

# Bachelorarbeit

Sylvia Sima

HRTF Measurements and Filter Design for a  
Headphone-Based 3D-Audio System

Sylvia Sima

HRTF Measurements and Filter Design for a  
Headphone-Based 3D-Audio System

Bachelorarbeit eingereicht im Rahmen der Bachelorprüfung  
im Studiengang Technische Informatik  
am Department Informatik  
der Fakultät Technik und Informatik  
der Hochschule für Angewandte Wissenschaften Hamburg

Betreuender Prüfer : Prof. Dr.-Ing. Wolfgang Fohl  
Zweitgutachter : Prof. Dr.-Ing. Andreas Meisel

Abgegeben am 6. September 2008

**Sylvia Sima**

## **Thema der Bachelorarbeit**

HRTF Messungen und Implementierung eines Kopfhörerbasierten 3D-Audiosystems

## **Stichworte**

HRTF, FIR, Kunstkopf , 3D, Digitale Audio Signalverarbeitung

## **Abstract**

Das menschliche Gehör ist in der Lage die Position und die Entfernung einer Schallquelle zu bestimmen. Auf dem Weg von der Schallquelle zum Ohr erfahren die Schallwellen eine Filterung. Dieser Filterungsprozess ist durch die Head Related Transfer Function(HRTF) beschrieben. Diese Arbeit behandelt die Implementierung eines HRTF-basiertes Systems, welches einen Virtuelles Sound generiert.

**Sylvia Sima**

## **Title of the paper**

HRTF Measurements and Filter Design for a Headphone-Based 3D-Audio System

## **Keywords**

HRTF, FIR, dummy head , 3D, Digital Audio Signal Processing

## **Abstract**

Humans are able to determine the position of a sound source by filtering sound waves as they travel from the sound source to the listener's ears. This filtering process is characterized by Head Related Transfer Functions (HRTF's).

The purpose of the thesis is to develop a HRTF based system that creates virtual sound. It aims to obtain a set of generalised HRTFs using the anthropomorphic dummy head. The system is expected to synthesize a binaural signal from any acoustic monaural signal using previously obtained HRTFs.

## **Acknowledgments**

I would like to thank a number of people who supported me during this work. First of all I'd like to thank my supervisor Dr. Wolfgang Fohl, who guided me and from whom I have learned a lot in terms of Digital Signal Processing. Many thanks for his assistance and for information he provided me with.

Many thank also to Law S. Ming for his technical assistance during measurements. Particularly for his readiness and patience. It is not really comfortable to spend all day in an anechoic chamber without sunshine and fresh air.

I'm grateful to the laboratory assistants K.D. Hempel and J. Neugebauer for supporting me in measurement setup. Finally ,I would like to thank my family and friends for their love and contribution to this work.

# Contents

- 1 Introduction 2**
  - 1.1 Premise . . . . . 3
  - 1.2 Related Work . . . . . 3
  
- 2 Fundamentals of Spatial Hearing 4**
  - 2.1 Physical Characteristics of Sound . . . . . 4
  - 2.2 The Hearing Organ . . . . . 6
  - 2.3 Characteristics of Human Hearing . . . . . 8
  - 2.4 Head Related Coordinate System . . . . . 8
  - 2.5 Inter-Aural Differences . . . . . 10
    - 2.5.1 Inter-Aural Time Differences . . . . . 10
    - 2.5.2 Inter-Aural Level Differences . . . . . 11
  - 2.6 Cone of Confusion and Monaural cues . . . . . 12
  - 2.7 Precedence Effect (Haas Effect, law of the first wave front) . . . . . 13
  - 2.8 Head Related Transfer Function . . . . . 13
  
- 3 Measurements 16**
  - 3.1 Overview . . . . . 16
  - 3.2 Measurement Technique . . . . . 16
    - 3.2.1 Loudspeaker Measurement . . . . . 16
    - 3.2.2 Analyzer Settings for Loudspeaker and HRTFs Measurements . . . . . 18
    - 3.2.3 HRTF Measurements . . . . . 22
  - 3.3 Analysis of HRTF curves . . . . . 25
  - 3.4 Headphones Measurements . . . . . 27
  
- 4 System Design 29**
  - 4.1 HRIR Filter . . . . . 29
  - 4.2 Headphone Equalization . . . . . 34
  
- 5 Test 36**
  - 5.1 Overview . . . . . 36
  - 5.2 Test Setup and Procedure . . . . . 36
  - 5.3 Test Results . . . . . 38

5.3.1	Test run with Sennheiser HD 270 and HD 600 for the Horizontal Plane	39
<b>6</b>	<b>Conclusion and Further Development</b>	<b>42</b>
<b>7</b>	<b>References</b>	<b>43</b>
	<b>List of Figures</b>	<b>46</b>
	<b>List of Tables</b>	<b>48</b>
<b>A</b>	<b>Matlab Scripts and Supporting Tool</b>	<b>55</b>

# 1 Introduction

For a long time researches have conducted on how people locate sound in a natural hearing environment. Our auditory system modifies the incoming sounds by filtering them depending on their direction. The modified sound involves a set of spatial cues used by the brain to detect the position of a sound.

Human hearing is binaural, that means we use both ears to perceive the sound-pressure signals. The term *binaural* literally means relating to or used with both ears. In 1948, Jeffress proposed the first binaural concept, a technique to detect inter-aural time differences (ITD), one of the most important cues for localization of sound sources. Since then binaural models have attracted the attention of scientists and have become more and more popular as an aid for understanding the human auditory system.

The binaural model makes use of cues for sound localization to create virtual sound that can be placed anywhere in the perceptual space(sound field) of a listener. So-called Head Related Transfer Function (HRTF) are used by a binaural model to generate a binaural sound.

The HRTF is a frequency response function of the ear. It describes how an acoustic signal is filtered by the reflection properties of the head, shoulders and most notably the pinna before the sound reaches the ear. This transfer function is the fundamental for the generation of three-dimensional sound for binaural models.

This project uses *binaural technology* that reproduces a binaural signal from a monaural one using a generalized set of HRTFs. The synthesized signal is then presented to the listener via headphones.

## 1.1 Premise

This document requires to be familiar with digital signal processing and digital filter design in general. Also it is preferable to read an introduction into basics of sound.

## 1.2 Related Work

There are previous, relative works which have used a *binaural technology* to implement a 3D simulation of a virtual sound scene over headphones. Two of them, will be presented in this chapter. The term *binaural technology* is defined as follow.

”Binaural technology is a body of methods that involves the acoustic input signals to both ears of the listener for achieving practical purposes, for example, by recording, analyzing, synthesizing, processing, presenting, and evaluating such signals” [2]

A comparable work was done in Massachusetts Institute of Technology (MIT) [13] and in frame of the Listen project [15]. Both projects have performed the HRIR/HRTF measurements but in different ways. MIT project has used the KEMAR dummy head microphone to obtain the left and right impulse response using maximum length(ML) pseudo-random binary sequences. Measurements of the Listen project were done using humans in blocked-meatus conditions which is apparently the common technique to get HRTF data . There, microphones were placed in the listener’s ears. As a measurement signal a 8192 points logarithmic sweep was used. This thesis also uses a binaural technology to implement a Headphone-Based 3D-Audio System.



## 2 Fundamentals of Spatial Hearing

This chapter provides relevant fundamental information pertaining to sound and spatial hearing from the perspective of physical and psycho-acoustics. It is organized as follows. In section 2.1 the basics of sound will be depicted. Section 2.2 is a compact depiction of the Hearing Organ and its components which are decisive for this thesis. Section 2.4 - 2.9 introduces cues for sound localization.

### 2.1 Physical Characteristics of Sound

The sound is the most important part in the acoustics, hence it is useful to depict the basics first. The sound is referred to as an any pressure variation in an elastic medium such as air, liquid or gases, which can be perceived by humans or animals. Sound waves can be propagated and perceived, if only there is medium between sound source and ear, so that the sound can propagate [14].

The ear differentiates between periodic and non-periodic motions. The periodic ones appear to us as pleasant and the non-periodic ones as bothersome. Particular periodic wave is harmonic and simple. The sine wave is periodic wave and of particular importance. Music and speech waves are complex and differ completely from simple sinusoid. However, mathematically every periodic wave even if it is not like sinusoid can be depicted as a combination of harmonic waves of various frequencies, amplitudes and phases . J.B. Fourier was the one who has invented this theorem, which is called Fourier theorem.

In general, humans perceive airborne sound which results from fluctuation of air pressure. The sound pressure  $p$  is a variation in the pressure generated by a sound wave and is superposed to constant atmospheric pressure(air pressure in absence of sound).

$$p = \frac{F}{A}$$

Thus, the sound pressure is  $F$ (force) divided by  $A$ (area). According to the SI(Systeme International) the unit of sound pressure is Pascal.

$$1Pa = \frac{1N}{m^2}$$

In the acoustics sound pressure can be given as root-mean-square or effective(rms). As a mea-

sured variable for the pressure is usually the sound level SPL which is given in dB.

$$SPL = 20 \lg \frac{p}{p_0}, \quad \text{whereas } p \text{ is the sound pressure}$$

$$p_0 = 2 * 10^{-5} = 20 \mu Pa$$

The standardized referenced value  $p_0$  is a pressure, which corresponds approximately to the human auditory threshold at 1000 Hz.

Loudness level is another term that is related to the characteristics of the human hearing. It measures how humans perceive the loudness of a sound which differs from human to human. Loudness is a subjective term and its unit is Sone. Sone is used to describe the difference in loudness between sounds. 1 Sone represents 40 Phons and is defined as a loudness of a 1 kHz tone at 40 dB SPL (sound pressure level). Loudness level is a physical term to express the loudness of a sound and its unit is Phon [14]. To get an idea of Phon the *Equal-loudness level contours* which was refined by Robinson and Dadson is used. Figure 1 depicts the human hearing response as a frequency-loudness level relation. The diagram shows how a pure tone of 1000 Hz must be adjusted (in terms of sound pressure level) until it is equally loud as a sound detected by listeners. Also, it shows, that the sound pressure level of a tone of 1 kHz and the loudness level are equal. For tones of frequencies aside from 1 kHz other sound pressure level are required to gain the same loudness impression [14].

The sound pressure is measured by a sound level meter. Usually a sound level meter is equipped with a filter whose frequency response is approximately like those of the human ear which is called A-weighting filter and corresponds to the inverse of the 40 dB (at 1 kHz) equal loudness level contour, see figure 2 (marked blue). A-weighting filter is used for noise measurements in audio systems. The abbreviation for A-weighted decibels is dBA. Aside from A-weighting filter there are B, C and D filter which are used for louder noises such as aircraft noise. When we talk about acoustic measurements the dB SPL (sound pressure level) is used which is referenced to  $20 \mu Pa = 0 \text{ dB SPL}$ . In frame of this thesis the Audio Analyzer R&S UPV was used for the sound pressure level measurement which also provides A-weighting filter.

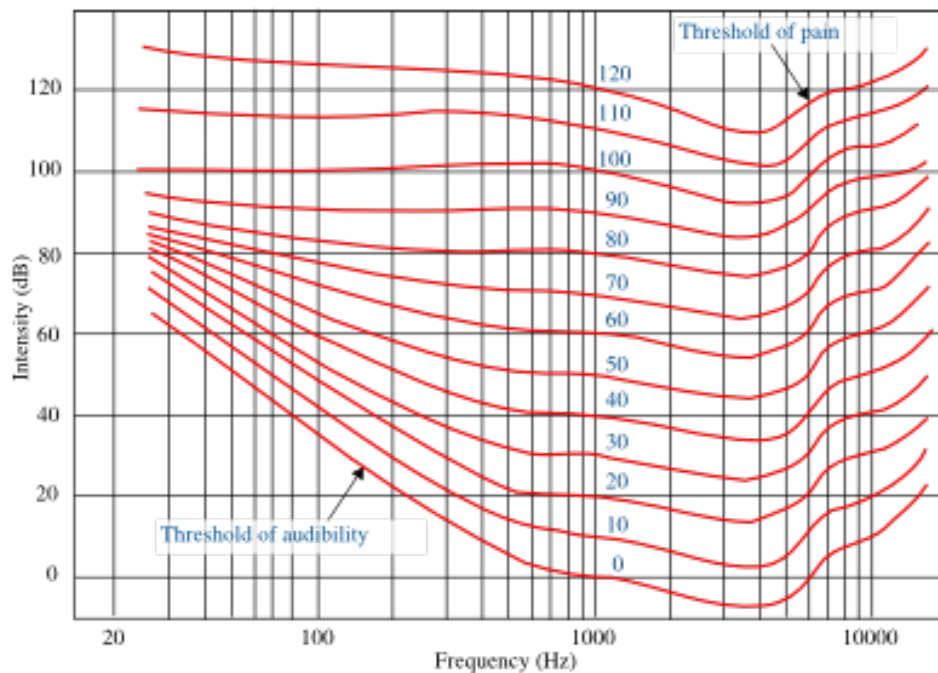


Figure 1: Equal-Loudness Contours of the Human Ear [1]

## 2.2 The Hearing Organ

The human hearing organ can be considered as an amazing signal processor due to its complex construction. It is divided into three sections the outer, middle and inner ear. Specific functions are assigned to each section. In this project however, the outer ear is of particular interest because it imposes filtering that helps to differentiate the direction of the sound source. The outer ear includes the pinna and ear canal ,see figure 3.

The pinna is decisive in terms of localization of a sound. It transmits sound through the ear canal to the ear drum. This process is very dependent on frequency, direction and distance of the incident sound. The pinna acts as a filter and causes distortions of a sound, which means that the amplitude and a phase spectrum of a sound are modified [3] . The pinna can be seen as a sound-gathering device.

Not only the pinna but also physical effects such as shadowing, reflection and diffraction of the head causes linear distortions. Furthermore, there are resonances in the pinna - ear canal - ear drum system. In technical terms these alteration effects can be described as Head Related Transfer Function's (HRTF) for the left and right ear respectively. The detailed description of HRTF will be presented in the section 2.9. The transfer function is mostly affected by the terminating impedance of the ear canal which is the impedance of the ear drum [3]. Section 3.3 will take a closer look on how the certain sections of the outer ear and the head affect a transfer

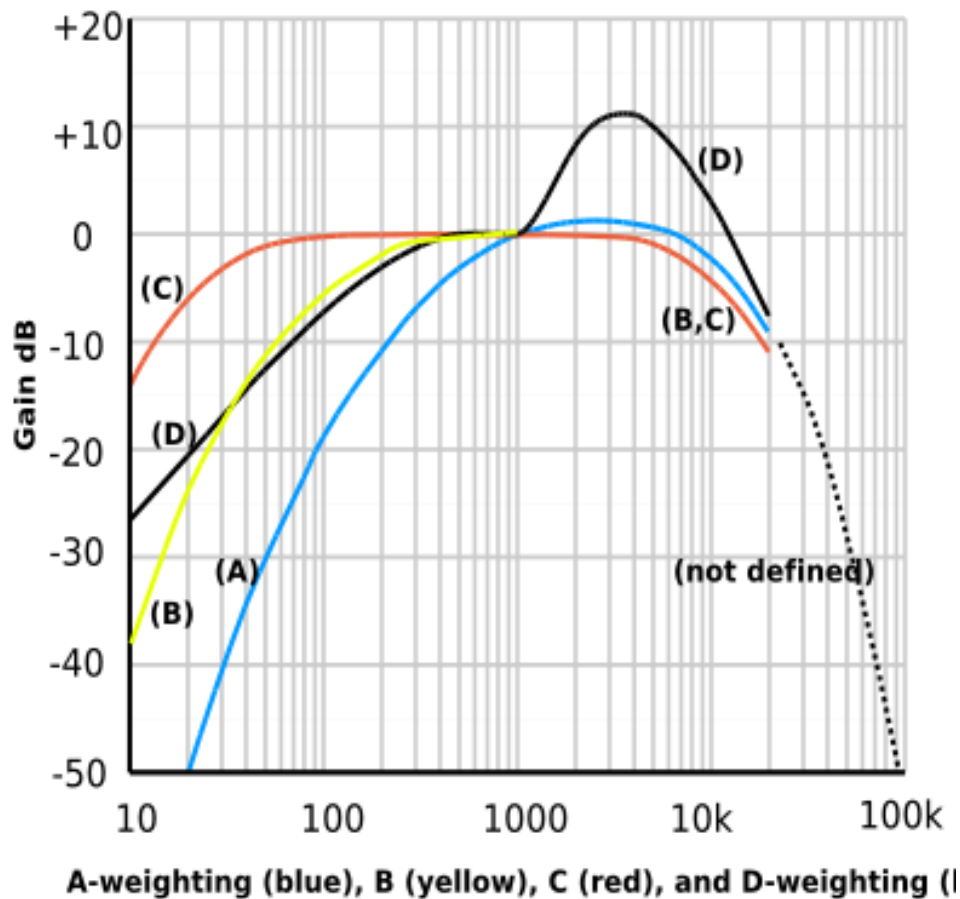


Figure 2: Weighting Filter Curves [26]

function.

The middle ear has a bunch of purposes such as propagating the sound waves from the outer ear to the inner ear. The propagation occurs due to the eardrum which is contained in the middle ear. If the air pressure in the compressed area of air molecules increases, the eardrum arches inwards and outwards, if the air pressure decreases. During the propagation alterations in sound-conducting media occur that cause changes of the magnitude of the impedance. The outer ear is filled with air and the inner ear with fluids. Due to the changes in media about 98 percent of a sound waves are reflected. In order to avoid the reflections, the low impedance in air must be transformed into the higher one of the fluids. This is actually the major function of the outer ear of such a transformer [4].

In terms of inner ear it is sufficient to mention that it includes the organ of hearing (cochlea). The inner ear converts mechanical vibrations into neural pulses. At this point the conversion between the time domain and frequency domain of the signal takes place.

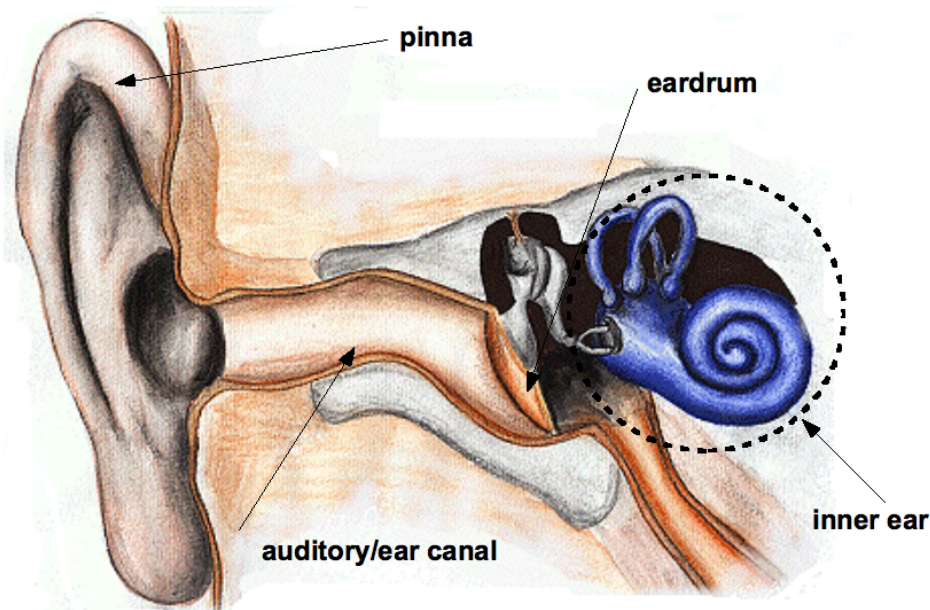


Figure 3: The Human Ear [28]

### 2.3 Characteristics of Human Hearing

Humans audible range involves frequencies from 20 Hz - 20 kHz . Figure 4 depicts the limits of human hearing. The sound pressure of 140 dB is referred as the threshold of pain. In contrary to the threshold of pain, the threshold of hearing is strongly frequency dependent. For each frequency there is another auditory threshold. The human ear is most sensitive around 3 kHz. For higher and lower frequencies the auditory threshold increases rapidly. To hear a sound of low frequency of 30 Hz the sound pressure must be around 50 dB.

### 2.4 Head Related Coordinate System

Before going into the details about spatial cues, some notes should be made on the head-related coordinate system, see figure 5 , which is being used throughout the following sections. The system is used to represent the position of an acoustic noise in space. There are three distinct planes: Horizontal, Median and Frontal. The origin of the system is placed on the inter-aural

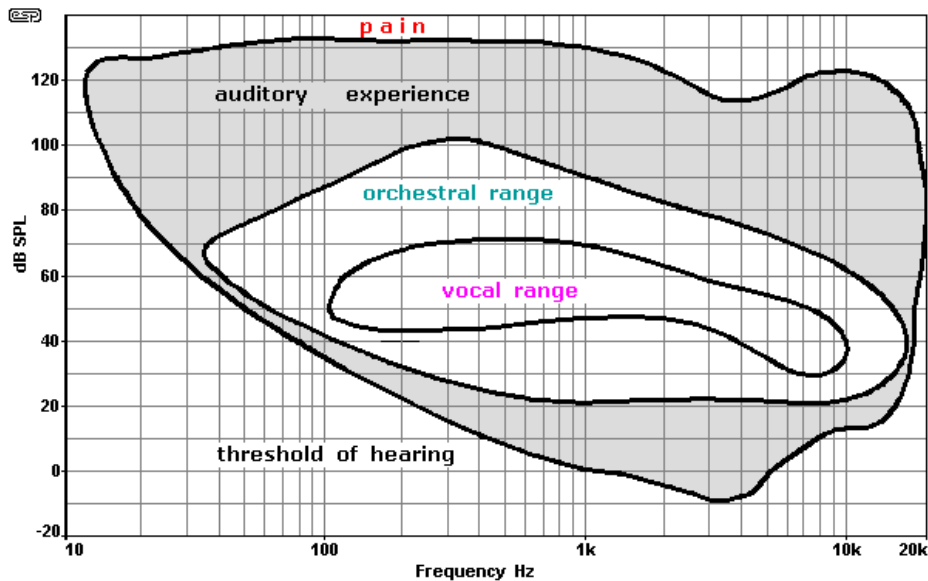


Figure 4: Human Hearing Range [2]

axis. It is the centre between the entrances to the ear canals. The inter-aural axis divides both eardrums.

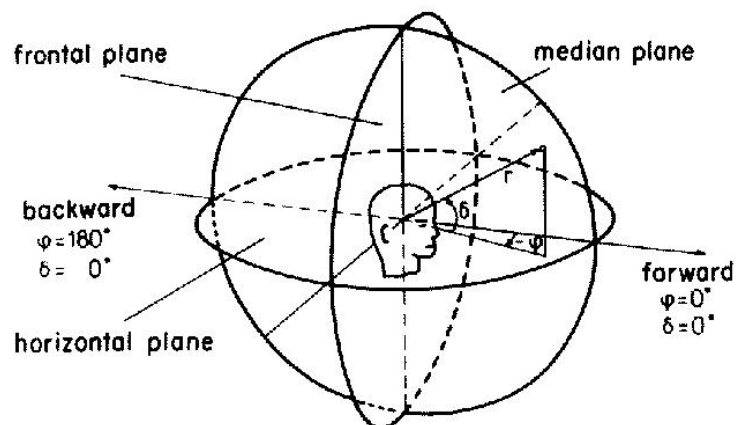


Figure 5: Head Related Coordinate System [14]

The position of the sound source is specified by: Azimuth  $\phi$ , Elevation  $\delta$  and Distance  $r$ , see Figure 5. The incident angle of the sound produced in the horizontal plane corresponding to the head is described by its azimuth  $\phi$ . It defines the movement of the sound anti- or clockwise through the horizontal plane.

The position of the sound in the median plane is described by elevation  $\delta$ . The median plane lies orthogonally on the horizontal plane as well as on the frontal plane. The frontal plane is vertical to the horizontal plane too. For example if  $\phi = 180$  and  $\delta = \delta$ , the sound source is directly behind the listener.

## 2.5 Inter-Aural Differences

According to the Duplex Theory which was invented by John Strutt, who is better known as Lord Rayleigh, there are two most important cues for sound localization in the horizontal plane: Inter-aural time Differences (or ITD) and Inter-aural level Difference (or ILD) [4] In the following these cues will be examined in detail.

### 2.5.1 Inter-Aural Time Differences

The time differences occur due to the difference in distance between the sound source and the two ears of the head. This difference is referred to as "inter-aural time difference", ITD. It is the difference in time that it takes for a sound wave to reach both ears. When the sound source is located to the right side of the head, it will reach the right ear earlier than the left one. The signal has further to travel the distance  $s$  to reach the left ear, see Figure 6.

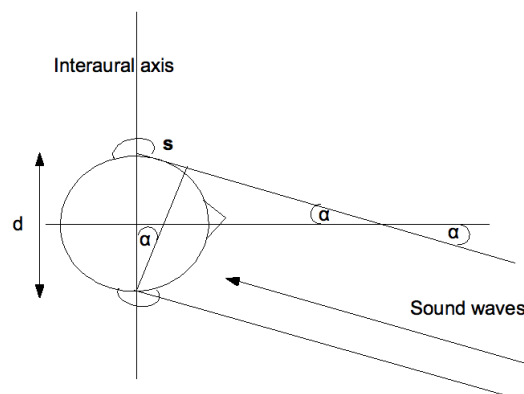


Figure 6: Inter-Aural Time Differences [28]

The ITD's are decisive in the low frequency range ( $< 1,5$  kHz). This is due to the relation between the wavelength and proportions of the head. In this case a wave length is larger than the diameter of the head. Sound waves then are not reflected by the head, see Figure 7.

This effect can be explained as follow, the formula for the wave length is:

$$\lambda = \frac{c}{f}$$

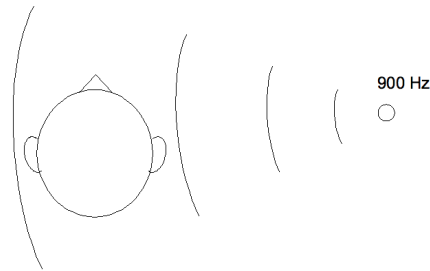


Figure 7: Waves of Low Frequencies

The distance a wave travels during one cycle period is a wavelength. Sound travels at a speed  $c$  of 343 m/s. The wave length for a frequency of 900 Hz is then 38,11 cm ,which is larger than a diameter of the head (18,5-23 cm).

Due to the physical dimensions of the head the maximum ITD is 0.63 ms when the sound source is placed at the azimuth  $-90^\circ$  or  $+90^\circ$ , left and right ear respectively [28]. Sound waves strikes a head at the certain angle. The incident angle of a sound  $\alpha$  is specified by the azimuth  $\phi$  according to the head related coordinate system, see figure 5. The ITD which is specified by  $\Delta s$  can be calculated as follow.

$$ITD = \frac{\sin \alpha}{d}$$

### 2.5.2 Inter-Aural Level Differences

The head is an attenuation factor that leads to interaural level differences of both ear signals. ILD (also frequently referred to as interaural intensity differences) is denoted as an amplitude difference between left and right ear by the sound's wave front. The hearing organ is able to perceive level differences in the whole audible frequency range to detect a direction of a sound waves. In the previous model for ILD estimation it has been alleged that the ILD is more decisive in the high-frequency range. The ILD becomes effective approximately above 1.5 kHz where the wave length is less than the diameter of the head. Due to diffraction of the head there are no perceptible level differences approx. below 500 Hz. This is called the head-shadow



effect, where the far ear is in the shadow of the head, see Figure 8.

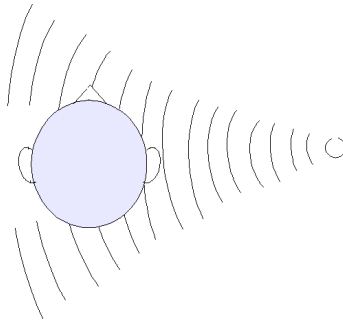


Figure 8: Waves of High Frequencies

According to Rayleigh's Duplex Theory ILD and ITD taken together contribute to sound localization throughout the audible frequency range. Also this modes asserts that there are no perceptible ITD and ILD when the sound arrives at the azimuth of  $0^\circ$ . However it has lately been disclosed that ITDs are important in the high-frequency range as well as ILDs [4].

Although the Duplex Theory is quite simple, it only asserts the perception of azimuth. There are situations where a sound has the same ILD and ITD for two points, which have equal distance from the listener's head. This is so-called "Cone of Confusion" and is explained in the next section.

## 2.6 Cone of Confusion and Monaural cues

The "Cone-of-Confusion" is critical in the median plane, see figure 4. From any point in this plane the interaural differences are at a minimum or zero for a perfect symmetrical head. Nevertheless the listener can distinguish sounds from points in the median plane, because the head is not perfectly symmetrical. The interaural differences can be identified, these are individual and dependent on elevation angle and in which half space the sound occurs(back or front) [3]. Figure 9 illustrates the concept of the "Cone of Confusion". Here, points 1 and 4, and points 2 and 3 have identical ITD and IID.

Several authors have supposed that monaural cues are used to locate sound sources in the median plane. For the monaural cues only one ear is sufficient to perceive them. Monaural cues are



flection properties of the head, shoulders and most notably the pinna before the sound reaches the ear. The outer ear has directional and nondirectional properties of spectral alterations in a certain frequency range. The torso and the head causes directional, the pinna and ear canal nondirectional modifications, which will be discussed in the Section 3.3 in detail. HRTFs are used to synthesize a monaural into binaural sound which will sound as if it comes from desired location. All linear characteristics of the sound transmission are contained in the HRTF. Localization cues such as ITD, ILD as well as monaural cues are included. They can be extracted from the HRTF, but the opposite is not the case [4].

Every human has his individual HRTFs dependent on his physical shape and his outer ear. To achieve the authentic auditory perception, an individualized set of HRTF should be used. Another way to synthesize a virtual sound is by using generalized set of HRTF acquired by measuring with the Dummy Head. Figure 10 depicts the HRTF in the frequency domains for the azimuth angle  $90^\circ$  and elevation  $0^\circ$ .

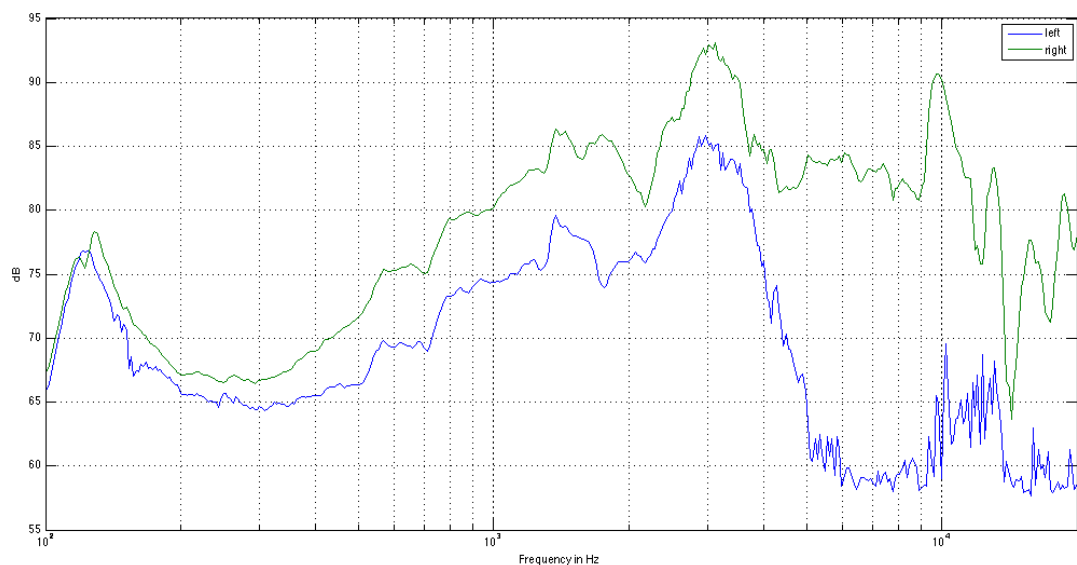


Figure 10: Frequency Response for Azimuth  $90^\circ$  and Elevation  $0^\circ$ .  $x$ [Hz],  $y$ [dB], left = blue, right = green

It is obvious that the sound reaches the ipsilateral ear (which is close to the sound-source) is stronger than the sound that reaches the opposite ear. In the frequency domain, the frequency response for the contra-lateral ear (left microphone of the dummy head) looks like a low-pass filter due to the shadowing effect of the head. The detailed description of the HRTF obtained in frame of this thesis is presented in the section 3.3.

Measurements of HRTF are often done from a fixed radius around the dummy head or listener's

head in an anechoic chamber. They are measured from different azimuths and elevations. Even when the measurements take place in an anechoic chamber errors can occur, because of unintended reflections of the equipment. Hence, it is desirable to reduce the amount of the equipment. The measurement setup for this thesis is described in the section 3.2. A conventional technique for measurement is to place microphones into a subject's ear and then playing a sweep signal through the speaker.

As mentioned in the section before instead of using a person it is more convenient to use an artificial head to perform the HRTF measurements. The dummy head is an approximate reproduction of the human head which has almost identical acoustical properties of a human. It is used to perform binaural recordings, whereas spatial effects of a sound are transformed into two electrical signals(also referred to as binaural signals ), which stand for the input of the right and left ear respectively. The recorded signals are not equivalent to those of the ear signals that the subject would have in the given situation. However, dummy head recordings are a good way to reproduce those signals authentically, even if individual deviations exist. Using the obtained HRTFs of a dummy head or of others, listeners are still capable of doing a correct localization assumptions.

## 3 Measurements

### 3.1 Overview

In this thesis a system has to be implemented which is based on existing methods and a newly obtained HRTF database. The aim was to simulate the transfer function of the outer ear as good as possible, using the dummy head and digital filters. The detailed description of the essential measurements gives section 3.2. In section 3.4 the implementation of the system will be examined. All required measurements were made at University of Applied Sciences in the anechoic chamber.

### 3.2 Measurement Technique

#### 3.2.1 Loudspeaker Measurement

A loudspeaker converts an electrical signal into sound pressure. As an initial step, the frequency response of the sound source in free field (a surroundings where no reflections exist) was measured. For all measurements the loudspeaker CANTON Plus XL.2 was used. This measurement was necessary because the response is not linear due to the complex mechanical structure of a loudspeaker. Not linear means, that different frequencies of the same input signal are differently reproduced by a loudspeaker. The frequency range of the loudspeaker corresponds to 40 Hz- 30 kHz. Figure 11 shows the measurement set-up.

The loudspeaker was placed at the incident angle  $0^\circ$  and the distance corresponds to 1.91 meters. The microphone and the loudspeaker were connected to the amplifier and the amplifier to the Audio Analyzer R&S UPV input and output respectively. The analyzer is running on Windows XP. The analyzer generated a sweep signal (from 40 Hz to 20 kHz) of 500 or 800 samples. This was required because in the first measurements, the phase difference between the right and left microphone of the dummy head were measured with 800 samples. For the measurement the free-field 1/2" microphone Type 4190 from Bruel & Kjaer was used, which has a flat frequency response. A microphone converts sound pressure into an electric signal. A free-field microphone is used to measure the sound pressure in a free field. The frequency response of the free field microphone is linear for the incident angle of  $0^\circ$ . The interference of a reflected sound which causes alteration of a sound pressure are compensated by the microphone. Figure 12 shows the frequency response of this microphone.

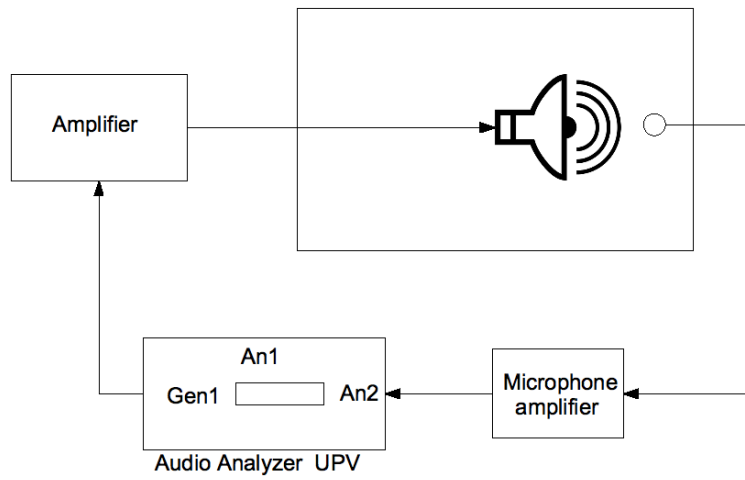


Figure 11: Loudspeaker Measurement Setup

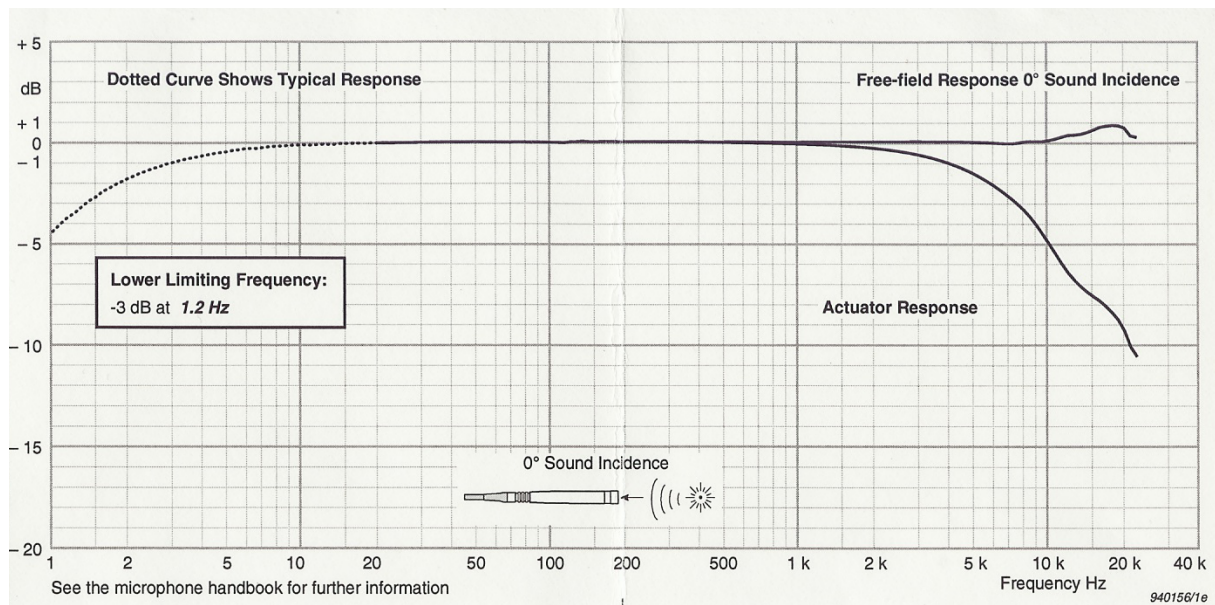


Figure 12: Free-Field Response 0° Sound Incidence

It is observable that it is flat in the frequency range of the loudspeaker. Before the actual measurement has started, the calibration of the microphone was needed. The calibration is the measurement with which the value of the microphone sensitivity is obtained. Sensitivity indicates how good the microphone converts sound pressure to output voltage. It should be denoted that the re-calibration of the microphones should be done for the next measurement. This is required because a mechanical wear and a drift in electronic components affect the accuracy of the measurements.

$$s_M = \frac{V}{p}$$

Wherein  $V$  is the effective or rms of the microphone output voltage and  $p$  is the rms of the sound pressure which produced the output voltage. The calibration is done by generating a sine wave signal with a common frequency of 1 kHz and sound pressure of 1Pa. The obtained microphone sensitivity of -33.96 dBV(referenced to 1 V = 0 dBV) minus 94 dB (127.96), was then entered in the Analyzer's Configuration Panel as a Referenced Value. This is necessary because the calibrator produces the pressure of 94 dB, thus the sound pressure can be displayed in dBSPL(dB referenced to 20  $\mu$  Pa) [21]. The full description of the settings of the analyzer will be discussed in the next section. The resulting frequency response is shown in figure 13.

It is observable that the response of the speaker is not linear. Especially low frequencies below 100 Hz cannot be reproduced by this loudspeaker because of its small membranes. The advantage of a loudspeaker with small membrane is that it can reproduce higher frequencies better than those with big membranes. The obtained frequency response was then used to equalize the HRTF measurements.

### 3.2.2 Analyzer Settings for Loudspeaker and HRTFs Measurements

First at all it should be denoted that the settings for all measurements can be saved and reused for the next measurement. This section describes the most important settings for the measurement of the loudspeaker response and HRTFs. The Audio Analyzer R&V UPV which was used for measurements has an analog interface, hence the setting for the **Instrument** should be *Analog*. The Channel 1 of the generator was used as an output for a sweep signal, see figure 14. (left panel), for HRTF measurement the output signal is the same for both channels (right panel).

The Bandwidth was set to 40 kHz and for the HRTF 22 kHz which means that maximum output frequency corresponds to 40 kHz and 22 kHz respectively. The **Output Type** was set to

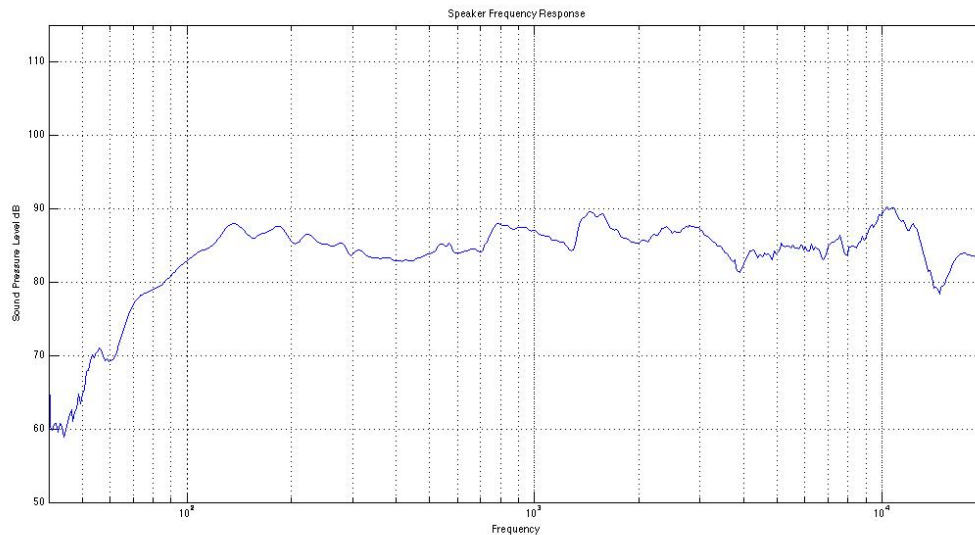


Figure 13: Loudspeaker Frequency Response.  $x[\text{Hz}]$ ,  $y[\text{dB}]$

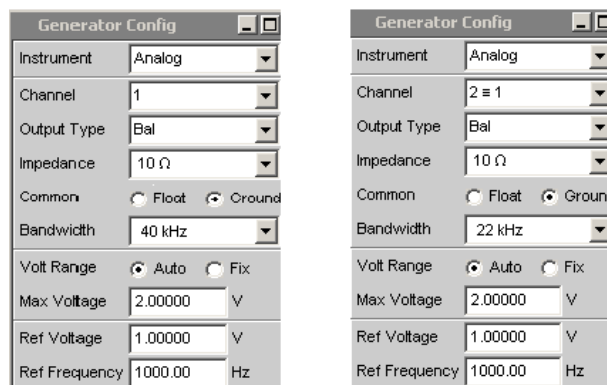


Figure 14: Generator Config Panel

*Bal*, which means that the output is symmetrical. This setting is preferable in audio technique because of unsymmetrical and long electrical cabling and many electronic devices, that are connected to a different power supply. This can lead to noise, distortion and ground loops. To avoid these unintended effects, XLR connectors are used. This connector enables further symmetrical processing by a generator or an analyzer. The advantage of the symmetrical processing is that the interferences which affect the signal can be avoided. In frame of this thesis XLR connectors were used for the generator output as well as for the analyzer input. Also the length of the cables was not that long to produce those unintended interferences. The **Common** option was set to *Ground*, however for audio measurements the *Float* is preferable to avoid the interference signal which occurs due to the electrical cabling and is added to the wanted signal [24], [21], thus *Float* should be used in the Generator Config as well as in Analyzer Config. As already mentioned, this may not affect the measurements that much because of the setting of the **Out-**



**put Type** and minimized amount of equipment. The Impedance should be set to  $10\Omega$  to keep the internal resistance as low as possible, which enables to produce a decent voltage source. **Max Voltage** should be adjusted individually to avoid harmonic distortions which are caused by an amplifier. In the Generator Function and Analyzer Function panel the settings for the loudspeaker and HRTF measurements are the same, see figure 15.

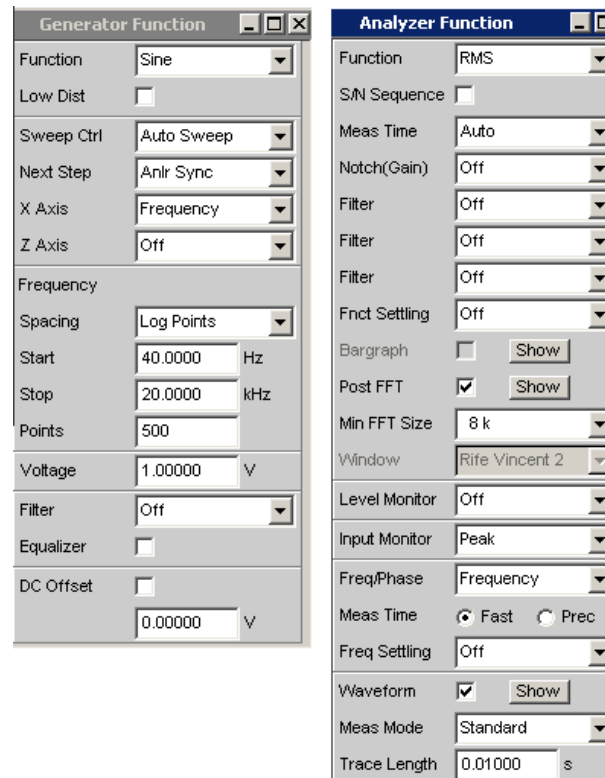


Figure 15: Generator and Analyzer Config

The **Function** is a *Sine* and in the menu **Sweep Ctrl** the *Auto Sweep* should be set, thus the parameter of the sweep can be configured, such as **Start** and **Stop** frequency. The **Start** frequency corresponds to  $40\text{ Hz}$  and the **Stop** to  $20\text{ kHz}$ . It is possible to fix the desired number of points of the sweep signal in the menu **Points**, a maximum of 1024 points is allowed. The higher the number of points the longer it takes to measure the responses. In this case 500 points are sufficient. In the Analyzer Function panel as the **Function** *RMS* should be set. Further settings have to be fixed in the Sweep Graph 1 Config, see figure 16. Depending on which Analyzer channel was used to connect the output of the microphone, the Reference value of  $-127.96\text{ dBV}$  should be set, see figure 16.

This is necessary because the Analyzer records AC rms voltage and display it as the absolute voltage level which is given as follow.

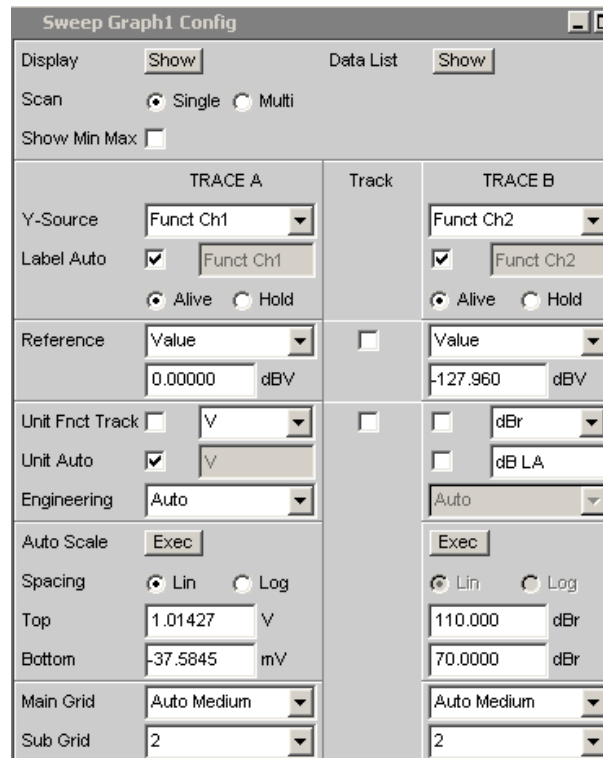


Figure 16: Sweep Config for Loudspeaker Measurement

$$20 \log \frac{V}{V_{ref}}, \quad \text{where } V_{ref} \text{ is a reference voltage.}$$

In menu **Unit Fct Track** *dBr* should be set, this allows using customized settings of the Reference value. The same settings should be done for the HRTF measurements. The obtained sensitivity for the left microphone of the dummy head corresponds to -38.100 dBV and of the right -38.500 dBV minus 94 dB of the calibrator it is equal to -132.100 and -132.500 respectively, see figure 16.

Also the difference in phase between the right and the left microphone signal was measured on the channel 2. As the reference the channel 1 was used, see figure 18. (right panel). It is possible to display the recordings in various formats such as degree or radiant. In frame of these measurements degree was chosen. To obtain the phase difference in the Analyzer Function panel, see figure 18(left panel), in the menu **Freq/Phase** the *Freq&Phase* should be set. All the recorded values can be saved as a trace file, see figure 17 at the very end.

**Sweep Graph1 Config**

Display  Data List

Scan  Single  Multi

Show Min Max

TRACE A		Track	TRACE B	
Y-Source	Funct Ch1		Funct Ch2	
Label Auto	<input checked="" type="checkbox"/> Funct Ch1		<input checked="" type="checkbox"/> Funct Ch2	
	<input checked="" type="radio"/> Alive <input type="radio"/> Hold		<input checked="" type="radio"/> Alive <input type="radio"/> Hold	
Reference	Value -132.100 dBV	<input type="checkbox"/>	Value -132.500 dBV	
Unit Fnct Track	<input type="checkbox"/> dBr	<input type="checkbox"/>	<input type="checkbox"/> dBr	
Unit Auto	<input type="checkbox"/> dB LA		<input type="checkbox"/> dB LA	
Engineering	Auto		Auto	
Auto Scale	<input type="button" value="Exec"/>		<input type="button" value="Exec"/>	
Spacing	<input checked="" type="radio"/> Lin <input type="radio"/> Log		<input checked="" type="radio"/> Lin <input type="radio"/> Log	
Top	116.000 dBr		116.000 dBr	
Bottom	60.0000 dBr		60.0000 dBr	
Main Grid	Auto Medium		Auto Medium	
Sub Grid	2		2	
Limit Upper	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Limit Lower	<input type="checkbox"/>		<input type="checkbox"/>	
Limit Shift	<input type="checkbox"/>		<input type="checkbox"/>	
X-Source	Sweep			
X-Axis	Frequency	Label Auto	<input checked="" type="checkbox"/> Frequency	
Unit	Hz	Unit Auto	<input checked="" type="checkbox"/> Hz	
Engineering	Auto	Ref Value	1000.00 Hz	
Scaling	<input type="radio"/> Manual <input checked="" type="radio"/> Auto	Spacing	<input type="radio"/> Lin <input checked="" type="radio"/> Log	
Left	40.0000 Hz	Right	20.0000 kHz	
Main Grid	On			
Sub Grid	Coarse			
Resolution	5dig / 0.001 dB			
Store Traces				
Store Trace as	Equalization List	Trace	<input type="radio"/> A <input checked="" type="radio"/> B	
Norm Freq	1000.00 Hz	Invert	<input checked="" type="checkbox"/>	Modify Equ. <input type="checkbox"/>
Store Trace to	d:\UPV\Setups\sima\SpeakerXPA225.veq			<input type="button" value="..."/>

Figure 17: Sweep Config for HRTFs Measurement

### 3.2.3 HRTF Measurements

First of all, the microphones of the dummy head HMS II.4 from HEAD acoustics were calibrated in the same way as the free-field microphone. Frequency and phase responses were obtained in the horizontal and median plane at the distance of approximately 1.91 meter from the sound source, see figure 19.

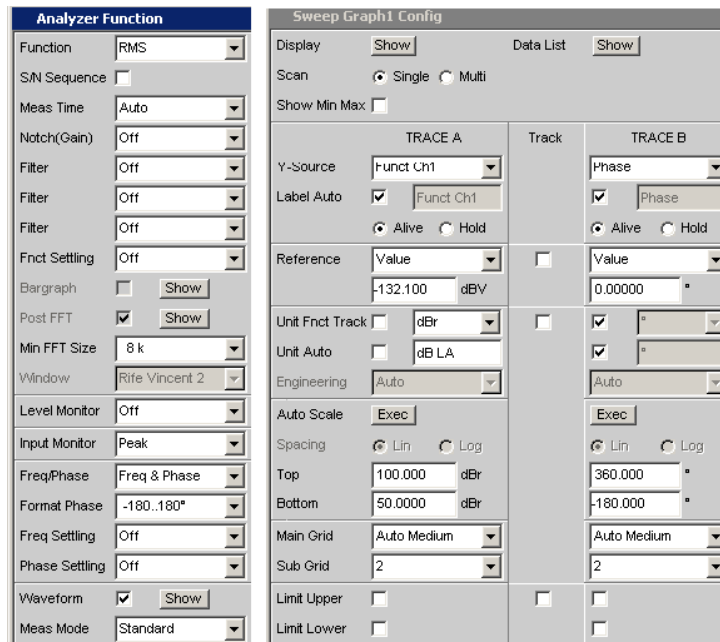


Figure 18: Analyzer Function and Sweep Graph Config Panel for Phase Measurement

In order to measure the phase response the left input of the analyzer was used as a reference signal. Altogether the measurements took place two times on different days. Table 1 shows all measured locations. The measurements were done only for the left area and for two positions of the right area, see figure 32, because it is symmetrical. To get the frequency response for the azimuth  $345^\circ$  the frequency response of  $15^\circ$  can be used. The frequency response of the right microphone is used for the left and vice versa.

Elevation	Number of Measurements	Azimuth Increment	Date of Measurement
$-41^\circ$	9	30	22.04.2008
$+45^\circ$	9	30	22.04.2008
$0^\circ$	9	30	22.04.2008
$+37.53^\circ$	15	15	14.07.2008
$0^\circ$	15	15	14.07.2008

Table 1: Measured Position in the Median and Horizontal Plane

The measurements in the horizontal plane and for positions of the positive elevation were done two times on different days and with different resolution for the azimuth. During measurements in the horizontal plane, the dummy head and the loudspeaker were placed on the floor, see figure 20, in the median plane, they alternatively were positioned at the platform, see figure 21.

For this measurements the same sweep signal (40 Hz - 20 kHz) was generated by the Audio Analyzer with 500 and 800 samples. The measurements contain the left and right HRTF re-

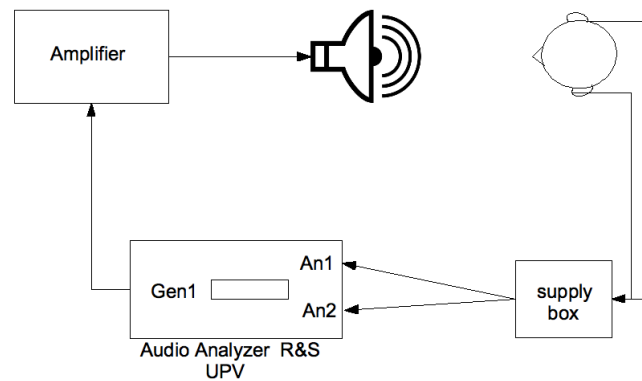


Figure 19: HRTF Measurement Setup

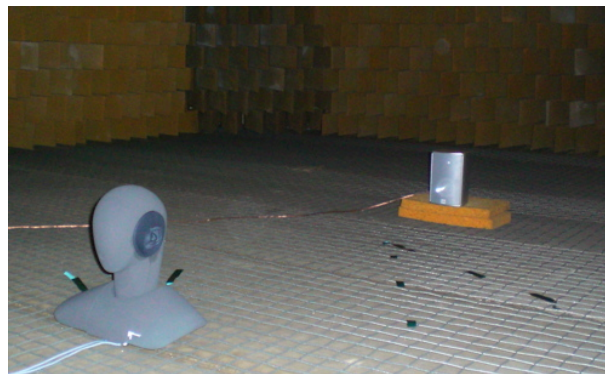


Figure 20: Measurement in the Horizontal Plane

spectively. Figure 22 shows the left HRTFs measured in the horizontal plane and figure 23 for the positions of  $+37.53^\circ$  elevation and azimuth increments of  $15^\circ$ .

Observable are strong sound pressure deviations in the frequency range between 2 and 4 kHz and slight differences at the low frequencies as expected. The reason for those deviations are the directional and the nondirectional spectral modifications due to the physical proportions of a dummy head, which will be discussed in the next chapter. It should be denoted that the measurements below 100 -150 Hz could not be reproduced due to the limited capability of a loudspeaker in the low frequency range.

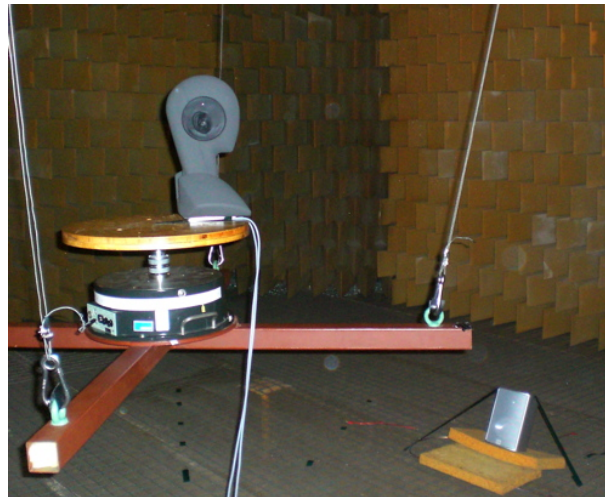


Figure 21: Measurement for Elevation of  $+37.53^\circ$ .  $x$ [Hz],  $y$ [dB],  $z$ [degree]

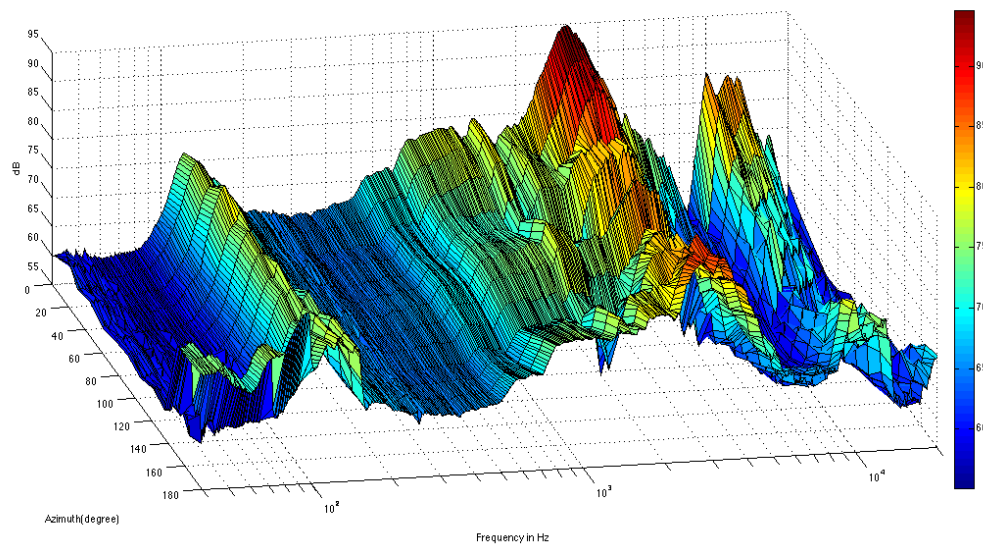


Figure 22: Left HRTFs measured in the Horizontal plane

### 3.3 Analysis of HRTF curves

In the chapter 2.9 it is denoted that the outer ear leads to directional and nondirectional modifications of an incoming sound. The HRTF of the dummy head at the incident angle of  $0^\circ$  in the horizontal plane will be discussed in detail.

The absence of the low frequencies, see figure 24 (green circle), is due to the limited capabilities of the loudspeaker, see section 3.2.1 and 3.2.3. The distortions caused by the head occur at the frequencies starting from 500 Hz, see figure 22 (pink circle) [3].

The acoustical behavior of the ear canal is given approximatively by the formula of an Helmholtz resonator. The Helmholtz - Resonator is used to determine a frequency where these resonances

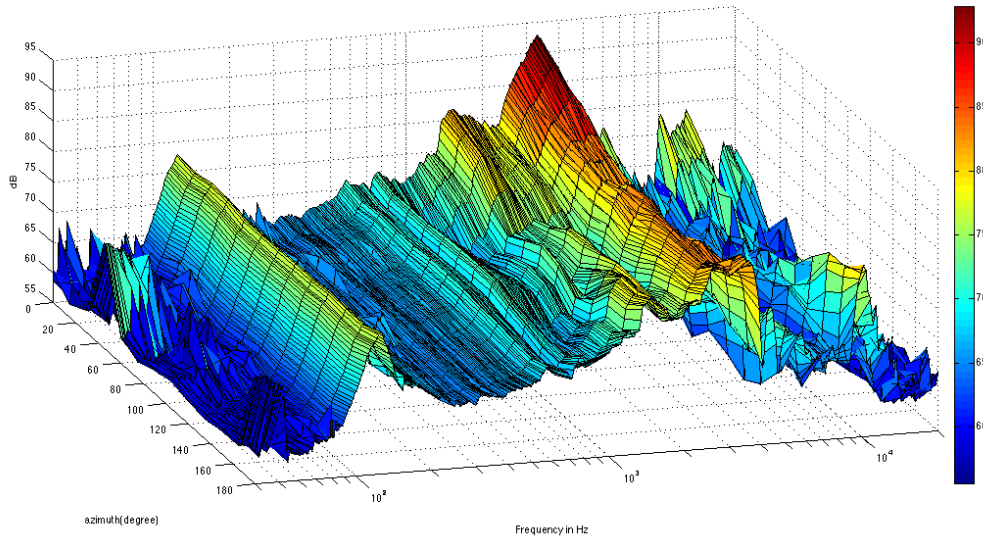


Figure 23: Left HRTFs measured for the Positions of  $+37.53^\circ$  Elevation and Azimuth of  $0^\circ$  to  $180^\circ$ . x[Hz], y[dB], z[degree]

occur. The equation for the Helmholtz-Resonator is given as follow.

$$f = \frac{c}{2\pi} \sqrt{\frac{\pi r^2}{V * l}}$$

Whereas  $r$  and  $l$  denotes the radius and the length of the ear canal respectively,  $c$  is the velocity of a sound and  $V$  is the cavity of the pinna. The radius of the ear canal corresponds to ca. 4 mm and its length is ca. 25 mm then volume of the pinna can be calculated as follow.

$$V = \frac{\pi r^2}{l * \left(\frac{f 2\pi}{c}\right)^2}$$

$$V = \frac{\pi 4\text{mm}^2}{25\text{mm} * \left(\frac{1500\text{Hz} * 2\pi}{343000\text{mm/s}}\right)^2} = 2663\text{mm}^3$$

In the figure 24 it is recognizable that amplifications of the pinna occur at the frequency range from 1 to 1.5 kHz (red circle). The ear canal causes amplifications due to its resonance at the frequency around 3 kHz. In addition the ear canal is similar to an organ pipe, because it terminates with the eardrum. The resonances are the same as those of the organ-pipe:  $\frac{\lambda}{4}$ ,  $\frac{3\lambda}{4}$  and  $\frac{5\lambda}{4}$ . The first resonance occurs at the frequency 3430 Hz (turquoise circle), assuming that the length of the ear canal averages 0.025 m which is equal to  $\frac{\lambda}{4}$ . From this it follows that the wavelength  $\lambda$  corresponds to 0.100 m [23].

$$f = \frac{c}{\lambda} = \frac{343m/s}{0.100m} = 3430Hz$$

The next resonance occurs at approximately  $\frac{343m/s}{0.033m} = 10394Hz$  (orange circle), and the last at  $\frac{343m/s}{0.020m} = 17150Hz$  (black circle), which can vary from human to human due to different physical shape. The deep notch around 8 kHz (big black circle) and serrated peaks at the very end occur due to the folds and ridges of the pinna [20]. More researches on human auditory system can be found here [17]. The analysis of the measurements has shown that the dummy head reproduces the anatomical model of the human auditory system as described above.

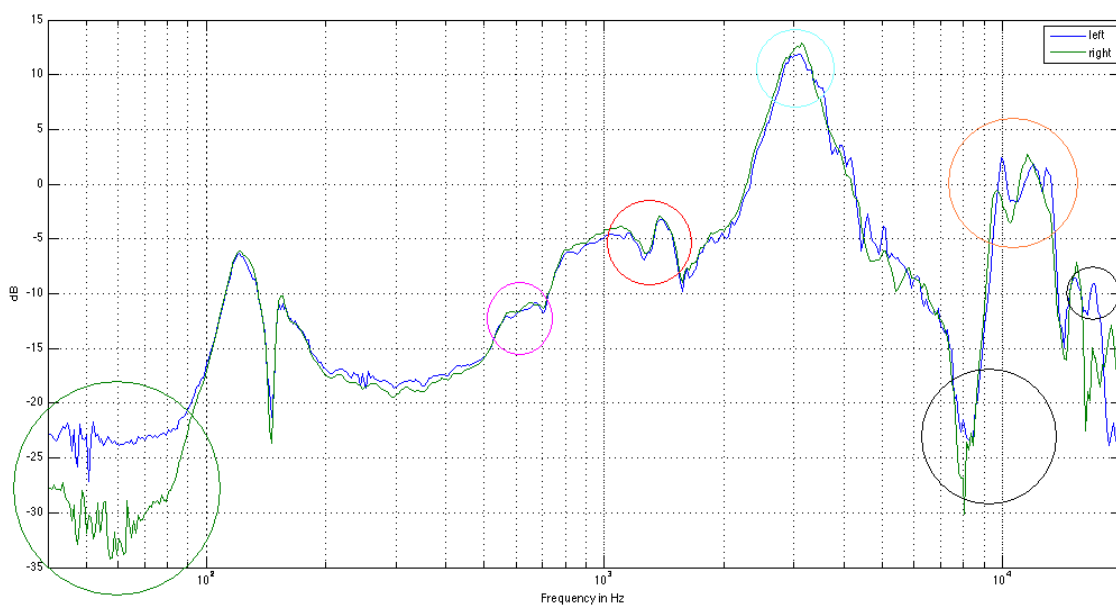


Figure 24: HRTF pair for 0° Elevation and Azimuth. x[Hz], y[dB]

### 3.4 Headphones Measurements

Additionally, various headphones from different manufacturer were measured placed on the dummy head. Table 2 depicts measured headphones. The settings of the Audio Analyzer are the same as those of the HRTF measurement, which are described in the section 3.2.1

The headphones causes distortions of a signal, therefore a pre-equalization of the sound for playback is required. To compensate those distortions an inverse filter was implemented, see section 4.2. Additionally the left and the right channel of some headphones cannot be reproduced equally, they differ from each other. Figure 25 shows the frequency response of the headphones Sennheiser HD 270 which was used for the system test. In the frequency range between 300 and 400 Hz of the right channel there is a valley and at 500 Hz there is a peak.



Measured Headphone
Sennheiser HD 500
Sennheiser HD 600
Sennheiser HD 270
Sennheiser CX300
Gembird AP 880
Manhattan
Philips

Table 2: Measured Headphones

In the high frequencies there are many serrated peaks and valleys, this effect is observed in all headphones measurements. This effect is caused by the pinna as discussed in section before.

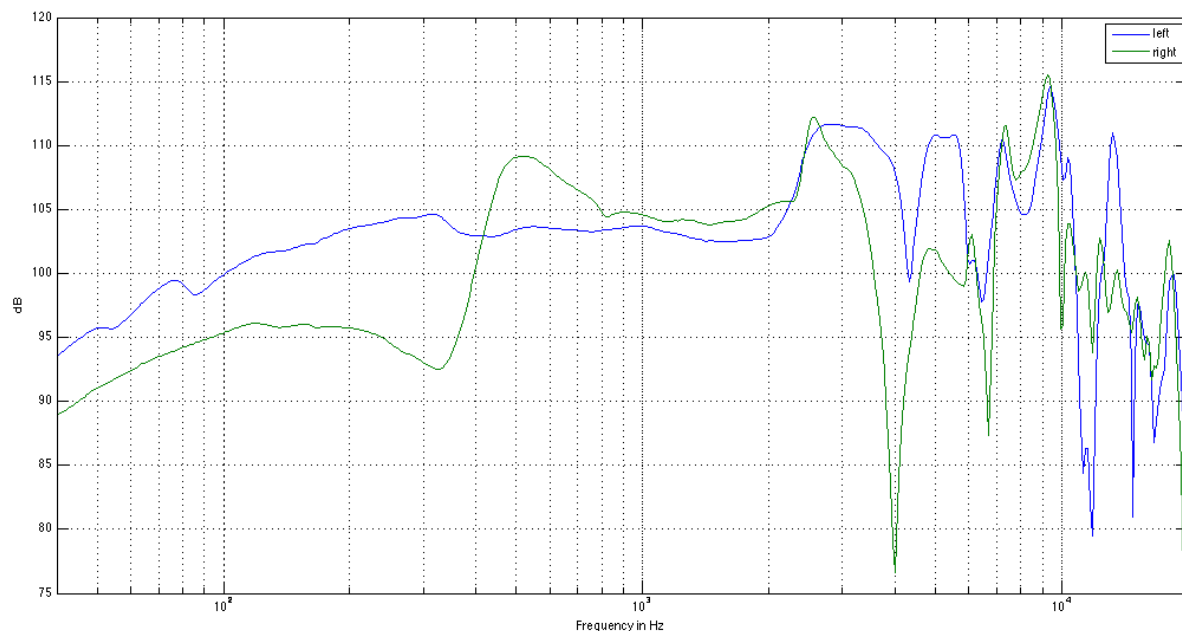


Figure 25: Frequency Response of the Sennheiser HD 270. x[Hz], y[dB]

Compared to the Sennheiser HD 270 the frequency response of the Sennheiser HD 600 looks better, see figure 26. In the low frequency range there are no peaks and valleys and the response is nearly even. The peak at 3 kHz is caused by the ear canal of the dummy head.

Figure 27 shows the frequency response of the cheap Philips earbud type headphones. Those type of headphones are especially tricky to measure because they don't have a good seal. For the system test, which is described in section 5, only Sennheiser HD 270 and HD600 headphones were used.

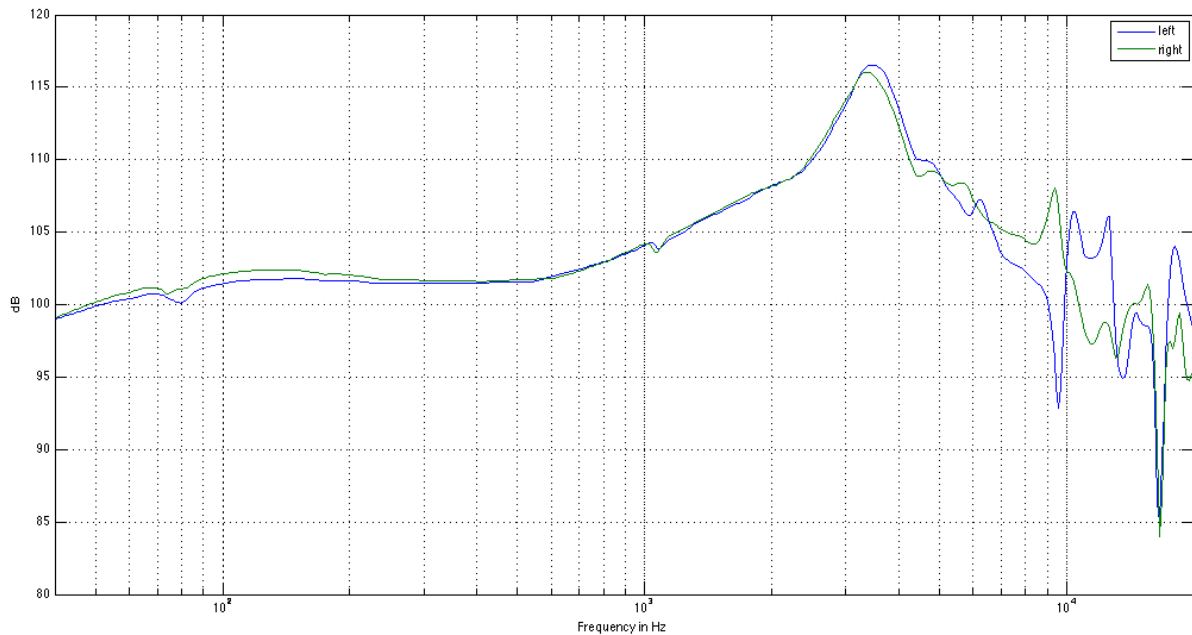


Figure 26: Frequency Response of the Sennheiser HD 600.  $x[\text{Hz}]$ ,  $y[\text{dB}]$

## 4 System Design

The designed system consists of a pair of digital filters that are characterized by an input and an output signal and its transfer function. The system is completely implemented in Matlab. Figure 28 depicts the structure of the system.

The frequency response of the dummy head was measured using the Audio Analyzer as described in the previous section. The obtained data was used to create a HRIR filter and a further filter to equalize the headphones. Section 4.1 gives a detailed description of the Head Related Impulse Response(HRIR) filter which is a Finite Impulse Response (FIR) filter. Section 4.2 describes the implementation of an inverse filter designed to equalize headphones.

### 4.1 HRIR Filter

There are two types of digital filters: Infinite Impulse Response(IIR) and Finite Impulse Response(FIR). In frame of this thesis the system was implemented as a FIR filter because of its linear phase response which is especially desired in digital audio processing [8]. Linear phase filter do not produce phase distortions of the signal, which means that the delay caused by a filter is the same at all frequencies. The advantage of the linear filter is that they can be implemented with a given frequency response and phase. The condition is that the desired response must include linear phase. The main disadvantage of FIR filters is that they require high pro-

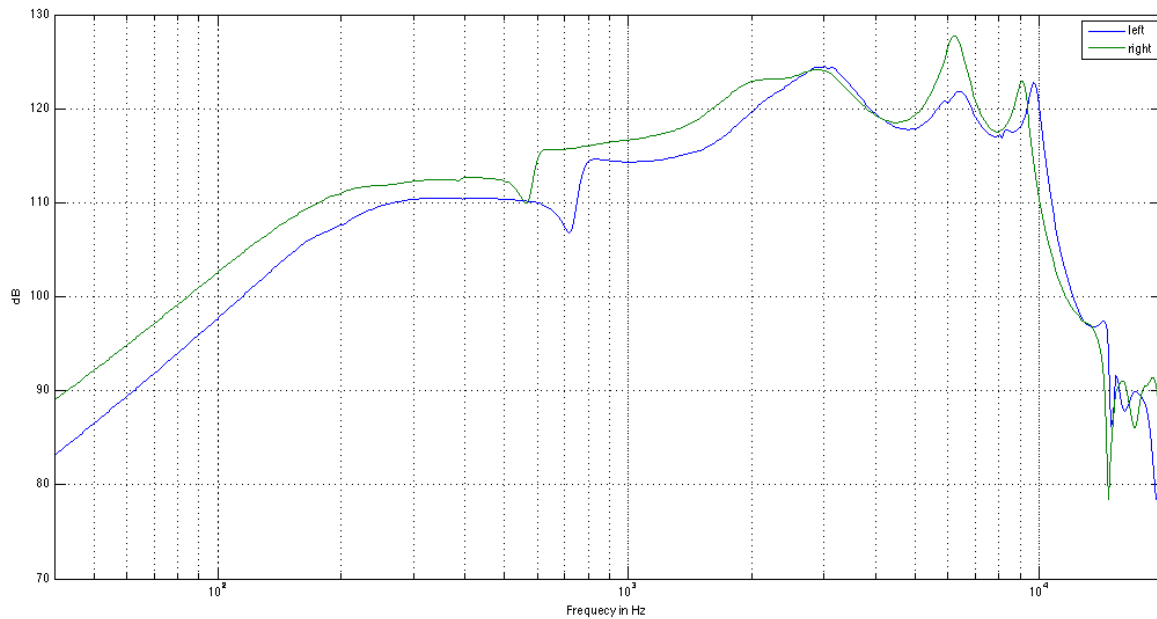


Figure 27: Frequency Response of the Philips Earbud Type Headphones.  $x$ [Hz],  $y$ [dB]

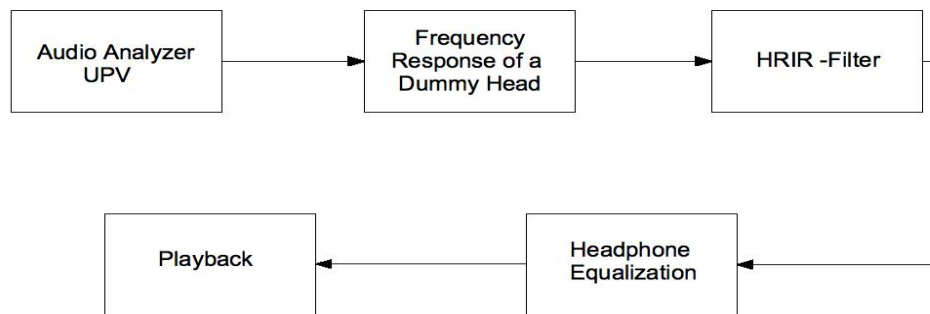


Figure 28: Structure of the System

cessing time and storage. In this thesis these parameters were not a part of the requirements. The FIR filter is implemented using the window method [8]. The window method is commonly used for FIR filter design. In this method the desired response of the filter is defined and the coefficients are obtained by evaluating the Inverse Fourier Transform (IFFT). The number of coefficients  $N_{FILT}$  corresponds to 512 which is sufficient to meet the requirements of the system. The code listing below means that we take 512 samples of the given frequency response at intervals  $k * f_s / N_{FILT}$ , where  $k = 0, 1, 2, \dots, N_{FILT} - 1$ . The sampling frequency  $f_s$  corresponds to 44,100 Hz.

```
ffreq = (0 : NFILT - 1) * fs / NFILT;
```

First of all, the data(frequency-, phase- and speaker-response) is imported into the workspace using the matlab script named `get_data.m`. The data is available as a m-File which was prepared for Matlab using the executable `PrepareData.jar`.

The next step is the interpolation of the speaker response at the frequency points of the dummy head response, see the code snippet below. The interpolation is required because the data contains different numbers of points(500 or 800).

```
ar_speaker = interp1(freq_speaker , ar_speaker , freq_ar);
```

As already mentioned in the previous chapter, the frequency response of the loudspeaker is non-linear and the frequencies below 100 Hz cannot be reproduced by the speaker at all, therefore the normalization of the level is required. The level of the frequency range below 100 Hz is replaced by the last index of the `rng_speaker_lo` and above 18 kHz by the first index of the `rng_speaker_hi`, see code snippet below.

```
rng_speaker_lo = find(freq_speaker < 100);
ar_speaker(rng_speaker_lo) = ar_speaker(rng_speaker_lo(end));
rng_speaker_hi = find(freq_speaker > 18000);
ar_speaker(rng_speaker_hi) = ar_speaker(rng_speaker_hi(1));
```

A simple subtraction is performed to equalize the frequency response of the dummy head. Whereas `ar_dB_l` and `ar_dB_r` are a frequency response of the left and right microphone respectively.

```
ar_dB_l_1 = ar_dB_l - ar_speaker;
ar_dB_r_1 = ar_dB_r - ar_speaker;
```

The next step is converting the frequency response from decibel into the magnitude and the phase response from degree to radiant.

```
ar_l = 10.^(ar_dB_l_1 / 20);
ar_r = 10.^(ar_dB_r_1 / 20);

pr_r = pr_deg_r * pi / 180;
```

The filter coefficients of a FIR filter are computed using IFFT. IFFT transforms a signal from frequency domain into the time domain. The fact is that the IFFT of a conjugate-symmetric vector is real [12]. By using this fact, mirroring makes sure that the transmitted signal is real. The following code lines shows the mirroring concept. First an upper and a lower half of the symmetric vector is defined.

```

rng_hi = NFILT : -1 : NFILT/2 + 1;
rng_lo = 1:NFILT/2;

```

In the following lines, only the desired frequency response of the lower half is defined because the upper half is a mirror of the lower half.

```

far_l(rng_lo) = interp1(freq_ar, ar_l, ffreq(rng_lo), 'linear', '
    extrapol');
far_r(rng_lo) = interp1(freq_ar, ar_r, ffreq(rng_lo), 'linear', '
    extrapol');

```

In the mirroring the lower half of the response is flipped around, complex conjugated and placed at the end of the lower half. The lines below show the flipped magnitude for the left and right channel respectively. The phase response `fpr_diff` represents the difference in phase between the left and the right channel.

```

far_l(rng_hi) = far_l(rng_lo);
far_r(rng_hi) = far_r(rng_lo);

fpr_diff = [fpr_diff; -fpr_diff(end:-1:1)];

```

Figure 29 depicts the magnitude after mirroring.

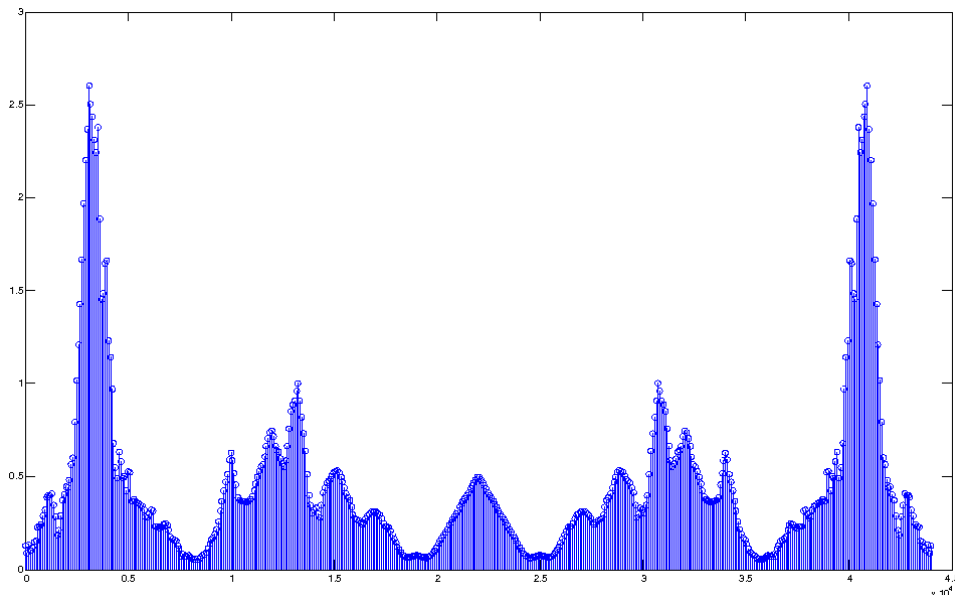


Figure 29: Magnitude after Mirroring

To realize the filter with a linear phase response its phase response must fulfill the following condition [8]:

$$\theta(\omega) = -\alpha\omega,$$

where  $\alpha$  is constant. The line below defines the linear phase response, where  $\alpha$  corresponds to  $-\pi$ . This means that the whole spectrum is delayed by  $(NFILT/2) - 1$ , because the  $NFILT$  is even. This step has two reasons: to obtain the linear phase characteristics, to delay a signal by half a filter length and then perform the window function (in this case the Hanning Window). The impulse response is shifted in order to avoid that the interesting parts of the impulse response are cropped by a window function.

$$\text{fpr\_base} = (0 : NFILT - 1)' * (-\pi);$$

Also the coefficients of the filter must be symmetrical around the center, that means, the first coefficient is the same as the last, the second as the next-to-last [19]. In this case the phase response is a function of the filter length [8], where  $\alpha = (NFILT/2) - 1$ :

$$h(n) = (NFILT - n - 1), n = 0, 1, \dots, (NFILT/2) - 1 \quad , \text{where } NFILT \text{ is even.}$$

The next step is to compute the phase response for the left and right channel, see the code below. Here the left and right signal is just shifted to the opposite directions.

$$\begin{aligned} \text{fpr\_l} &= \text{fpr\_base} - 0.5 * \text{fpr\_diff}; \\ \text{fpr\_r} &= \text{fpr\_base} + 0.5 * \text{fpr\_diff}; \end{aligned}$$

The coefficients of the filter are computed using the IFFT which represents the impulse response of the filter.

$$\begin{aligned} \text{b\_l} &= \text{real}(\text{ifft}(\text{far\_l} .* \exp(j * \text{fpr\_l}))); \\ \text{b\_r} &= \text{real}(\text{ifft}(\text{far\_r} .* \exp(j * \text{fpr\_r}))); \end{aligned}$$

The coefficients represent a selected set of time-domain samples.

$$\text{wnd} = \text{hanning}(NFILT);$$

$$\begin{aligned} \text{b\_l} &= \text{b\_l} .* \text{wnd}; \\ \text{b\_r} &= \text{b\_r} .* \text{wnd}; \end{aligned}$$

The impulse responses which is referred to as Head Related Impulse Response (HRIR) for the left and right ear are then saved in a Matlab m-file using the matlab script `write_coeff.m`. Figure 30 shows HRIR for the left and right ear of the dummy head at the incident angle  $90^\circ$  in the horizontal plane. As expected the signal arrives at the right ear before the signal at the left ear.

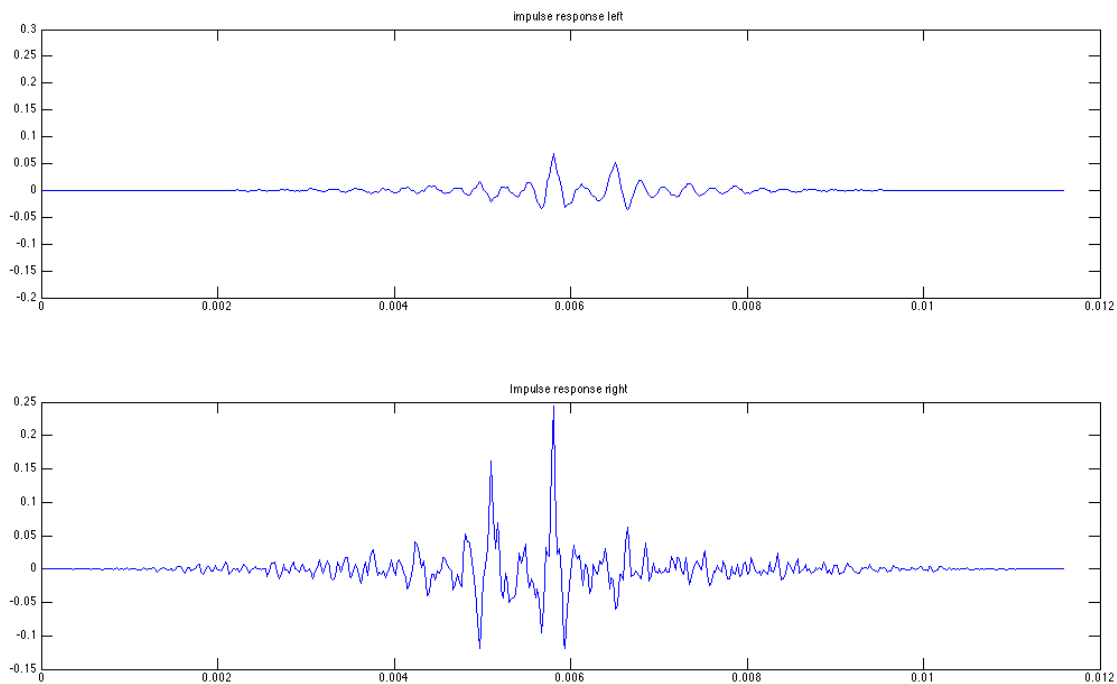


Figure 30: Left and Right HRIR for incident angle of  $90^\circ$  in the horizontal plane

## 4.2 Headphone Equalization

As already mentioned, sound samples need to be pre-equalized for playback. The distortions caused by headphones are compensated using an inverse filter with linear characteristics. The process of inverse filtering is also referred to as deconvolution. Using deconvolution a signal can be compensated for an undesired convolution. The aim is to reproduce the signal as it existed before the convolution took place. In order to design an inverse filter which recreates the original signal, the frequency response of the filter must be inverse of the response of the headphones. That is done by inverting the frequency spectrum of headphones, see code snippet below.

```
ar_dB_hp = - ar_dB_hp;
```

The inverse filter was designed in the same manner as the HRIR filter using the window method to compute filter coefficients, see the code snippet below.

```
far_hp(rng_hi) = far_hp(rng_lo);  
  
b_hp = real(iffc(far_hp' .* exp(j * fpr_base)));  
  
b_hp = b_hp .* wnd;  
  
write_headphones_coeff(b_hp, headphones, headphonesHRIR);
```



## 5 Test

### 5.1 Overview

The hearing test is necessary to validate (Did we realize the right system?) and verify (Are we satisfied with the results?) the implemented system. The effects of varying parameters for the playback system will be analyzed. There are two parameters, the HRTF-Measurement and the headphones. The results will show how these parameters affect the perception of a sound.

### 5.2 Test Setup and Procedure

The sound samples were presented to the listener through two different headphones in varying environments respectively. It was not possible to perform the test in the same environment because some of the subjects could not arrange to come to the university. Sound samples were generated for different azimuth and elevation, see table 5.2. Speech was used as a testing-signal because it is familiar to humans, hence it is easier to localize.

Elevation	Number of Measurements	Azimuth Increment	Date of Measurement
-41°	9	30	22.04.2008
+37.53°	15	15	14.07.2008
0°	15	15	14.07.2008

Table 3: Positions for Generation of Sound Samples

The headphones Sennheiser HD 600 and Sennheiser HD 270 were used for the reproduction of sound samples and played back using iTunes. The test procedure is shown below in chronological order. A piece of paper was given to the testing subject with a circle on it. The circle is divided in many sections, see figure 31.

1. Briefing
2. Adaptation phase
3. Sound Playback
4. Interview

The lines represent all possible directions the sound can come from. The sound samples were numbered in the sequential order for 15° resolution and 30° resolution, see figure 32. The testing subject listened to the samples and put the number around the circle. If the direction of the speech could not be specified, it was played again.

**Test Record**

<b>Name:</b>	<b>Age:</b>	<b>Gender:</b>
<b>Elevation:</b>	<b>Date:</b>	
<b>Headphone Type:</b>		

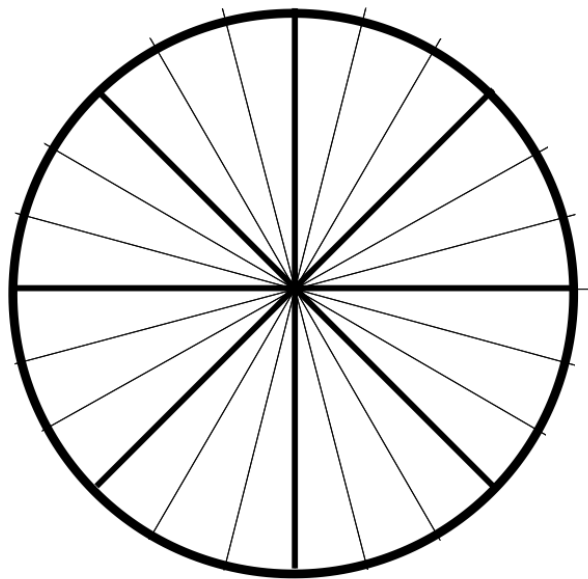


Figure 31: Testrecord

In the adaption phase a testing subject listened to several testing sound samples. The person was told to listen carefully and pay attention to changes between speech samples. Most testing-persons who have attend the test have alleged, that it was difficult to determine the direction of the speech due to the fact, that they do not really know how it would sound in a real situation. Before the test had begun, another person went around the testing subject and said something. The subject closed the eyes and pointed to the direction where the person was present. All of testing subject could determine the position accurately.

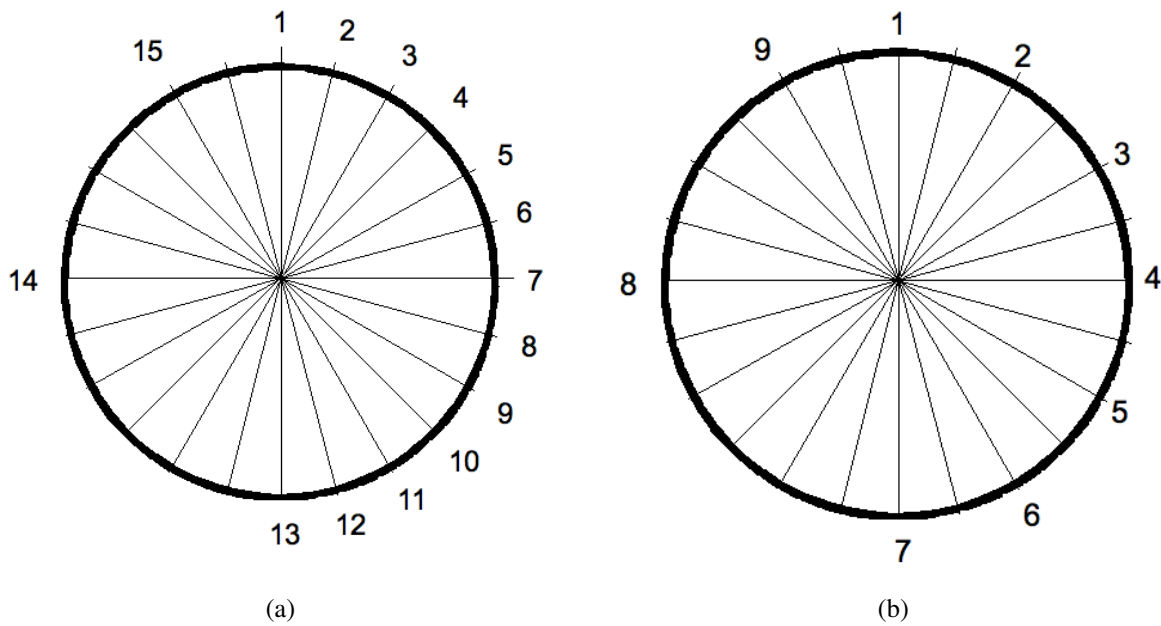


Figure 32: Numbered Positions for 15°(a) and 30°(b) Resolution

Section Number	Azimuth in degree
1	0°, 15°, 30°, 45°, 60°, 75°, 90°
2	105°, 120°, 135°, 150°, 165°, 180°
3	195°, 210°, 225°, 240°, 255°, 270°
4	285°, 300°, 315°, 330°, 345°

Table 4: Sections with corresponding angles

### 5.3 Test Results

Thirteen persons have participated in the test. The age of the testing subjects ranges from 23 to 56 years. The average age is 34.8 and The test results were gathered and merged for headphones type and elevation angle respectively, see figure 36 e.g.

After every test run the subject was asked to determine the elevation of sound samples. Table 5.3 shows the results .The elevation 0° was identified in most cases correctly with Sennheiser HD 600 as well as with HD 270. For some persons it seemed as if the speech came from different elevation even if it was the same. The majority of testing subjects have said that high frequency sounds appear more above and low frequency sounds below the eye level. The elevation 37° has been determined properly by seven persons with the HD 270 and HD 600. Only four subjects could determine the elevation -41° with HD 270 and three subjects with HD 600.

The "Front-Back Confusion" effect could be observed. Most of the testing subjects have confused the front with the back. Some have said that the speech sample no.1 (0°) and no.13 (180°)

No.	Gender	Age	HD 270 elev 0°	HD 270 elev 370°	HD 270 elev -41°	HD 600 elev 0°	HD 600 elev 37°	HD 600 elev -41°
1	m	23	n	y	n	y	y	y
2	m	24	y	y	n	n	n	n
3	m	24	y	n	n	y	n	n
4	m	25	y	y	n	y	n	n
5	m	25	n	n	y	n	y	y
6	m	26	y	y	n	y	y	n
7	f	26	y	y	y	n	n	n
8	m	26	y	n	n	y	y	n
9	m	27	y	n	n	y	n	n
10	m	36	n	n	y	y	n	n
11	m	48	y	n	n	y	y	n
12	f	50	n	y	y	y	y	y
13	m	56	y	y	n	y	y	n

Table 5: Evaluation of the Elevation Angle, y = yes, n = no, m = male, f = female

sound equally, so that both numbers were placed at 0° or 180° of the circle 36. In test runs for the horizontal plane with the Sennheiser HD 270 and HD 600, two of the testing subjects could not localize speech sample no.1 that came from the front. The testing subjects had the impression that the speech appear inside the head, see figure 36, 39. Sample no. 1 was placed in the middle of the circle to denote that it appears inside the head. This effect is normal and occurs when headphones are used to play back a sound. This happens because of missing information about the distance of an incoming sound. A further reason is that humans move the head to detect the sound source which is not possible when a sound is reproduced by headphones [18].

For instance, in the next sections only test runs for the horizontal plane will be discussed.

### 5.3.1 Test run with Sennheiser HD 270 and HD 600 for the Horizontal Plane

Although some of the testing persons have complained about poor sound quality of the Sennheiser HD 270, the majority have localized the sample no.13 in the horizontal plane correctly, see figure 36 and also in the test run with HD 600. Sample no.1(0°) is scattered around the left half of the circle, but in some cases it was assigned to the right position. For most testing persons sample no.15 (330°) seems to appear behind the head and no.14(270°) at the front, see figure 36, 39. Testing subjects asserted that a sound that contains more low frequencies appear behind the head and high frequency sound more in the front. Figure 33 shows that the transfer function of the left and right HRIR filter of 330° incident angle looks like a low pass filter. This could be an explanation why the filtered speech sample sounds deep.

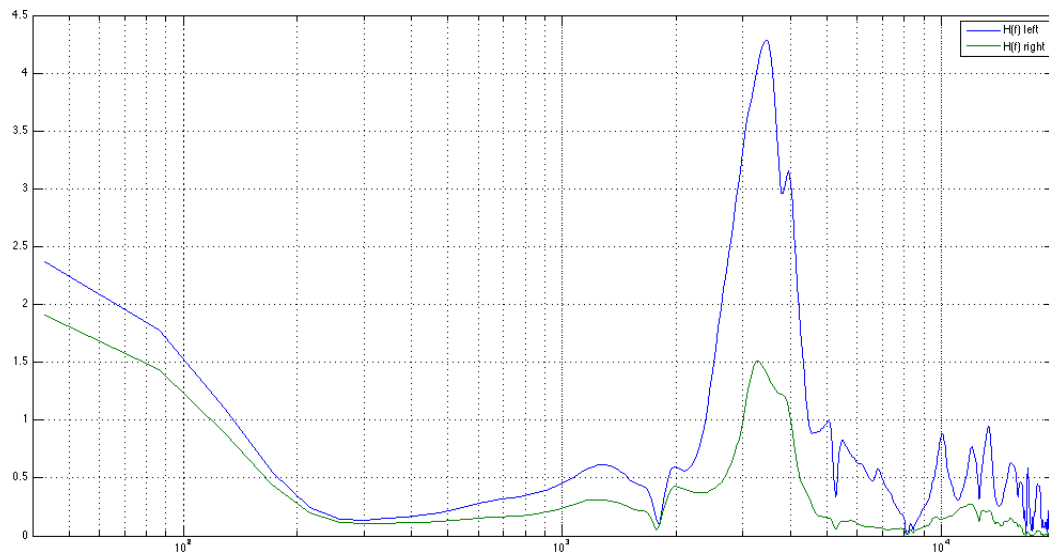


Figure 33: Transfer Function of the Left and Right HRIR Filter 330°.  $x[\text{Hz}]$ ,  $y[\text{dB}]$

The frequency response of the left HRIR filter of 270° incident angle behaves as a high-pass filter, see figure 34. In this case the filtered sound seemed to appear in the front and above the subject, although the elevation corresponds to 0°.

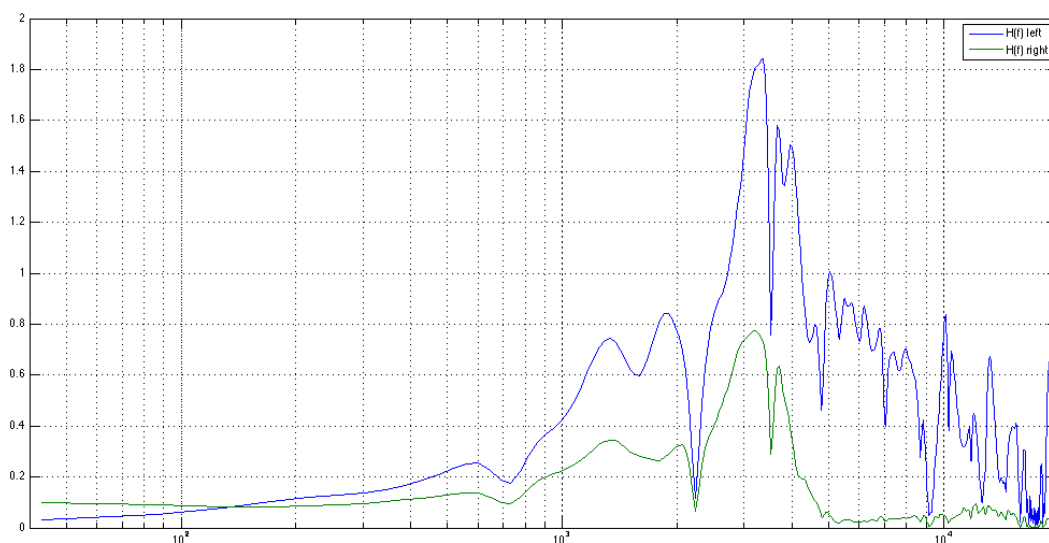


Figure 34: Transfer Function of the Left and Right HRIR Filter 270°.  $x[\text{Hz}]$ ,  $y[\text{dB}]$

Sample no. 2(15°) using the HD 270 was localized quite accurately in many cases unlike when the HD 600 was used. Here, the sample is scattered around the right half of the circle and gathered at the 135°. In this case one can assume that the transfer function of the HRIR filter must act as a low pass filter. On the picture 35 one can see that low frequencies are passed through the filter, this could be an explanation for the test results. Good results could be achieved with the samples no.4,5,6,7 in the test runs with the HD 270 as well as with the HD

600.

According to the "Cone of Confusion" samples no.6 and no.8 for instance in the horizontal plane can appear from the same direction, see figure 9. In fact some of testing subjects put those numbers very close to each other or sometimes exactly at the same position. The same effect must occur with sample no 7 for  $37^\circ$  and no.4 for  $-41^\circ$  elevation. Here one of the user was asked to try to determine exactly the elevation, but this was not the case.

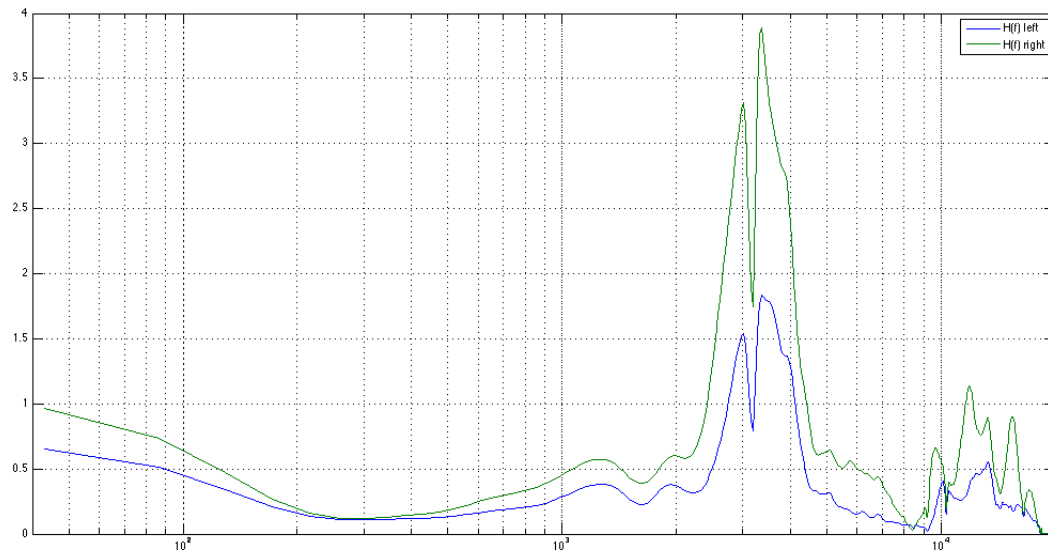


Figure 35: Transfer Function of the Left and Right HRIR Filter  $15^\circ$ .  $x[\text{Hz}]$ ,  $y[\text{dB}]$

Generally it appears that better test results in the horizontal plane could be achieved with the Sennheiser HD 270. It should be denoted that a higher number of testing persons is required to test the system extensively. Nevertheless, the test has shown that the measurement procedure and the FIR filter approach is a right way to develop a decent 3D Audio System and can be used for further works. In this test the samples were played back in sequential order. In the next test, samples can be played back randomly just to see if other results will come up. Also the test could be repeated with the same testing subjects. The measurements done on 22.04.2008 were not tested completely this can be done in further works. Another approach is to use the data measured by MIT and IRCAM for testing and compare it with the results of this project.

## 6 Conclusion and Further Development

The purpose of this thesis was to design and develop a 3D Audio System. This project involves the reproduction of a virtual auditory space in headphones using head related transfer functions(HRTF). It is a fact that this thesis is based on related works, but it has been shown that it is possible to realize such a system with less equipment and some effort. This project is an initial step in using head related transfer functions to create acoustic events. Further the system can be analyzed in terms of solving the "front-back confusion" problem using the following work [5]. To extend existing HRTF databases new measurements are required with a better resolution of azimuth and elevation or by using interpolation to obtain a desired HRTF. In this thesis a FIR filter approach was used, but also an IIR approach filter could be considered to see if the system could be enhanced. The project provides a basis for further developments like head-tracking or the simulation of room acoustics.

Overall this project has provided a method for a filter design and HRTF measurements.

## 7 References

- [1] ANSWERS.COM. *Frequency-Loudness Level Relationship*. [http://content.answers.com/main/content/wp/en-commons/thumb/4/4b/400px-FletcherMunson\\_ELC.svg.png](http://content.answers.com/main/content/wp/en-commons/thumb/4/4b/400px-FletcherMunson_ELC.svg.png)
- [2] BLAUERT. *Head Related Coordinate System*. <http://web.media.mit.edu/~raffik/zero-g/aup/Figure3.jpg>
- [3] BLAUERT, Jens: *Räumliches Hören*. 1. S. Hirzel Verlag Stuttgart, 1985. – ISBN 3-7776-0410-0
- [4] BLAUERT, Jens: *Communication Acoustics*. 1. Springer Berlin Heidelberg New York, 2005. – ISBN 3-540-22162-X
- [5] CHO, Sang J. ; OVCHARENKO, A. *Front-Back Confusion Resolution in 3D Sound Localization with HRTF Databases*. <http://ieeexplore.ieee.org/Xplore/login.jsp?url=/iel5/4107286/4107287/04107369.pdf&isnumber=4107287&prod=CNF&arnumber=4107369&arSt=239&ared=243&arAuthor=Sang+Jin+Cho>
- [6] ELIJAH BREDER, David M.: Csound - based Auditory Localization. In: *The Csound book: perspectives in software synthesis, sound design, signal processing, and programming*, MIT Press, 2000, S. 493 – 503
- [7] ELLIOTT, Rod. *Human Hearing Range*. <http://sound.westhost.com/articles/fadb-fl.gif>
- [8] EMMANUEL C. IFEACHOR, Barrie W.: *Digital Signal Processing*. 2. Prentice Hall, 2002. – ISBN 0-201-59619-9
- [9] EVEREST, F. E.: *Master Handbook of Acoustics*. 4. McGraw-Hill, 2001. – ISBN 0-07-136097-2
- [10] E.ZWICKER, M.Zollner: *Elektroakustik*. 1. Springer Berlin Heidelberg New York Tokyo, 1984. – ISBN 3-540-13509-X
- [11] FLIEGE, Norbert: *Systemtheorie*. 1. B.G. Teubner Stuttgart, 1991. – ISBN 3-519-06140-6
- [12] FOHL. *email conversation*
- [13] GARDNER, Bill ; MARTIN, Keith. *HRTF Measurements of a KEMAR Dummy-Head Microphone*. <http://sound.media.mit.edu/KEMAR.html>



- 
- [14] GÜNTER BOYE, Urbi F. H.: *Handbuch der Elektroakustik*. 3. Hüthig Buch Verlag Heidelberg, 1989. – ISBN 3-7785-1575-6
- [15] IRCAM. *HRTF Measurements*. <http://recherche.ircam.fr/equipements/salles/listen/context.html>
- [16] KJAER, Bruel · *Free-field response 0° Sound Incidence of the Free Field Microphone*. Brochure
- [17] LOPEZ-POVEDA. *Peer-reviewed papers*. <http://web.usal.es/~ealopezpoveda/publications.htm>
- [18] MARKUS LINDENBERG, Nina van Ackern u. *Räumliches Hören*. <http://www.uni-mannheim.de/fakul/psycho/irtel/lehre/seminararbeiten/w96/Hoeren1/Hoeren1.html>
- [19] MATTHIES, Jörn: *Entwicklung von VST-Plugins zum automatischen Entwurf von frequenzgangkorrigierenden FIR-Filtern aus Hörtest-Daten*, Hochschule für Angewandte Wissenschaften Hamburg, Bachelorarbeit, 2005
- [20] ROOM, Head. *Frequency Response of Headphones*. <http://www.headphone.com/technical/product-measurements/how-to-interpret-headphone-data/>
- [21] SCHWARZ, Rohde · *Calibrating Measuring Microphones and Sound Sources for Acoustic Measurements with Audio Analyzer R&S UPV*
- [22] VEIT, Ivar: *Technische Akustik*. 6. Vogel Buchverlag, 2005. – ISBN 3-8343-3013-2
- [23] WIKIPEDIA. *Äußerer Gehörgang*. [http://de.wikipedia.org/wiki/Äusserer\\_Gehörgang](http://de.wikipedia.org/wiki/Äusserer_Gehörgang)
- [24] WIKIPEDIA. *Brummschleife*. <http://de.wikipedia.org/wiki/Brummschleife>
- [25] WIKIPEDIA. *Relation Between the Sone and Phon*. [http://upload.wikimedia.org/wikipedia/de/c/c7/Akustik\\_phon2sone3.jpg](http://upload.wikimedia.org/wikipedia/de/c/c7/Akustik_phon2sone3.jpg)
- [26] WIKIPEDIA. *Weighting Filter Curves*. [http://en.wikipedia.org/wiki/Weighting\\_filter](http://en.wikipedia.org/wiki/Weighting_filter)
- [27] W.SMITH, Steven. *The Scientist and Engineer's Guide to Digital Signal Processing*. <http://www.dspguide.com/pdfbook.htm>
- [28] WUPPERTAL, Uni. *Physiologie des Ohres*. [http://www.dasp.uni-wuppertal.de/ars\\_auditus/physiologie/inhaltphysio.htm](http://www.dasp.uni-wuppertal.de/ars_auditus/physiologie/inhaltphysio.htm)

- 
- [29] ZÖLZER, Udo: *DAFX Digital Audio Effects*. 1. John Wiley & Sons, Ltd England, 2002.  
– ISBN 0-471-49078-4

## List of Figures

1	Equal-Loudness Contours of the Human Ear [1] . . . . .	6
2	Weighting Filter Curves [26] . . . . .	7
3	The Human Ear [28] . . . . .	8
4	Human Hearing Range [2] . . . . .	9
5	Head Related Coordinate System [14] . . . . .	9
6	Inter-Aural Time Differences [28] . . . . .	10
7	Waves of Low Frequencies . . . . .	11
8	Waves of High Frequencies . . . . .	12
9	The Cone of Confusion . . . . .	13
10	Frequency Response for Azimuth 90° and Elevation 0°. x[Hz], y[dB], left = blue, right = green . . . . .	14
11	Loudspeaker Measurement Setup . . . . .	17
12	Free-Field Response 0° Sound Incidence . . . . .	17
13	Loudspeaker Frequency Response. x[Hz], y[dB] . . . . .	19
14	Generator Config Panel . . . . .	19
15	Generator and Analyzer Config . . . . .	20
16	Sweep Config for Loudspeaker Measurement . . . . .	21
17	Sweep Config for HRTFs Measurement . . . . .	22
18	Analyzer Function and Sweep Graph Config Panel for Phase Measurement . . . . .	23
19	HRTF Measurement Setup . . . . .	24
20	Measurement in the Horizontal Plane . . . . .	24
21	Measurement for Elevation of +37.53°. x[Hz], y[dB], z[degree] . . . . .	25
22	Left HRTFs measured in the Horizontal plane . . . . .	25
23	Left HRTFs measured for the Positions of +37.53° Elevation and Azimuth of 0° to 180°. x[Hz], y[dB], z[degree] . . . . .	26
24	HRTF pair for 0° Elevation and Azimuth. x[Hz], y[dB] . . . . .	27
25	Frequency Response of the Sennheiser HD 270. x[Hz], y[dB] . . . . .	28
26	Frequency Response of the Sennheiser HD 600. x[Hz], y[dB] . . . . .	29
27	Frequency Response of the Philips Earbud Type Headphones. x[Hz], y[dB] . . . . .	30
28	Structure of the System . . . . .	30
29	Magnitude after Mirroring . . . . .	32

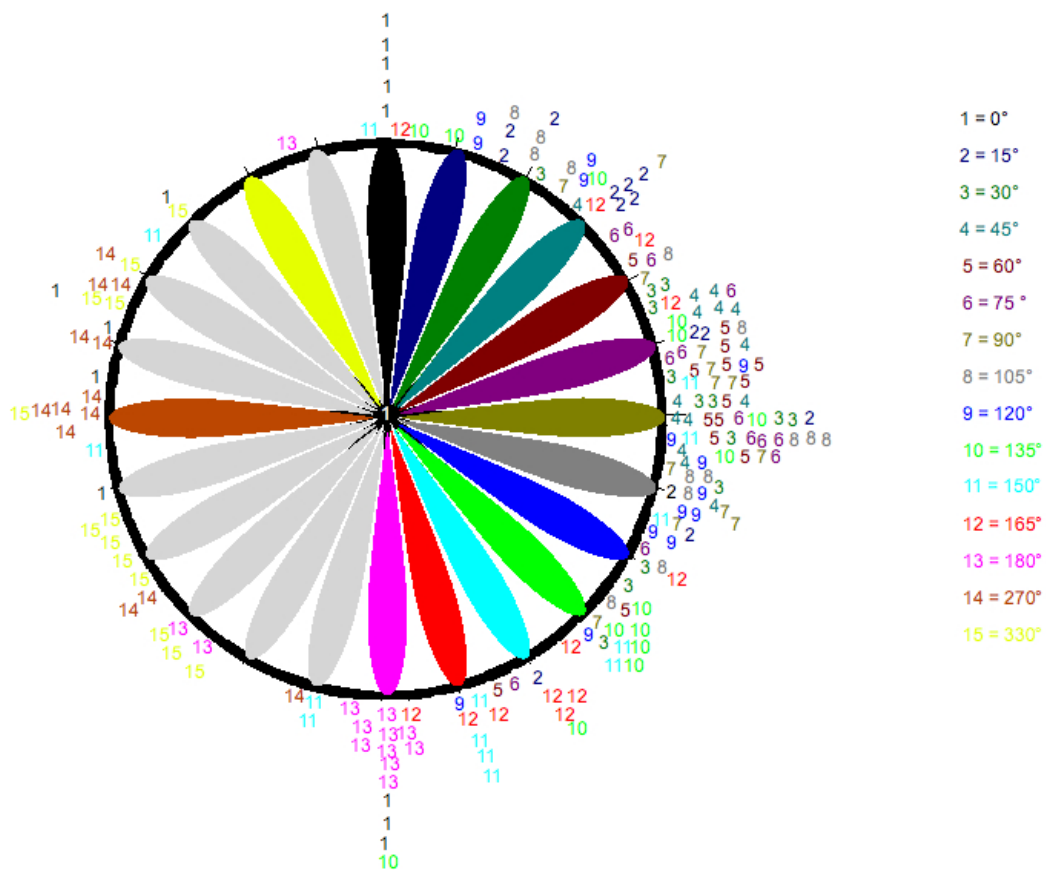
---

30	Left and Right HRIR for incident angle of $90^\circ$ in the horizontal plane . . . . .	34
31	Testrecord . . . . .	37
32	Numbered Positons for $15^\circ$ (a) and $30^\circ$ (b) Resolution . . . . .	38
33	Transfer Function of the Left and Right HRIR Filter $330^\circ$ . x[Hz], y[dB] . . . . .	40
34	Transfer Function of the Left and Right HRIR Filter $270^\circ$ . x[Hz], y[dB] . . . . .	40
35	Transfer Function of the Left and Right HRIR Filter $15^\circ$ . x[Hz], y[dB] . . . . .	41
36	Test with Sennheiser HD 270 $0^\circ$ elevation . . . . .	49
37	Test with Sennheiser HD 270 $37^\circ$ elevation . . . . .	50
38	Test with Sennheiser HD 270 $-41^\circ$ elevation . . . . .	51
39	Test with Sennheiser HD 600 $0^\circ$ elevation . . . . .	52
40	Test with Sennheiser HD 600 $37^\circ$ elevation . . . . .	53
41	Test with Sennheiser HD 600 $-41^\circ$ elevation . . . . .	54

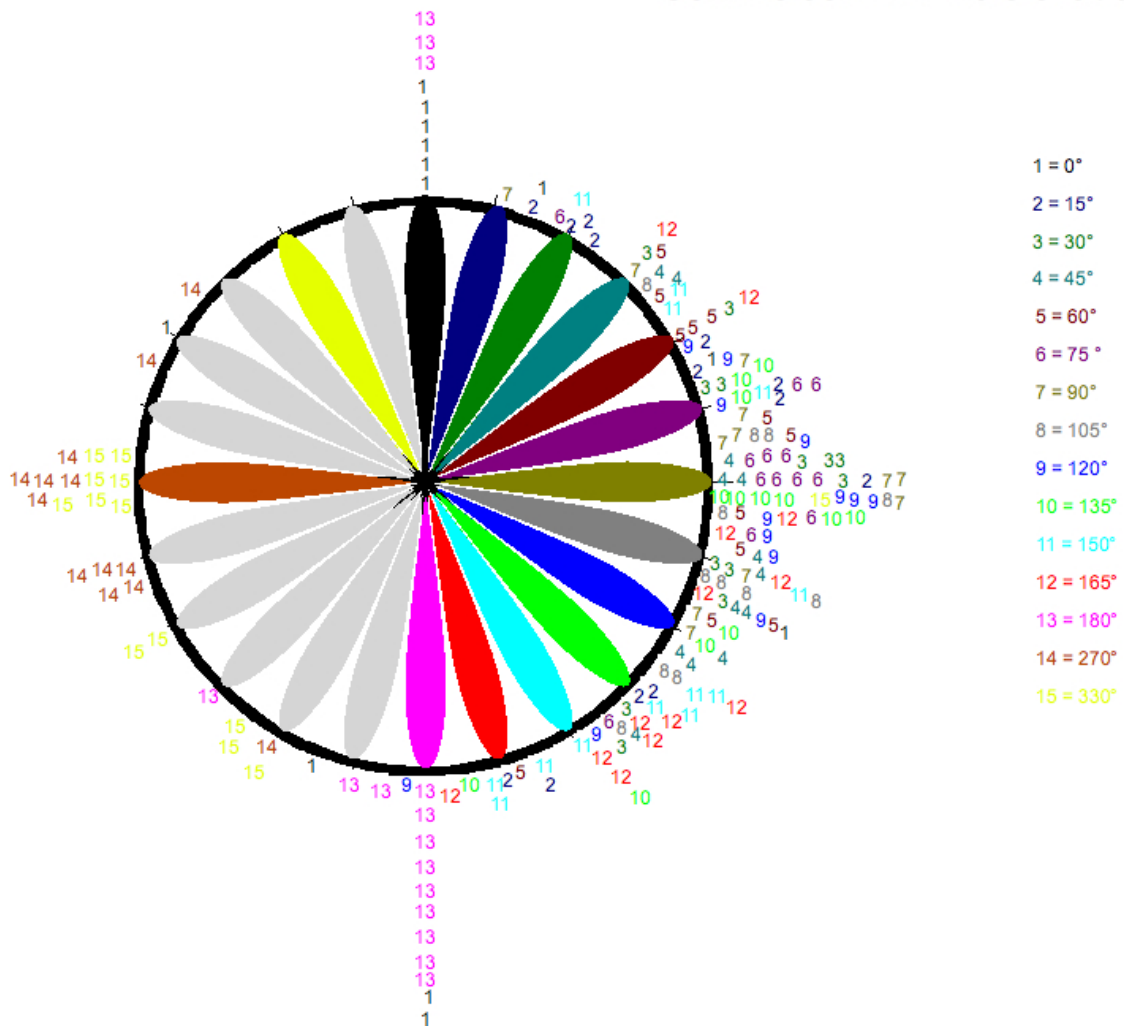
## List of Tables

1	Measured Position in the Median and Horizontal Plane . . . . .	23
2	Measured Headphones . . . . .	28
3	Positions for Generation of Sound Samples . . . . .	36
4	Sections with corresponding angles . . . . .	38
5	Evaluation of the Elevation Angle, y = yes, n = no, m = male, f = female . . . . .	39

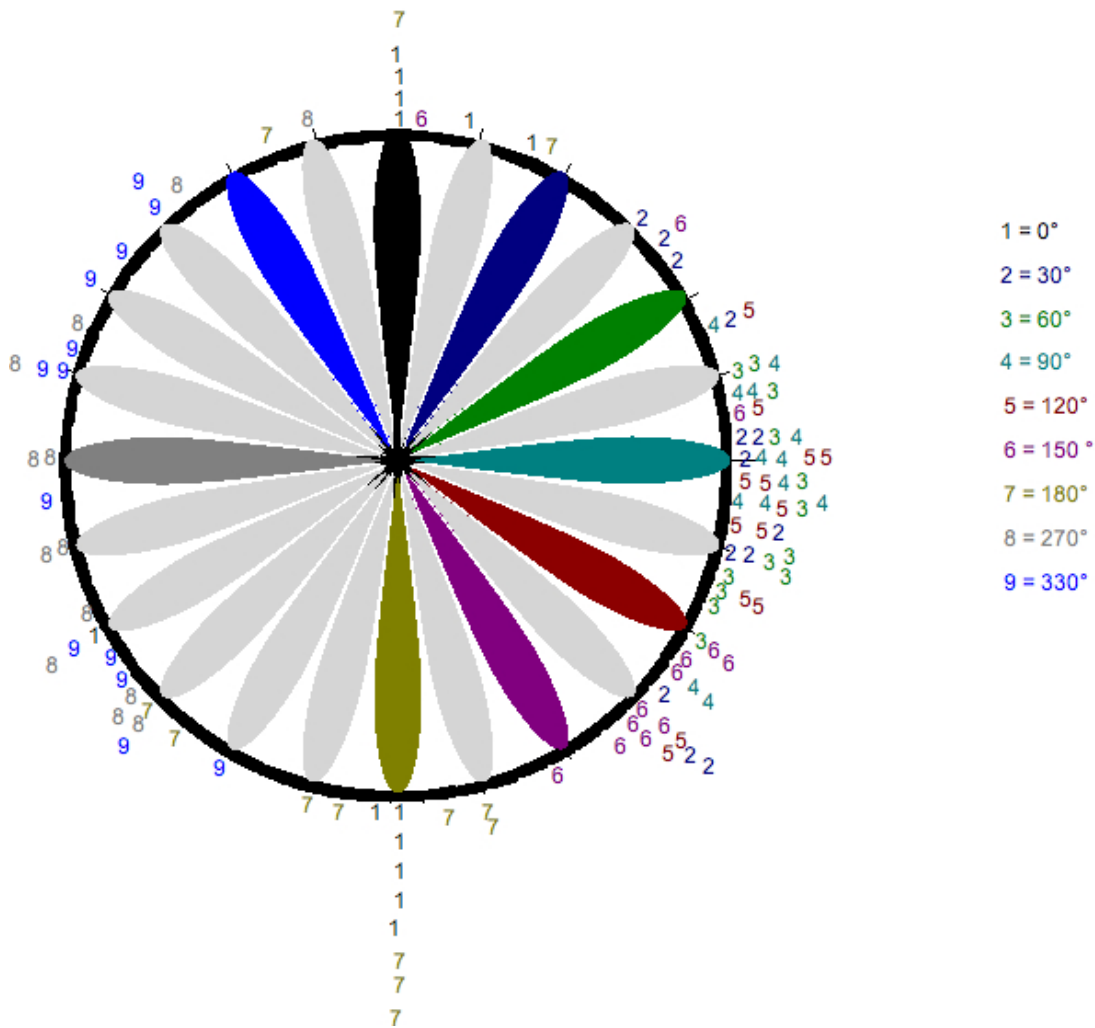
Sennheiser HD 270 elevation 0°



Sennheiser HD 270 elevation 37°

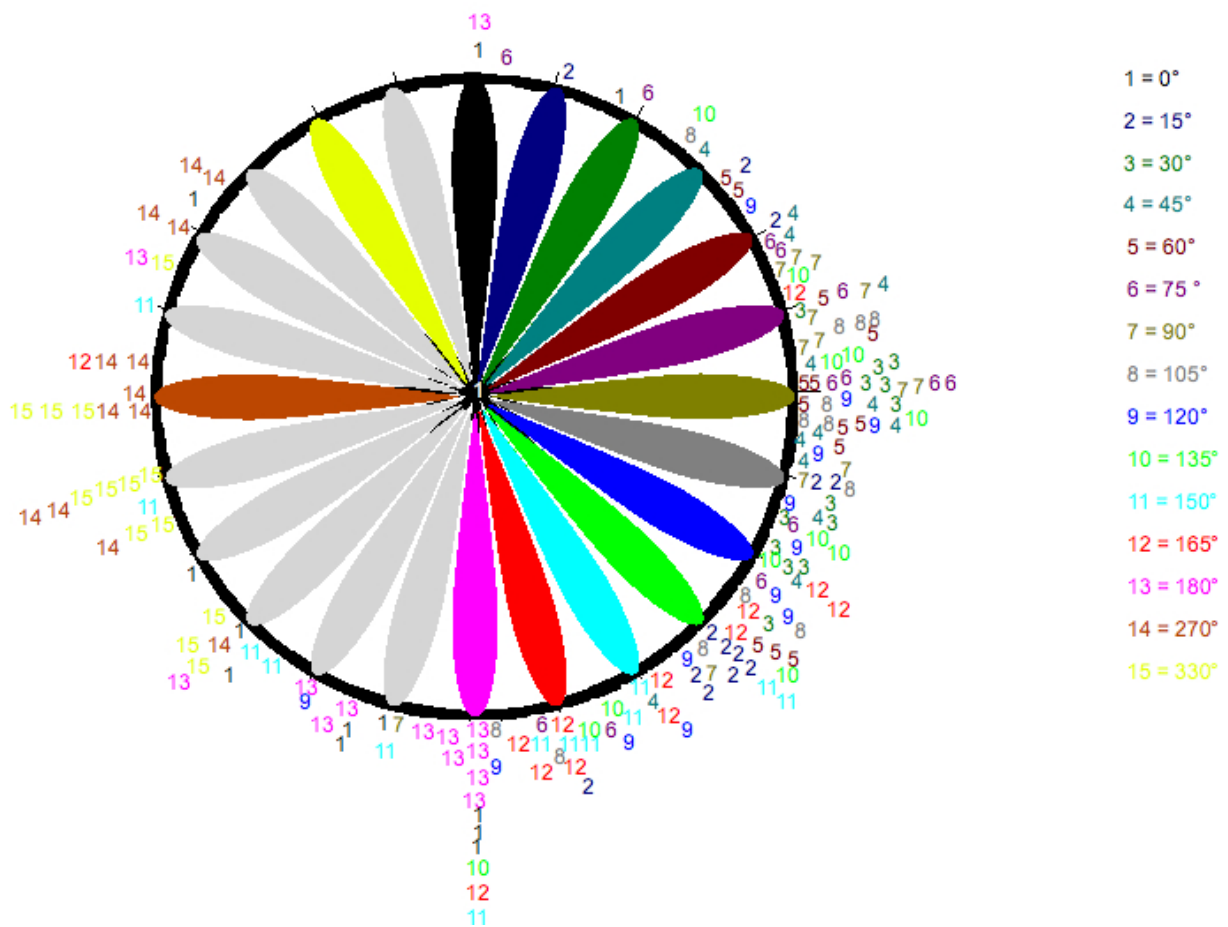


Sennheiser HD 270 elevation -41°

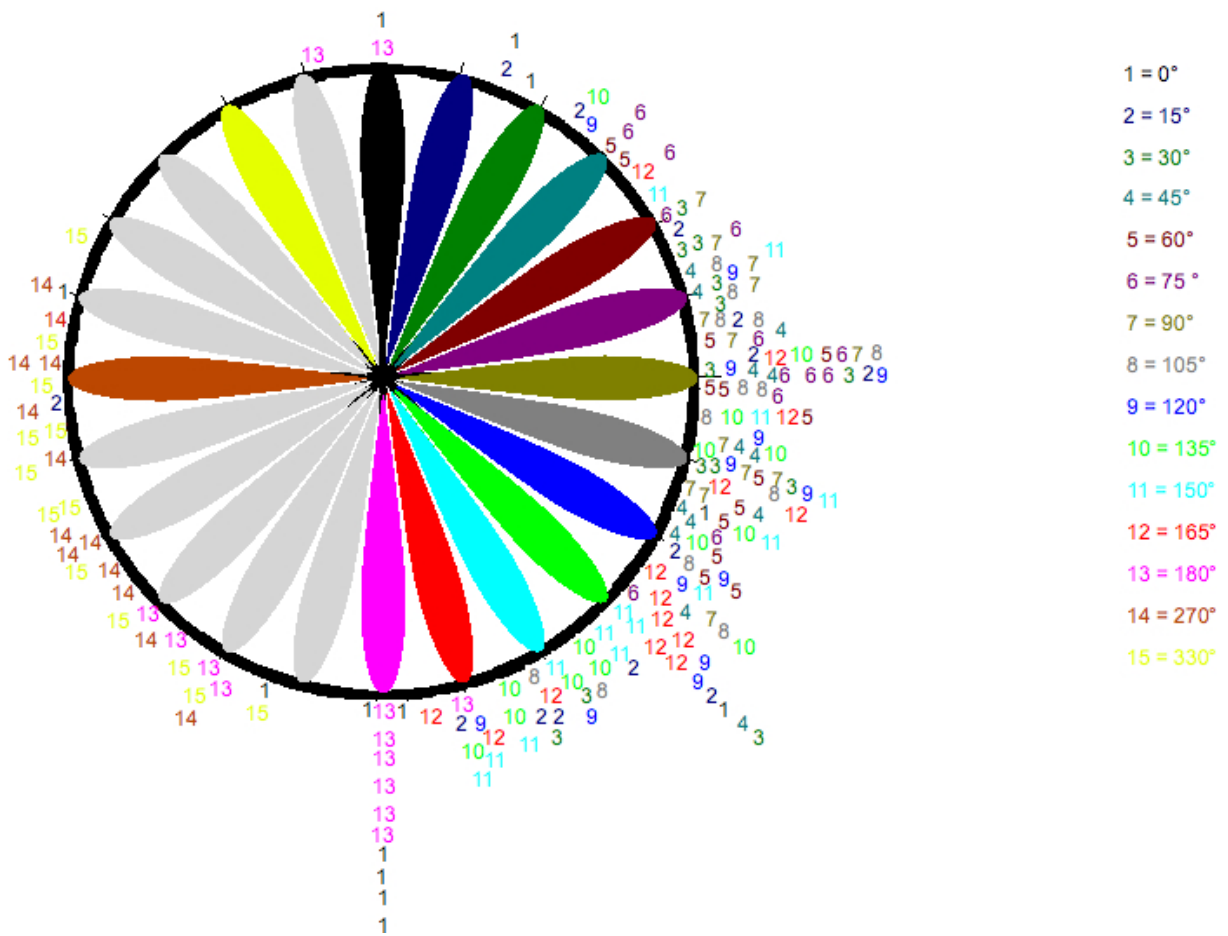




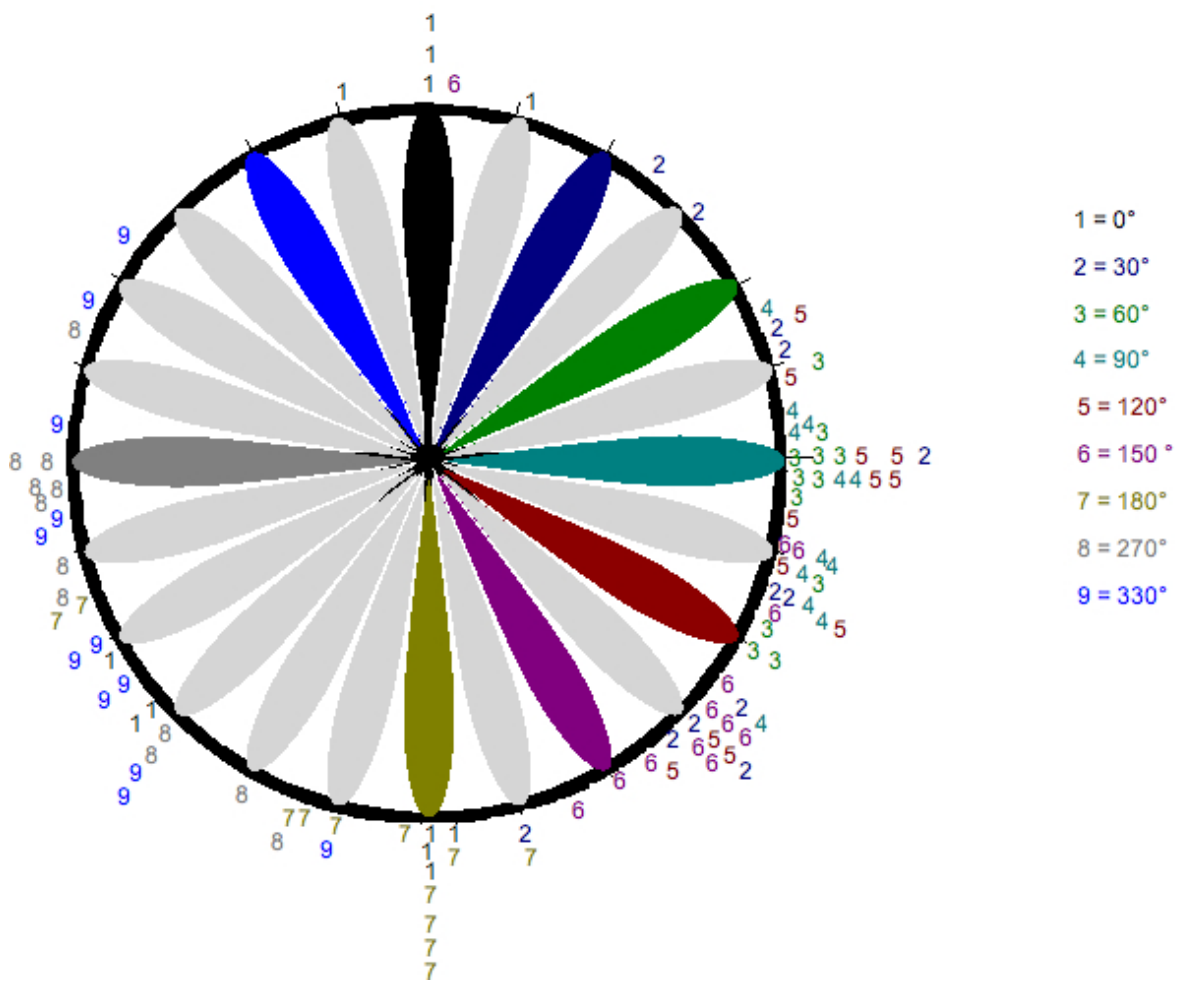
Sennheiser HD 600 elevation 0°



Sennheiser HD 600 elevation +37°



### Sennheiser HD 600 elevation -41°



## A Matlab Scripts and Supporting Tool

All tools, scripts and data is available on the attached cd.

**PrepareFiles.jar** : This is the executable java jar file which converts the analyzer's data into a Matlab file format. Please type following line to execute the application:

```
java -jar PrepareFiles.jar fullpath
```

**get\_data.m** : This function loads R&S Analyzer Data which was prepared as a m- File with PrepareFiles.jar application.

**calculate\_hrir.m** : This function computes left and right HRIR filter coefficients.

**equalize\_headphone.m** : Returns filter coefficients for a headphone equalization.

**write\_coeff.m** : Writes HRIR filter coefficients into a Matlab file.

**lininterp.m** : Interpolates filter coefficients for a desired position.

**write\_headphones\_coeff.m** : Writes filter coefficients for the headphone equalization into a Matlab file.

**plot\_three\_d.m** : Plots an 3D graph of a measured frequency response for a desired elevation and azimuth.

**filter\_sound.m** : Filters any sound for a desired position.

**main.m** : All parameters are defined in this file to compute the filter coefficients.

## Versicherung über Selbstständigkeit

Hiermit versichere ich, dass ich die vorliegende Arbeit im Sinne der Prüfungsordnung nach §24(5) ohne fremde Hilfe selbstständig verfasst und nur die angegebenen Hilfsmittel benutzt habe.

Hamburg, September 6, 2008

---

Ort, Datum

---

Unterschrift