incidence and especially at frequencies above 1 kHz the pinna behaves like a reflector as it directs sound waves toward the ear canal. One of the most important acoustical parameters of the external ear is the physical dimensions of the concha that connects the pinna to the entrance of the ear canal. The pinna and the concha cause phase cancellation, or a "pinna notch", which is usually located at frequencies between 7 kHz – 10 kHz [10, 11], where the difference in the total path length between the direct sound entering the ear canal and the sound reflected from the pinna is half a wavelength.

An average human ear canal is approximately 26 mm in length and 7 mm in diameter. The canal usually has two bends close to the entrance as depicted in Fig. 2.1. The tube-like structure and small dimensions of the canal consequently makes it an acoustical waveguide for sound waves up to the highest audible frequencies. The ear canal is terminated by the eardrum, which means that the ear canal behaves acoustically as a quarter-wave resonator having its first resonance approximately at 3 kHz for adult humans.

In addition to the frequency content of the sound entering the ear canal, both the length of the ear canal and the compliance[1] of the eardrum are very important factors in the formation of the pressure frequency response at the eardrum P_{freq}, as reported in PI. The length of the ear canal determines the frequencies of the quarter-wave resonances. The compliance of the eardrum has a strong effect on the sharpness of these resonances and the quarter-wave minima found in frequency responses

[1]The inverse of mechanical stiffness and acoustical impedance.

measured along the ear canal. A rigid eardrum yields sharp resonances whereas a compliant eardrum results in a smooth P_{freq}. The frequencies of the quarter-wave minima depend not on the length of the ear canal but on the distance from the point of measurement to the eardrum.

When the ear canal is blocked with a headphone that is inserted into the ear canal (insert headphone), the quarter-wave resonances disappear and half-wave resonances emerge instead, the frequencies of which depend on the residual length of the ear canal. In the measurements for this thesis (PI, PIII, PVI) these resonances have been found at approximately 6 kHz, 12 kHz, and 18 kHz. However, the strength and sharpness of the resonances depend not only on the compliance of the eardrum but also on the acoustic impedance of the insert headphone used. Hence, the acoustics of an ear canal blocked with an insert headphone should always be studied together with the acoustic characteristics of the headphone.

The eardrum, or tympanic membrane, is a thin membrane that transmits sound from the ear canal to the ossicles. The height of the eardrum is approximately 10 mm and the width 8 mm. There is no definite border between the eardrum and the ear canal wall; instead, the eardrum can be considered as being a continuation of the ear canal wall that forms the innermost part of the ear canal. Often the eardrum curves inward by a few millimeters. The handle of the malleus ossicle, the Umbo, forms a small bulge outward at the center of the eardrum [12]. According to Fay *et al.* [13], the complicated structure of the eardrum serves to transfer more energy to the ossicles as compared to a flat structure, especially at high frequencies. The angle between the bottom part of the ear canal wall and the eardrum is on average approximately 40°, as depicted in Fig. 2.1. Hence, the area of the eardrum is larger than the cross-sectional area of the ear canal, which can increase sound transmission to the ossicles. In addition, its asymmetric shape creates multiple resonances at high frequencies and smoothens the power flow to the ossicles [13].

There are large differences between the impedances of human ear canals [14], and the topic has been studied extensively, e.g., for hearing aid fitting purposes [15, 16, 17]. The large interindividual variability of the impedance of the ear canal is an essential topic in this thesis. Acoustic impedance measurements of the human ear canal are useful not only with respect to headphones and hearing aids but also in the recognition of possible dysfunctions and in the studying of the function of the middle and the inner ear [18].

2.3 Binaural hearing

The ability of humans to perceive the surrounding soundscape in three dimensions is based on binaural hearing, i.e., using two ears for listening. Direct sounds, early reflections and reverberation of the space provide the auditory system with information on the direction, distance, and loudness of sound sources. The binaural and monaural cues used to localize sound sources are based on the magnitudes and phases of P_{freq} at both eardrums. The interaural differences in magnitude and phase of P_{freq} are interpreted by the auditory system to formulate the binaural cues while the monaural cues are based on the instantaneous magnitudes of P_{freq} [3].

The binaural cues are commonly categorized to interaural time difference (ITD) and interaural level difference (ILD) [19]. The ITD is the difference in the amount of time it takes for a sound signal to reach the eardrums. This difference is sometimes referred to as interaural phase difference (IPD) – especially when periodic signals are discussed – in which case the IPD is determined as the difference in the phases of the two signals at the eardrums [20]. The ILD, i.e., the level difference between the signals at the two eardrums, is very much dependent on frequency. Hence, it describes the difference between the magnitudes of P_{freq} at the two eardrums. With a plane wave the ILD is caused by the shadowing effect of the head and has typically larger values at high frequencies and values close to zero at the lowest audible frequencies. The importance and dominance of the ITD and ILD cues are affected by the familiarity of the sound source, the frequency content of the source, the plausibility of the cues, as well as the coherence of the cues across different frequencies [21]. However, ITD is usually considered the dominant cue at low frequencies, whereas ILD is considered more salient at high frequencies.

The magnitude of P_{freq} provides the auditory system with information on the elevation angle of sound incidence and helps to differentiate auditory events in front from those behind the head. If there is no interaural difference in the phase or the magnitude of P_{freq}, the auditory event will then appear in the median plane or close to it. In such situations, the magnitude response only is used in solving for the elevation angle of sound incidence [3]. In this thesis, the magnitude frequency responses at the human eardrum are discussed with respect to elevation perception and front-back confusion in binaural reproduction (PIV, PV, and PVI).

2.4 Head-related transfer function

The head-related impulse response (HRIR) describes how a sound wave excited by a point sound source in free field is filtered by the reflections and diffractions of the torso, head, and the external ear before reaching the eardrum. The Fourier transform of the HRIR is defined as the head-related transfer function (HRTF). The HRTF can also be characterized as the complex ratio between P_{freq} in free-field conditions and the free-field pressure frequency response in the same position (without the subject), when both are evoked by the same point sound source.

One problem with the definition of the HRTF is that the pressure over the surface of the eardrum is not constant at high frequencies [13]. Hence, the expression “pressure at the eardrum” is ambiguous. From a purely hearing perspective knowing the frequency response of the signal that is transmitted by the eardrum to the ossicles and further to the auditory system is important – this question, however, falls beyond the scope of this thesis.

In free-field conditions with one point sound source, the auditory system formulates the binaural and monaural cues based on the HRTFs. Hence, in such circumstances the ILD cues and the ITD cues are generated based on the magnitude responses and the phase responses of the HRTFs of both ears, respectively. In a reverberant environment, however, not only the HRTFs but also the acoustics of the space contributes to the formation of the ITD and ILD cues.

In a study by Fels *et al.* [22] the effects of various anthropometric parameters (excluding the ear canal) on the individual HRTFs are reported. The depth and the breadth of the concha, the distance between the ear and the shoulder, and the breadth of the head were shown to be important parameters with regard to differences between individual HRTFs. The height of the head and the pinna were found to have only a very slight influence on the HRTFs, which is congruent with the findings presented in PI, Section 3.1.3. The usage of HRTFs in binaural reproduction is discussed in Section 4.

3. Obtaining pressure signals at the eardrum

For successful implementation of binaural reproduction (BR), all that is needed are a binaural recording or HRTFs as well as the headphone transfer function (HpTF) of the headphone used in the reproduction. The point of reference of the recording, the HRTFs, and the HpTFs can be the eardrum, the blocked ear canal entrance, or some other point in the ear canal. Different methods of obtaining these pressure signals are presented in this section, the focus being on obtaining HRTFs and HpTFs at the eardrum. The practical steps in the reproduction are discussed in Section 4.

If individual measurements are not feasible, HRTF and HpTF filter banks [23, 11] can also be used in BR, but perceptually authentic reproduction is best achieved with individual HRTFs and HpTFs [24, 25, 26, 27]. These non-individual HRTFs and HpTFs can be measured using human subjects or dummy heads, or they can be constructed using computational modeling or parametric HRTF models [28].

In the reminder of this thesis, both the free-field HRTFs at the eardrum and the HpTFs at the eardrum are denoted, when possible, as P_{D}. The difference between P_{freq}, which was presented earlier, and P_{D} is that the responses of the source(s) and the space are included P_{freq}, but not in P_{D}.

3.1 Measurements and recordings

Because of the large interindividual variety in the acoustical features of the external ear, individual HRTFs, HpTFs, and binaural recordings should be preferred to HRTF databases or, e.g., dummy head recordings in BR. The differences in the magnitudes of HRTFs and HpTFs between different individuals can be as high as 20 dB at high frequencies. Using non-individual recordings or HRTFs may result in localization errors regarding the direction and distance of the sound sources [24]. It has also been suggested that improved quality in BR can be achieved with "optimal", but non-individual HRTFs, at least in terms of localization accuracy [29]. No conclusive results supporting this theory have been presented, however.

3.1.1 HRTFs and HpTFs

In practice, individual free-field HRTFs are obtained by measuring[1] P_{freq} and dividing it with the free-field reference that includes the loudspeaker and the microphone frequency responses only. HpTFs are obtained in a similar manner, but only the response of the microphone is removed from the measurement result. Regardless of the choice of HRTF and HpTF measurement technique, the HpTFs need to be measured using the same point of measurement and the same measurement setup used in the HRTF measurements. Failure to do so leads to spectral coloration in BR.

The problems included in measuring P_{D} with a probe tube microphone [33] or a probe microphone has long been recognized [34, 35]. If the probe measurement is made at a point further away than approximately 3 mm from the innermost part of the ear canal, i.e., the lower part of the eardrum, the reflection from the eardrum starts to attenuate the measured response at audible frequencies as discussed in PI and PIV. According to Stinson *et al.* [36], even pressure minima can exist on the surface of the eardrum. In addition, the introduction of a probe tube microphone at the eardrum can disturb the sound fields close to the eardrum and thereby cause measurement error [37].

Consequently, probe microphone or probe tube microphone measurements are not generally regarded as reliable above approximately 4 kHz. However, if the distance between the microphone and the innermost part of the ear canal can be estimated, the reflection from the eardrum can be compensated to some extent as presented in PIV.

3.1.2 Binaural recording

The simplest method for obtaining binaural audio for reproduction is to do the recording with microphones positioned at the ears of a real person or an acoustic manikin (dummy head). The differences in the spectral

[1] Frequency response measurements are most commonly performed using the swept-sine technique [30, 31, 32].

structure of individual binaural recordings are large when the recording is made at the eardrums or some other positions in the open ear canals [24, 38]. Hence, if a binaural recording is intended to be used by more than one individual, the blocked ear canal recording method should be preferred. In addition, due to the problems related to probe microphone positioning and the quality of such microphones, a high-fidelity individual binaural recording is probably easiest to obtain by using the blocked ear canal method.

3.2 Physical simulation

It is convenient to use dummy heads and ear canal simulators as replacements for human test subjects for many reasons. Measuring P_{D} of live subjects is difficult and risky, and the result of the measurement is unreliable due to positioning problems. Physical (tangible) simulators can never mimic the human external ear precisely, but they are useful in the design of headphones and hearing aids.

3.2.1 Ear simulators

Ear simulators designed to mimic the acoustical impedance of real ears were originally meant for calibration of hearing aids and audiometric headphones. Simulators, or couplers, that only provide simple volumetric coupling are normally small volumes with a reasonable load for hearing aid or insert headphone measurements. The pressure frequency response evoked by a headphone at the eardrum position of the simulator is usually not what it would be at a human eardrum. For example, ear simulator measurements with insert headphones often exaggerate low-frequency levels because the insert headphone can be coupled to the simulator tighter than to human ears, which excludes the effect of leakage.

Since the compliance of the ear canal walls have only a minor effect on P_{D} [18], a straight rigid-wall tube, i.e., an acoustical waveguide, can be used to simulate the ear canal with reasonable accuracy. Instead of using a rigid termination, an eardrum model the impedance of which is similar to that of the human eardrum can be used, which improves the simulator significantly. Most of the ear simulator measurements for this thesis were performed with a purposely-built ear canal simulator, the “adjustable ear canal simulator” (Adecs) [39]. The device has rigid canal walls and both damped and rigid ‘eardrums’. Ear simulators are described in more detail in the ANSI S3.25 [40] and IEC60711 [41] standards as well as in PI.

3.2.2 Acoustical dummy heads

Acoustic manikins (artificial heads, dummy heads) have been used, e.g., for binaural recordings for at least 70 years [42, 43]. According to Firestone [20], an acoustic dummy head was used as a recording device for the first time at the end of the 1930’s.

The HRTFs of acoustical dummy heads are commonly used in BR. One of the greatest advantages with using dummy head HRTFs is that the measurement grid can be made extremely dense, since there is basically no upper limit to the duration of the measurements. A dummy head built as a copy of an individual human can be almost as good for individual binaural recordings as the original human head in terms of the spectral content of the recording [44, 45].

Many of the reports on the effects of variable headphone positioning on P_{D} and headphone characteristics in general are based, at least partially, on measurements with standard dummy heads [26, 46]. The applicability of these measurements to human ears is not self-evident as the ear canal of a dummy head rarely mimics exactly the ear canal of the average or individual human subject. For example, pressure division ratio values (see Section 4.3.1) of headphones are most likely very much different with dummy heads as compared to humans.

An acoustic dummy head was built specifically for some of the measurements presented in this thesis. The “dummy head with adjustable ear canals and interchangeable pinnae” (Dadec) is presented in PI. Thorough and enlightening reports on dummy heads and binaural recordings have been presented by Vorländer [47] and Paul [42].

3.3 Computational modeling

Physics-based computational models are often used in parallel with actual acoustical measurements. There are many benefits involved in having an accurate computational model of the external ear that can be used together with a tangible simulator or with measurements from human test subjects. A computational model can save time and money as the behavior of an acoustic system can be first investigated computationally before conducting laborious measurements.

The distributed-element-based modeling methods are used widely in the field of acoustics. These methods include, amongst others, the finite element method (FEM), the boundary element method (BEM), and the finite difference time-domain modeling (FDTD). The methods can be used to estimate HRTFs using 3-D models of the head and the pinna and to study the characteristics of HRTFs in detail, or to visualize the diffractions and reflections of the wavefronts on the pinna surface [10]. One additional advantage of these methods is that it is possible to change the shape of the external ear or the head in ways that are not possible with real subjects and to study the effects of the deformations analytically.

3.3.1 Lumped elements

One straightforward method of modeling an acoustical system is to transform the system under investigation to a lumped electrical circuit. Physical components with acoustical properties can be modeled as their electrical analogies such as capacitors, resistors, and inductors or simple combinations of these. Lumped element models are considered valid whenever $L_c << \lambda$, where L_c is the characteristic length of the circuit, and λ is the operating wavelength of the circuit. For example, Bravo *et al.* [48], using a lumped parameter approach, modeled the ear simulator specified in IEC 60318-1 [49], which has been designed to mimic normal human ears in terms of the acoustic load, or acoustic impedance, of the human ear when coupled with a headphone. They showed that the lumped element model becomes less valid as the acoustic wavelength approaches the characteristic dimensions of the simulator, which supports the principle that the external ear should be modeled as a distributed rather than a lumped system for improved accuracy especially at high frequencies.

3.3.2 Ear canal simulator as waveguide

In physics-based modeling of the ear canal it is important to take into account that its size is of the same magnitude as the sound wavelengths of interest and that its geometry can be complex. A sophisticated individualized distributed element model of the ear canal, however, can be used for a large frequency range with good accuracy.

The sound pressure and volume velocity at different points inside a tube with constant cross-sectional area can be estimated using transmission line modeling, where the transmission line is an acoustic waveguide [50, 51]. A waveguide is a duct with a small diameter compared to the wavelength of interest, e.g., an ear canal simulator with constant cross-sectional area A. A computational model of such an ear canal simulator is composed of a termination impedance, a lossless acoustic transmission line, and an external pressure sound source with an internal acoustic impedance. In the case of an unblocked simulator the source impedance is equivalent to the radiation impedance of the simulator's canal entrance due to the reciprocity principle. When an insert headphone is used as the sound source in the model the source parameters need to be estimated separately as presented in PI and PIII.

An ear canal or an ear canal simulator with variable A can also be modeled using the transmission line approach, if the model is constructed of several consecutive short waveguides with different A. In order to model an individual ear canal using the transmission line approach the area function of the ear canal and the impedance of the eardrum need to be estimated first as discussed below.

In this thesis, the transmission line equations have been used in PI and PII to model an ear canal simulator having constant cross-sectional area and a rigid termination. In PIV the transmission line equations were applied only to probe measurements made with a probe microphone very close to the eardrum, namely, to compensate for the attenuation caused by the (small) distance between the probe microphone and the eardrum at frequencies above approximately 10 kHz.

3.3.3 3-D imaging

Individual 3-D models of the ears, the heads, and the torsi of human subjects can improve significantly the accuracy of computational models as compared to non-individual models. One method that has been used for 3-D imaging of external ears is X-ray computed tomography (CT) [18]. The usage of CT scans for research purposes in living humans, however, has been limited due to the high exposure to radiation [52, 53].

Laser scanning [10], on the other hand, is an accurate method of obtaining a 3-D model of the pinna, head and torso, and it involves practically no risk to the health of the test subject. The greatest disadvantage with laser scanning is that it is impossible to use it directly for ear canal and eardrum imaging. In addition, the rendering of narrow regions, such as the cavities in the pinna and the volume between the pinna and the head, may result in poor resolution unless optimal triangulation techniques are used [54]. The ear canal and the eardrum can be added to the laser-scanned 3-D model of the head and torso if an alternative imaging method is applied to the ear canal. The 3-D model of the replica of the human ear presented in PIV, Section IIIB was constructed by laser scanning a silicone casting of the pinna and the mold of an ear canal and combining these into a single 3-D model using CAD software.

Magnetic resonance imaging (MRI) has also been used for imaging of the head, torso, and external ear for modeling purposes. The greatest advantage with MRI is that a model of the complete ear canal can be obtained, assuming the proper techniques are used. One of the future directions of the research of the present author is to use complete 3-D models that include the ear canal and the eardrum for modeling HRTFs with the BEM modeling technique.

3-D images of the pinna can also be constructed from photographs or from video sequences using shape from shading [55], but constructing a high-resolution model from two-dimensional images is far from trivial.

3.3.4 Finite element method

In FEM modeling, the acoustical system to be solved is divided into small elements each of which describes a simple mechanical system such as a mass and spring. These physically simple elements are linked together and the conditions at their boundaries are solved, which results in a complete FEM model of the system. FEM models have commonly been used to solve mechanical and aeronautical problems, but nowadays also in acoustics, especially to solve problems related to low frequencies. FEM requires a large amount of computational power since the boundary conditions between elements are difficult to solve and due to the large amount of elements needed.

FEM was used by Fay *et al.* [13] in their computer simulation of the ear canal, eardrum, and ossicles for the analysis of the significance of anatomical features of the human eardrum. In their model they used not only finite elements but also rigid bodies and asymptotic methods to ensure accuracy over a frequency range of 200 Hz to 20 kHz. Other successful studies using FEM include models of the human ear canal and middle ear cavity as presented by Gan *et al.* [56], the IEC 60318-1 artificial ear by Bravo *et al.* [48], and the Brüel&Kjær Type 4157 occluded ear simulator by Jonsson *et al.* [57]. In addition, Schmidt *et al.* [17] have recently studied the accuracy of measurements of the acoustic impedance of the ear canal by using a FEM model consisting of the ear canal and the coupling tubes used in the actual measurements.

3.3.5 Boundary element method

In BEM modeling, a computational mesh of the surfaces of the target model is required only, i.e., free air does not need to be modeled. This makes BEM often more efficient than other methods, including finite elements, in terms of computational resources for problems where there is a small surface-to-volume ratio.

BEM has been used for modeling of HRTF by several scientists. In 2007 Huttunen *et al.* [59] made a simulation of the acoustics of a head-and-torso model for frequencies ranging from 20 Hz to 20 kHz. The simulation tool was based on the parallelized ultra-weak variational formulation method and the unbounded physical problem was truncated with perfectly matched layers. In a later study they used a model of the head and torso of the present author, which included the ear canals all the way to the eardrums. The head, shoulders, and the pinnae were laser scanned and the ear canals were added to the model after laser scanning molds of the canals. The computational mesh used consisted of 24,000 elements and 12,000 nodes, which was dense enough for modeling up to 6 kHz. Fig. 3.1 depicts the computational mesh as well as the computed pressure distribution on the surface of the model at 3 kHz and 6 kHz when a plane wave approaching from the right side was used as sound source.

Katz *et al.* have also presented interesting computations of individual HRTFs obtained with BEM [60, 61]. New fast multipole BEM methods and the applying of the reciprocity principle by placing a source in the ear

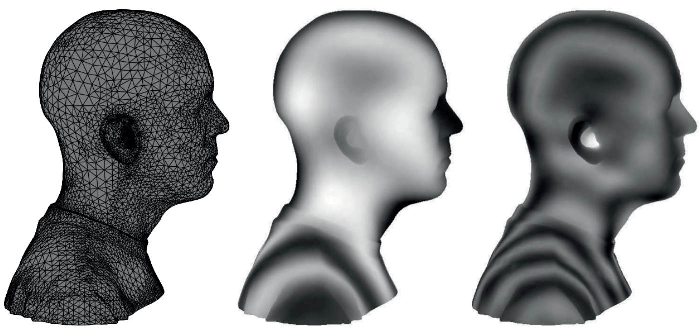


Figure 3.1. From left to right, the BEM mesh on the surface of a model of the head and torso of the present author; pressure magnitude distribution at 3 kHz and 6 kHz. Reprinted with permission [58].

and computing the sound field outside of the head are among the recent directions of development in head-related BEM technology [54, 62].

3.4 Estimation of pressure at the eardrum

The new method of estimating P_{D} presented in this thesis is based on measurements of acoustic pressure and particle velocity at the ear canal entrance. Other methods to achieve the same goal have been presented, but none of these have utilized direct measurements of particle velocity. In most of the solutions presented, however, estimates of the volume velocity or impedance at the ear canal entrance have been used to solve P_{D}. The motivation of an accurate estimation method is in the unreliability and risks related to probe microphone measurements at the eardrum, as described in Section 3.1.1.

One common approach for estimating P_{D} is to first solve the individual ear canal parameters, i.e., the area function [63, 53] and the length [35] of the ear canal and the impedance of the eardrum. The pressure at the eardrum can then be solved with physics-based computational modeling as presented in [51]. Hudde has presented a variety of methods to solve the ear canal parameters, such as the estimation of the area function by sound pressure measurements [64], and measurement of the eardrum impedance of human ears [65] using, for instance, an acoustic measuring head [66]. Voss *et al.* [14] as well as Keefe *et al.* [67, 16] have also published important results on eardrum and ear canal impedance and reflectance measurements, as have Ćirić *et al.* [15], who used a four-microphone impedance tube coupled to the human ear canal to measure the impedances the of ear canals. In addition, sound intensity and the forward wave component of the sound pressure wave have been used to estimate P_{D} [68]. These methods include separating the incident acoustic intensity in the ear canal from the reflected intensity [69], or using forward pressure level to avoid the effect of standing waves [70].

According to Hudde *et al.* [51], direct measurements of both the pressure and the volume velocity at the ear canal entrance can be used to solve the chain matrices (PI, Eq. 1; PIV, Eq. 3) of individual ear canals and thereby also P_{D}. Particle velocity measurements at the ear canal entrance could be applied to other methods to solve P_{D} than the energy-based method presented in this thesis. For example, applying pressure-velocity measurements to the pressure-minima or the reflectance phase methods [51] is an interesting task for the future.

The acoustics of the occluded ear canal has been studied mostly from a hearing aid point of view [71, 72], albeit not all hearing aids occlude the ear canal completely. The goal there is to calibrate the hearing aid individually to achieve best possible correction for the hearing impaired. In the design of insert headphones the goal is often to produce the best possible overall perceptual quality of sound. Hence, knowing the acoustic behavior of the occluded ear canal is very beneficial in the design of inserts as well as in binaural reproduction with inserts. To mention a few examples of studies by other scientist in the field, Voss *et al.* [73] published results related to the ear-canal sound pressures generated by earphones and Sankowsky-Rothe *et al.* [74] used a joint model of the source and the individual ear to predict P_{D} in occluded human ears with an accuracy of approximately ± 5 dB in the frequency range of 100 Hz – 10 kHz.

A detailed presentation of the energy-based estimation method used in this thesis can be found in PIV, Section IIB. The method is based on the assumption that the total acoustic energy density, which is the sum of the kinetic energy density and potential energy density, is preserved in an acoustic waveguide with constant A.

$$D = D_{\mathrm{k}} + D_{\mathrm{p}} = \frac{1}{2}\rho|\vec{u}|^2 + \frac{|p|^2}{2\rho c^2} \tag{3.1}$$

Because of the relatively high impedance of the eardrum the particle velocity is very small at close proximity to the eardrum, hence,

$$\frac{1}{2}\rho|\vec{u}|^2 << \frac{|p|^2}{2\rho c^2}, \tag{3.2}$$

which means that most of the energy density is concentrated on the potential energy component. If pressure and particle velocity is measured at the ear canal entrance with a pressure-velocity sensor [75], the magnitude of the pressure can be estimated using Eq. (6) or (7) in PIV. The applicability of the method to human ears with variable cross-sectional areas was studied in PIV by comparing the estimation results to measurements from the eardrum with a probe microphone.

Because the estimation method yields only the magnitude of the pressure, the phases of the velocity and the pressure at the ear canal entrance can be disregarded. This makes the method more robust than the transmission line equations (Section 3.3.2), e.g., when the estimated source parameters of an insert headphone are used to compute the velocity at the ear canal entrance as in PIII.

4. Binaural reproduction

The process of reproducing a binaural recording or simulated 3-D audio through headphones is called binaural reproduction (BR). Hence, the term binaural audio may be defined as being two-channel 3-D audio intended for reproduction over headphones. However, binaural audio may be reproduced over loudspeakers as well, but in most cases it is self-evident that headphones are used when speaking of BR. The binaural sound to be reproduced is obtained by recording or through binaural synthesis. Binaural recordings are usually made using acoustic dummy heads that have microphones in both ears or with microphones attached to the ears of an actual human listener. In binaural synthesis a multi-channel audio signal is processed so as to mimic a two-channel binaural recording.

4.1 Multichannel vs. binaural reproduction

There are several techniques for the creation of sound fields such that the listener experiences being in a real 3-D sound environment. The goal with spatial sound reproduction is most commonly to give the listener the impression of being in the space and position where the recording of the original sound was made. A straightforward approach to 3-D audio reproduction is to use multiple loudspeakers and pan the audio signals in to the desired direction. A well-known example of multichannel reproduction with multiple loudspeakers is the 5.1 surround sound system that can be found in many home theaters [76]. The number of loudspeakers used in multichannel reproduction can be anything from two upward. Therefore, techniques, such as Ambisonics [77, 78, 79] and Directional Audio Coding (DirAC) [80, 81] that allow reproduction of multichannel audio with an arbitrary number of loudspeakers have been developed.

Synthesizing a perceptually realistic 3-D soundscape using multichannel audio and multiple loudspeakers is technically easier compared to a single pair of speakers or a pair of headphones. One of the most important goals of this research is to develop techniques that enable BR with the same level of quality that is achieved with multiple loudspeaker systems.

4.2 Binaural synthesis

Binaural synthesis can be defined as the process where the audio to be reproduced is processed in such a way that a 3-D soundscape is generated when the sound is played back over headphones. If the binaural audio signals are designed to be played back over loudspeakers, a more descriptive term would be transaural synthesis [82, 83].

In binaural synthesis, the most straightforward technique to produce a perceptually plausible virtual 3-D soundscape is to filter a multichannel sound signal with appropriate HRTFs according to the nominal angle of sound incidence of each channel. The filtered channels are then rendered to two stereo channels for reproduction.

Recently Laitinen *et al.*[84], presented the technique to implement BR for DirAC. In their method, a plausible spatial impression was achieved by simulating virtual loudspeakers using HRTFs and head tracking.

HRTF filters for BR can be obtained through computational modeling or real measurements with a dummy head or a human test subject as described in Section 3. With measured HRTFs the alternatives are to use blocked ear canal HRTFs or HRTFs measured at the eardrum. Depending on the quality of the HRTF measurement, smoothing of the magnitude response might be required before filter design [85].

A method commonly used in the design of HRTF filters is to decompose the HRTF into a minimum-phase representation using the cepstrum analysis and then model the phase, or the ITD, as a constant delay [86, 87]. One method of estimating the ITD is to calculate the delay corresponding to the maximum of the cross-correlation of the original left and right ear HRIRs [88]. Frequency sampling finite impulse response (FIR) filter design is one applicable method of constructing filters from HRTF measurements [86, 89]. The order of the filter is determined by windowing the measured HRIR with a desired window.

A comparison of FIR and infinite impulse response (IIR) filter design methods was presented by Sandvad *et al.* [90]. The FIR filters were generated directly from measured HRIRs and the IIR filters using a modified Yule-Walker algorithm. Listening tests showed that an FIR of order 72, equivalent to a 1.5 ms impulse response, was capable of retaining all of the desired localization information, whereas an IIR filter of order 48 was needed for the same localization accuracy. Small order filters can reduce computational load and thus be useful in binaural reproduction with mobile audio devices [88].

If the angle between adjacent HRTFs measured is large and therefore a small number of HRTF filters are used in binaural synthesis, smooth and continuous movement of sound sources cannot be achieved. Interpolation of HRTF filters for obtaining new filters at intermediate directions is then necessary [87]. Most of the previously reported methods use the magnitude responses of minimum-phase reconstructed HRTFs in the interpolation after which a new HRTF filter is designed from the magnitude response obtained [91]. Linear [92], spline [93], or triangular interpolation can be used to obtain an interpolated magnitude response between the magnitude responses of two or several HRTFs. The perceptual consequences of different HRTF interpolation methods in 3-D sound reproduction have been studied by Wenzel *et al.* [92], Huopaniemi *et al.* [94], and Runkle *et al.* [95].

4.3 Headphones in binaural reproduction

Having recorded or synthesized high-quality individual binaural audio for reproduction does not guarantee the success of the reproduction. There are several technical details to be accounted for in the actual reproduction phase, namely, proper equalization of the headphones, the effects of headphone positioning, and the overall characteristics of the headphones used.

4.3.1 Headphone characteristics

The choice of headphone type is an essential task in BR since the quality of the reproduction is affected by the acoustical coupling of the headphone to the ear canal. A method for obtaining HpTFs for one headphone type might not be applicable to others. Circumaural headphones, or full size headphones, encompass the ears completely whereas supra-aural headphones only cover the pinnae [96]. Circumaural and supra-aural headphones may be designed with closed or open backs of the earcups.

Open-back circumaural headphones are often preferred in BR using blocked ear canal HRTFs, since they can have almost free-air equivalent coupling to the ear (FEC). With FEC-type headphones, the pressure division of the Thévenin pressure at the open ear canal entrance is not affected by the headphone, i.e., the pressure division ratio (PDR) is close to one [97, 98]. According to the results published by Møller *et al.* [98], headphones that are completely free from the ear are likely to have a PDR close to unity in a large frequency range. The open-back circumaural headphones tested could also be considered to have FEC properties within a reasonable tolerance. However, headphones that are mounted closer or inside the ear cannot be considered having FEC properties.

A major part of this thesis (PI, PIII, PVI) concentrates on reproduction with headphones that are inserted to the ear canal and thus seal the ear canal completely. This headphone type is called “insert headphones” (inserts), “in-ear headphones”, “in-ear phones” [99], and “in-ear monitors”. In reproduction with inserts, the occlusion of the ear canal results in a good bass response due to the pressure chamber effect [96, 99] and it allows low sound levels to be used because of increased background noise attenuation. As a result, the sound quality of inserts can be relatively good even in the more affordable models.

4.3.2 Equalization of headphones

In order to achieve high perceptual plausibility, the headphones used in BR must be calibrated, i.e., equalized individually [26, 100, 101]. With appropriate individual HpTFs available, the headphones are equalized using the HpTFs as filters to produce a flat frequency response at the eardrum. The filtering process is similar to the HRTF filtering described in Section 4.2. In other words, the binaural signal is filtered with the HpTF filters. Consequently, an audio signal having a flat spectrum will produce a flat frequency response at the eardrum after the HpTF filtering and playback through the headphones in question.

If blocked ear canal HRTFs and HpTFs are used, the ear canal responses are not included, but their filtering effects are naturally added in sound reproduction through headphones, assuming the headphones used have FEC as described above. Hence, with blocked HRTFs and HpTFs the headphones are equalized with the blocked ear canal entrance as the point of reference instead of the eardrum. If the headphones used in BR do not have completely FEC and blocked ear canal HRTFs are used, measured

PDR of the headphone can be used to compensate for the errors caused by the headphone's coupling to the ear [98, 102].

4.3.3 Effect of headphone positioning

Especially at frequencies above 4 kHz, the HpTFs are very much dependent on the positioning of the headphone and even a small displacement of the headphone after equalization causes errors in the frequency response at the eardrum [100, 46]. According to Wightman *et al.* [103] and Pralong *et al.* [26], these errors can be significant even at 200 Hz and above. However, recent studies indicate that the standard deviation of the HpTFs are significant generally only above 8 kHz – 9 kHz [104, 105], which is consistent with the results in this thesis (PV, Section 2.1). The errors are caused mainly by the movement of sharp dips in the HpTFs, which means that frequency smoothing can be used to remove a portion of these errors. In addition, the type of headphone used has a strong influence on the magnitudes of the errors.

Martin *et al.* [106] have studied the effect of headphone positioning on localization accuracy by using individual equalization based on a single HpTF measurement. The results suggested that headphone positioning has only a small influence on localization accuracy. Their test subjects were able to localize virtual sound sources accurately with eight different consecutive headphone placements. Hence, even though the changing of headphone position also changes the position of sharp dips in the frequency response, these changes do not seem to deteriorate the ability to localize virtual sound sources.

4.4 Perceptual aspects

If a binaural recording or synthesized binaural signal is reproduced over headphones so that the time-domain pressure signals at the eardrums are exactly the same as they would have been in the location in space where the original recording was made, all spectral attributes as well as binaural cues would also be identical. In BR it is often difficult to verify that the eardrum signals are exactly what they are expected to be, in which case the evaluation needs to be made using listening tests, as in this thesis (PV and PVI). In binaural reproduction-related research, listening test are commonly used to study aspects related to overall quality, timbre and localization [107].

Timbre is a very important attribute in audio reproduction, closely related to the magnitude frequency response at the eardrum. Timbre can be defined as the attribute of an auditory event that can help the listener to differentiate two sound events with similar pitch and loudness [108]. In BR the goal is to produce the same frequency response at the eardrum of the listener as the original sound source, and failure to do so causes a perception of timbral coloration. In binaural synthesis, the use of HRTF filters is reported to be a source of timbral coloration. A method of reducing these artifacts without losing accuracy in localization has been presented by Merimaa *et al.* [109, 25]. In this thesis, the differences in the magnitude frequency responses of different HRTF filters and the effect of these differences on perceived timbral coloration is discussed in PV. Smoothed frequency responses, i.e., frequency resolutions similar to that of the auditory system, are used, which means that very narrow dips in the frequency responses are smoothed out [110].

It has been shown that similar localization performance with binaurally reproduced virtual sources can be achieved as with real sources [111, 112, 106]. There are numerous unanswered questions related to sound source localization, however, and the subject has been studied by several scientists. These reports are related to, e.g., multichannel reproduction [113, 114, 115], to localization of several sound sources [19, 116], and to the necessity to use individual HRTFs and HpTFs in BR [112, 27, 106, 45]. In PVI, the effect of the magnitude frequency response at the eardrum on localization is studied as it has been reported that erroneous frequency responses at the eardrum may produce errors in localization [117, 26]. As suggested by Griesinger [44], the spectral maps used for localization are learned over time and cannot be adapted in a short period of time. This would mean that for accurate localization of sound sources the timbre must be correct.

In natural circumstances people use head movement to achieve more accurate sound source localization. It has been shown that using head tracking in BR reduces significantly front-back and back-front confusion and increases the probability of externalization of auditory events [118, 84].

5. Summaries and most important results of the publications

The peer-reviewed international journal and conference articles included in this thesis are presented separately in this section. First, a brief overview of the backgrounds of the publications is given.

5.1 Timeline and backgrounds of the publications

The first article, PI, was published in the Journal of the Audio Engineering Society (Vol. 58, No. 4) in April 2010. A preliminary version of the study [119] was published at the 126th Convention of the Audio Engineering Society in Munich, Germany, in May 2009, where it was awarded the Student Technical Papers Award. The research leading to the publication was a continuum to the Master's Thesis of the present author. The studies were carried out between May 2008 and November 2009 under the supervision of the late prof. Matti Karjalainen († May 30th 2010).

PII was presented at the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics in New Paltz, NY, in October 2009. Inspired by a talk by Dr. David Griesinger, it was Dr. Ville Pulkki who came up with the idea to use a miniature pressure-velocity sensor for measurements at the ear canal entrance. The research project started under the supervision of Dr. Pulkki in February 2009 and its first results are presented in this publication.

PIII was presented at the IEEE International Conference on Acoustics, Speech and Signal Processing in Prague, Czech Republic, in May 2011. The research, which builds on PI, was carried out in the fall of 2010.

PIV was published in the Journal of the Acoustical Society of America in May 2012. The topic of the article had been discussed earlier in a paper presented at the 6th Forum Acusticum in Aalborg, Denmark, in June 2011 [120]. The research that led to these result is a continuum to PII.

While the four aforementioned publications focus on measurements and modeling, PV is the first of two publications in this thesis that include results from listening tests. The article was presented at the AES 45th international conference in Helsinki, Finland, March 2012. The research, which was built on measurements made in PIV, was performed between May and October 2011.

PVI was presented at the AES 132nd Convention, Budapest, Hungary, in April 2012. Listening test results are presented as verification of the methods discussed in PIII. The research was carried out during the fall of 2011.

5.2 Publication I: Modeling of external ear acoustics for insert headphone usage

This publication focuses on the effects of the length of the ear canal and the eardrum impedance on sound pressure at the eardrum and at the ear canal entrance during insert headphone listening. Special ear canal simulators with adjustable length and eardrum impedance as well as an acoustical dummy head and insert headphones with in-ear microphones were custom-made for this study. Standard ear canal simulators or dummy heads were not suited for the purposes of this study. Measurements were carried out both with open ear canals and ear canals occluded with insert headphones. Eardrum measurements were made using the simulators and the dummy head only, whereas human subjects could be used for the measurements at the ear canal entrance. Computational modeling was applied to the ear canal simulator with rigid termination only using a termination impedance of 200 MΩ.

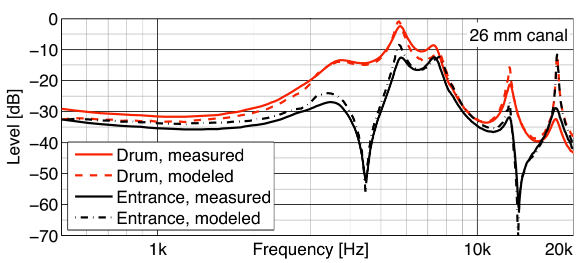


Figure 5.1. Modeled and measured frequency responses at the rigid eardrum and at the entrance of an ear canal simulator evoked by an insert headphone. Canal length: 26 mm, canal diameter: 8.5 mm, termination impedance: 200 MΩ.

One important result of this study is the successful estimation of the Thévenin electroacoustic source model for one of the insert earphones used in the measurements. The Thévenin source model enables the modeling of the pressure at the entrance and at the eardrum of the ear canal simulator when an earphone is attached to the simulator, as depicted in Fig. 5.1. It could also be confirmed that both the length of the ear canal and the impedance of the eardrum can be more important factors in the formation of the frequency response measured at the eardrum, than the shape and size of the pinna (in free-field conditions).

The figures in PI that contain results from measurements and modeling are presented in color and larger size in the appendix after the actual publication in this thesis in order to facilitate the interpretation of the results.

The main results of PI can be summarized as follows:

- Computational modeling with transmission line equations of an ear canal simulator with constant cross-sectional area and rigid termination in free-field is accurate up to the highest audible frequencies.
- Estimation of an insert headphone as a Thévenin equivalent electroacoustic pressure source can be accomplished using pressure measurements and several long open tubes with different cross sectional areas.
- Computational modeling of the pressure at the eardrum and at the entrance of an ear canal simulator using the Thévenin source model of the insert headphone and transmission line equations is accurate up to the highest audible frequencies.

5.3 Publication II: Estimating pressure at eardrum with pressure-velocity measurement from ear canal entrance

The sound pressure signal at the eardrum is of great interest in binaural reproduction and in audiological applications. Direct measurement from the eardrum is unreliable and involves risk to the health of the test subject. In this study, both sound pressure and particle velocity are measured at the ear canal entrance of an ear canal simulator and a dummy head using a miniature-sized PU sensor. The magnitudes of the pressures at the eardrums are estimated from these measurements using transmission line equations (with and without phase information) and a novel energy based method that does not require the length or the cross-sectional diameter of the ear canal as parameters.

The method is validated first by using a computational model of an ear canal simulator with constant cross sectional area and a rigid termination impedance of approximately 200 MΩ. Real measurements with the ear canal simulator and a dummy head also showed that the estimation was accurate, as depicted in, e.g., Fig. 5.2. No human test subjects were included in the study.

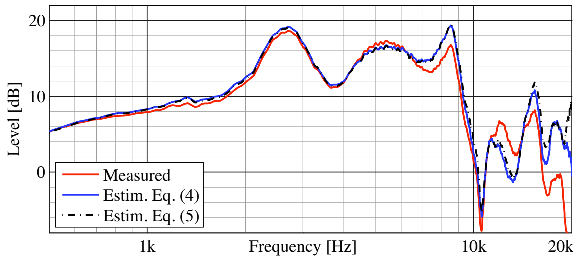


Figure 5.2. Measured and estimated pressure frequency responses at the damped eardrum of a dummy head with canal length of 25 mm in free-field conditions. The estimations are based on real measurements together with transmission line modeling as well as the new energy-based estimation method (Eq. (4) and Eq. (5) respectively). (azimuth $90°$, elevation $0°$)

Main results:

- The pressure at the eardrum of an ear canal simulator and a dummy head can be estimated using a pressure-velocity measurement at the entrance of the ear canal.
- The method is applicable both for damped and rigid eardrums.
- Both transmission line equations and a novel energy-based approach yield accurate results up to approximately 16 kHz.
- The energy-based estimation technique can be used without information of the length and cross sectional area of the ear canal.

5.4 Publication III: Estimating pressure and volume velocity in the ear canal for insert headphones

The HpTFs of insert headphones when worn by human subjects are extremely difficult to measure. Therefore, no individual equalization method for insert headphones is commonly known.

In this study, a method that enables the determination of both the pressure and the volume velocity evoked by an insert earphone in the human ear canal is developed. The method, which is partially based on the most important findings in PI and PII, can be used to accurately estimate the pressure frequency response at the eardrum of human subjects. An insert earphone equipped with an in-ear microphone is modeled as a Norton equivalent electroacoustic volume velocity source using five resistive load tubes. A custom-made new miniature acoustical particle velocity sensor is used in the estimation of the Norton equivalent velocity source parameters. For the estimation, the velocity responses of the headphone inside five tubes with different diameters were measured at a distance of approximately 7 mm from the headphone. This position mismatch was removed by taking the minimum phase versions of the impulse responses obtained.

The success of the estimation is verified through probe microphone measurements at the eardrums of the test subjects as well as through pressure and particle velocity measurements in ear canal simulators with different ear canal lengths. The measurements verify that the estimation method presented can be accurate at frequencies up to and well above 10 kHz for human subjects, as shown in Fig. 5.3 and Figures 4.1 – 4.5 in the appendix after the actual publication in this thesis. The method can be applied to binaural reproduction as well as to audiological measurements.

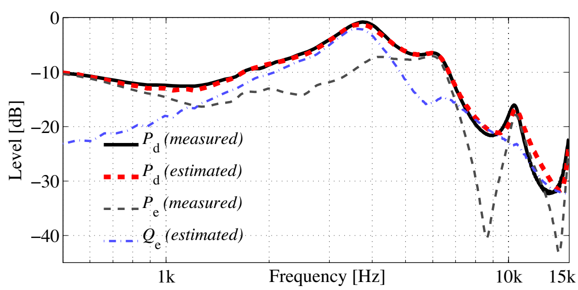


Figure 5.3. Measured and estimated pressure P_{d} at a point close to the eardrum of one test subject; measured pressure P_{e} and estimated volume velocity Q_{e} at the entrance of the ear canal.

The main results of this publication:

- A Norton type electroacoustic volume velocity source model of an insert headphone can be estimated using particle velocity measurements and long open-end tubes with different cross sectional areas.
- The Norton source model and pressure measurement with the in-ear microphones of the insert headphones at the ear canal entrance can be used to estimate volume velocity at the ear canal entrance.
- Energy-based estimation of the pressure at the eardrum yields accurate results for ear canal simulators and human test subjects.

5.5 Publication IV: Estimating head-related transfer functions of human subjects from pressure-velocity measurements

The most commonly used method of measuring individual human HRTFs is that in which the ear canal is blocked with an earplug and the frequency responses are measured in front of the earplug. This paper presents a method that allows obtaining individually correct magnitude frequency responses of HRTFs at the eardrum using pressure-velocity measurements at the ear canal entrance with a miniature PU sensor.

The HRTFs of 25 test subjects with nine directions of sound incidence were estimated using real anechoic measurements and an energy-based estimation method. To validate the approach, measurements were also conducted with probe microphones near the eardrums as well as with microphones at the blocked ear canal entrances.

The comparisons between the measurements made at the eardrum, with blocked ear canals, and with the PU sensor proved that the method presented is a valid and reliable technique for obtaining magnitude frequency responses of HRTFs (and HpTFs) as can be seen from Fig. 5.4. In addition, the HRTF filters designed using the PU measurements deviated less from the reference than the filters designed using blocked ear canal measurements.

HRTFs and HpTFs with the eardrum as point of reference are applicable to all types of headphones, not only FEC type of headphones. The method presented is an accurate and straightforward technique that does not require knowledge of the length or the area function of the ear canal. It can be used to obtain the magnitudes of the HRTFs at the eardrum with good accuracy for frequencies at least up to 10 kHz. One of the greatest advantages of the method is that the HRTFs can be obtained without having to introduce a probe microphone at the eardrum.

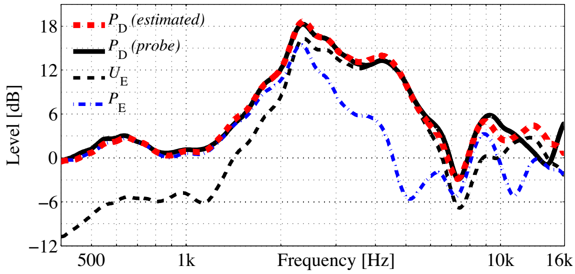


Figure 5.4. Measured and estimated pressure, P_D, at the right eardrum of one subject; measured pressure P_E and velocity U_E at the open ear canal entrance. Direction of sound incidence is azimuth $0°$, elevation $0°$.

The main results are as follows:

- HRTF's at the eardrums of human subjects can be estimated from measurements with the PU probe at the ear canal entrance.
- The method is applicable to ear canals with non-uniform cross-sectional area and a compliant eardrum.
- The magnitudes of the frequency responses of the HRTF filters designed using the PU-measurements are more correct than those designed using blocked ear canal HRTF measurements.

5.6 Publication V: Audibility of coloration artifacts in HRTF filter designs

The amount of coloration introduced by HRTF filters in binaural synthesis was studied in this work through perceptual evaluation. The goal was to study the differences between HRTF filter design methods using 1) measurements at the blocked ear canal entrance and 2) pressure-velocity measurements at the open ear canal entrance. Both individual and non-individual HRTFs and HpTFs were used in the listening test. Filters designed using individual probe microphone measurements at the eardrum were used as reference.

The differences in the amount of perceived coloration were significant. The method using pressure-velocity measurements at the open ear canal entrance and the energy-based estimation technique caused least coloration while the most coloration was caused by the methods using non-individual blocked ear canal measurements. Non-individual HRTFs combined with individual or non-individual HpTFs were found as most colored.

The results show that the least amount of coloration in binaural synthesis is obtained when individual HRTFs and HpTFs with the eardrum as point of reference are used in the design of HRTF filters.

The main results of this publication are:

- HRTF filter design using the energy-based individual HRTF estimation method causes less coloration than methods utilizing individual or non-individual blocked ear canal measurements.
- Individual headphone equalization does not reduce the amount of coloration if non-individual HRTFs are used in the reproduction.

5.7 Publication VI: Localization in Binaural Reproduction with Insert Headphones

Individual equalization of the headphones used in binaural reproduction is essential. The suitability of insert headphones for binaural reproduction after individual equalization using the method presented in PIII was studied. The ability of 14 test subjects to localize virtual sound sources as well as externalization in binaural listening with insert headphones and high-quality circumaural headphones was evaluated through a listening test. The samples used were bursts of pink noise recorded individually with microphones at the eardrum and at the blocked ear canal entrance in an anechoic chamber. The samples were reproduced with and without proper headphone equalization.

The overall performance regarding externalization and the likelihood of front-back confusion using the different methods of equalization were almost similar. Significant differences were found in the accuracy of the elevation perception, as depicted in Fig. 5.5. Hence, individual equalization of insert headphones proved to improve performance in the localization task.

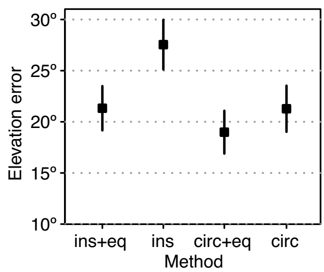


Figure 5.5. Marginal means and 95 % confidence intervals of the different methods in terms of accuracy of elevation perception. From left to right: Equalized and unequalized inserts; equalized and unequalized circumaural headphones.

The main results of this publication:

- Equalization of insert headphones using the Norton equivalent velocity source model and the energy-based eardrum pressure estimation method facilitates accurate localization of virtual sound sources.
- No significant difference in the accuracy of localization and the likelihood of externalization is found between equalized inserts and high-quality circumaural headphones.

6. Overview of the research, summary of methods, future directions and conclusions

To give the reader a clearer overall picture of the results reported in this thesis, the two main paths of the research are presented separately. First, the findings related to measurements and estimation of HRTFs in general are discussed. In the second section the research leading to the new findings in binaural reproduction with insert headphones is presented as a complete entity. Finally, future directions are discussed and conclusions are drawn.

6.1 New approach to HRTF measurements

The goal of the research was to find a solution for measuring HRTFs with the eardrum as the point of reference without the need to resort to probe or probe tube microphones. It was suggested that the task could be accomplished through measurements with the Microflown PU sensor, which is a relatively new device able to measure acoustical particle velocity by utilizing hot-wire anemometer technology [75].

The idea was first tested using a standard Microflown PU sensor and an ear canal simulator with constant cross-sectional area as depicted in Fig. 6.1. After careful analysis of the pressure and velocity signals measured the conclusion was that computation of the energy density at the entrance of the simulator could be used to estimate the magnitude of the pressure at the eardrum of the simulator. The idea is based on the assumption that very close to the eardrum, most of the sound energy is concentrated on the pressure component due to the high impedance of the eardrum.

Since the structure of the standard version of the PU probe is such that it cannot be used for measurements inside a human ear canal, two custom made probes were designed in cooperation with Microflown technologies

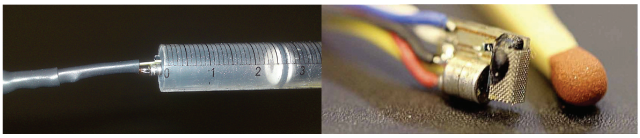


Figure 6.1. On the left: a standard Microflown PU probe and the custom-made ear canal simulator. On the right: a custom-made Microflown PU probe.

(Fig. 6.1). The probes were equipped with a mesh that protects the very thin platinum wires from breakage from contact with hair.

Measurements with 25 human subjects were carried out in order to test whether the method was applicable to human ear canals, the length and cross-sectional area of which were unknown.

The estimated pressures at the eardrums of the simulators as well as at the eardrums of human subjects were compared to actual microphone measurements at the eardrums. In addition, the results were verified through a listening test where the audibility of coloration caused by different HRTF filter design methods were investigated. The HRTF filter design methods tested applied pressure measurements at the eardrum and the blocked ear canal entrance as well as pressure-velocity measurements at the open ear canal entrance.

The comparisons between the eardrum measurements and the estimations showed that the magnitude of the pressure at the eardrum of an ear canal simulator, a dummy head, and human test subjects can be estimated using a pressure-velocity measurement at the entrance of the ear canal. The method also works with various eardrum impedances and non-uniform cross-sectional ear canal areas and not only with completely rigid eardrums and straight ear canals. The eardrum pressure can be estimated using the transmission line equations, too, but the energy-based estimation is applicable without accurate information on the length and cross sectional area of the ear canal.

The magnitudes of the HRTF filters designed using the PU measurements are closer to those designed from eardrum probe measurements than the ones designed using blocked ear canal HRTF measurements. The listening test confirmed the findings as the filters designed using the energy-based estimation caused less coloration than the method utilizing blocked ear canal measurements. Part of these findings can be explained with the fact that the open-back circumaural headphones used in the study do not have completely free-air equivalent coupling to the ear.

Figure 6.2. Insert headphones used in the research. On the left is the Philips SHN2500 (INS_1) and on the right the Nokia HS-86 (INS_2).

Regarding HRTF filter design, the approach presented proved to be at least as effective as the traditional blocked ear canal method.

The new method of measuring HRTFs and HpTFs presented in this thesis is a safe and reliable alternative to probe microphone measurements of pressure at the eardrum. The price of the PU sensor is currently significantly higher than the price of a single pressure transducer. However, since the manufacturing costs of a velocity transducer are even smaller than those of pressure transducers the situation will probably change if the technique is taken into wider use.

6.2 Individual equalization of insert headphones for binaural reproduction

Even though insert headphones are becoming more and more popular, they have not been commonly used in binaural reproduction, where circumaural headphones are more popular instead. In this section, a brief summary of the research related to insert headphone equalization for binaural reproduction is presented.

The first insert headphones used in the research were Philips SHN2500 PI (INS_1) and the second inserts were Nokia HS-86 headphones (INS_2) (PIII, PVI). Knowles FB-series miniature microphones were added in front of the transducer gates of the inserts as depicted in Fig. 6.2. The positioning of the in-ear microphones was improved in INS_2 for better usability with human test subjects. Pressure frequency responses inside the ear canal of simulators and human test subjects were measured with the in-ear microphones.

One[1] Thévenin electroacoustic source model of INS_1 was estimated by measuring the pressure signals evoked by the headphone inside five tubes with various diameters and solving the source parameters using Eq. (3) in PI. The pressures were measured with a separate probe microphone.

[1]Only the left earphone of INS_1 was modeled.

INS_2 was modeled as two Norton equivalent velocity sources by first measuring the acoustic particle velocities evoked by the headphones inside the above-mentioned tubes and then solving the source parameters using Eq. (3) in PIII. One of the new PU probes presented in PIV was used in the measurements.

Using INS_1 it was possible to compute the pressure at the eardrum of an ear canal simulator with rigid termination and constant cross-sectional area. Pressure measurement at the ear canal entrance and the transmission line equations were used in the computation, which means that the length and diameter of the simulator were used as variables in the computation. Pressure measurements at the eardrum of the simulators were used to validate the results. Estimating the pressure at the eardrum evoked by an insert headphone, i.e., the HpTFs of the insert headphones, is more complicated with human subjects than with a simulator due to the unknown length and cross-sectional area of the ear canal. Nevertheless, the estimation was successful with INS_2 by using the energy-based estimation technique presented first in PII. The validation of the results were made by comparing the estimated frequency responses to those measured with a probe microphone at the eardrums of the test subjects.

The success of the HpTF estimation was also tested with a localization listening test where INS_2 was equalized individually for each test subject. The equalization improved significantly the ability of the test subjects to localize virtual sound sources to a level as good as when high-quality circumaural headphones were used.

The results presented in this thesis are useful in the design of inserts and hearing aids. In addition, this thesis presents a method that enables individual in-situ equalization of inserts for binaural reproduction. The methods can also be used to study the noise exposure in music playback over insert headphones [121].

6.3 Summary of methods

A number of different measuring and modeling techniques have been implemented in the publications in this thesis. In this section, these techniques are revisited briefly. More detailed descriptions of the methods are presented in the publications.

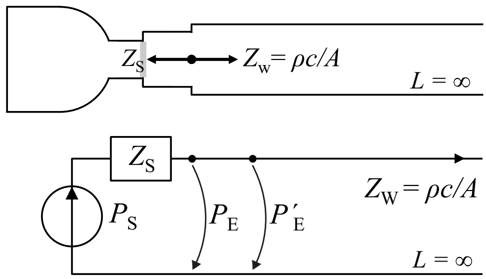


Figure 6.3. Upper part: A schematic illustration of an insert headphone connected to a tube load of infinite length L. The impedance to the right from the connection point is Z_W and the impedance to the left is Z_S. Lower part: A circuit model of the same situation. The pressure impulse response at the point of measurement, which is 7 mm from the connection point, is $P'_E = P_E + \delta t$, where δt is the time delay caused by this 7 mm distance and P_E is the response at the connection point.

6.3.1 Estimation of Thévenin and Norton source parameters

The estimation of the Thévenin and Norton source parameters of the insert headphones presented in PI and PIII is accomplished by using 3-meter long open-end pneumatic tubes with diameters of 5, 6, 7, 8, and 10 mm. The inserts are rigidly connected to these tubes by gluing the interchangeable rubber caps of different sizes to the front ends of the tubes. The inner diameter of the interchangeable rubber caps, i.e., the tube that connects the transducer port and the load tubes, is 6 mm in all cases. The earphones are then put to place for the measurements by connecting them to the rubber caps and sealing possible leaks between the rubber caps and the earphones with blu-tack. Hence, the impedance towards the source, as seen at the connection point, remains constant with each of the tubes, which would not necessarily be the case if a compressed flexible seal were used in the connection point. In the study by, e.g., Huang *et al.* [122] such a flexible seal was used, which changed the impedance at the connection point when using tubes with different diameters because: *“the seal is less compressed because it fills a larger diameter”*.

Another significant difference to be noted is that the method commonly found in literature includes short closed-end tubes with rigid terminations. In the present method, the back reflection from the open end of the tubes is removed by temporal windowing of the impulse response obtained with the miniature microphone (or the PU probe in PIII). Hence, the windowing removes the effects of standing waves in the tubes.

The point of measurement is 7 mm away from the insert headphone in order to avoid near-field effects. At this distance a forward traveling plane wave has already been formed. Having the point of measurement not exactly at the transducer port does not mean that the short part of the tube between the point of measurement and the headphone should be considered to be a part of the source impedance Z_{S}. The measured impulse response P'_{E} (in Figure 6.3) represents perfectly the impulse response P_{E} at the connection point in the theoretical situation that all near-field effects were removed. The only difference is that a small delay has been added to P'_{E}. This delay is removed by using minimum phase representations of the obtained impulse responses in the calculations of the source parameters. Hence, in the estimation of the Thévenin or Norton source parameters the impedance of the load is in each case $Z_{\mathrm{W}} = \rho c/A$ and the impedance of the source Z_{S} remains constant.

In PIII the impulse responses at the entrances of the tubes were measured with a miniature particle velocity sensor instead of a microphone to estimate the Norton source parameters of the inserts. This approach is supported by Hudde *et al.* [51], who stated: *"the volume velocity q_0 is more appropriate than the corresponding pressure because the earphones behave more like ideal velocity sources than ideal pressure sources."*

The accuracy of the Thévenin and the Norton parameters (PI, PII, and PIII) were verified, just as in, e.g., [122] by using short closed-end tubes with known theoretical impedance.

6.3.2 Ear canal simulators

It is customary to use standard ear simulators or couplers in studies related to the behavior of headphones. In the research presented in PI, PII, and PIII a custom-made ear canal simulator was used instead. One motivation here was that the acoustical characteristics (and impedance) of the simulator with constant cross sectional area and rigid termination could easily be modeled computationally and thereby be used for the validation of the Thévenin and Norton parameters as described above. In addition, the simulator was specifically designed for the custom made inserts with in-ear microphones used in PI and PIII. The goal was to study the effect of the length of the ear canal and the impedance of the termination of the ear canal simulator on the frequency responses measured with the in-ear microphones of the inserts as well as with the eardrum microphone of the simulator. The adjustable impedance of the artificial eardrum was used to find a match between the frequency responses measured with the in-ear microphones at the ear canal entrances of human test subjects and the simulator.

6.3.3 Estimation vs. computational modeling

In the publications, purely computational modeling, not including measured data, was applied to the simulator with constant cross sectional area and rigid termination only. Computational modeling of individual human ear canals is best accomplished using, e.g., FEM or BEM modeling as discussed in Section 3.3.

In the energy-based estimation method presented first in PII, measured pressure as well as estimated (PIII) or measured (PII) particle velocity at the ear canal entrance is used to obtain the magnitude of the pressure frequency response at the eardrum. Again, purely computational modeling, to verify the estimation equation, was applied to the simulator with constant cross sectional area and rigid termination only. Hence, it is important to distinguish between pure computational modeling and the estimation methods that use measured data with, e.g., human test subjects.

6.3.4 Obtaining velocity at the ear canal entrance

When the particle velocity is measured at the open ear canal entrance, information of the cross sectional area and the length of the ear canal are not used as parameters in the energy-based estimation method presented in this thesis. However, when the particle or volume velocity in front of an insert headphone is estimated using pressure measurements with in-ear microphones of the inserts, an individual estimation of the cross-sectional area of the ear canal is needed. E.g., in PIII, the wave impedances ($Z_{\mathrm{w}} = \rho c/A$) of the individual ear canals were computed individually by using estimates of the cross-sectional areas (A) of the ear canals close to the ear canal entrance. As reported in PIV, the varying cross-sectional area of a human ear canal and the impedance of the human eardrum do have an effect on the accuracy of the energy-based estimation method, but when compared to probe microphone measurements from the eardrum, the method yields accurate results at least up to 10 kHz.

6.3.5 High-frequency correction in probe microphone measurements

In PIII and PIV a high-frequency correction was performed for the probe microphone measurements from positions close to the eardrum of human test subjects. This technique is presented in more detail in Section IVA in PIV. The main principle, however, is that an estimate of the effective distance between the eardrum and the probe microphone is made first, after which its effect on the frequency response is cancelled out. This cancellation is achieved by making an approximation of the transfer function from the point of measurement to the eardrum using a simple transmission line model consisting of an average ear canal diameter and a simple eardrum model. The closer the probe microphone is to the eardrum, the smaller the effect is on the overall frequency response. For most test subjects, the high-frequency correction is needed at frequencies above 10 kHz only. Hence, the probe microphone measurements presented are uncorrected for most test subjects up to at least 8 kHz.

However, the correction does not have any effect on the HRTF filters designed using the measurements, since both the HRTFs and the HpTFs used in the individual filter designs undergo the same high-frequency correction. The effect of the correction is canceled out it the computation of the magnitude response of the HRTF filter.

6.4 Future directions

Further analysis of the estimation methods presented in this study is needed. Combining PU measurements and estimates of the area function of the ear canal would most likely yield interesting results.

The PU probes used in the research are small enough to fit into the ear canal of most human subjects. However, smaller PU probes than the ones presented here could be used, e.g., to measure the impedance of the ear canals of small children. Smaller and further developed probes might also be easier to use for ear measurements in general. A thorough assessment of the effect of the PU probe in the canal on the frequency response at the eardrum would also be valuable. According to Blau *et al.* [123] there can be effects even at frequencies below 6 kHz with obstacles of approximately the same size as the PU probe.

In the research related to timbral aspects in HRTF filter design, the studies presented in this thesis should be expanded to include all directions and a larger variety of samples. In addition, using a larger selection of different headphone types, such as headphones with completely FEC characteristics and headphones that are connected closer to the ear, would give valuable information of the usability of the present energy-based estimation method.

A more profound understanding of the sound quality and the usability of inserts in high-fidelity binaural reproduction is called for. To design a reliable listening test for the analysis of timbral aspects in music reproduction with inserts would be of great interest. Such a listening test would be applicable for further verification of the calibration method presented in this thesis. In addition, a localization test including proper visual cues would probably yield results with smaller errors in elevation and azimuth and increase the likelihood of externalization. Furthermore, using HRTF filters and binaural synthesis instead of binaural recordings, as in PVI, would improve sound quality and thus create a more immersive audiovisual environment.

Many reports that utilize BEM modeling of HRTFs have already been published, but there are still new areas in that field to be covered. Accurate models that include the torsi, the heads as well as the complete external ears of real human subjects would be useful. With the complete ear canal included in the model, thorough investigations of the acoustical behavior of circumaural and insert headphones become possible.

Furthermore, an interesting topic would be to make dummy head models of actual human subjects and use them for binaural recording in order to study the benefits of individual recordings in binaural reproduction.

6.5 Conclusions

This thesis presents significant new research results in the field of reproduction of 3-D audio over headphones. Firstly, a new method of obtaining individual filters needed for 3-D sound synthesis is revealed. Secondly, a new method that allows insert headphones to be individually equalized in-situ for binaural reproduction is presented.

According to the results presented in this thesis, individual magnitudes of HRTFs and HpTFs of humans with the eardrum as point of reference can be estimated from measurements made at the ear canal entrance using a pressure-velocity sensor. The measurements and a new energy-based estimation method yield the magnitude frequency responses at the eardrums. HRTF filters designed using the estimates are suitable for binaural reproduction using headphones that do not have FEC properties, e.g., insert headphones.

It is also shown that by using the methods presented in this thesis insert headphones can be equalized individually for each user, e.g., to obtain realistic 3-D audio reproduction. After the equalization, insert headphones can perform equally well as high-quality circumaural headphones in binaural reproduction.

Errata

Publication III

Section 2.2: In the analogue model of the ear used in the study by Hammershøi and Møller (ref. 6) the ear canal is presented as a two-port but not as a transmission line.

Publication IV

Section II A: In the analogue model of the ear used in the study by Hammershøi and Møller (ref. 2) the ear canal is presented as a two-port but not as a transmission line.
Section III C: The Sennheiser HD 590 headphones were not included in the study by Møller *et al.* (ref. 5).

Bibliography

[1] T. Rossing, P. Wheeler, and R. Moore, *The Science of Sound*. San Francisco, CA, USA: Addison Wesley, 3rd ed., 2002.

[2] J. Blauert, ed., *Communication Acoustics*. Berlin, Germany: Springer-Verlag, 2005.

[3] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization*, pp. 36–50. Cambridge, MA, USA: MIT Press, revised ed., 1997.

[4] L. Beranek, *Acoustics*. MI, USA: American Institute of Physics for the Acoustical Society of America, 3rd printing ed., 1988.

[5] M. Karjalainen. Espoo, Finland: Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing.

[6] G. V. Békésy, *Experiments in hearing*. New York, NY, USA: Mcgraw Hill, 1960.

[7] M. Russo, N. Rožić, and M. Stella, "Biophysical cochlear model: Time-frequency analysis and signal reconstruction," *Acta Acustica United with Acustica*, vol. 97, no. 4, pp. 632–640, 2011.

[8] E. Zwicker, "Subdivision of the audible frequency range into critical bands," *Journal of the Acoustical Society of America*, vol. 33, no. 2, p. 248, 1961.

[9] J. Merimaa, *Analysis, Synthesis, and Perception of Spatial Sound - Binaural Localization Modeling and Multichannel Loudspeaker Reproduction*. PhD thesis, Helsinki University of Technology, Espoo, Finland, 2006.

[10] P. Mokhtari, H. Takemoto, R. Nishimura, and H. Kato, "Pinna sensitivity patterns reveal reflecting and diffracting surfaces that generate the first spectral notch in the front median plane," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, (Prague, Czech Republic), pp. 2408–2411, May 2011.

[11] S. Spagnol, M. Hiipakka, and V. Pulkki, "A single-azimuth pinna-related transfer function database," in *Proc. 14th International Conference on Digital Audio Effects (DAFx-11)*, (Paris, France), Sep 2011.

[12] L. Alvord and B. Farmer, "Anatomy and orientation of the human external ear," *Journal of the American Academy of Audiology*, vol. 8, no. 6, pp. 383–390, 1997.

[13] J. Fay, S. Puria, and C. Steele, "The discordant eardrum," *Proceedings of the National Academy of Sciences*, vol. 103, Dec 2006.

[14] S. Voss and J. Allen, "Measurement of acoustic impedance and reflectance in the human ear canal," *Journal of the Acoustical Society of America*, vol. 95, no. 1, pp. 372–384, 1994.

[15] D. Ćirić and D. Hammershøi, "Acoustic impedances of ear canals measured by impedance tube," in *Proc. International Congress on Acoustics*, (Madrid, Spain), pp. 1–6, Sep 2007.

[16] D. Keefe, J. Bulen, K. Arehart, and E. Burns, "Ear canal impedance and reflection coefficient in human infants and adults," *Journal of the Acoustical Society of America*, vol. 94, no. 5, pp. 2617–2638, 1993.

[17] S. Schmidt and H. Hudde, "Accuracy of acoustic ear canal impedances: Finite element simulation of measurement methods using a coupling tube," *Journal of the Acoustical Society of America*, vol. 125, no. 6, pp. 3819–3827, 2009.

[18] J. Fels, J. Paprotny, and L. Feickert, "Ear canal impedances of children and adults – investigations with simulation and measurements," in *Proc. 19th International Congress on Acoustics*, (Madrid, Spain), Sep 2007.

[19] R. Stern, G. Brown, and D. Wang, *Binaural Sound Localization*, ch. 5 in book *Computational Auditory Scene Analysis* (ed. D. Wang and G. Brown). John Wiley Sons, Inc., 2005.

[20] F. Firestone, "The phase difference and amplitude ratio at the ears due to a source of pure tone," *Journal of the Acoustical Society of America*, vol. 2, no. 2, pp. 260–270, 1930.

[21] R. Gilkey and T. Anderson, eds., *Binaural and Spatial Hearing in Real and Virtual Environments*, ch. 28. Lawrence Erlbaum Associates, 1997.

[22] J. Fels and M. Vorländer, "Anthropometric parameters influencing head-related transfer functions," *Acta Acustica united with Acustica*, vol. 95, no. 2, 2009.

[23] D. T. R. Algazi, R. Duda and C. Avendano, "The CIPIC HRTF database," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, (New Paltz, NY, USA), pp. 99–103, Oct 2001.

[24] H. Møller, M. Sørensen, C. Jensen, and D. Hammershøi, "Binaural technique: Do we need individual recordings," *Journal of the Audio Engineering Society*, vol. 44, no. 6, pp. 451–469, 1996.

[25] J. Merimaa, "Modification of HRTF filters to reduce timbral effects in binaural synthesis, Part 2: Individual HRTFs," in *Proc. 129th AES Convention*, (San Francisco, CA, USA), Nov 2010.

[26] D. Pralong and S. Carlile, "The role of individual headphone calibration for the generation of high fidelity virtual auditory space," *Journal of the Acoustical Society of America*, vol. 100, no. 6, pp. 3785–3793, 1996.

[27] E. Wenzel, M. Arruda, D. Kistler, and F. Wightman, "Localization using nonindividualized head-related transfer functions," *J. Acoust. Soc. of Am.*, vol. 94, no. 1, pp. 111–123, 1993.

[28] Y. Kim, J. Kim, and S. Ko, "A parametric model of head-related transfer functions for sound source localization," in *Proc. 122nd AES Convention*, (Vienna, Austria), May 2007.

[29] D. Brungart and G. Romigh, "Spectral HRTF enhancement for improved vertical-polar auditory localization," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, (New Paltz, NY, USA), pp. 305–308, Oct 2009.

[30] A. Farina, "Simultaneous measurement of impulse response and distortion with a swep-sine technique," in *Proc. 108th AES Convention*, (Paris, France), Feb 2000.

[31] S. Müller and P. Massarani, "Transfer-function measurement with sweeps," *Journal of the Audio Engineering Society*, vol. 49, no. 6, pp. 443–471, 2001.

[32] G.-B. Stan, J.-J. Embrechts, and D. Archambeau, "Comparison of different impulse response measurement techniques," *J. Audio Eng. Soc*, vol. 50, no. 4, pp. 249–262, 2002.

[33] E. Villchur and M. Killion, "Probe-tube microphone assembly," *J. Audio Eng. Soc*, vol. 57, no. 1, pp. 238–240, 1975.

[34] R. McCreery, A. Pittman, J. Lewis, S. Neely, and P. Stelmachowicz, "Use of forward pressure level to minimize the influence of acoustic standing waves during probe-microphone hearing-aid verification," *Journal of the Acoustical Society of America*, vol. 126, no. 1, pp. 15–24, 2009.

[35] J. Chan and C. Geisler, "Estimation of eardrum acoustic pressure and of ear canal length from remote points in the canal," *Journal of the Acoustical Society of America*, vol. 87, no. 3, pp. 1237–1247, 1990.

[36] M. Stinson, "The spatial distribution of sound pressure within scaled replicas of the human ear canal," *Journal of the Acoustical Society of America*, vol. 78, no. 5, pp. 1596–1602, 1985.

[37] H. Hudde and S. Schmidt, "Sound fields in generally shaped curved ear canals," *Journal of the Acoustical Society of America*, vol. 125, no. 5, pp. 3146–3157, 2009.

[38] H. Møller, "Fundamentals of binaural technology," *Applied Acoustics*, vol. 36, no. 3-4, pp. 171–218, 1992.

[39] M. Hiipakka, "Measurement apparatus and modelling techniques of ear canal acoustics," Master's thesis, Helsinki University of Technology, Espoo, Finland, Dec 2008.

[40] ANSI S3.25-1989, *American national standard for an occluded ear simulator*. American National Standards Institute (ANSI), 1989.

[41] IEC60711, *Occluded-ear simulator for the measurement of earphones coupled to the ear by ear inserts*. International Electrotechnical Commission (IEC), 1981.

[42] S. Paul, "Binaural recording technology: A historical review and possible future developments," *Acta Acustica united with Acustica*, vol. 95, no. 5, pp. 767–788, 2009.

[43] A. Wilska, *Untersuchungen uber das Richtungshören*. PhD thesis, University of Helsinki, Helsinki, Finland, 1938.

[44] D. Griesinger, "Frequency response adaptation in binaural hearing," in *Proc. 126th AES Convention*, (Munich, Germany), May 2009.

[45] P. Minnaar, S. Olesen, F. Christensen, and H. Møller, "Localization with binaural recordings from artificial and human heads," *J. Audio Eng. Soc*, vol. 49, no. 5, pp. 323–336, 2001.

[46] F. Toole, "The acoustics and psychoacoustics of headphones," in *Proc. 2nd AES International Conference*, (Anaheim, CA, USA), May 1984.

[47] M. Vorländer, "Past, present and future of dummy heads," in *Proc. FIA conference and EAA Symposium*, (Guimaraes, Portugal), Sep 2004.

[48] A. Bravo, R. Barham, M. Ruiz, J. López, and G. De Arcas, "A new 3D finite element model of the IEC 60318-1 artificial ear," *Metrologia*, vol. 45, pp. 448–458, 2008.

[49] IEC60318-1, *Electroacoustics – Simulators of human head and ear – Part 1: Ear Simulator for the Calibration of Supra-aural Earphones*. International Electrotechnical Commission (IEC), 1 ed., 1998.

[50] A. Møller, "An experimental study of the acoustic impedance of the middle ear and its transmission properties," *Acta Oto-Laryngol*, vol. 60, pp. 129–149, 1965.

[51] H. Hudde, A. Engel, and A. Lodwig, "Methods for estimating the sound pressure at the eardrum," *Journal of the Acoustical Society of America*, vol. 106, no. 4, pp. 1977–1992, 1999.

[52] D. Brenner and E. Hall, "Computed tomography – an increasing source of radiation exposure," *The New England Journal of Medicine*, vol. 357, no. 22, pp. 2277–2284, 2007.

[53] D. Rasetshwane and S. Neely, "Inverse solution of ear-canal area function from reflectance," *Journal of the Acoustical Society of America*, vol. 130, no. 6, pp. 3873–3881, 2011.

[54] N. Gumerov, A. O'Donovan, R. Duraiswami, and D. Zotkin, "Computation of the head-related transfer function via the fast multipole accelerated boundary element method and its spherical harmonic representation," *Journal of the Acoustical Society of America*, vol. 127, no. 1, pp. 370–386, 2010.

[55] S. Cadavid and M. Abdel-Mottaleb, "3-D ear modeling and recognition from video sequences using shape from shading," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 4, pp. 709–718, 2008.

[56] R. Gan and Q. Sun, "Finite element modeling of human ear with external ear canal and middle ear cavity," in *Proc. Second Joint EMBS/BMES Conference*, (Houston, TX, USA), Oct 2002.

[57] S. Jønsson, B. Liu, A. Schuhmacher, and L. Nielsen, "Simulation of the IEC 60711 occluded ear simulator," in *Proc. 116th AES Convention*, (Berlin, Germany), May 2004.

[58] T. Huttunen, A. Vanne, T. Avikainen, and L. Kärkkäinen, "Pilvilaskenta akustisessa mallinnuksessa," in *Proc. Akustiikkapäivät*, (Tampere, Finland), Acoustical Society of Finland, May 2011.

[59] T. Huttunen, E. Seppälä, O. Kirkeby, A. Kärkkäinen, and L. Kärkkäinen, "Simulation of the transfer function for a head-and-torso model over the entire audible frequency range," *Journal of Computational Acoustics*, vol. 15, no. 4, pp. 429–448, 2007.

[60] B. Katz, "Boundary element method calculation of individual head-related transfer function," *Journal of the Acoustical Society of America*, vol. 110, no. 5, pp. 2440–2448, 2001.

[61] R. Greff and B. Katz, "Round robin comparison of HRTF simulation systems: Preliminary results," in *123rd AES Convention*, (New York, NY, USA), Oct 2007.

[62] Z. Chen and W. Kreuzer, "A fast multipole boundary element method for calculating HRTFs," in *Proc. 122nd AES Convention*, (Vienna, Austria), May 2007.

[63] M. Stinson and B. Lawton, "Specification of the geometry of the human ear canal for the prediction of sound-pressure level distribution," *Journal of the Acoustical Society of America*, vol. 85, no. 6, 1989.

[64] H. Hudde, "Estimation of the area function of human ear canals by sound pressure measurements," *Journal of the Acoustical Society of America*, vol. 73, no. 1, pp. 24–31, 1983.

[65] H. Hudde, "Measurement of the eardrum impedance of human ears," *Journal of the Acoustical Society of America*, vol. 73, no. 1, pp. 242–247, 1983.

[66] H. Hudde, A. Lodwig, and A. Engel, "A wide-band precision acoustic measuring head," *Acustica*, vol. 82, no. 6, pp. 895–904, 1996.

[67] D. Keefe, R. Ling, and J. Bulen, "Method to measure acoustic impedance and reflection coefficient," *Journal of the Acoustical Society of America*, vol. 91, no. 1, pp. 470–485, 1992.

[68] S. Neely and M. Gorga, "Comparison between intensity and pressure as measures of sound level in the ear canal," *Journal of the Acoustical Society of America*, vol. 104, no. 5, pp. 2925–2934, 1998.

[69] B. Farmer-Fedor and R. Rabbitt, "Acoustic intensity, impedance and reflection coefficient in the human ear canal," *Journal of the Acoustical Society of America*, vol. 112, no. 2, pp. 600–620, 2002.

[70] R. Scheperle, S. Neely, J. Kopun, and M. Gorga, "Influence of in situ, sound-level calibration on distortion-product otoacoustic emission variability," *Journal of the Acoustical Society of America*, vol. 124, no. 1, pp. 288–300, 2008.

[71] J. Kates, "A computer simulation of hearing aid response and the effects of ear canal size," *Journal of the Acoustical Society of America*, vol. 83, no. 5, pp. 1952–1963, 1988.

[72] R. Paulsen, *Statistical Shape Analysis of the Human Ear Canal with Application to In-the-Ear Hearing Aid*. PhD thesis, Technical University of Denmark, 2004.

[73] S. Voss, J. Rosowski, C. Shera, and W. Peake, "Acoustic mechanisms that determine the ear-canal sound pressures generated by earphones," *Journal of the Acoustical Society of America*, vol. 107, no. 3, pp. 1548–1565, 2000.

[74] T. Sankowsky-Rothe, M. Blau, E. Rasumow, H. Mojallal, M. Teschner, and C. Thiele, "Prediction of the sound pressure at the ear drum in occluded human ears," *Acta Acustica united with Acustica*, vol. 97, no. 4, pp. 656–668, 2011.

[75] H.-E. de Bree, "An overview of microflown technologies," *Acta Acustica united with Acustica*, vol. 89, no. 1, pp. 163–172, 2003.

[76] G. Lorho and N. Zacharov, "Subjective evaluation of virtual home theatre sound systems for loudspeakers and headphones," in *Proc. AES 116th Convention*, (Berlin, Germany), May 2004.

[77] M. Gerzon, "Surround sound psychoacoustics," *Wireless World*, vol. 80, pp. 483–486, Dec 1974.

[78] P. Fellgett, "Ambisonics. part one: General system description," *Studio Sound*, vol. 17, no. 8, 1975.

[79] M. Gerzon, "Ambisonics. part two: Studio techniques," *Studio Sound*, vol. 17, no. 8, 1975.

[80] V. Pulkki, "Spatial sound reproduction with directional audio coding," *Journal of the Audio Engineering Society*, vol. 55, no. 6, pp. 503–516, 2007.

[81] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *Journal of the Audio Engineering Society*, vol. 45, no. 6, pp. 456–466, 1997.

[82] P. Damaske, "Head-related two-channel stereophony with loudspeaker reproduction," *Journal of the Acoustical Society of America*, vol. 50, no. 4, pp. 1109–1115, 1971.

[83] D. Cooper and J. Bauck, "Prospects for transaural recording," *J. Audio Eng. Soc*, vol. 37, no. 1, pp. 3–19, 1989.

[84] M.-V. Laitinen and V. Pulkki, "Binaural reproduction for directional audio coding," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, (New Paltz, NY, USA), Oct 2009.

[85] J. Huopaniemi, N. Zacharov, and M. Karjalainen, "Objective and subjective evaluation of head-related transfer function filter design," *J. Audio Eng. Soc*, vol. 47, no. 4, pp. 218–239, 1999.

[86] D. Kistler and F. Wightman, "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction," *Journal of the Acoustical Society of America*, vol. 91, no. 3, pp. 1637–1647, 1992.

[87] H. Hachabiboglu, B. Günel, and A. Kondoz, "Head-related transfer function filter interpolation by root displacement," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, (New Paltz, NY, USA), pp. 134–137, Oct 2005.

[88] J. Huopaniemi, *Virtual Acoustics and 3-D Sound in Multimedia Signal Processing*. PhD thesis, Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, 1999.

[89] D. Begault, "Challenges to the successful implementation of 3-d sound," *Journal of the Audio Engineering Society*, vol. 39, no. 11, pp. 864–870, 1991.

[90] J. Sandvad and D. Hammershøi, "Binaural auralization. comparison of FIR and IIR filter representation of HIRs," in *Proc. 96th AES Convention*, (Amsterdam, Netherlands), Feb 1994.

[91] P. Minnaar, J. Plogsties, and F. Christensen, "Directional resolution of head-related transfer functions required in binaural synthesis," *J. Audio Eng. Soc*, vol. 53, no. 10, pp. 919–929, 2005.

[92] E. Wenzel and S. Foster, "Perceptual consequences of interpolating head-related transfer functions during spatial synthesis," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, (New Paltz, NY, USA), pp. 102–105, Oct 1993.

[93] T. Nishino, S. Kajita, K. Takeda, and F. Itakura, "Interpolating head related transfer functions in the median plane," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, (New Paltz, NY, USA), pp. 167–170, Oct 1999.

[94] J. Huopaniemi and M. Karjalainen, "Review of digital filter design and implementation methods for 3-D sound," in *Proc. AES 102nd Convention*, (Munich, Germany), Mar 1997.

[95] P. Runkle, M. Blommer, and G. Wakefield, "A comparison of head related transfer function interpolation methods," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, (New Paltz, NY, USA), pp. 88–91, Oct 1995.

[96] C. Poldy, *Headphones*, ch. 14 in Loudspeaker and Headphone Handbook (ed. J. Borwick). Oxford, UK: Focal Press, 3rd ed., 2001.

[97] H. Møller, C. Jensen, D. Hammershøi, and M. F. Sørensen, "Design criteria for headphones," *J. Audio Eng. Soc*, vol. 43, no. 4, pp. 218–232, 1995.

[98] H. Møller, D. Hammershøi, C. Jensen, and M. Sørensen, "Transfer characteristics of headphones measured on human ears," *J. Audio Eng. Soc*, vol. 43, no. 4, pp. 203–217, 1995.

[99] G. Wersényi, "On the measurement and evaluation of bass enhanced in-ear phones," in *Proc. 20th International Congress on Acoustics, ICA*, (Sydney, Australia), Aug 2010.

[100] B. Masiero and J. Fels, "Perceptually robust headphone equalization for binaural reproduction," in *Proc. AES 130th Convention*, (London, UK), May 2011.

[101] D. Griesinger, "Binaural techniques for music reproduction," in *Proc. 8th AES International Conference*, (Washington D.C., USA), May 1990.

[102] D. Ćirić and D. Hammershøi, "Coupling of earphones to human ears and to standard coupler," *Journal of the Acoustical Society of America*, vol. 120, no. 4, pp. 2096–2107, 2006.

[103] F. Wightman and D. Kistler, "Headphone simulation of free-field listening. I: Stimulus synthesis," *J. Acoust. Soc. Am*, vol. 85, no. 2, pp. 858–867, 1989.

[104] K. McAnally and R. Martin, "Variability in the headphone-to-ear-canal transfer function," *Journal of the Audio Engineering Society*, vol. 50, no. 4, pp. 263–266, 2002.

[105] A. Kulkarni and H. Colburn, "Variability in the characterization of the headphone transfer function," *Journal of the Acoustical Society of America*, vol. 107, no. 2, pp. 1071–1074, 2000.

[106] R. Martin, K. McAnally, and M. Senova, "Free-field equivalent localization of virtual audio," *Journal of the Audio Engineering Society*, vol. 49, no. 1/2, pp. 14–22, 2001.

[107] S. Bech and N. Zacharov, *Perceptual Audio Evaluation – Theory, Method and Application*. Chichester, England: John Wiley & Sons, Ltd, 2006.

[108] B. Moore, *An Introduction to the Psychology of Hearing*. Academic Press, 4th ed., 1997.

[109] J. Merimaa, "Modification of hrtf filters to reduce timbral effects in binaural synthesis," in *Proc. AES 127th International Convention*, (New York, NY, USA), Oct 2009.

[110] R. Bücklein, "The audibility of frequency response irregularities," *J. Audio Eng. Soc*, vol. 29, no. 3, pp. 126–131, 1981.

[111] P. Zahorik, F. Wightmann, and D. Kistler, "On the discriminability of virtual and real sound sources," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, (New Paltz, NY, USA), Oct 1995.

[112] A. Bronkhorst, "Localization of real and virtual sound sources," *J. Acoust. Soc. Am*, vol. 98, no. 5, pp. 2542–2553, 1995.

[113] V. Pulkki, M. Karjalainen, and J. Huopaniemi, "Analyzing virtual sound source attributes using a binaural auditory model," *Journal of the Audio Engineering Society*, vol. 47, no. 4, pp. 203–217, 1999.

[114] V. Pulkki, "Localization of amplitude-panned virtual sources II: three-dimensional panning," *Journal of the Audio Engineering Society*, vol. 49, no. 9, pp. 753–767, 2001.

[115] V. Pulkki and M. Karjalainen, "Localization of amplitude-panned virtual sources I: Sterephonic panning," *Journal of the Audio Engineering Society*, vol. 49, no. 9, pp. 739–752, 2001.

[116] O. Santala and V. Pulkki, "Directional perception of distributed sound sources," *Journal of the Acoustical Society of America*, vol. 129, no. 3, pp. 1522–1530, 2011.

[117] F. Asano, Y. Suzuki, and T. Sone, “Role of spectral cues in median plane localization,” *Journal of the Acoustical Society of America*, vol. 88, no. 1, pp. 159–168, 1990.

[118] D. Begault, E. Wenzel, and M. Anderson, “Direct comparison of the impact of head-tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source,” *Journal of the Audio Engineering Society*, vol. 49, no. 10, pp. 904–916, 2001.

[119] M. Hiipakka, M. Tikander, and M. Karjalainen, “Modeling of external ear acoustics for insert headphone usage,” in *Proc. 126th AES Convention*, (Munich, Germany), May 2009.

[120] M. Hiipakka, T. Kinnari, and V. Pulkki, “HRTF Measurements With Pressure-velocity Sensor,” in *6th Forum Acusticum*, (Aalborg, Denmark), Jun 2011.

[121] S. Oksanen, M. Hiipakka, and V. Sivonen, “Estimating individual sound pressure levels at the eardrum in music playback over insert headphones,” in *Proc. 47th AES International Conference*, (Chicago, IL, USA), Jun 2012.

[122] G. Huang, J. Rosowski, S. Puria, and W. Peake, “A noninvasive method for estimating acoustic admittance at the tympanic membrane,” *Journal of the Acoustical Society of America*, vol. 108, no. 3, pp. 1128–1146, 2000.

[123] M. Blau, T. Sankowsky, A. Stirnemann, H. Oberdanner, and N. Schmidt, “Acoustics of open fittings,” in *Proc. Acoustics08*, (Paris, France), pp. 711–716, Jun 2008.

Aalto-DD 149/2012

ISBN 978-952-60-4864-2
ISBN 978-952-60-4865-9 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934
ISSN 1799-4942 (pdf)

Aalto University
School of Electrical Engineering
Department of Signal Processing and Acoustics
www.aalto.fi