

Measurement and evaluation of room impulse response
measurements performed with a flexible highly directive source

Master Thesis

B.Sc. Ander Gaspar Perez Palacios
Matrikelnummer: 327884
Elektrotechnik - Toningenieur V066413

May 2, 2014

Betreuer:
Dipl.-Ing. Martin Pollow
Dipl.-Ing. Johannes Klein
DI Dr.rer.nat. Franz Zotter

Beurteiler: Univ. Prof. Dr. Robert Höldrich

IEM Institut für Elektronische Musik und Akustik
Universität für Musik und darstellende Kunst

In Zusammenarbeit mit
ITA Institut für Technische Akustik
Rheinisch-Westfaelische Technische Hochschule Aachen

Contents

1	Introduction	7
2	Directivity Measurement	10
2.1	Source characterization	10
2.2	Spherical Harmonics Transform of Angularly Continuous Spherical Functions	13
2.2.1	Orthogonality of Spherical Harmonic Functions	14
2.2.2	Completeness of Spherical Harmonics Transform of Angularly Continuous Spherical Functions	14
2.2.3	Energy of Spherical Harmonics	14
2.3	Discrete Spherical Harmonic Transform	14
2.3.1	Matrix / vector notation of the Spherical harmonic transform	15
2.4	Sampling on a Sphere	16
2.5	Definition of the directivity synthesis	17
3	Measurement of Loudspeaker Directivities	18
3.1	Measurement procedure	18
3.1.1	Choosing the right measurement signal	19
3.1.2	Interleaved exponential sweeps	20
3.1.3	Measurement	22
3.2	Windowing the directivity function	23
3.3	Combining the hemispheres	24
3.4	Removal of redundant measurement nodes	25
3.5	Analysis of the spherical harmonics' energy	35
3.6	Spherical harmonic decomposition using experimental order limitation	39
4	Measurement Set up Preparation	44
4.1	Swivel arm positioning error	44
4.2	Reproducibility	46
4.2.1	Swivel arm positioning correction	48
5	Radiation pattern control	52
5.1	Inverse problems	53
5.1.1	Minimum Norm Least Squares	53
5.1.2	Tikhonov Regularisation	53
5.2	The weighting vector as a holographic filter	54
5.3	Spatial aliasing	58
5.4	Synthesis of a Dirac delta function	64
6	Outlook	67
7	Anhang	68

7.1	Coordinate Systems	68
7.2	Spherical harmonic expanded data	69
7.3	Radiation pattern control	69
	Bibliography	73

Abstract

The digital era has opened a number of fields in acoustics like the auralisation of enclosures. Usually, room impulse responses measured with a dodecahedron loudspeaker are convolved with a dry music or speech signal to create the auralize the dry signal. Since the room impulse response is primarily defined by the source, the synthesis of room impulse responses for arbitrary sources is a major research topic in acoustics. In this work, a new method of room impulse response synthesis developed at the Institute of Technical Acoustics in Aachen has been evaluated. The hardware consists of a wooden spherical measurement source, on which surface 28 transducers are placed. The radiation characteristics of each of the transducers was measured in a semi anechoic chamber to analyse the behaviour of the measurement source. Based on these results, the feasibility of the synthesis was studied.

Kurzfassung

Mit dem Aufkommen des digitalen Zeitalters, haben sich in der Akustik neue Felder eröffnet, beispielsweise die computergestützte Auralisation von Räumen. Üblicherweise benutzt man dafür Raumimpulsantworten, die mit einem Dodekaeder Lautsprecher gemessen worden sind. Da die Raumimpulsantwort primär von der Quelle abhängt, sind diese Raumimpulsantworten nicht ideal wenn Instrumente auralisiert werden. Die Synthese von Raumimpulsantworten beliebiger Schallquellen ist deswegen ein immer wichtigeres Forschungsfeld in der Akustik

In dieser Arbeit ist eine Messquelle untersucht worden, mit der die Synthese von Raumimpulsantworten besonders effizient berechnet werden kann. Die Messquelle besteht aus 28 verschiedenen Treibern, die einzeln in einem reflexionsfreien Raum gemessen wurden. Nach der Analyse der Messergebnisse wurden die tatsächlichen Synthesemöglichkeiten, die die Quelle bietet, analysiert.

Pledge of Integrity

I hereby certify that the work presented in this thesis is my own, that all work performed by others is appropriately declared and cited, and that no sources other than those listed were used.

Place:

Date:

Signature:

Eidesstattliche Erklärung

Ich versichere ehrenwörtlich, dass ich diese Arbeit selbständig verfasst habe, dass sämtliche Arbeiten von Anderen entsprechend gekennzeichnet und mit Quellenangaben versehen sind, und dass ich keine anderen als die angegebenen Quellen benutzt habe.

Ort:

Datum:

Unterschrift:

Acknowledgements

This thesis has been performed thanks to the cooperation between the Institute of Electronic Music and Acoustics (IEM) of the University of Music and Performing Arts Graz in Austria and the Institute for Technical Acoustics (ITA) of the Rheinisch-Westfaeliche Technische Hochschule Aachen in Germany. I want to thank Prof. Behler and Prof. Vorländer for their help to ensure my stay in Germany after September 2013. Without their trust in this thesis, much of it could have not been possible.

I want to thank Johannes Klein and Martin Pollow, who during my stay in Aachen, enriched my work through their experience, encouragement and guidance. I want to thank Franz Zotter for... for a lot of things. Firstly, because of his inspiring words and contagious enthusiasm. Secondly, his deep knowledge and challenging questions are essential to this thesis and to my understanding of Fourier acoustics. I want to thank the people I had the pleasure of getting to know during my stays in Graz and Aachen.

I dedicate this thesis to my family and Jose Ignacio Ostolaza, who have given me support and courage to reach for all that I am passionate about. Thank you so much.

Was aus Liebe getan wird, geschieht immer jenseits von Gut und Böse

Friedrich Nietzsche, *Jenseits von Gut und Böse*, Aphorismus 153

*Ez esan inoiz
hua bukatua dago.
Ez inoiz esan bidea amaitu da
Nik ez diat inoiz esango
hire egin beharra zein den
nik aspaldi eginela aukera.
Jo Ta Ke, erabazi arte!*

Su Ta Gar, aus dem Lied *Jo Ta Ke*

1 Introduction

Each sound source radiates in an unique frequency dependent radiation pattern referred to as directivity. Therefore, any room impulse response measured with a non-omnidirectional sound source will contain the characteristics of both the room and the source. This effect is not desired if acoustical parameters are to be derived from the room impulse response [8].

Auralization is the virtual recreation of acoustical scenarios. A room acoustical auralization consists of the convolution of a *dry signal*, ideally recorded in an anechoic chamber, and the room impulse response. The most accurate auralization contains the room impulse response measured with the source which appears in the dry signal. Since room impulse response measurements using real sources are not feasible, the first challenge that arises to measure such an impulse response correctly using an sound source with a variable directivity. In the previous years, numerous projects at the ITA have been carried out to derive a suitable measurement system. Martin Pollow focused on the suitability of a dodecahedron loudspeaker to radiate with variable directivities [11] and Martin Kunkemöller derived a set of measurement strategies based on the rotation of a dodecahedron and the superposition of weighted measured room impulse responses. He also derived an inverse problem to compute the gain vector, which contains the weighting coefficients for the room impulse responses superposition . [10]. Johannes Klein then derived a spherical multi channel measurement source optimized to synthesize different radiation patterns [8] .

The aim of the current work is to test the suitability of the optimized measurement system in acoustically interesting environments such as concert halls. To do so, first the directivity of each transducer mounted on the optimized measurement system were measured.

Chapter 2 presents a characterization of the source and the basic knowledge on spherical acoustics, especially the properties of the spherical harmonic decomposition. Chapter 4 describes the measurement set up to measure the directivities and its inherent technical problems and contains the measurement results as well as the spherical harmonic expansion of the measured data. such as the swivel arm positioning error. The proposed solutions to these problems are also included in this chapter. Chapter 5 deals with the radiation pattern control and the different ways to assess it.

Nomenclature

Constants

c speed of sound in $\frac{m}{s}$

Variables

p scalar pressure in Pa

f frequency in Hertz

L total number of transducers

l transducer index

p scalar pressure in Pa

rot number of rotations

N maximal order

Q order of analysis

n order index

m degree index

r radius in m

ms millisecond

φ azimuthal angle

ϑ elevation angle

α turntable azimuthal rotation angle

Notation

x scalar

\mathbf{x} vector

$\hat{\mathbf{x}}$ spherical harmonic coefficient vector

$\overline{\mathbf{A}}$ transposed matrix

$\hat{\mathbf{A}}^+$ generalized Inverse matrix

$\hat{\mathbf{A}}^\dagger$ Moore - Penrose Inverse matrix

\mathbf{A}^\oplus Tikhonov Inverse matrix

\mathbf{A}^H Hermitian transposed matrix

Y_m^n spherical harmonic coefficient

P_m^n associated Legendre Polynomial of the first kind

Q_m^n associated Legendre Polynomial of the second kind

\mathbf{I} Identity matrix

\mathbf{O} Zero matrix

SNR signal to noise ratio in dB

MIMO Multiple Input Multiple Output System

2 Directivity Measurement

2.1 Source characterization

The spherical multi channel measurement source consists of a wooden sphere of radius $r = 0.2m$, mounted on an iron pipe (500 mm) using thrust bearings [8]. The iron pipe pierces the sphere on the east and west pole ($\vartheta = +\frac{\pi}{2}, \vartheta = \frac{\pi}{2}$). This mounting strategy makes it possible to rotate the sphere around its own axis. On the surface of the sphere 28 loudspeakers, grouped in 3 different sizes, were mounted. 4 low - middle frequency range transducers , 12 high - middle frequency range transducers and 12 high frequency range transducers. In figure 2.1, the flexible highly directive source is depicted:



Figure 2.1: Flexible highly directive source

The 12 high - middle frequency range loudspeakers were changed before the measurement so an

impedance measurement of all transducers was performed. The correct behaviour of each transducer can be derived from the impedance curve by either analysing the resonance frequency and its amplitude or by checking the nominal impedance, which is the lowest impedance measured for any frequency greater than the resonance frequency. Following EN 60268 - 5, no impedance value throughout the whole frequency range can be less than 20% of the nominal impedance. [26]. Out of the impedance measurements, the nominal impedance was calculated. The results are shown in figures 2.2 to 2.4:

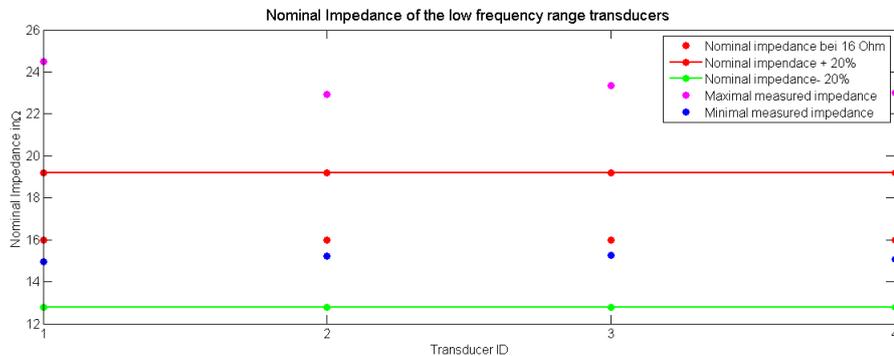


Figure 2.2: Low - middle frequency transducers

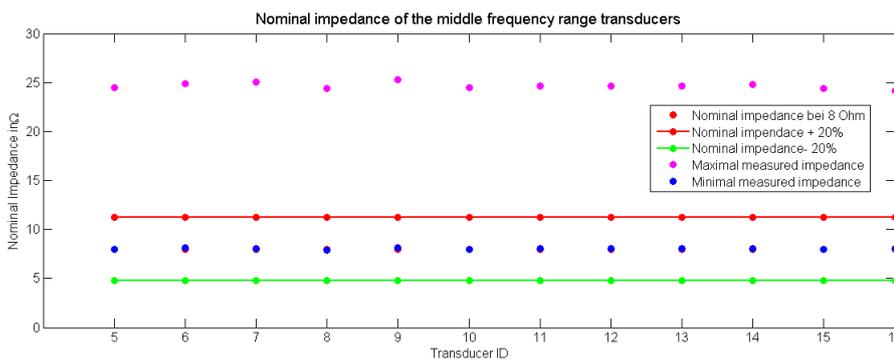


Figure 2.3: High - middle frequency transducers

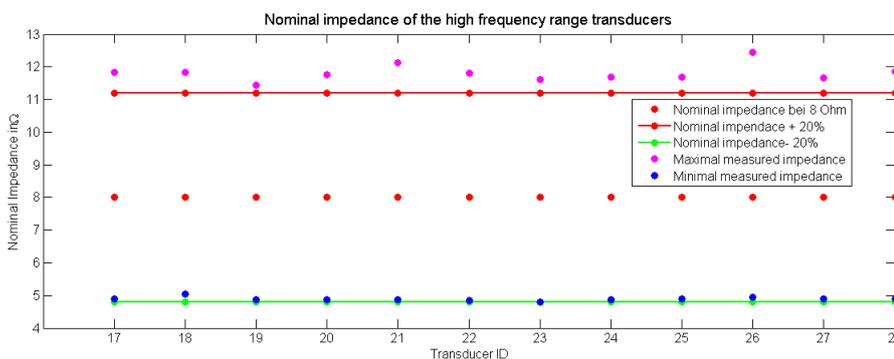


Figure 2.4: High frequency transducers

In all three transducer groups no transducer stands out for atypical values which is a sign for the correctness of the measurement and rules out malfunctioning transducers. In fact, the minimal measured impedance values of all transducers do lie within the required interval. The maximal

impedance values are not critical since a high electrical impedance does imply that the amplifier needs to supply less current. However, when the impedance at a certain frequency is low, the current that is requested from the power amplifier rises. To prevent the power amplifier from getting damaged, the minimal impedance of the transducer needs to satisfy the safety norm EN 60268 - 5. For the sake of completeness, the resonance frequency of each transducer has been listed:

Loudspeaker	Resonance Frequency in Hz
1	146
2	125.8
3	123.1
4	139.3
5	113.7
6	115.7
7	103.6
8	107.7
9	113.7
10	104.3
11	106.3
12	108.3
13	109.7
14	107
15	100.3
16	112.4
17	228.8
18	239.6
19	236.2
20	217.4
21	232.8
22	236.2
23	246.3
24	232.8
25	218
26	251
27	226.8
28	221.4

Table 2.1: Resonance Frequency of all 28 loudspeakers

In terms of directivity synthesis, this measurement source is revolutionary: the optimized transducer placement concept allows the synthesis of much complexer radiation functions than any other common spherical array, such as dodecahedron loudspeakers. Connecting the mathematical formulation of the directivity synthesis with the theory of spherical harmonic expansion, the benefits of this measurement source can be clarified.

2.2 Spherical Harmonics Transform of Angularly Continuous Spherical Functions

The spherical shape of the source implies that its radiation can be expanded in a set of *spherical base functions* called spherical harmonic functions. As it will get clear in 2.5, operating with spherical harmonic expanded data poses a set of advantages concerning the synthesis of arbitrary directivity functions.

Spherical harmonics are a set of orthogonal functions on the two-dimensional sphere \mathbb{S}^2 [12], defined by order n and degree m :

$$Y_n^m(\vartheta, \varphi) = \sqrt{\frac{(2n+1)(n-|m|)!}{4\pi(n+|m|)!}} P_n^{|m|}(\cos \vartheta) e^{j|m|\varphi} \quad (2.1)$$

Analogue to an orthonormal vector base, they form a complete and orthonormal base of the Hilbert space $L^2(S^2)$. In figure 2.5, the spherical harmonic functions up to order $N = 2$ are depicted (phase is colour coded):

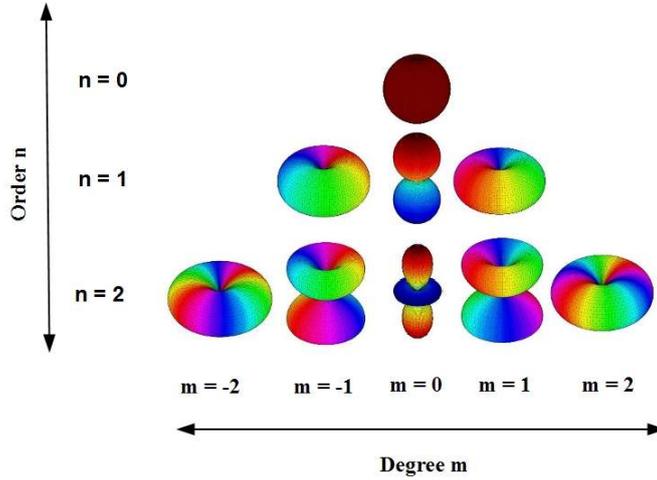


Figure 2.5: Spherical harmonic coefficients up to order $N = 2$

Recalling the orthogonality property of the associated Legendre function $P_n^{|m|}(x)$, for a given order n , a whole set of orthogonal functions is given. Spherical harmonic functions can be regarded as a set of vectors on which a function on a sphere, in this case, a sound pressure distribution, can be projected on. This mathematical operation is the *spherical harmonic transformation SHT* :

$$f_{nm} = \int_{\varphi=0}^{2\pi} \int_{\vartheta=0}^{\pi} f(\vartheta, \varphi) Y_n^{m*}(\vartheta, \varphi) \sin(\vartheta) d\vartheta d\varphi \quad (2.2)$$

where f_{nm} is the angular spectrum. Since the spherical harmonic transformation shows completeness, the *inverse spherical harmonic transform ISHT* allows to represent any square integrable function $f(\vartheta, \varphi)$ defined on the surface of a sphere as a series expansion:

$$f(\vartheta, \varphi) = \sum_{n=0}^{\infty} \sum_{m=-n}^n f_{nm} Y_n^m(\vartheta, \varphi) \quad (2.3)$$

Equation 2.3 is often referred to as spherical harmonic expansion, expression that will be used throughout this work.

2.2.1 Orthogonality of Spherical Harmonic Functions

Spherical harmonics are orthonormal to one another. Thus, integration of two different harmonics along a sphere equals always zero and integration of two equal harmonics equals one

$$\int_0^{2\pi} \int_0^{\pi} Y_n^m(\vartheta, \varphi) Y_{n'}^{m'*}(\vartheta, \varphi) \sin \vartheta d\vartheta d\varphi = \delta_{nn'} \delta_{mm'} \quad (2.4)$$

where $\delta_{nn'}$ and $\delta_{mm'}$ are Kronecker delta functions which are used to show the orthogonality property of spherical harmonic functions, as it can be derived from equation 2.5 [4]

$$\delta_{nn'} = \begin{cases} 1, & n = n' \\ 0, & n \neq n' \end{cases} \quad (2.5)$$

2.2.2 Completeness of Spherical Harmonics Transform of Angularly Continuous Spherical Functions

Completeness can be shown by proving the equivalence of the spherical harmonic expanded transform of a function to its original: $\mathcal{ISHT}\{\mathcal{SHT}\{f(\vartheta, \varphi)\}\} = f(\vartheta, \varphi)$

2.2.3 Energy of Spherical Harmonics

As a direct consequence of the completeness property of the spherical harmonic transform, the energy in the spherical harmonic domain (\mathcal{SHT}) remains the same as in the angular domain (\mathcal{ISHT}). The mathematical formulation of this statement in signal processing is known as the *Parseval's theorem* [8].

$$\int_{\varphi=0}^{2\pi} \int_{\vartheta=0}^{\pi} |f(\vartheta, \varphi)|^2 \sin(\vartheta) d\vartheta d\varphi = \sum_{n=0}^{\infty} \sum_{m=-n}^n |f_{nm}|^2 \quad (2.6)$$

2.3 Discrete Spherical Harmonic Transform

The discrete spherical harmonic transform (\mathcal{DSHT}) is the practical implementation of the \mathcal{SHT} and the \mathcal{ISHT} . Thus, a finite set of discrete nodes is regarded, which gives rise to further considerations. Firstly, 2.3 has to be rewritten to:

$$f(\vartheta_i, \varphi_j) = \sum_{n=0}^N \sum_{m=-n}^n f_{nm} Y_n^{m*}(\vartheta, \varphi) \quad (2.7)$$

so that $f(\vartheta_i, \varphi_j)$ is no longer a continuous function but a discrete one. Since $f(\vartheta, \varphi)$ and therefore also the spherical harmonic functions Y_n^m depend on discrete values, the discrete spherical harmonic transform can be defined as an algebraic problem. To simplify things, a linear indexation is defined:

$$q = n^2 + n + m + 1 \quad (2.8)$$

Consider a vector \mathbf{f} holding L samples. Defining $q = nm$, the benefits of linear indexation become clear:

$$\mathbf{f}_q = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ \vdots \\ f_{(N+1)^2} \end{pmatrix} = \mathbf{f}_{nm} = \begin{pmatrix} f_{0,0} \\ f_{-1,1} \\ f_{0,1} \\ f_{1,1} \\ f_{-2,2} \\ \vdots \\ f_{N,N} \end{pmatrix} \quad 2n+1$$

2.3.1 Matrix / vector notation of the Spherical harmonic transform

Equation 2.7 can be rewritten as a matrix - vector multiplication. The discrete spherical harmonics Y_n^m for each position are stored in a matrix \mathbf{Y} , which throughout this work will be defined as:

$$\mathbf{Y} = \begin{pmatrix} Y_0^0(\varphi_0, \vartheta_0) & Y_{-1}^1(\varphi_0, \vartheta_0) & Y_0^1(\varphi_0, \vartheta_0) & Y_1^1(\varphi_0, \vartheta_0) & \cdots & Y_m^N(\varphi_0, \vartheta_0) \\ Y_0^0(\varphi_1, \vartheta_1) & Y_{-1}^1(\varphi_1, \vartheta_1) & Y_0^1(\varphi_1, \vartheta_1) & Y_1^1(\varphi_1, \vartheta_1) & \cdots & Y_m^N(\varphi_1, \vartheta_1) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ Y_0^0(\varphi_L, \vartheta_L) & Y_{-1}^1(\varphi_L, \vartheta_L) & Y_0^1(\varphi_L, \vartheta_L) & Y_1^1(\varphi_L, \vartheta_L) & \cdots & Y_m^N(\varphi_L, \vartheta_L) \end{pmatrix} \quad (2.9)$$

A function $f(\varphi, \vartheta)$ sampled at L points can be written as a vector \mathbf{f} :

$$\mathbf{f} = \begin{pmatrix} f(\varphi_0, \vartheta_0) \\ f(\varphi_1, \vartheta_1) \\ f(\varphi_2, \vartheta_2) \\ \cdots \\ f(\varphi_L, \vartheta_L) \end{pmatrix} \quad (2.10)$$

Consequently, the discrete spherical harmonic transform is defined as:

$$\mathbf{f} = \mathbf{Y} \cdot \hat{\mathbf{f}} \quad (2.11)$$

The discrete spherical harmonic expansion becomes:

$$\hat{\mathbf{f}} = \mathbf{Y}^\dagger \cdot \mathbf{f} \quad (2.12)$$

N is the maximal achievable spherical harmonic order for the given amount of measurement points, which sets the first boundary condition on the *DSHT*:

Since completeness has to be guaranteed, the Parseval's theorem has to be satisfied. In order to achieve a maximum order N for a given number of points, several sampling grid strategies have been developed, as it will be explained in section 2.4.

2.4 Sampling on a Sphere

The number of measurement nodes required to achieve an order N depends heavily on the sampling grid used [12]. The most popular sampling grids are: *Gaussian sampling*, *Equiangular sampling* and *Hyperinterpolation sampling*. Within each sampling strategy lies a trade off between the sampling grid's *efficiency* $\eta = \frac{q_{max}}{L}$ and its technical feasibility. q_{max} is the number of spherical harmonic coefficients needed to decompose a function up to a maximum order of n_{max} . L points are needed to sample this maximum order exactly. Thus, q_{max} is given by the chosen order and can not be changed by the sampling strategy while L is variable.

- *Hyperinterpolation*: Hyperinterpolation sampling strategies are based on the calculation of exact sampling points applying polynomial interpolation on the sphere instead of approximations such as the least squares or the quadrature approach. Since no approximation errors are done for the $L = (N + 1)^2$ sampling points, hyperinterpolation enables the calculation of a well - conditioned full ranked matrix \mathbf{Y} [12]. In other words, the hyperinterpolation sampling is the most effective sampling, $\eta = 1$ Since the sound pressure (or the sound velocity) has to be measured at the exact sampling points, hyperinterpolation sampling strategies are often referred to as critical sampling [12]. Thus, the technical challenge lies in the accuracy of the measurement system.
- *Equiangular sampling*: The equiangular sampling strategy is a weighted quadrature based sampling method. Quadrature sampling strategies are the most computationally effective sampling methods in terms of matrix inversion [12], but their efficiency η is always lower than 1. Though the equiangular sampling's efficiency is rather low with $\eta = \frac{1}{4}$, the azimuthal and elevation angles are equally spaced so the technical implementation can be easily performed. Moreover, quadrature sampling grids are more robust against placement errors than hyperinterpolation.

$$f_{nm} = \sum_0^{2N+1} \sum_0^{2N+1} w_j \cdot f(\vartheta_j, \varphi_k) \cdot Y_n^{m*}(\vartheta_j, \varphi_k) \quad (2.13)$$

- *Gaussian sampling*: The Gaussian sampling strategy is a weighted quadrature based sampling method. It constitutes a trade off between the high efficiency of the hyperinterpolation method and the robustness inherent to quadrature samplings. The azimuth angle is sampled with $2(N+1)$ like in the equiangular sampling, but the elevation angle is only sampled at $(N+1)$ almost equally spaced positions [16]. Therefore, its efficiency is $\eta = \frac{1}{2}$.

$$f_{nm} = \sum_{j=0}^N \sum_{k=0}^{2N+1} w_j \cdot f(\vartheta_j, \varphi_k) \cdot Y_n^{m*}(\vartheta_j, \varphi_k) \quad (2.14)$$

In case of the Gaussian sampling, w_j are weights based on the areas of Voronoi cells associated with the sampling nodes, which ensure the orthogonality among the spherical harmonic coefficients. Defining a vector \mathbf{w} holding the coefficients w_j , \hat{f} can be calculated without explicitly inverting \mathbf{Y} . L is amount of sampling points. This characteristic of weighted quadrature sampling strategies makes them computationally very effective.

$$\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_{\eta \cdot L} \end{pmatrix}$$

$$\hat{\mathbf{f}} = \bar{\mathbf{Y}} \cdot \text{diag}\{\mathbf{w}\} \cdot \mathbf{f} \quad (2.15)$$

In order to exploit the benefits of the Gaussian sampling, the transducers have been placed in arcs of different orders: while the low - middle frequency range transducers have been placed on physical arc of $N = 3$, the high - middle frequency range and the high frequency range transducers have been placed on an arc of $N = 11$.

2.5 Definition of the directivity synthesis

The representation of the source radiation pattern in spherical harmonic functions is very helpful in the context of directivity function synthesis since, by weighting the different spherical harmonic functions of the directivity function of the source $\hat{\mathbf{D}}$, a *target directivity* $\hat{\mathbf{d}}_T$ can be calculated. Matrix $\hat{\mathbf{D}}$ holds the spherical harmonic expanded radiation pattern data for each transducer at every frequency. In 2.16, the structure of the directivity matrix $\hat{\mathbf{D}}$ is shown for one frequency.

$$\hat{\mathbf{D}} = \begin{pmatrix} p_{01}^0 & p_{-11}^1 & p_{01}^1 & p_{11}^1 & \cdots & p_{N1}^N \\ p_{02}^0 & p_{-12}^1 & p_{02}^1 & p_{12}^1 & \cdots & p_{N2}^N \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{028}^0 & p_{-128}^1 & p_{028}^1 & p_{128}^1 & \cdots & p_{N28}^N \end{pmatrix} \quad (2.16)$$

As already introduced in equation 2.12, all variables of the form $\hat{\mathbf{x}}$ hold data expanded in spherical harmonic functions. The mathematical formulation of the target directivity synthesis is given as follows:

$$\hat{\mathbf{d}}_T = \hat{\mathbf{D}} \cdot \mathbf{g}_T \quad (2.17)$$

For a given target directivity the weighting vector \mathbf{g}_T can be found by multiplying the target directivity with the generalized inverse of the directivity matrix $\hat{\mathbf{D}}^+$.

$$\mathbf{g}_T = \hat{\mathbf{D}}^+ \cdot \hat{\mathbf{d}}_T \quad (2.18)$$

Equations 2.17 and equations 2.18 hold the key challenges faced in this work: To obtain a weighting vector \mathbf{g}_T , the radiation pattern of the measurement source as well as the target directivity pattern have to be known, which implies a directivity measurement of the source as well as the measurement or the simulation of a desired radiation pattern. Moreover, the computation of the generalized inverse of $\hat{\mathbf{d}}$ has a direct effect on the weighting vector, which has to be studied in order to calculate sensible weights. After having outlined the principals of spherical harmonic expansion and its practical implementation, it becomes clear that the synthesis of desired radiation patterns can effectively be carried out by calculating a weighting vector \mathbf{g}_T , which shapes the directivity $\hat{\mathbf{d}}$ to $\hat{\mathbf{d}}_T$. Using this knowledge, the aim of an optimal synthesis source can be formulated as follows: Since it has to be able to synthesize arbitrary sources, the radiation pattern of the measurement source has to excite *all spherical harmonic functions within all orders N at every frequency*. As it will become clear throughout this work, this condition is neither feasible nor even needed to guarantee a correct synthesis.

3 Measurement of Loudspeaker Directivities

In section 2.5, the term directivity function, also called radiation pattern, has been introduced. In the literature, different definitions can be found. Assuming far field conditions, radiated waves can be treated as plane waves [15], so that a distance independent *directivity factor* D_0 can be defined as [24]:

$$D_0(\vartheta, \varphi) = \frac{p(r, \vartheta, \varphi)}{p(r, \vartheta_0, \varphi_0)} \quad (3.1)$$

The directivity factor relates the radiation in any given direction to the radiation in a reference direction. A further description of the radiating characteristics can be given by normalizing to the average energy radiated by the source, denoted as $\langle |p(r, \vartheta, \varphi)|^2 \rangle_{(\vartheta, \varphi)}$:

$$D_m(\vartheta, \varphi) = \frac{|p(r, \vartheta, \varphi)|^2}{\langle |p(r, \vartheta, \varphi)|^2 \rangle_{(\vartheta, \varphi)}} \quad (3.2)$$

The directivity factor expressed in decibels is commonly known as the directivity D_{Lm} [24]:

$$D_{Lm}(\vartheta, \varphi) = 10 \log \frac{|p(r, \vartheta, \varphi)|^2}{\langle |p(r, \vartheta, \varphi)|^2 \rangle_{(\vartheta, \varphi)}} \quad (3.3)$$

In a more general sense, the directivity function can be regarded as the representation of a sound pressure distribution over an arbitrary surface. In this case, the arbitrary surface is a spherical surface. Therefore, the directivity functions that will be discussed and analysed throughout this work do not correspond to the definitions presented in equations 3.1 and 3.2 because no reference direction or reference energy has been used. The directivity is therefore defined as:

$$D_0(\vartheta, \varphi) = p(r, \vartheta, \varphi) \quad (3.4)$$

3.1 Measurement procedure

The loudspeaker directivity measurements presented in this work were carried out in the semi anechoic chamber of the Institute of Technical Acoustics. Figure 3.1 depicts the measurement set up as flow diagram:

The source is placed on top of a turntable, which can be rotated $\varphi = 360^\circ$ on the horizontal plane xy . The swivel arm can describe circular paths up to $\vartheta = 120^\circ$ with a radius $r = 2$ m. In order to place the source at the centre point of the circular paths, the vertical centre of the source must also be placed at 2 m height. At the end of the swivel arm a microphone is placed. Interleaved exponential sweeps were used as excitation signals, allowing a fast measurement of the response of each transducer. The benefits of this technique will be discussed in sections 3.1.1 and 3.1.2. For

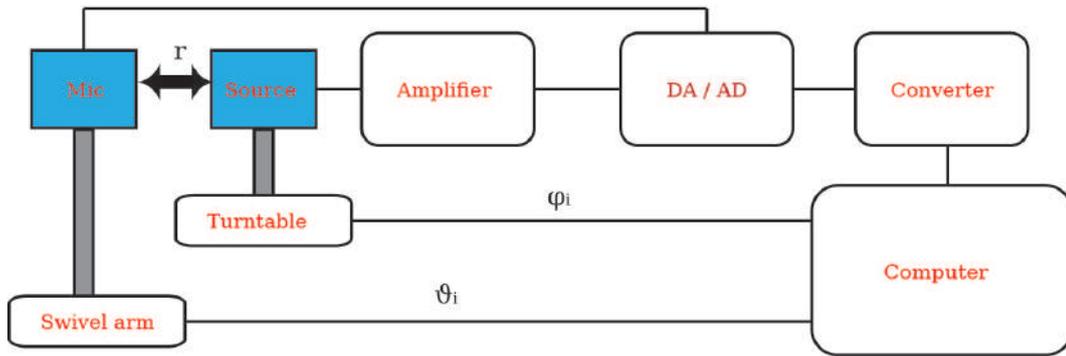


Figure 3.1: Directivity measurement set up

weighted quadrature samplings, the set of elevation angles correspondent to each azimuthal angle remains the same, as it can be seen in figure 3.2. By each azimuthal rotation, the source points towards a different azimuthal angle φ . For each azimuthal angle φ , the swivel arm is moved to the whatever elevation angles along an arc defined on the sphere of $r = 2\text{m}$ around the source, enabling the measurement of a full sampling. Since the radiation pattern of the transducers is calculated out of the result at each measurement point, the semi anechoic chamber has to be time invariant. This procedure is shown in figure 3.2.

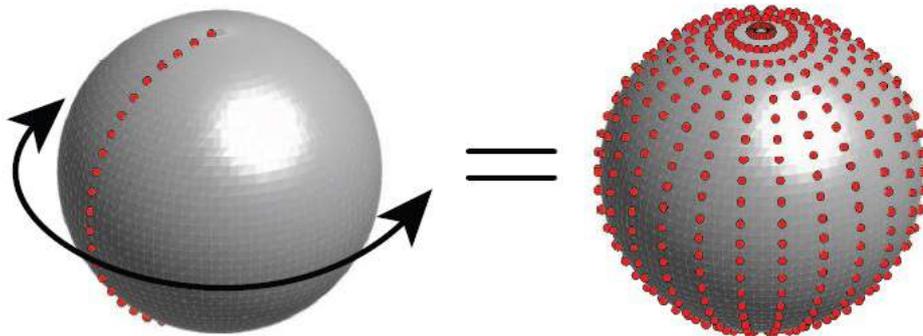


Figure 3.2: Full sampling by rotation [8]

The swivel arm cannot be rotated below an angle of $\vartheta = 120^\circ$, so both hemispheres have to be measured separately. To measure the lower hemisphere, the source is rotated by $\vartheta_{rot} = -180^\circ$ around its own axis (x axis). This way, the set of measurement points used to measure the upper half of the source is equal to the one used to measure the lower half of it. The measurement signal, the excitation of each transducer, the recording of each response and the vertical and the horizontal movement are controlled by a computer outside the chamber.

3.1.1 Choosing the right measurement signal

The measurement signal respectively the measurement procedure have to be chosen carefully to achieve a sensible measurement time without any quality loss. An ideal measurement signal should

exhibit the following properties:

Crest factor:

Defined as the ratio between the peak and the mean value of a function, the crest factor of an ideal measurement signal should theoretically be 0 dB [9]. Every measurement setup (amplifier, DA / AD...) and loudspeaker has a voltage limit. Thus, the peak value of the measurement signal has to be normalized to this value. Hence, the mean value (RMS) will be lower. In other words, the crest factor shows how much energy can not be transferred compared to a signal with crest factor = 0 dB [9]. Since mostly all loudspeakers show non linear behaviour, a compromise must to be made to minimize non linearities and still extract as much information from the loudspeaker as possible.

Time Variance:

Among the possible acoustical systems, mostly rooms are commonly defined as linear, time invariant (LTI) systems due to their inherent rather simple mathematical definition. Since the acoustical properties of the air enclosed in the room do vary in time, it is important that the measurement signal processing strategy is robust against time variations.

Non linearities:

Devices which are commonly modelled as weakly non linear systems are for example loudspeakers [29]. Since the measurement signal will be fed to a loudspeaker, it is sensible to minimize the effect of non linearities by choosing an optimal output amplification. In [20], an exponential sine signal with exponentially increased frequency, commonly known as *exponential sweep*, was firstly proposed as an alternative to maximum length sequences (MLS), the state of the art in room acoustical measurements at that time [8]. While maximum length sequences are rather defenceless against both time variations and non linearities, exponential sweeps are less vulnerable against both of them [20, 9]. Concerning the crest factor, the exponential sine sweep presents a crest factor of 3.01 dB against the 5-8 dB crest factor of an MLS signal [9]. The room impulse response is extracted from a maximum length sequence measurement by performing an autocorrelation between the input and the measured signal. On the other side, exponential sweeps rely on the *deconvolution* to obtain the transfer function of the loudspeaker: The spectrum of the device under test within the frequency range contained in the measurement signal is calculated by dividing the output spectrum (recorded signal) by the input spectrum (measurement signal),

3.1.2 Interleaved exponential sweeps

In section 3.1.1, the exponential sweep has been characterized as a reliable and robust measurement signal. Regardless of the measurement signal, the measurement time is a critical parameter. Considering a multi channel source with L sources, the measurement time sums up to [22]

$$\tau_{meas,total} = L \cdot (\tau_{meas} + \tau_{stop})$$

τ_{meas} is the actual measurement time and τ_{stop} is the time frame in which the system decays. In [21] a less time consuming measurement technique was proposed, based on the semi parallel (interleaved) playback of exponential sweeps. For the multi channel source directivity measurement and the room impulse measurement, a similar approach developed at the ITA, was used. An in depth approach as well as the general formulation for the measurement time using interleaved sweeps can be found in [22]:

$$T_{meas,total} = (L - 1)\tau_w + \tau_{meas} + \tau_{stop}$$

τ_w is the waiting period between the sweeps, τ_{meas} and τ_{stop} remain the same. The measured room impulse responses have the shape shown in figure 3.3

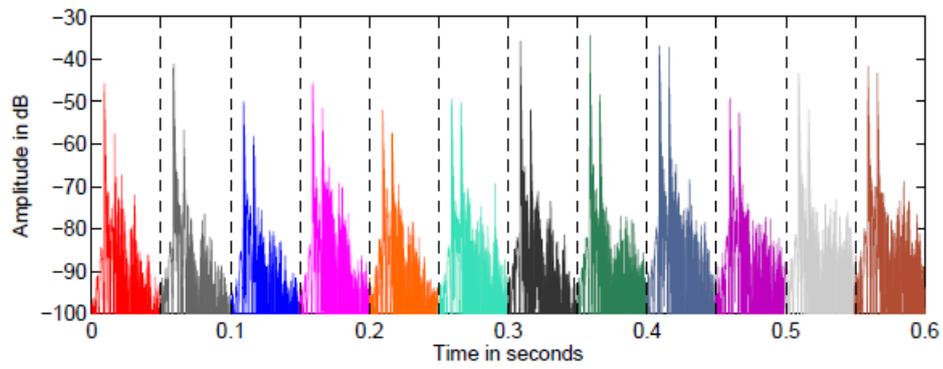
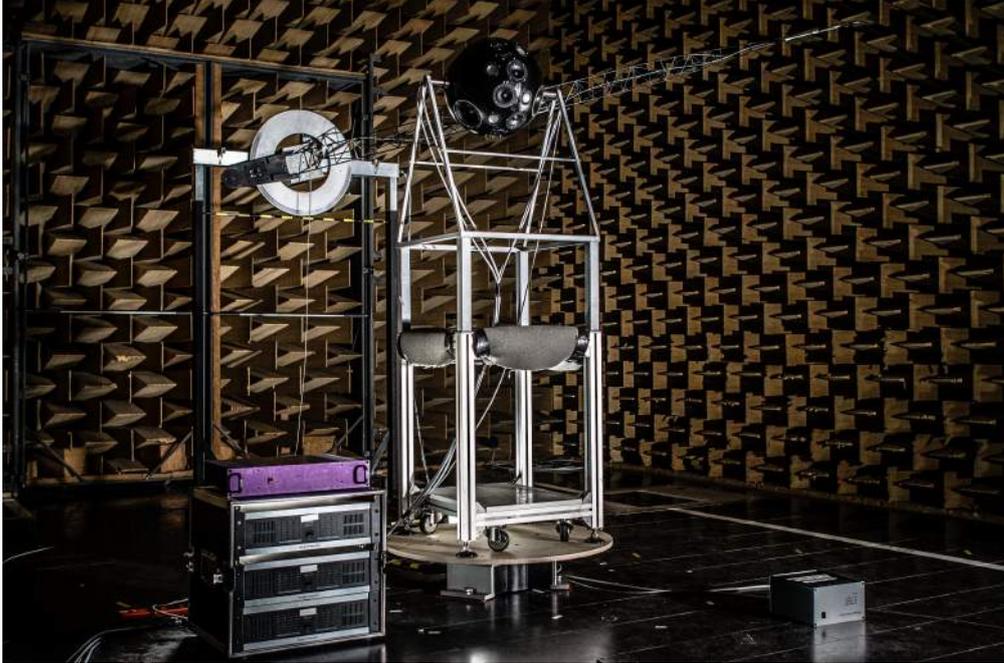


Figure 3.3: Measured impulse responses using interleaved sweeps [10]

3.1.3 Measurement

The directivity measurement of each transducer was conducted in the semi-anechoic chamber at the ITA following the measurement procedure described in section 3.1:



The measurement sampling was a Gaussian sampling of order $N = 82$, which generates 13778 points along a sphere of a given radius, see [2.4]. Interleaved sweep signals were used as measurement signals. The parameters needed to measure with interleaved sweep signals were calculated using the optimizer described in [22]. The nomenclature chosen for table 3.1 is consistent with the one employed in [22]:

τ_{meas}	τ_{IR}	τ_{RIR}	τ_{stop}	k_{max}	f_{min}	f_{max}	SNR
0.3715 s	0.004 s	0.04 s	0.001 s	5	60 Hz	22000 Hz	70 dB

Table 3.1: Input parameters for the optimizer

The measurement equipment used was:

- PC: Intel(R) Core(TM) i7-2600, 16 GB RAM 64-bit
- Software: MATLAB with ITA-Toolbox
- Interface: 1 x RME HDSPe RayDAT
- AD/DA: 4 x Behringer ADA 8000
- Microphone: 1 x Sennheiser KE4

- Amplifier: 1 x IT1
- Amplifier: 3 x Monacor STA - 1508

3.2 Windowing the directivity function

Semi-anechoic rooms are rooms in which all areas but the floor are highly absorbent. Therefore, the sound reflected by the floor needs to be removed. Since the amount of measurement nodes is rather high, the arrival time of the direct sound and the floor reflection at the microphone varies. In order to find an optimal window, the arrival time of both the direct sound and the floor reflection were studied for each measurement point. As a result of this study, a *Hann window* was chosen to remove the signal components after 12.5 ms and before 3.4 ms:

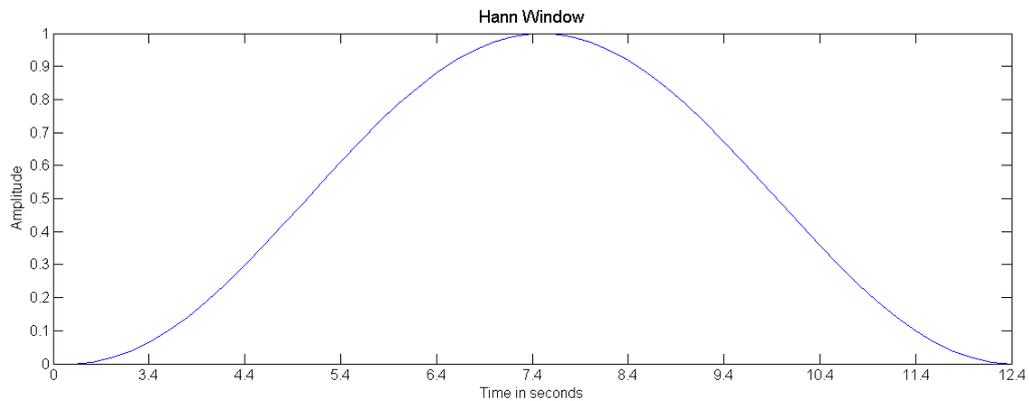


Figure 3.4: Hann window

The effect of the window can be seen in figures 3.5 and 3.6:

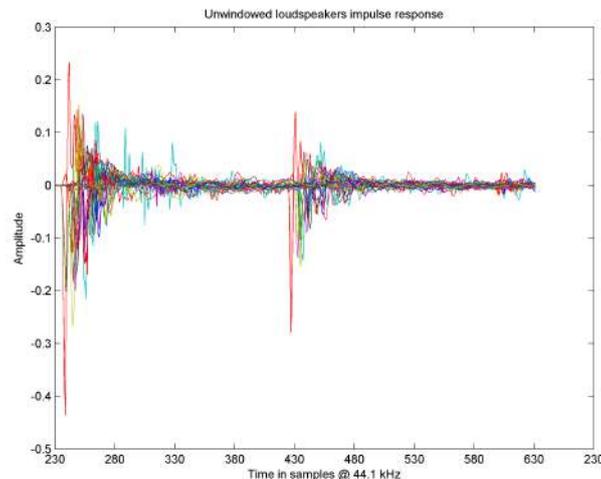


Figure 3.5: Unwindowed impulse response of the 28 transducers

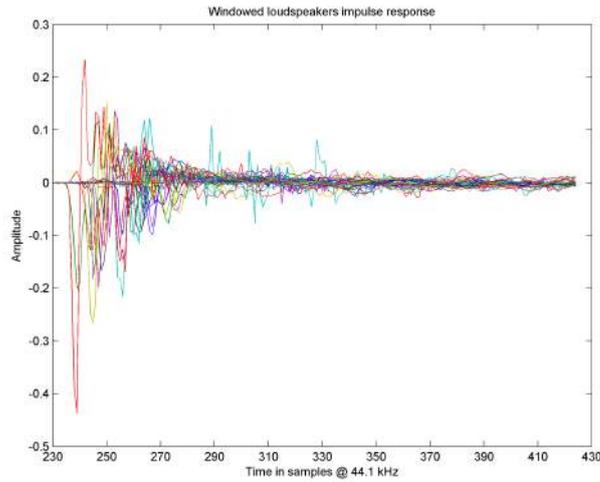


Figure 3.6: Windowed impulse response of the 28 transducers

3.3 Combining the hemispheres

As explained in section 3.1, the measurement of the lower half of the source is performed by rotating it by $\vartheta_{rot} = -180^\circ$ around its own axis. The first step before combining the measurements for the lower and the upper half spheres consists of rotating the measurement results $\vartheta_{rot} = 180^\circ$. The combination of both hemispheres is compromised by two different factors. First, changes in the temperature over the whole measurement time can become so severe, that the sound velocity variation causes a noticeable phase shift. Secondly, the positioning error, discussed in chapter 4, varies over time. Thus, the swivel arm position for measurement nodes present in both hemisphere measurements can vary for both hemispheres. From previous experience, the semi anechoic chamber of the ITA can be regarded as time invariant, leaving the time varying swivel arm positioning error as the only compromising factor for the combination of the hemispheres. Since the wavelength decreases with the frequency, the hemisphere mismatch becomes more noticeable for increasing frequencies: A hemisphere mismatch of $\Delta = 1\text{mm}$ represents a phase mismatch of $\Delta\varphi = 0.106^\circ$ for 100 Hz and $\Delta\varphi = 5.8^\circ$ for 5.5 kHz. Since the time dependence of the temperature in the semi anechoic is not known, 2 different possibilities to combine the hemispheres were considered:

- Removal of redundant measurement nodes
- Cross fading of redundant measurement nodes applying a windowing function

3.4 Removal of redundant measurement nodes

Each half sphere was measured with 9130 points, starting at an elevation angle $\vartheta = 120^\circ$. The amount of redundant measurement nodes p_{red} was calculated as:

$$p_{red} = \begin{cases} 9130 - 6972 = 2158 & \text{if } \vartheta > 90^\circ \text{ (Northern Hemisphere)} \\ 9130 - 6806 = 2324 & \text{if } \vartheta \geq 90^\circ \text{ (Southern Hemisphere)} \end{cases}$$

Directivity plots are a common way to characterize and check the behaviour of a loudspeaker. They show either the magnitude or the phase values at a given frequency of the radiation pattern. Figures 3.7 to 3.17 show the magnitude of the measured directivity function for different frequencies of transducers 1 (high-low frequency range transducer) and 11 (low-middle frequency range) and 28 by combining the spheres, free from redundant measurement nodes:

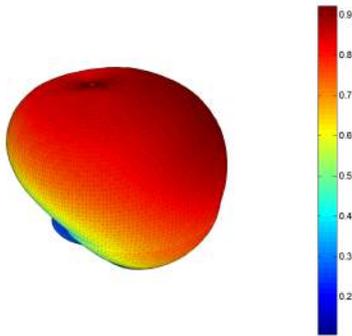


Figure 3.7: Transducer 1 for $f = 500$ Hz

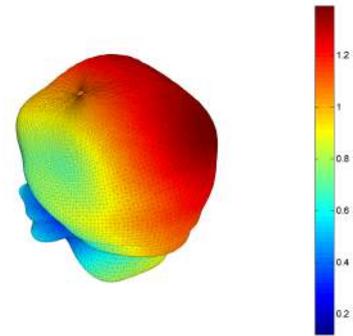


Figure 3.8: Transducer 1 for $f = 1$ kHz

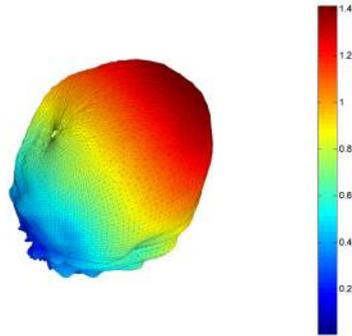


Figure 3.9: Transducer 1 for $f = 2$ kHz

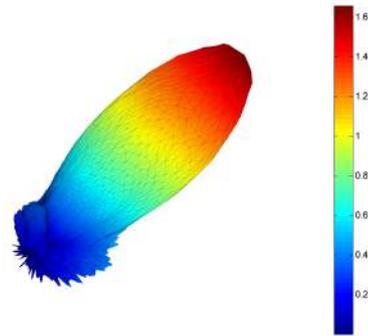


Figure 3.10: Transducer 1 for $f = 5500$ Hz

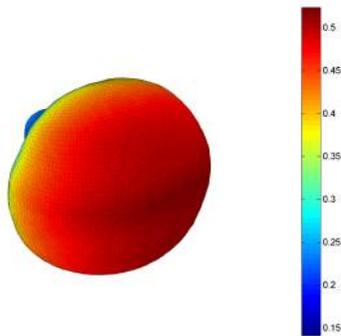


Figure 3.11: Transducer 11 for $f = 500$ Hz

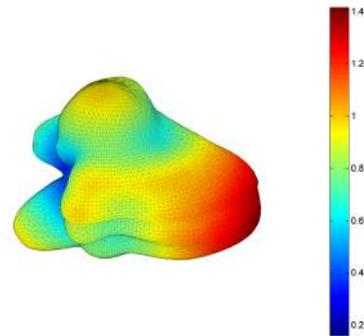


Figure 3.12: Transducer 11 for $f = 1$ kHz

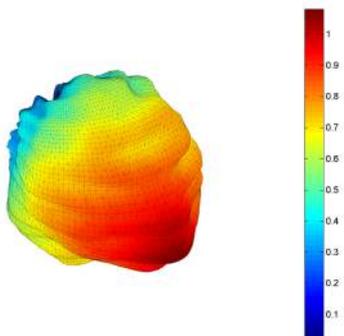


Figure 3.13: Transducer 11 for $f = 2$ kHz

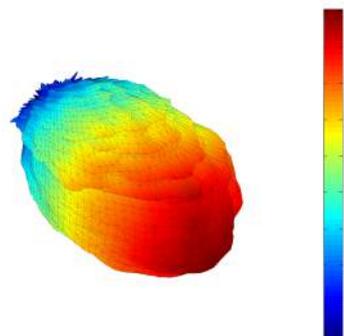
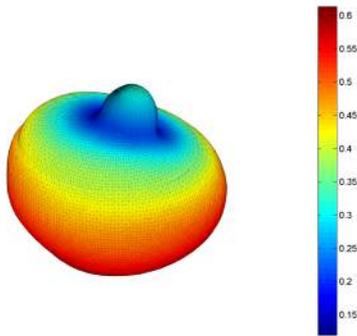
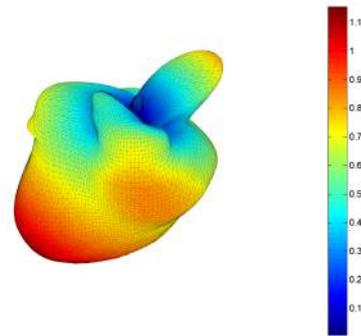
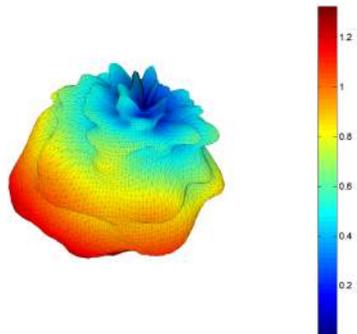
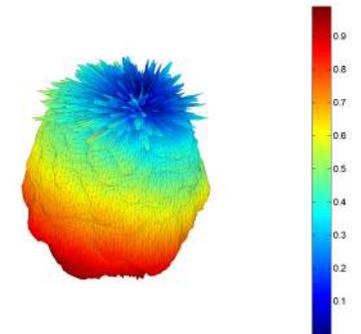


Figure 3.14: Transducer 11 for $f = 5500$ Hz


 Figure 3.15: Transducer 28 for $f = 500$ Hz

 Figure 3.16: Transducer 28 for $f = 1$ kHz

 Figure 3.17: Transducer 28 for $f = 2$ kHz

 Figure 3.18: Transducer 28 for $f = 5500$ Hz

Figures 3.7 to 3.17 show that the chosen to combine both half spheres is valid since only at 5500 Hz a slight position mismatch can be noticed. Therefore, the geometrical cross fading of both measurements was not regarded as necessary. To study the directivity of every transducer type for every single frequency would be too time consuming, so an alternative way to visualize the performance of each transducer type over the whole frequency range has been used: power within each order. This concept was introduced in equation 2.6.

$$\mathbf{e}_{n_0} = \sum_{m=-n}^n |f_{nm}|^2 \quad (3.5)$$

Since the measured data corresponds to a sound pressure distribution along a spherical surface, its energy has to be regarded as *potential energy*. The *kinetic energy* is the energy within the sound velocity field. Pictures 3.19 to 3.21 show the frequency dependent order energy for the three different transducer types. Transducer 1 is one of the 4 low - middle frequency range transducers, which explains why most of the energy its radiated within a lower frequency range than in pictures 3.20 and 3.21, in which the radiation of a high - middle frequency range transducers and a high frequency range transducer are depicted. Common to all three transducer types is the linear growth of the energy over the frequency and the presence of the greatest order energies in low orders at low frequencies. That for higher frequencies, higher orders get excited, can be explained by the fact that every transducer starts focusing the energy at a certain frequency. Therefore, radiation patterns for higher frequencies require higher spherical harmonic functions to describe the geometrical structure of the radiation pattern.

The energy held in orders higher than the ones underneath the linear growth have been labelled as

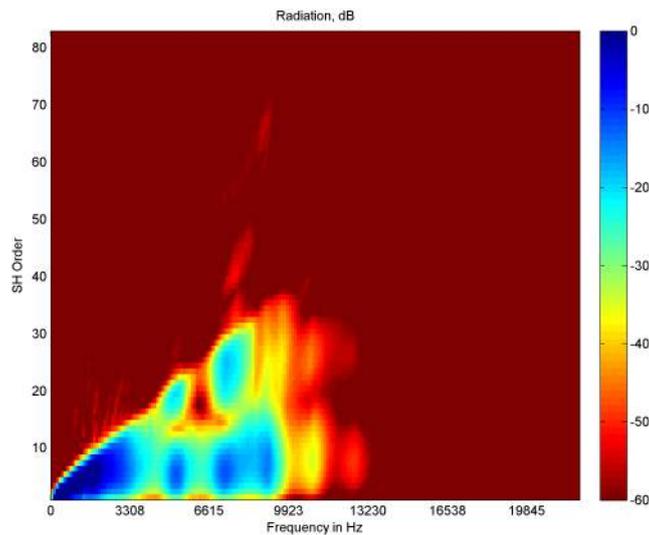


Figure 3.19: Measured order dependent radiated energy of transducer 1

error energy. Since this energy is spread differently for each transducer, it is clear that this energy is caused by a measurement error, which cause has been investigated. Three possible error sources were considered:

- the measurement room
- the measurement periphery
- the measurement signal

A reflection travelling through the room can not be separated from the real response of the transducers if they arrive at the same time at the microphone. Thus, it can not be filtered away. Reflections of the measurement periphery should depend only on the geometrical placement of the transducers, the error energy should therefore vary for each transducer. An error in the measurement signal would imply that the impulse responses interfere with each other, which would have devastating effects. The frequency dependent order energy has been studied for every physical transducer to find the possible cause. Since some transducers did not show any noticeable error energy, as it can be seen in figure 3.22, the reflections on the periphery seem to be the most probable error energy source.

To investigate the effect of the error energy on the whole array, the combined radiation of all 28 transducers was calculated by adding the energy of equal spherical harmonic coefficients of each transducer and subsequently calculating the energy held within every order:

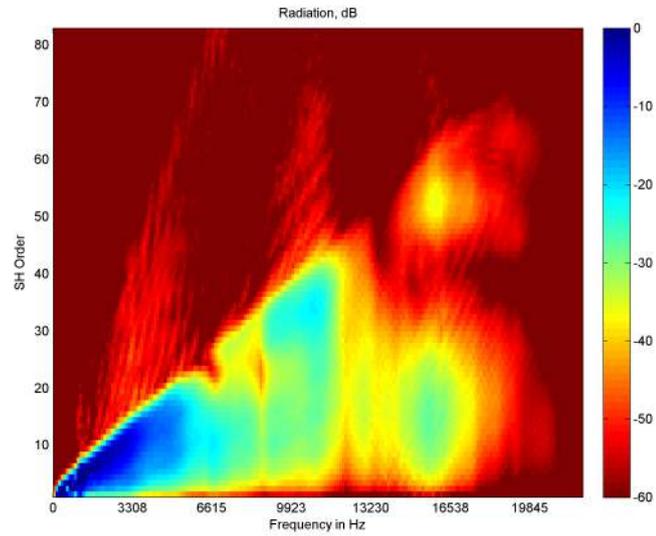


Figure 3.20: Order dependent radiated energy of transducer 5

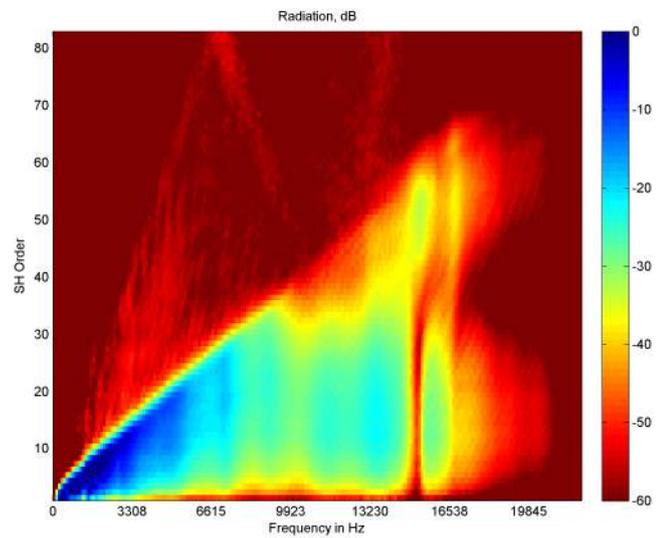


Figure 3.21: Order dependent radiated energy of transducer 28

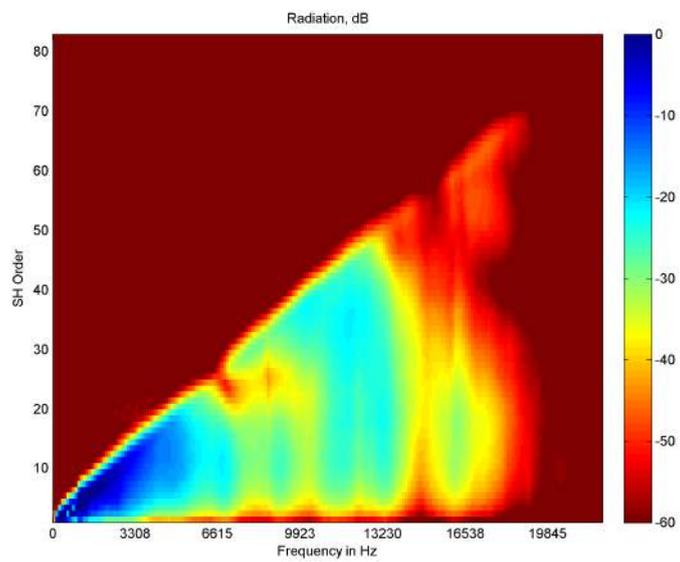


Figure 3.22: Order dependent radiated energy of transducer 12

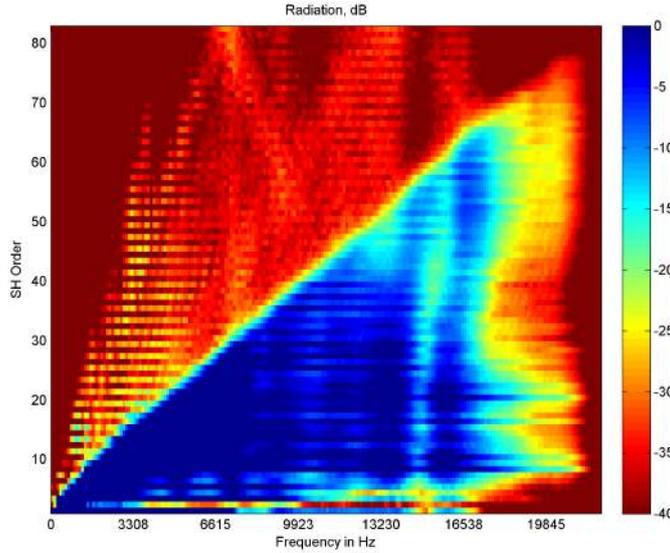


Figure 3.23: Order dependent radiated energy of all transducers

Till approximately 5200 Hz, the error energy contributions are highest. The higher the frequency is, the greater is the range of orders than are excited by this energy. Nevertheless, the measurement source shows a constant amount of energy till approximately $f = 13$ kHz, coded in blue on picture 3.23, which shows the suitability of the transducer choice made in [8]

To find the object most likely to cause the reflection, certain investigations can be carried out. As stated in [19], $He = 1$ distinguishes acoustical characteristic ranges. The influence of an object on the sound field depends on the ratio between its dimensions and the regarded wavelength:

$$He = \begin{cases} \ll 1, & x \ll \frac{\lambda}{6} & \text{Diffraction range} \\ \approx 1, & x = \frac{\lambda}{2\pi} \approx \frac{\lambda}{6} & \text{Border range} \\ \gg 1, & x \gg \frac{\lambda}{6} & \text{Reflection range} \end{cases}$$

The measurement source is suspended by thrust bearings with a diameter of approximately 1 cm, so the reflection range of the thrust bearings starts approximately at $f = 5716.7Hz$. Since frequencies which wavelengths are less than $\lambda = \frac{c}{f} = \frac{343 \frac{m}{s}}{5716.7Hz} = 0.06m$ are (partially)reflected towards the source, the sound field arriving at the microphone can be considered as distorted and can not be compensated. Since the major contributions of the error energy occur under 5 kHz, it is plausible that the sound is mainly reflected on other items of the room. As explained in section 3.2, the measured time signal was windowed between 3.4 ms and 12.5 ms so any reflections disturbing the measurement need to arrive at the microphone placed at the end of the swivel arm within that time frame. Considering that the sound covers a distance of 4.28 meters in 12.5 ms, reflections on the swivel arm and the thicker thrust bearings holding the source, which can be seen on picture 3.1.3, could eventually be responsible for the error energy.

To show the effect of the unwanted reflections, the time samples 1 to 162 of the measured signals for transducers 1, 11 and 28 have been depicted:

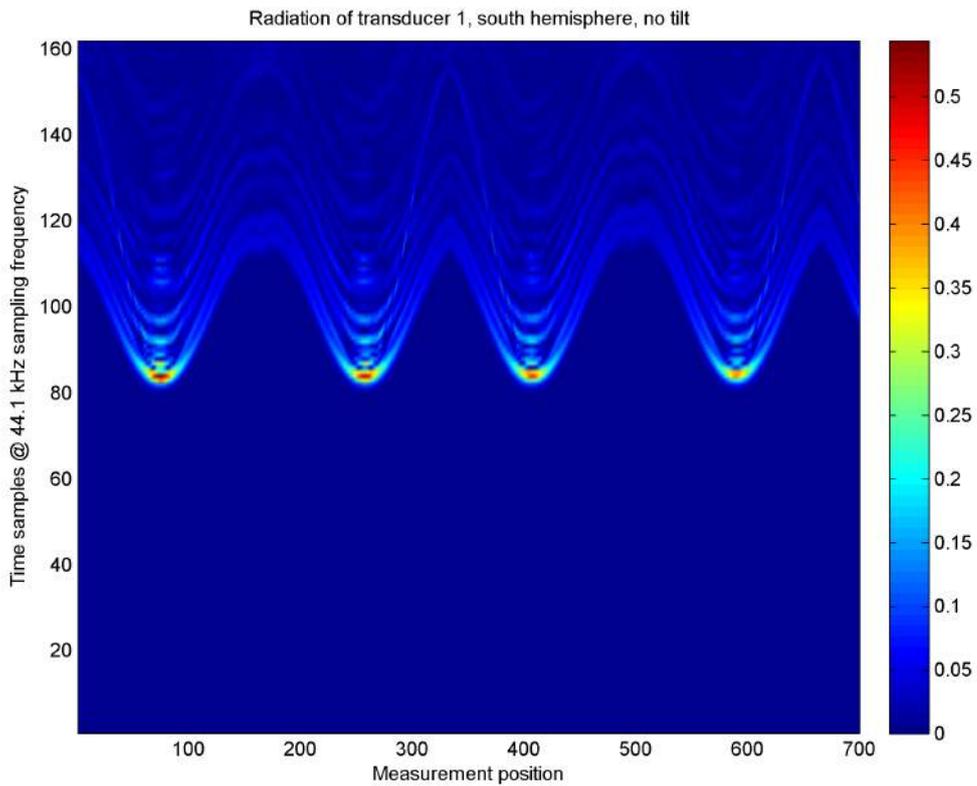


Figure 3.24: Radiation of transducer 1, No Tilt, Lower Hemisphere

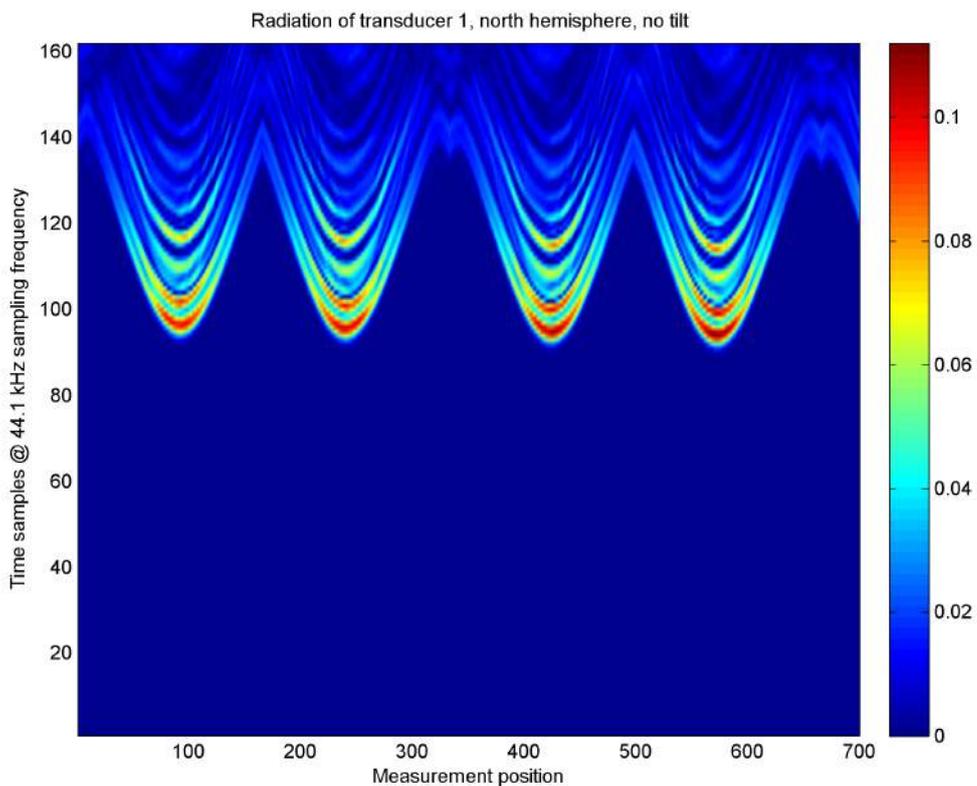


Figure 3.25: Radiation of transducer 1 in dB, No Tilt, Upper Hemisphere

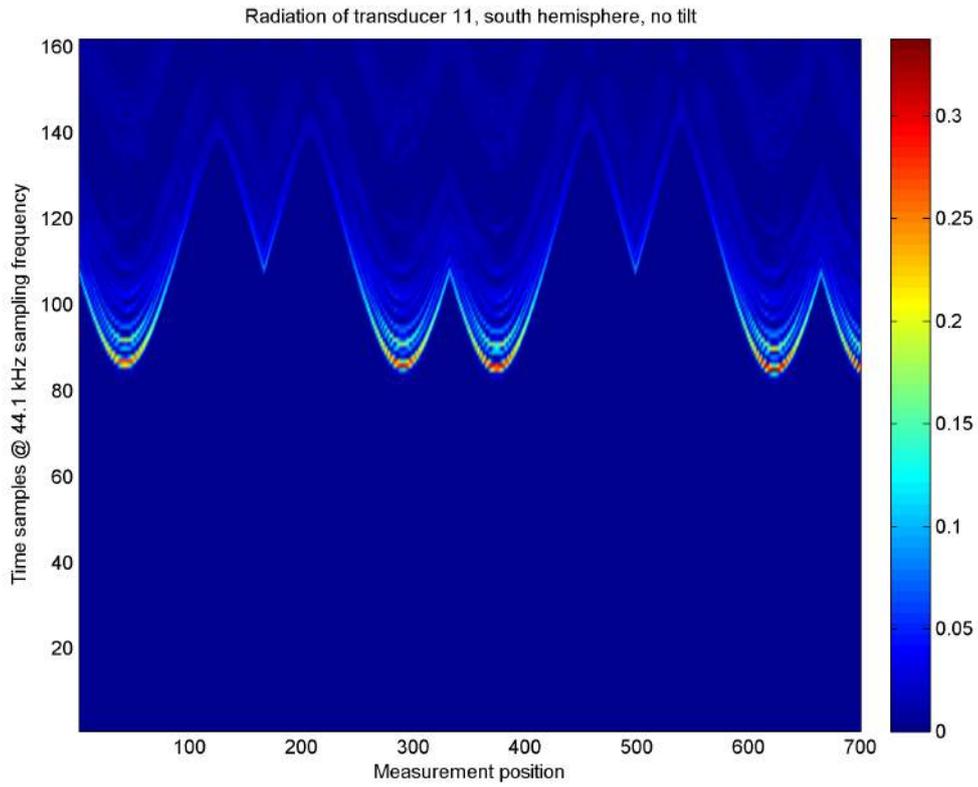


Figure 3.26: Radiation of transducer 11 in dB, No Tilt, Lower hemisphere

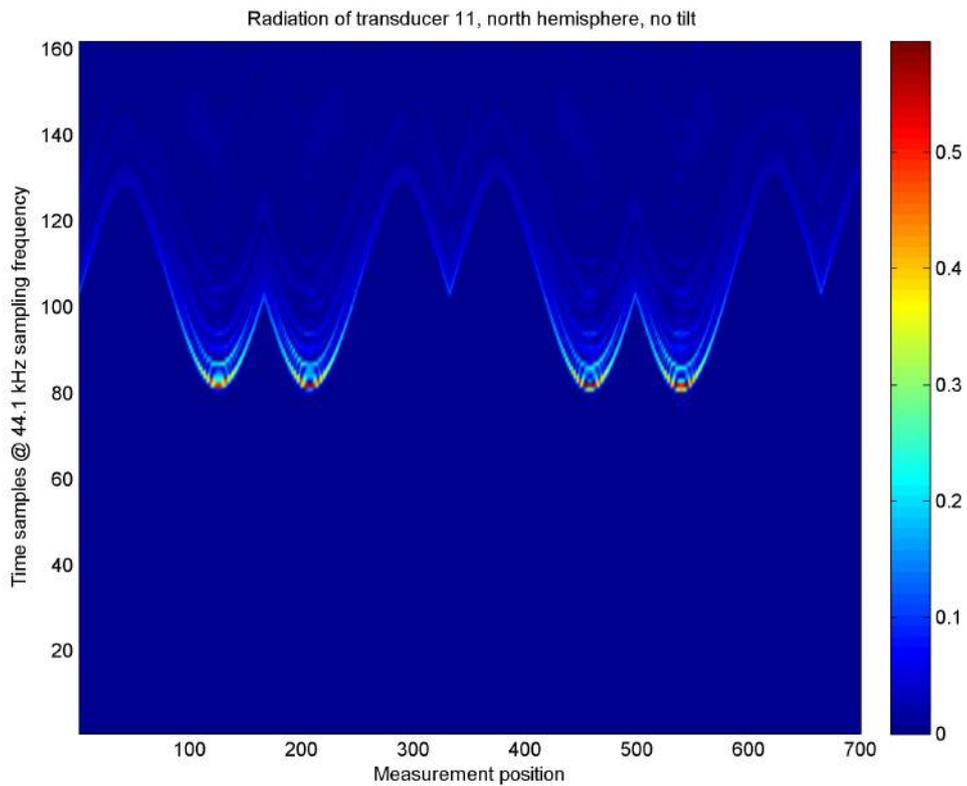


Figure 3.27: Radiation of transducer 11 in dB, No Tilt, Upper hemisphere

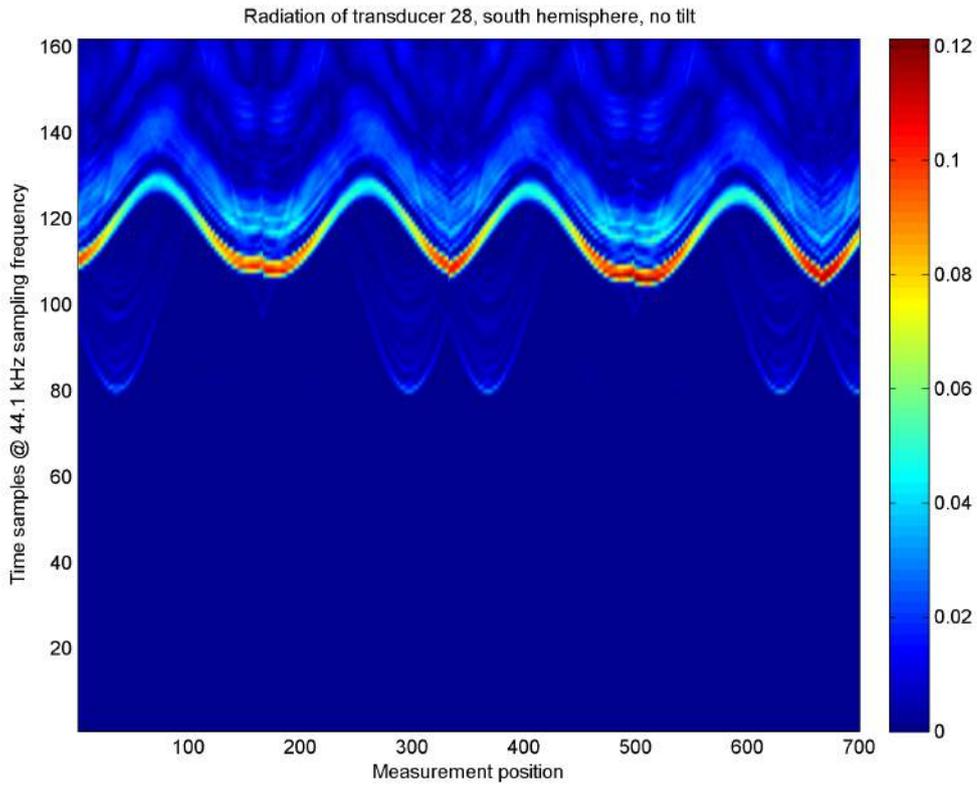


Figure 3.28: Radiation of transducer 28 in dB, No Tilt, Lower Hemisphere

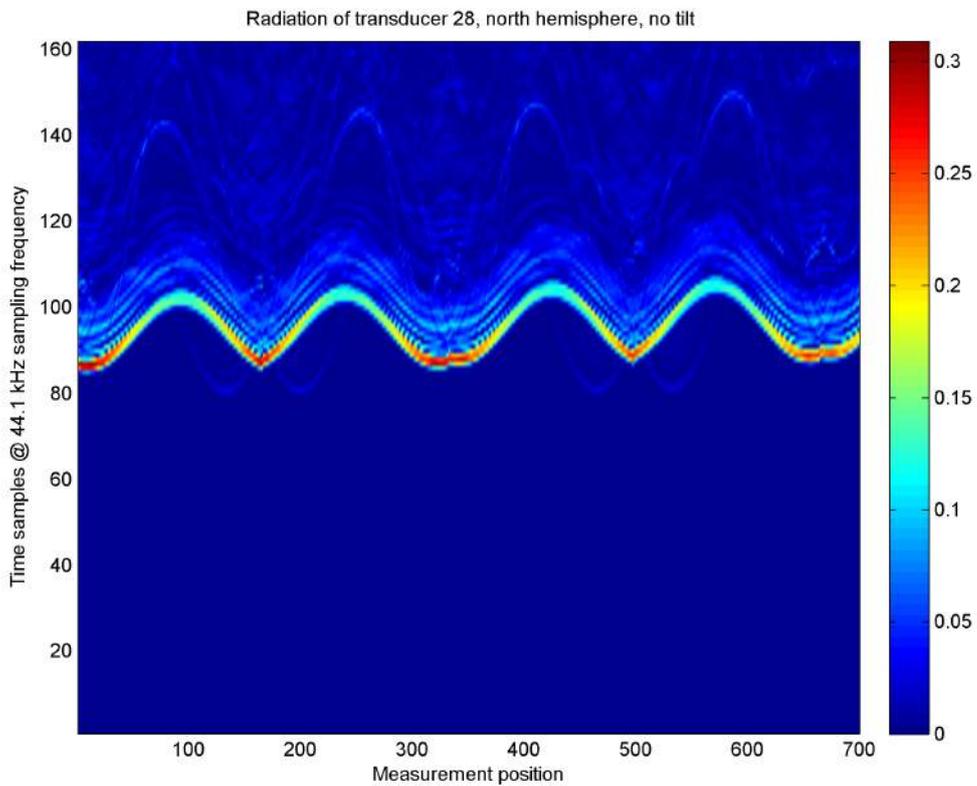


Figure 3.29: Radiation of transducer 28 in dB, No Tilt, Upper Hemisphere

In pictures 3.24 to 3.29 two different facts become clear: The measurement of the lower hemisphere shows a greater corruption than the upper hemisphere. This effect is a consequence of the non symmetric thrust bearing structure, which acoustically affects the lower hemisphere greater than the upper hemisphere, as it can be seen in figure 2.1. Moreover, the number of reflections grows with each measurement so, although the interleaved sweep measurement technique takes into account the decay of the room and the impulse response time of the transducer itself, it can not window the reflections on measurement periphery (thrust bearing, swivel arm, amplifiers...)away. Therefore, the signal corruption increases in each measurement.

3.5 Analysis of the spherical harmonics' energy

The energy within the different spherical harmonic orders gives a first glimpse about the frequency dependent radiation. By means of the order energy, the suitability of the measurement procedure as well as the suitability of the measurement results has been verified. But the order energy does not reveal the energy of each spherical harmonic. Since the radiation of an arbitrary directivity pattern is created by scaling the spherical harmonics, it is necessary to study the energy of each spherical harmonic to verify if the radiation pattern of the source indeed excites a large number of spherical harmonics within the orders free from error energy. Furthermore, determining either if the error energy is spread over all spherical harmonics or just over specific ones delivers an elegant and solid proof for the assumption that the error energy is produced by the reflections of the thrust bearing holding the source.

It becomes clear in picture 3.23, that the error energy follows a certain pattern: Some orders are intensely excited while others are not, which would imply that the error energy comes from specific directions. To determine these directions and the intensity of the involved spherical harmonics ,a spherical harmonic based compressor was developed. In the first step, the energy of each spherical harmonic is divided by the energy held within the order it belongs too. Equation 3.6 expresses this operation in dB:

$$e_{nm,ref} = 10 \cdot \log \frac{|f|_{nm}}{\sum_{m=-n}^n |f_{nm}|^2} \quad (3.6)$$

Each of the values for $e_{nm,ref}$ is stored in a matrix $E_{nm,ref}$, which is computed for each frequency present in the radiation function of each transducer:

$$E_{nm,ref} = \begin{pmatrix} e_{00,f_{min},ref} & e_{-11,f_{min},ref} & e_{01,f_{min},ref} & \cdots & e_{N-1N-1,f_{min},ref} \\ e_{00,f_2,ref} & e_{-11,f_2,ref} & e_{01,f_2,ref} & \cdots & e_{N-1N-1,f_2,ref} \\ e_{00,f_3,ref} & e_{-11,f_3,ref} & e_{01,f_3,ref} & \cdots & e_{N-1N-1,f_3,ref} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ e_{00,f_{max},ref} & e_{-11,f_{max},ref} & e_{01,f_{max},ref} & \cdots & e_{N-1N-1,f_{max},ref} \end{pmatrix} \quad (3.7)$$

The energy of each spherical harmonic has to be referenced to the total energy within its order because each order holds a different number of coefficients. If the value of $e_{nm,ref}$ is less than a given threshold in dB, this particular energy is discarded. The threshold therefore expresses the desired dynamic range which is left uncompressed. The highest value for $e_{nm,ref}$ is always 0 dB for $e_{00,ref}$, since order $N = 0$ only hold 1 spherical harmonic. The performance of the compressor has been tested for 4 different thresholds. The maximum value has been normalized to 1 and the minimum to 0.

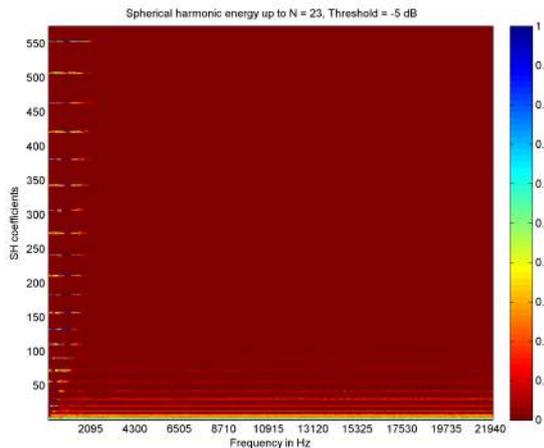


Figure 3.30: Threshold = -5 dB

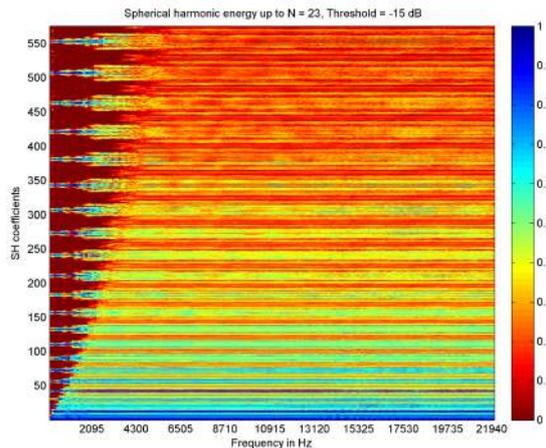


Figure 3.31: Threshold = -15 dB

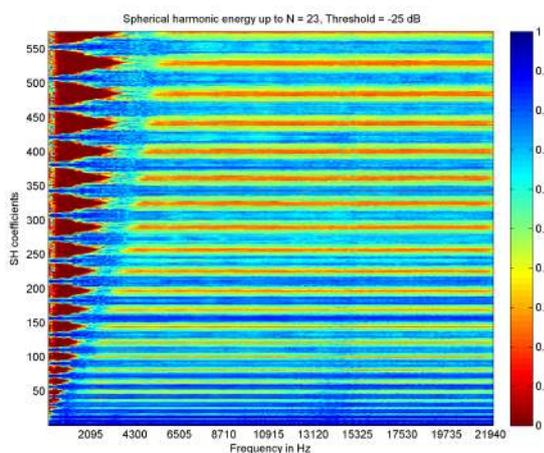


Figure 3.32: Threshold = -25 dB

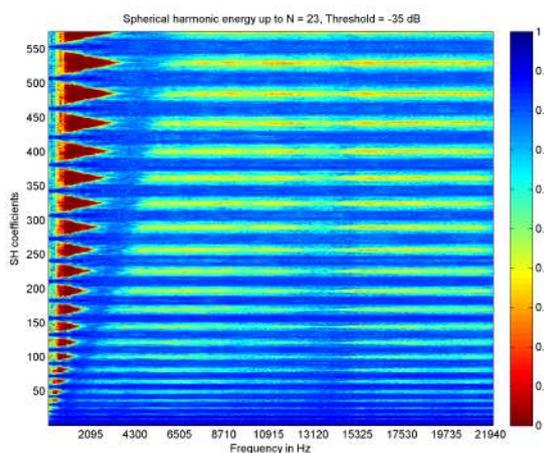


Figure 3.33: Threshold = -35 dB

The normalization of the data could induce to error since the values for $e_{nm,ref}$ increase for growing dynamics. Recalling equation 3.6, it becomes clear that it is not the energy of the spherical harmonics that rises, which would not make sense, but their relative strength. By increasing the dynamic (lowering the threshold), higher spherical harmonics are taken into account. Although the directivity measurement was performed up to order $N = 82$, only coefficients up to $N = 23$ have been regarded since the measurement source theoretically only allows directivity control up to order $N = 23$ at a frequency of 8 kHz due to constructional constraints [7]. Figure 3.32 shows especially clearly the strength of this optimized source: All coefficients within orders up to $N = 23$ are excited, which meets the requirement posed in section 2.5 on a practical level: Not all spherical harmonic functions within all orders N at every frequency are excited by this source, but all coefficients within a number of orders in which no measurement error at every frequency can be warranted.

The results obtained by using the spherical harmonic based compressor can be also be used to describe the error energy: In the low frequency range, the error energy is present even if only spherical harmonics within 5 dB dynamic range are regarded (picture 3.30). In higher frequencies, the presence of single energy rich coefficients is not tragic since, as it can be seen in pictures 3.31 to 3.33, the source excites these spherical harmonics by itself. Nevertheless, it becomes clear that the error energy is not caused by the excitation of all spherical harmonics within the orders but much more due to the excitement of a few of them.

To identify these coefficients, the matrices $E_{nm,ref}$ for each transducer have been converted into binary matrices: 1 if the value of $e_{nm,ref}$ lies within the dynamic range set by the threshold, 0 if

lies outside. The binary representation of the spherical harmonics' energy has been calculated for orders up to $N = 4$ (thresholds -5 dB) and orders up to $N = 9$ (threshold = -15 dB).

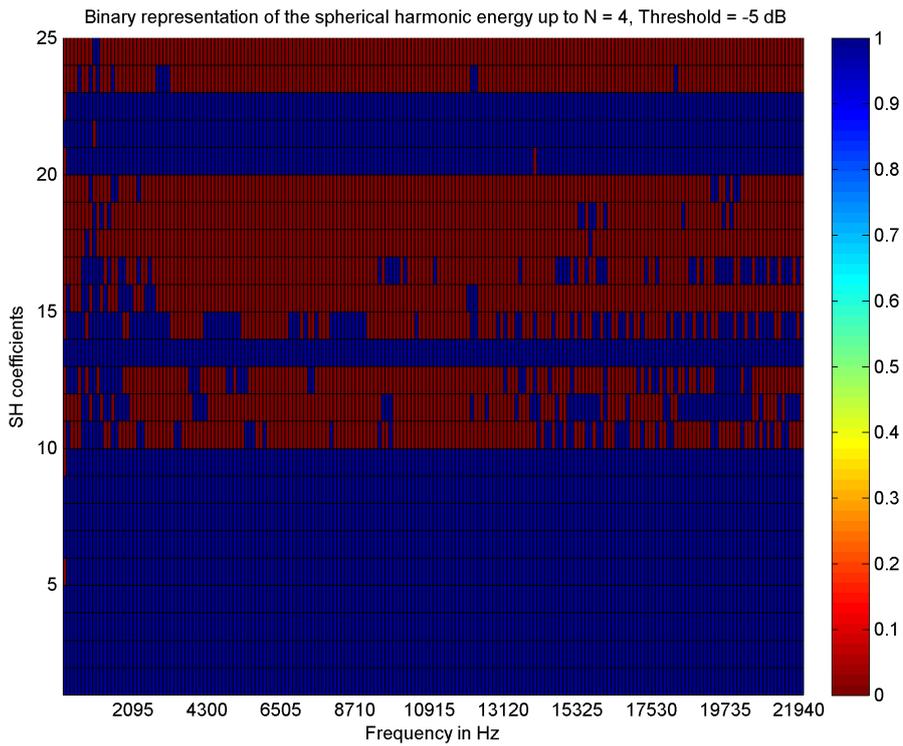


Figure 3.34: Binary representation of the energy within a dynamic range of 5 dB

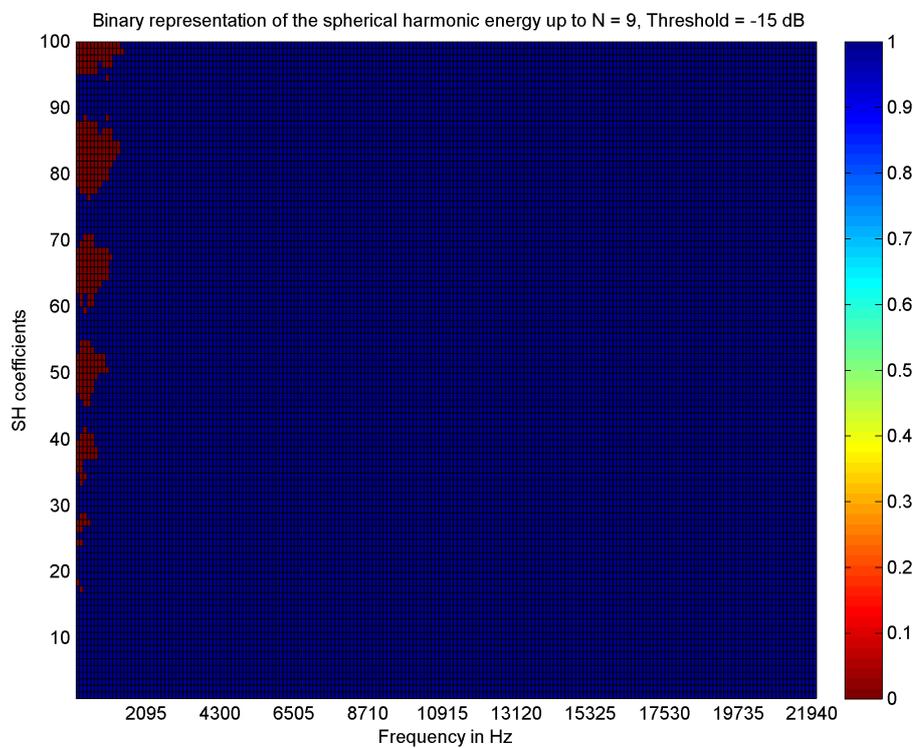


Figure 3.35: Binary representation of the energy within a dynamic range of 15 dB

The coefficients up to order $N = 2$ (1-10) are all excited by the source while the energy of the coefficients in higher orders is concentrated in coefficients of low degree, as it can be seen in figure 3.34. This behaviour is also observable in higher orders when the dynamic range is widened up to 15 dB. In figure 3.36 the referenced energy within a dynamic range of 15 dB is shown for frequencies up to 4430 Hz. It can be clearly seen that, especially at low frequencies and for lower orders, that

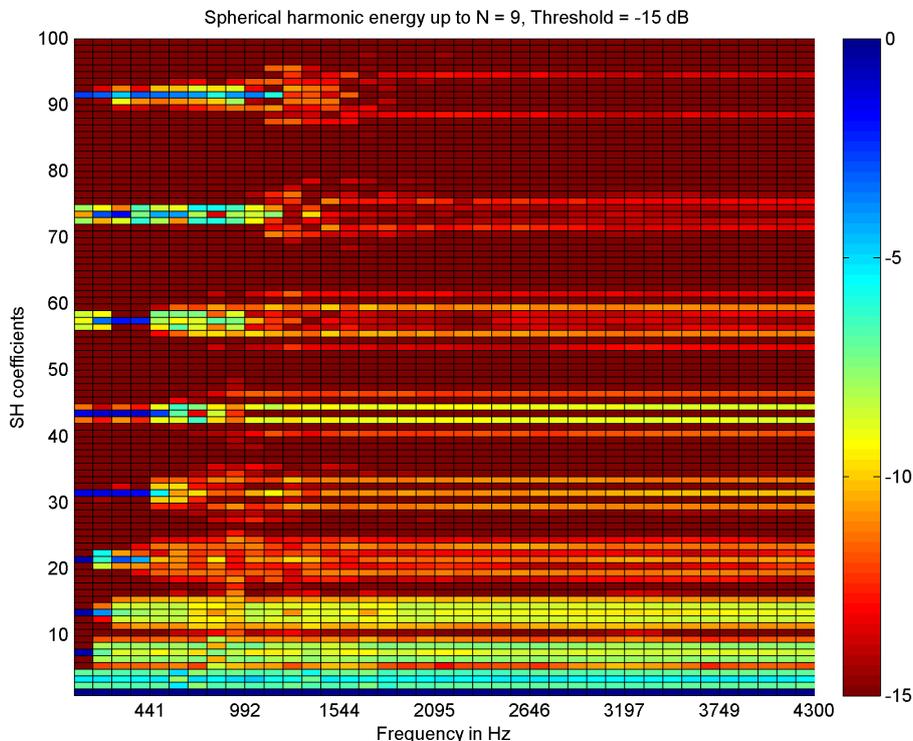


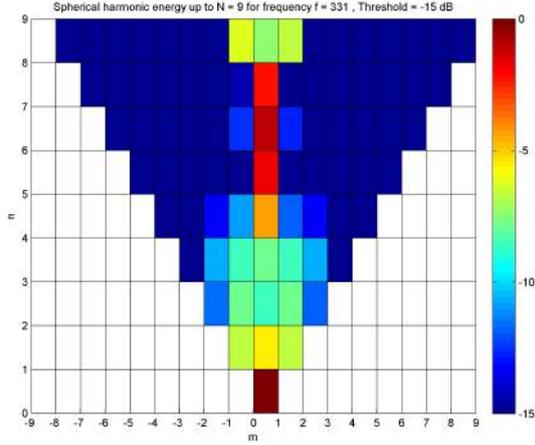
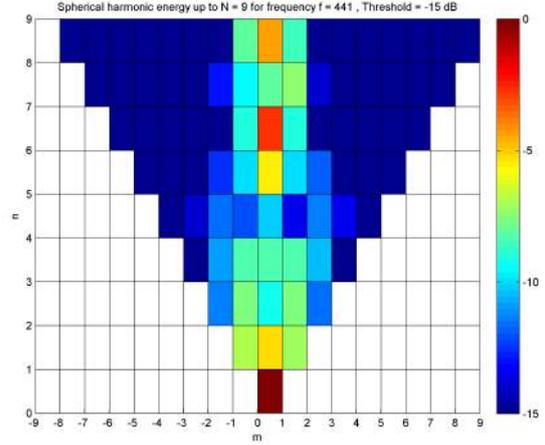
Figure 3.36: Energy within a dynamic range of 15 dB

coefficients f_{n0} show the greatest energy. This fact is shown for frequencies $f = 331$ Hz and $f = 441$ Hz in figures 3.37 and 3.38

As shown in picture 2.5, spherical harmonics of degree $m = 0$ are functions pointing in the positive and negative vertical direction. Since the existence of a ceiling reflection is not probable (semi anechoic chamber!), the only option left is the considering reflections on the periphery holding the source.

3.6 Spherical harmonic decomposition using experimental order limitation

Up to this point, the energy caused by reflections has been regarded as a negative consequence of the measurement set up. Nevertheless, it could be argued that this extra energy is actually beneficial to later on perform a directivity synthesis: Since the target directivity $\hat{\mathbf{d}}_T$ can be arbitrarily chosen, the measurement source should ideally hold energy in every spherical harmonic coefficient within any order. Although this provocative approach may seem tempting, it needs to be considered that


 Figure 3.37: $f = 331$ Hz

 Figure 3.38: $f = 441$ Hz

the reproducibility of this results has not been proven. Therefore, taking into account energy which distribution underlies unknown conditions would mean to give up control over the synthesis process. Hence, the question at this point is how to estimate the maximal number of orders necessary to describe the spatial structure of each frequency.

As exposed in [32], the *DSHT* poses a trade off between the angular resolution and the error magnitude. Moreover, dependant on the order N and the available computational resources, a matrix inversion can easily become an impossible task to handle. In [18], a frequency dependent experimental order limitation was introduced, based on the approach presented in [32]. As a rule of thumb, the truncation order should be roughly equal to the wave number $\nu = \frac{1}{\lambda}$:

$$N_{trunc}(N_{add}) = \lfloor kr_{min} \rfloor + N_{add} \quad (3.8)$$

r_{min} is the radius of the minimal sphere that, from the centre of the source, encompasses the source and all secondary scattering objects. Choosing $\lfloor kr_{min} \rfloor$ as the truncation order is based on the fact that the higher the radiated frequency, the more complex becomes its radiation pattern. Nevertheless, N_{add} was introduced in [18] to ensure the truncation order is high enough and was set to $N_{add} = 2$. Using this technique, the dimensions of matrix \mathbf{Y} become frequency dependent $[(N + 1)^2, L] \rightarrow [(N_{trunc} + 1)^2, L]$

In equation 3.8, r_{min} is the minimal radius of a virtual sphere encompassing all active sources. The minimal radius r_{min} has to be chosen greater than the source radius r_s since the thrust bearings are reflecting objects that need to be considered. Moreover, to show the impact of the parameter $N_{reserve}$, the order dependent radiation for all transducers was calculated using $N_{reserve} = 2$ and $N_{reserve} = 3$. To show the effect of the radius, the order dependent radiation for all transducers has been computed using $r_{min} = r_s = 0.2m$ and $r_{min} = 0.3m$:

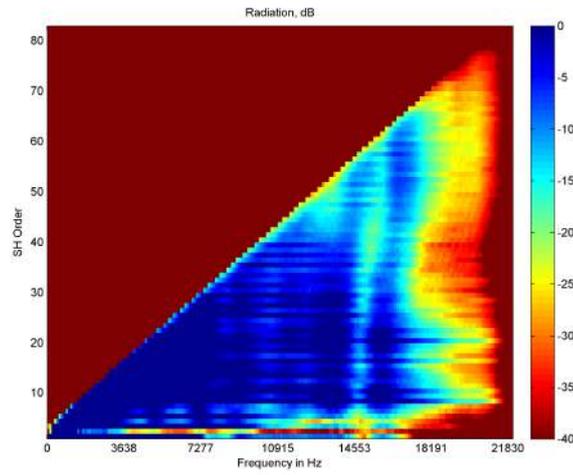


Figure 3.39: $N_{reserve} = 2$, $r_{min} = 0.2m$

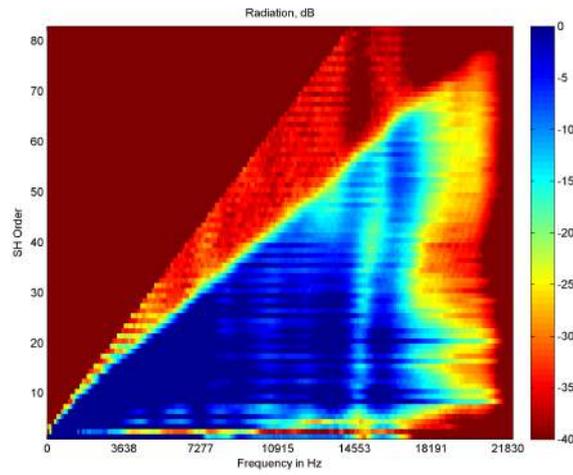


Figure 3.40: $N_{reserve} = 2$, $r_{min} = 0.3m$

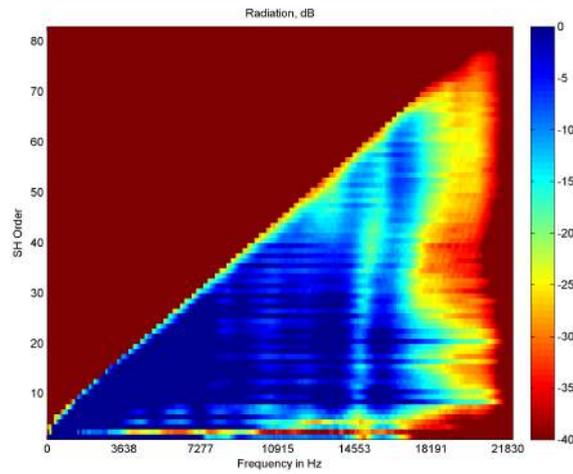
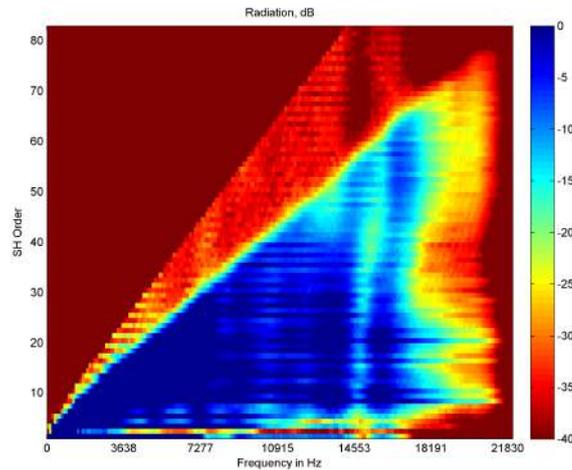


Figure 3.41: $N_{reserve} = 3$, $r_{min} = 0.2m$


 Figure 3.42: $N_{reserve} = 3$, $r_{min} = 0.3m$

Figures 3.40 and 3.42 show the great effect the radius has on the suitability of the kr-limit: A variation of just 0.1 m shows a large increase of the error energy in the higher orders all over the frequency range while an increase of $N_{reserve}$ is barely noticeable. What this figures reveal is that a correct spherical harmonic decomposition can not be performed by using all coefficients to decompose the directivity function at every frequency. The kr-limit can be regarded as a spherical harmonic low pass function, which instead of being defined by a cut off frequency f_{cut} is defined by a maximal order N.

To show how good the kr-limit performs, the radiation patterns shown in figure 3.7 to 3.17 have been expanded into spherical harmonics by using equation 3.8. The frequency dependent number of coefficient used to decompose the measured directivity functions were calculated using $r_{min} = 2$ and $N_{add} = 2$.

f in Hz	r_{min} in m	$N_{trunc}(N_{add} = 2)$
500	0.2	3
1000	0.2	5
2000	0.2	9

 Table 3.2: kr - limit, $r_{min} = 0.2$, $N_{add} = 2$

Pictures 3.43 to 3.47 show the spherical harmonic expanded and the measured data for $f = 2$ kHz. It can be clearly seen, that the kr-limit is a rather precise rule of thumb to correctly expand radiation patterns into spherical harmonics. In chapter 7 the spherical harmonic expanded data that corresponds to the measured radiation patterns displayed in figures 3.7 to 3.17 can be found. No scale has been given on these figures since the measurements were performed uncalibrated: While the absolute values of the sound pressure level distribution are of great interest when it comes to plan a loudspeaker system, the important aspect of this measurement was to determine as detailed as possible how the frequency dependent radiation pattern of each transducer looks like.

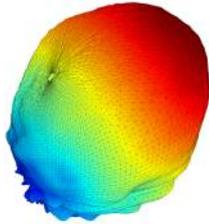


Figure 3.43: Transducer 1 for $f = 2$ kHz

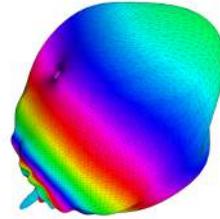


Figure 3.44: Transducer 1 for $f = 2$ kHz, SH

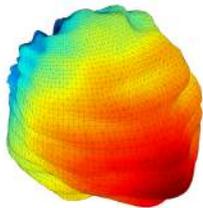


Figure 3.45: Transducer 11 for $f = 2$ kHz

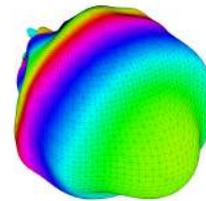


Figure 3.46: Transducer 11 for $f = 2$ kHz, SH

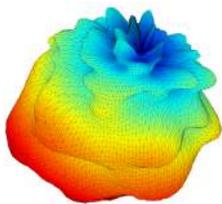


Figure 3.47: Transducer 28 for $f = 2$ kHz

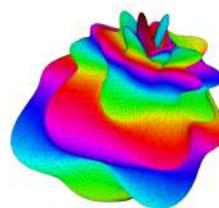


Figure 3.48: Transducer 28 for $f = 2$ kHz, SH

4 Measurement Set up Preparation

The directivity measurement for each transducer mounted on the optimized measurement system has been carried out using the measurement set up explained in chapter 3. The measurement set up is highly elaborate and therefore tainted with many different error sources. In [27], the impact of geometrical errors was analysed and method to estimate and minimize it was proposed for arbitrary measurement set ups. It motivated the development of a method capable of minimizing the geometrical measurement errors based on empirical geometrical error

4.1 Swivel arm positioning error

A strategy to empirically evaluate the position error of the swivel arm was developed: Using the measurement set up described in [10], a laser was mounted at the place of the loudspeaker. The laser, mechanically connected to the base of the swivel arm, rotated with it. At the arm's end, a piece of millimetre paper with a two dimensional Cartesian coordinate system depicted on it was mounted. The vertical component was named dy' and the horizontal component was named dx' . This way, the deviation from the ideal point (the origin of the coordinate system depicted in the millimetre paper) could be empirically measured for the *local coordinates* dy' and dz' . Since the swivel arm rotates along several elevation angles, the origin of the Cartesian coordinate system and the axes orientation vary with each elevation angle. Each of these coordinate systems is a *local coordinate system*. The radius was given by the laser itself. The swivel arm was rotated from $\vartheta = 90^\circ$ to $\vartheta = 0^\circ$ in 10° steps and the meter-reading was performed for each elevation angle. The evaluation points have been defined on a global coordinate system in which the three different components are denoted as dx , dy and dz . The components dx , dy and dz present the same orientation as the axis x , y and z in figure 7.2

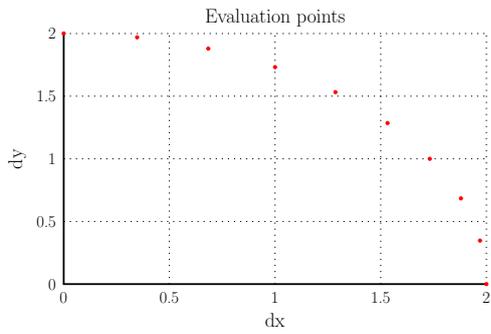


Figure 4.1: Evaluation points

The laser was mounted in a way that the deviation is zero at $\vartheta = 90^\circ$. The swivel arm used in the measurements was developed after the original swivel arm could not longer be used due to an accident. The error evaluation results

The first evaluation was carried out by hanging 3 objects of different weights on the swivel arm: a loudspeaker (920 g), a normed weight (500 g) and a dummy microphone (64 g). Pictures 4.2 to 4.4 show the measured deviation produced by each object along the three global axis, which are given in the range of centimetres.

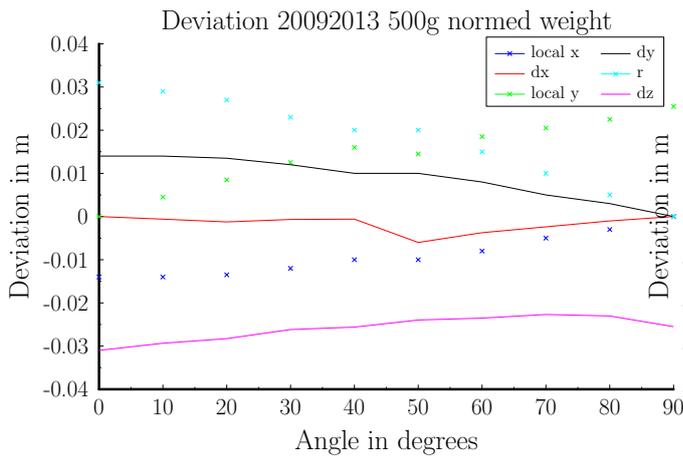


Figure 4.2: Normed weight

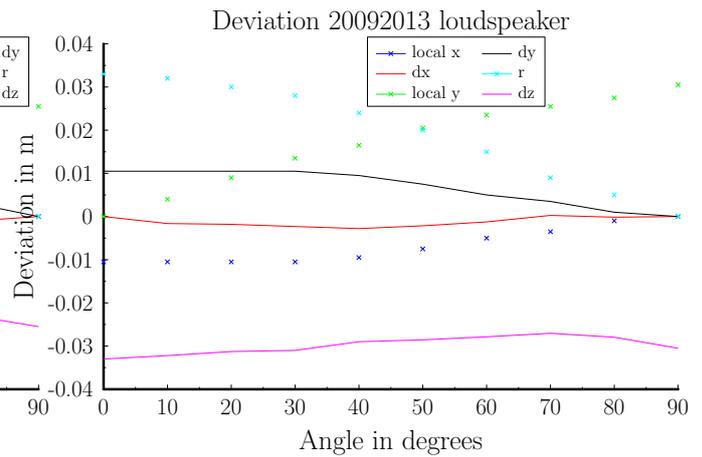


Figure 4.3: Loudspeaker

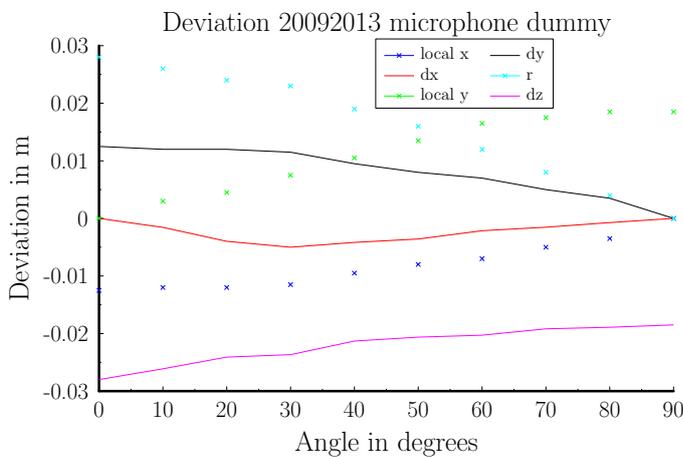


Figure 4.4: Microphone dummy

To discuss this results in a suitable manner, it is necessary to take into account the drawbacks of this measurement method. The laser pointer is originally intended to perform distance measurements, so the width of the laser beam is too wide. In order to be consistent, the point in the middle of the beam was chosen to be the error evaluation point. Nevertheless, before considering the quantitative data, it is necessary to prove the measured data on physical plausibility. The radial component r in the figures 4.2 to 4.4 reveals the same trend for all 3 weights: At elevation angle $\vartheta = 0^\circ$, the sagging of the swivel arm is greatest and so is the measured error too. The absolute value of the error diminishes with the elevation angle since the ideal measurement position was defined at $\vartheta = 90^\circ$. This proves for the plausibility of the measured data. A further proof of the reliability of this error evaluation method can be found by comparing the deviation calculated on the global z axis, which represents the vertical deviation of the swivel arm: it is possible to observe, that the microphone dummy causes the lowest sagging over all the evaluated positions, since it is the lightest of all three objects. Although all 3 measured deviations along the z axis have a similar trend and a similar value range within millimetres, the impact of the testing object's weight can be clearly seen. The error measured along the global y coordinate, which was also referenced to the elevation angle $\vartheta = 90^\circ$, also shows a plausible trend, but remains much more stable due to the construction restrictions of the swivel arm. Though being almost zero on most of the evaluated points, the azimuthal error can not be neglected: If only the elevation angle is corrected at each measurement point, the different values for the azimuthal error would corrupt the circular shape of the measurement position arc, which has been depicted in 4.1. In figure 4.2 the global x coordinate value varies strongly between the angles 40° and 50° , which is certainly forced by an evaluation error.

4.2 Reproducibility

At this point, the empirical evaluation of the deviation has been confirmed suitable. To validate not only the trend within the measured deviations but also the amount of the deviation, a second evaluation was carried out.

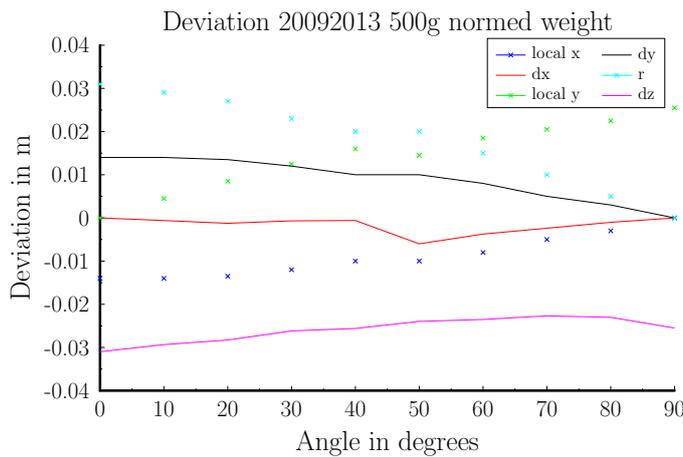


Figure 4.5: Normed weight, first measurement

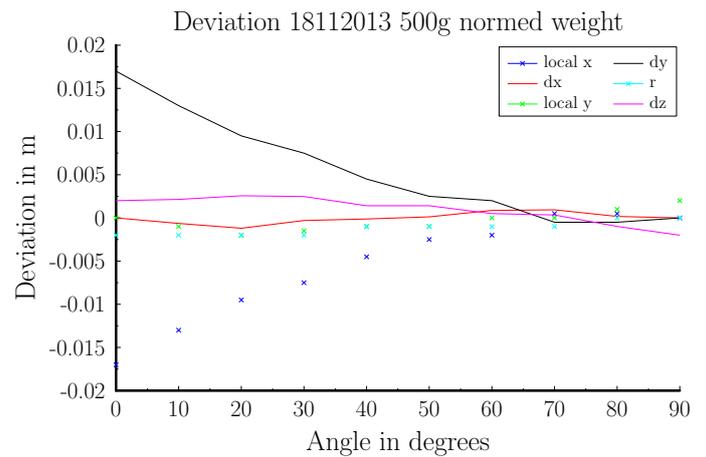


Figure 4.6: Normed weight, second measurement

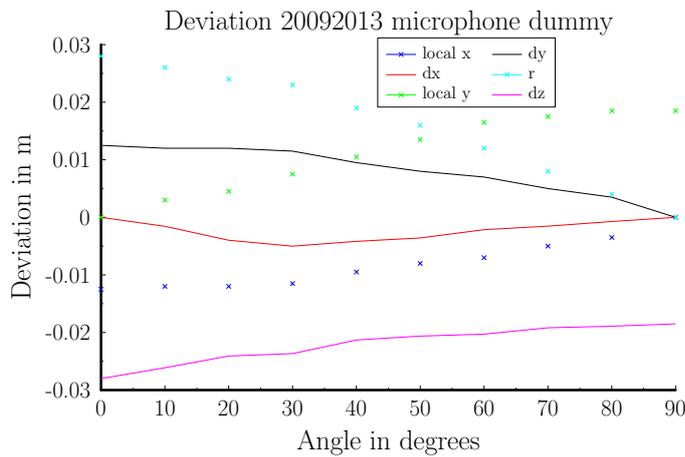


Figure 4.7: Microphone dummy, first measurement

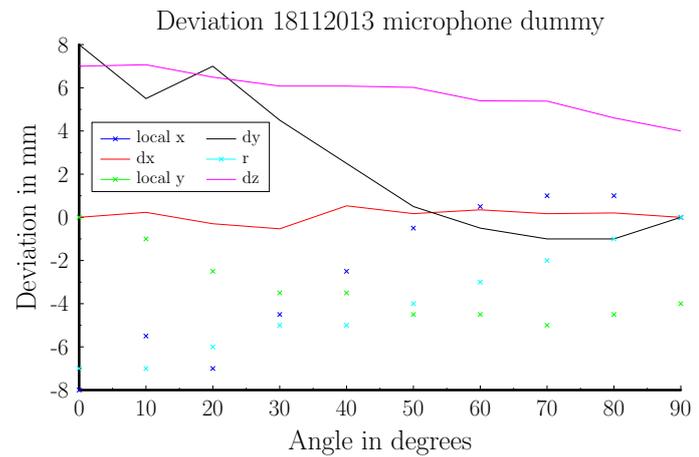


Figure 4.8: Microphone dummy, second measurement

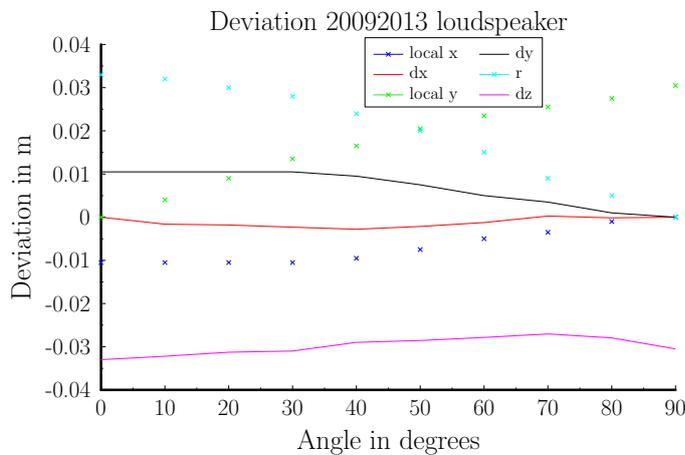


Figure 4.9: loudspeaker, first measurement

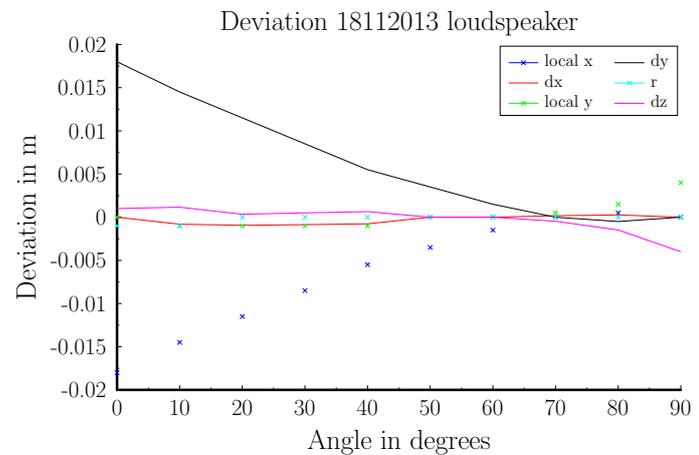
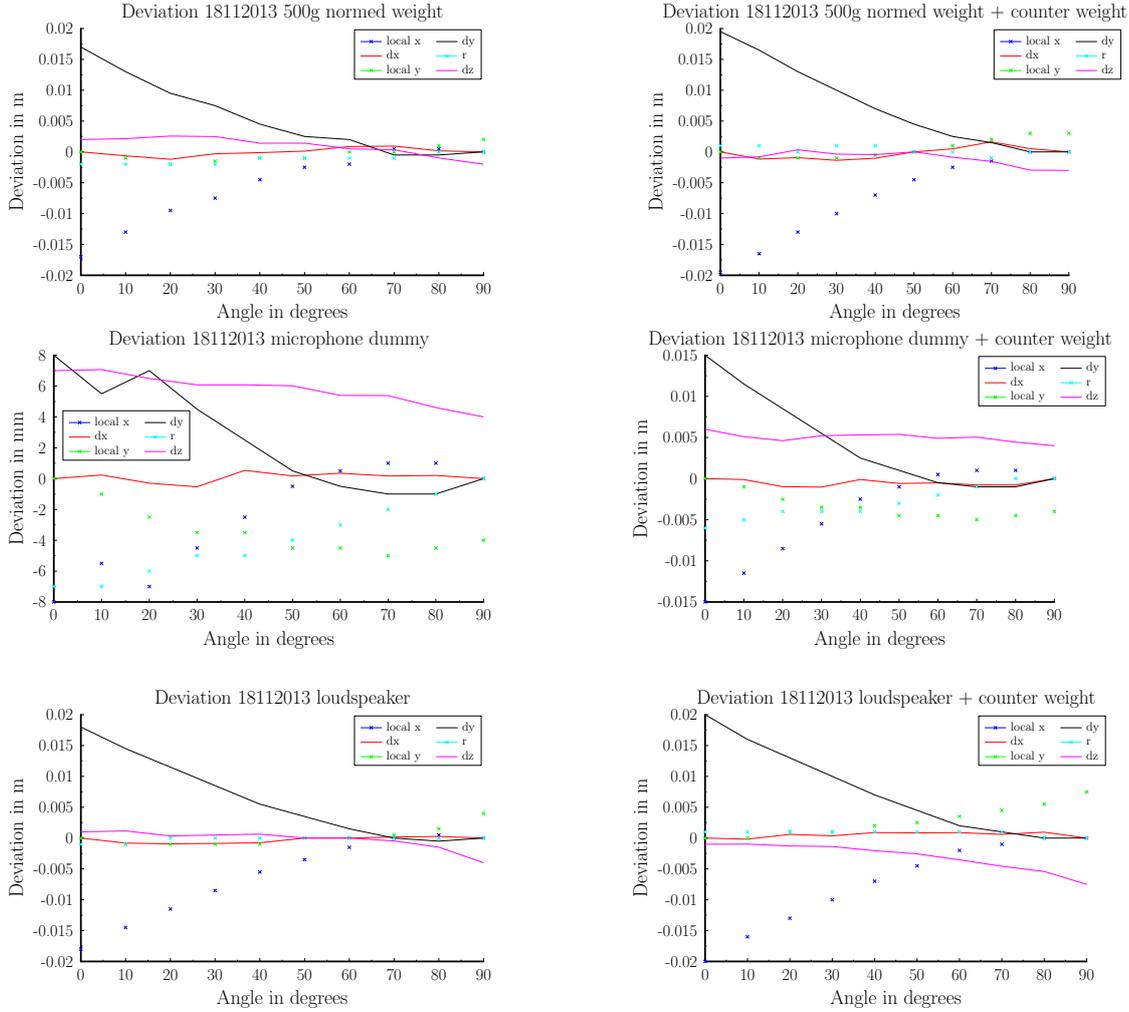


Figure 4.10: loudspeaker, second measurement

Figures 4.5 and 4.6 show the deviation measured in the local coordinate system and its transform into the global coordinate system. The comparison between both measurements shows that positioning errors are not exactly reproducible, but the trends are similar: The deviation along the global x axis does not change while the deviation along the y axis shows different shapes along the evaluation points: In figure 4.6, the deviation along the y axis decreases much sooner than in the first measurement, depicted in 4.5, where the deviation in the y axis remains constant. This effect becomes clear by comparing with the evaluations using the 500g normed weight: In the first measurement, the error in dy at $\vartheta = 60^\circ$ is still above 1 centimetre while in the second measurement, the error at the same position is below 0.25 centimetres. The comparison of the evaluated error in the vertical axis z does not show any sign of reproducibility. During the second evaluation, the effect of a counter weight at the beginning of the arm was also studied. The counter weight helps stabilizing the arm and minimizing the impact of the sagging. Before using it in a measurement, its effect on the deviation pattern had to be studied. The stabilizing effect of the counter weight can be clearly seen along the y axis error throughout the measurement, especially for greater elevation angles. On the one hand, the measurement with the microphone dummy and the counter weight shown in figure 4.2 does not present the error a measurement error that can be seen in figure 4.8 for y . The error along z does not seem to vary much by the use of the counter weight, it even slightly grows when the loudspeaker is mounted at the end of the swivel arm, as it can be seen in figure 4.2. All in all, the effect of the counter weight was proven to be beneficial and therefore the transducer directivity measurements were conducted using it. Summing up, both error evaluations show a



lack of reproducibility that gave rise to the creation of a swivel arm position correction function using the ITA Toolbox. This algorithm was integrated into the measurement routine and has been tested on a Head Related Transfer Function measurement. It has to be taken into account that the distance between the evaluation point and the laser is 2 meters. Therefore, an evaluation error of 2 mm is equivalent to an error of less than 1° in the elevation angle ϑ , which can be neglected in most cases.

4.2.1 Swivel arm positioning correction

As stated in 4.1, at each elevation angle the orientation of the microphone mount changes. This fact implies that the empirically measured error along the local coordinates is measured using a different local coordinate system at every measurement position. In order to perform a correction of the positioning, the first step consists of projecting the data back to the global coordinate system, which is the Cartesian coordinate system presented in 7.1. Based on the picture 4.11, a rotation matrix \mathbf{R} was defined [13]:

$$\mathbf{R} = \begin{pmatrix} 0 & -\cos \vartheta & -\sin \vartheta \\ -1 & 0 & 0 \\ 0 & -\sin \vartheta & -\cos \vartheta \end{pmatrix} \quad (4.1)$$

Picture 4.11 shows how the local coordinate systems were defined: By performing the multiplication

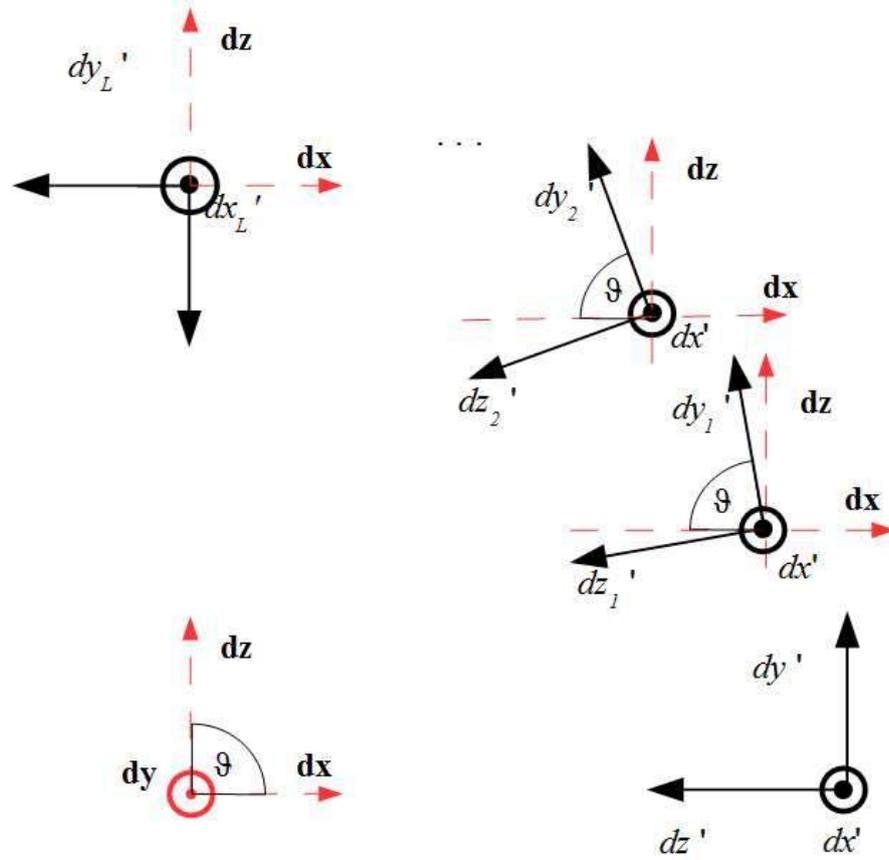


Figure 4.11: Local coordinate systems for $\vartheta = 0^\circ$ to $\vartheta = 90^\circ$

of the local components dx' , dy' , dz' with 4.1, the deviations measured on the local coordinate systems are projected onto the global coordinate system dx , dy , dz :

$$\mathbf{R} \cdot \begin{pmatrix} dx' \\ dy' \\ dz' \end{pmatrix} = \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} \quad (4.2)$$

After having projected the data onto the global coordinate system, the projection is used to create a polynomial of order 3 for each coordinate to estimate the error for all measurement points included in a desired measurement scheme. This way, the evaluation of a small amount of well distributed measurement points is sufficient to define the necessary error functions for each point along the circular path described by the swivel arm. After having evaluated the error for each elevation angle, a matrix \mathbf{P} is created. It contains the real positions of the swivel arm approximated by the polynomials, representing $\mathbf{p}(\vartheta)$ a vector containing the components for the three axis.

$$\mathbf{P} = \begin{pmatrix} \mathbf{p}(\vartheta_0) \\ \mathbf{p}(\vartheta_1) \\ \vdots \\ \mathbf{p}(\vartheta_{N-1}) \end{pmatrix} = \begin{pmatrix} p_x(\vartheta_0) & p_y(\vartheta_0) & p_z(\vartheta_0) \\ p_x(\vartheta_1) & p_y(\vartheta_1) & p_z(\vartheta_1) \\ \vdots & \vdots & \vdots \\ p_x(\vartheta_{N-1}) & p_y(\vartheta_{N-1}) & p_z(\vartheta_{N-1}) \end{pmatrix} \quad (4.3)$$

The polynomial order was set to 3 to prevent over fitting and were tested with the error evaluations discussed in 4.1: Figures 4.2.1, 4.2.1 and 4.2.1 show the approximation based on the polynomials for the evaluated angles. The robustness against over fitting can be clearly seen in 4.2.1 in which the evaluation error is not approximated.

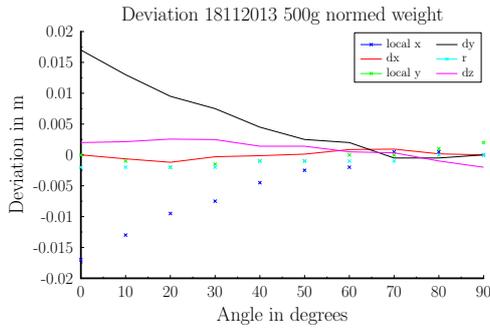


Figure 4.12: Measured deviation for the 500 g normed weight

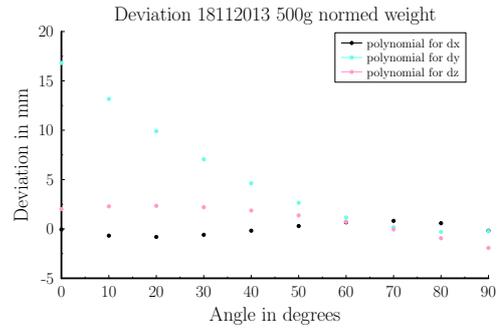


Figure 4.13: Polynomial approximation

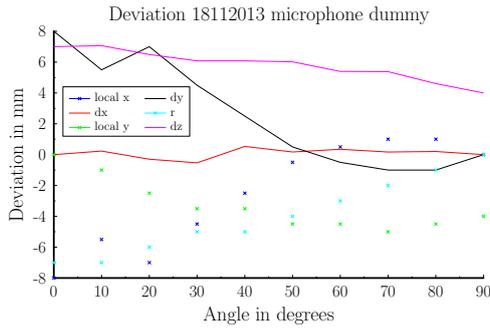


Figure 4.14: Measured deviation for the dummy microphone

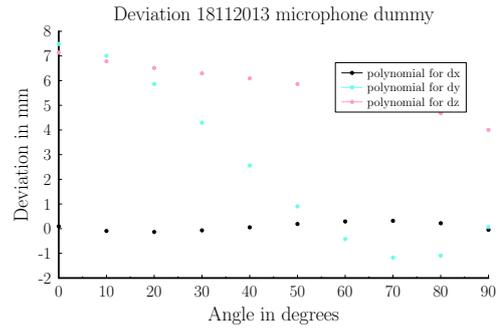


Figure 4.15: Polynomial approximation

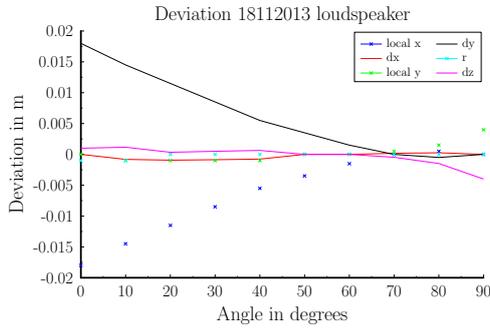


Figure 4.16: Measured deviation for the loudspeaker

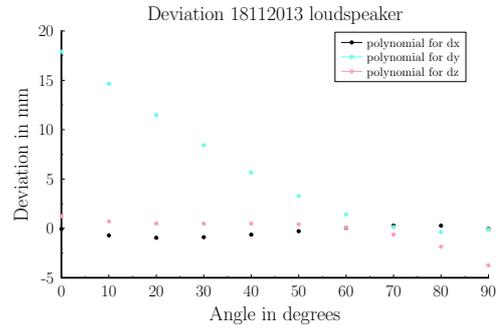


Figure 4.17: Polynomial approximation

Figure 4.18: Deviation in global and local coordinates for 3 different weights

To perform the correction in terms of the azimuth and the elevation angle, \mathbf{P} has to be rewritten in spherical coordinates:

$$\mathbf{P} = \begin{pmatrix} p_r(\vartheta_0) & p_\varphi(\vartheta_0) & p_\vartheta(\vartheta_0) \\ p_r(\vartheta_1) & p_\varphi(\vartheta_1) & p_\vartheta(\vartheta_1) \\ \vdots & \vdots & \vdots \\ p_r(\vartheta_{N-1}) & p_\varphi(\vartheta_{N-1}) & p_\vartheta(\vartheta_{N-1}) \end{pmatrix} \quad (4.4)$$

By projection of the real swivel arm position onto the ideal swivel arm position, assuming that φ is zero, a correction term for the 3 coordinates of every angle can be given. The projection is performed by subtracting each vector $\mathbf{p}(\vartheta)$ by the ideal position vector $\mathbf{p}_{ideal}(\vartheta)$. Therefore, a

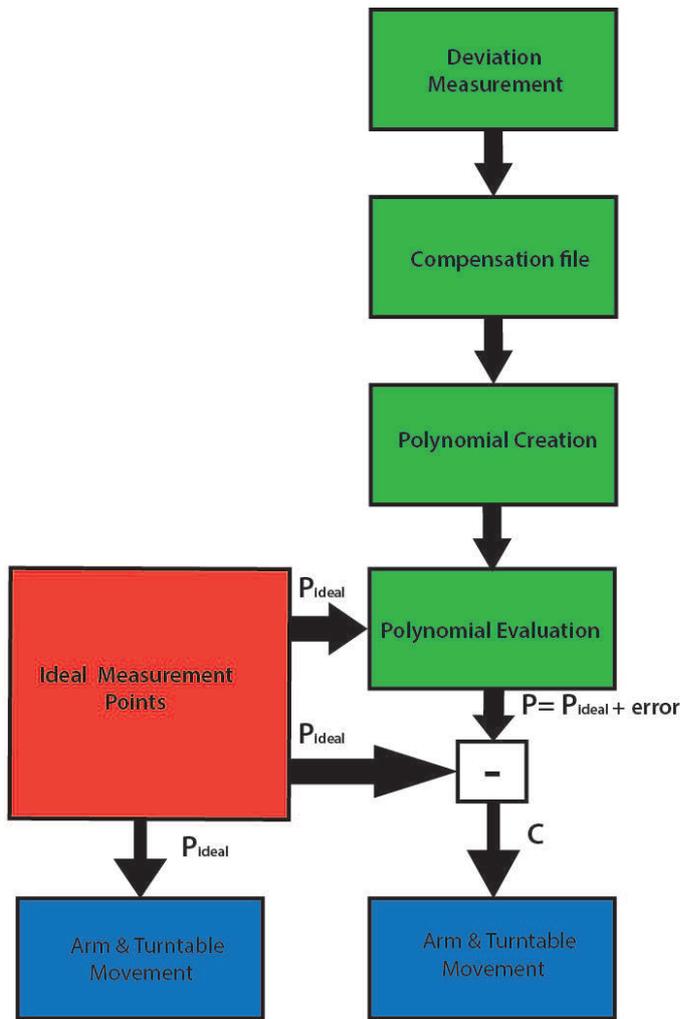


Figure 4.19: Stages of the measurement points correction strategy

matrix $\mathbf{P}_{\text{ideal}}$ has been defined in 4.5

$$\mathbf{P}_{\text{ideal}} = \begin{pmatrix} p_{\text{ideal},r}(\vartheta_0) & p_{\text{ideal},\varphi}(\vartheta_0) & p_{\text{ideal},\vartheta}(\vartheta_0) \\ p_{\text{ideal},r}(\vartheta_1) & p_{\text{ideal},\varphi}(\vartheta_1) & p_{\text{ideal},\vartheta}(\vartheta_1) \\ \vdots & \vdots & \vdots \\ p_{\text{ideal},r}(\vartheta_N) & p_{\text{ideal},\varphi}(\vartheta_N) & p_{\text{ideal},\vartheta}(\vartheta_N) \end{pmatrix} \quad (4.5)$$

Using 4.3 and 4.5, a correction matrix \mathbf{C} has been defined as follows:

$$\mathbf{C} = \mathbf{P} - \mathbf{P}_{\text{ideal}} \quad (4.6)$$

The measurement position correction explained in this section is depicted as a flow graph in figure 4.2.1: In figure [4.2.1] the movement of the swivel arm and turntable has been split into two steps to improve its comprehension: Firstly, the swivel arm and the turntable move, yielding the ideal position defined in the ideal sampling. Then, the swivel arm and the turntable are moved by the specific correction value stored in \mathbf{C} . In practice, the swivel arm and the turntable are moved directly to the corrected position. The measurement points correction algorithm was used for the measurement of a full set of Head Related Transfer Functions (HRTF) and yielded a good result. Time dependency is a factor that can not be neglected when the usage of this strategy is being considered: The swivel arm positioning error varies with time, so the measured error will not be valid if the swivel arm is used a few days later.

5 Radiation pattern control

In equations 2.17 and 2.18, the radiation pattern synthesis as well as the calculation of the weighting vector were introduced. At that stage, only the 28 physical transducers were considered to be used to synthesize an arbitrary radiation pattern. In section 2.4 it was shown that the number of spherical harmonic functions that can be correctly calculated depends on the number and *the distribution* of the sampling nodes, which now will be called *microphone positions*. Applying the reciprocity principle it becomes clear that the spherical harmonic radiation control depends on the transducer number on the spherical surface, the *transducer positions*. By taking into account the 28 physical transducers, spherical harmonic functions up to order $N = 4$ can be controlled, assuming every spherical harmonic function can be controlled, which depends on the sampling nodes distribution. By rotating the spherical source by a certain azimuthal angle φ , the physical transducers are located at new points on the sphere. They are called *virtual transducers* on the spherical surface. By superposing the physical transducers with the virtual transducers, a *virtual source* with $rot \cdot 28$ transducers is created. In 8, two modes are suggested to perform the rotation. The first mode implies rotating the source by multiples of $\frac{\pi}{12}$, yielding therefore a Gaussian sampling of $N = 11$. The second mode consists of rotating the source by multiples of $\frac{\pi}{24}$, firstly without tilting the sphere and then tilting the sphere by $\frac{4}{180} \cdot \pi$. This way, a Gaussian sampling of $N = 23$ can be generated. Due to computational restrictions, only the first mode has been investigated, experimenting with different rotation angles. At this point, the benefits of the spherical harmonics expansion become clear. The spherical harmonic decomposition releases the data from the constraints of the measurement points. Considering the dimensions of $\hat{\mathbf{f}}$ in 2.12, this important characteristic becomes clear:

$$\underbrace{\hat{\mathbf{f}}}_{\text{SH coefficients, 1}} = \underbrace{\mathbf{Y}^\dagger}_{\text{SH coefficients, discrete measurement positions}} \cdot \underbrace{\mathbf{f}}_{\text{discrete measurement positions, 1}} \quad (5.1)$$

This way, the rotation of $\hat{\mathbf{f}}$ can be done by manipulating the spherical harmonic decomposed directivity function measured at *one* orientation. Otherwise, the directivity of the measurement source had to be measured for each orientation, which would be very time consuming. The arbitrary rotation of spherical harmonic decomposed functions requires a *Wigner D rotation* [12]. Consequently, $\hat{\mathbf{D}}$ has to be rewritten as:

$$\hat{\mathbf{D}}_{\text{rot}} = [\hat{\mathbf{D}}_1 \quad \hat{\mathbf{D}}_2 \quad \hat{\mathbf{D}}_3 \quad \cdots \hat{\mathbf{D}}_O] \quad (5.2)$$

The mathematical formulation of the target directivity synthesis has to be rewritten to:

$$\hat{\mathbf{d}}_T = \hat{\mathbf{D}}_{\text{rot}} \cdot \mathbf{g}_T \quad (5.3)$$

The weighting vector \mathbf{g}_T is consequently defined as:

$$\mathbf{g}_T = \hat{\mathbf{D}}_{\text{rot}}^+ \cdot \hat{\mathbf{d}}_T \quad (5.4)$$

Equation 5.4 is an *inverse problem* in which the inverted matrix is not quadratic. Next section deals with the methods that can be applied to find valid solutions for these kind of mathematical problems.

5.1 Inverse problems

A set of linear equations can be expressed in terms of a multiplication between \mathbf{A} and the vector \mathbf{x} :

$$\mathbf{b} = \mathbf{A} \cdot \mathbf{x} \quad (5.5)$$

\mathbf{x} holds all unknown variables so the matrix \mathbf{A} needs to be inverted in order to find a solution. The matrix inversion is the most critical step to solve \mathbf{x} since an unique solution can only be calculated for full ranked, square matrices. Since $\hat{\mathbf{D}}_{\text{rot}}$ is a non - square matrix, the inverse matrix can only be approximated. To do so, the most common approaches will be presented in this section

5.1.1 Minimum Norm Least Squares

A widely used approximation method is based on the least squares method, which minimizes the error energy. Recalling equation 5.5, the error energy can be defined as follows [8]:

$$\|\mathbf{e}\|^2 = \|\mathbf{Ax} - \mathbf{b}\|^2 \rightarrow \min \quad (5.6)$$

Since the linear equation system has infinitely many valid solutions (overdetermined equation system), a valid vector \mathbf{x} with the minimum 2 norm has to be chosen. The *Moore - Penrose Inverse* \mathbf{A}^\dagger , commonly known as pseudo-inverse, performs both steps with a single calculation [8]:

$$\mathbf{x} = (\mathbf{A}^H \cdot \mathbf{A})^{-1} \cdot \mathbf{A} \cdot \mathbf{b} = \underbrace{\mathbf{A}^\dagger}_{\text{pinv}} \cdot \mathbf{b} \quad (5.7)$$

5.1.2 Tikhonov Regularisation

The Moore - Penrose Inversion is commonly used in signal processing topics since many mathematical problems are assumed to be *well posed* and its' computation can be easily implemented using *singular value decomposition*. As soon as this assumption is not valid (condition number $\kappa \gg 1$), which would occur if the measurement positions are irregularly distributed, \mathbf{A}^\dagger would become unstable. In the context of the discrete spherical harmonic transform, the orthogonality assumption of the spherical harmonic coefficients contained in \mathbf{Y} would not be valid any more [12]. Three questions come up at this point:

- To which extent do the measurement errors affect the inverse of matrix \mathbf{A}
- To which extent is the gain vector \mathbf{x} affected
- How sensitive is the inverse of matrix \mathbf{A} to different result vectors \mathbf{x}

The error sensitivity of a matrix is commonly defined by the *condition number* $\kappa = \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\|$ [28] for any matrix A . Considering a square matrix \mathbf{A} , its inverse \mathbf{A}^{-1} can be easily computed, yielding

that κ denotes the ratio between the greatest and the lowest eigenvalue. In eq.5.8, the *relative error* in the calculated vector x can be related to κ and the relative error of the result vector \mathbf{b} (without proof) [28]

$$\frac{\|\Delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \kappa \cdot \frac{\Delta \mathbf{b}}{\mathbf{b}} \quad (5.8)$$

$\frac{\|\Delta \mathbf{x}\|}{\|\mathbf{x}\|}$ and $\frac{\Delta \mathbf{b}}{\mathbf{b}}$ are the relative errors vectors \mathbf{x} and \mathbf{b} . Equation 5.8 poses a limit on the error which can be found on \mathbf{x} based on the error variation of \mathbf{b} , κ being sort of a gain factor that depends on the properties of the \mathbf{A}^{-1} : High values for \mathbf{A}^{-1} are given if \mathbf{A} is nearly singular, which means that its eigenvalues are very close to zero. Thus, inverting matrix \mathbf{A} could lead to numerical instability. Hence, the ideal condition number is 1 since it means that no information gets lost by inverting the matrix. In 5.3, the spherical harmonic decomposition performed with coefficients up to order $N=10$ yields a condition number $\kappa = 1$ while the one performed with $N = 15$ yields a condition number of $\kappa = 1.1118 \cdot 10^{17}$, which is explained by the big energetic differences among the spherical harmonic coefficients of the transformed function.

The *Tikhonov regularization* consists of adding a regularization matrix $\mathbf{\Gamma}$, which is conveniently scaled by the positive real valued factor α , to increase the condition number of the matrix \mathbf{A} . In this work, $\mathbf{\Gamma}$ has been defined as the identity matrix \mathbf{I} . Equation 5.6 has to be rewritten as:

$$\|\mathbf{e}\|^2 = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 + \|\mathbf{\Gamma}\|^2 \rightarrow \min \quad (5.9)$$

Analogously, equation 5.7 is redefined as:

$$\mathbf{x} = (\mathbf{A}^H \cdot \mathbf{A} + \alpha \cdot \mathbf{\Gamma})^{-1} \cdot \mathbf{A} \cdot \mathbf{b} = \mathbf{A}^{\oplus} \cdot \mathbf{b} \quad (5.10)$$

It can be seen that the Moore-Penrose Inverse is equivalent to choose a scaling factor $\alpha = 0$.

5.2 The weighting vector as a holographic filter

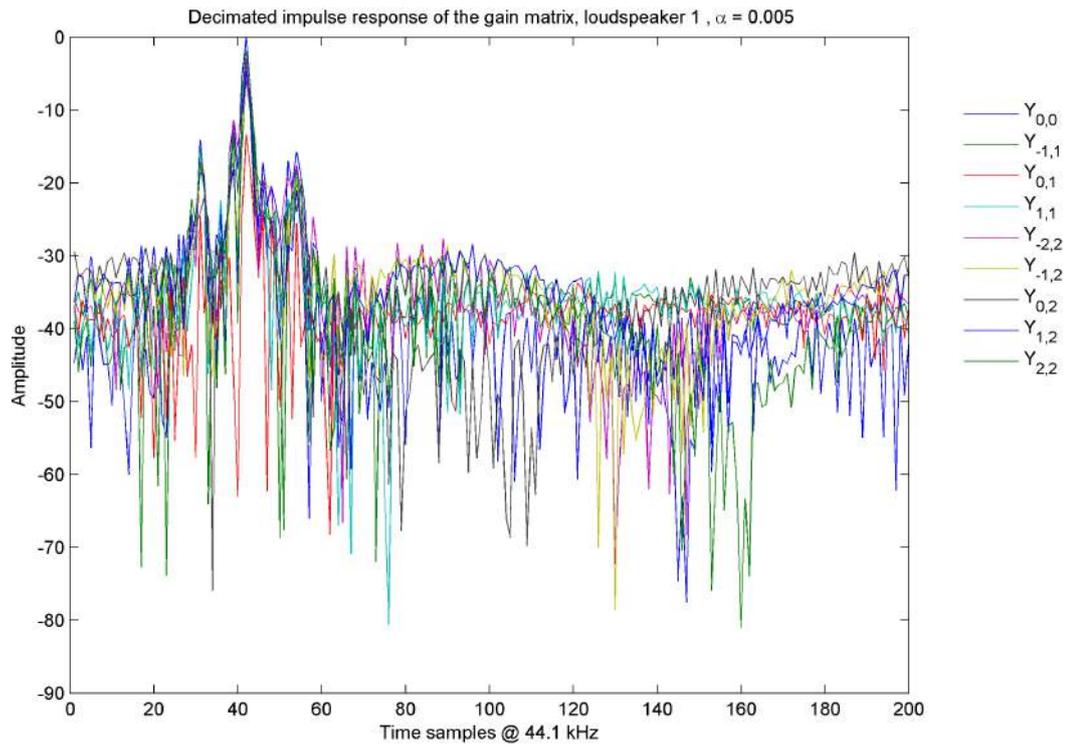
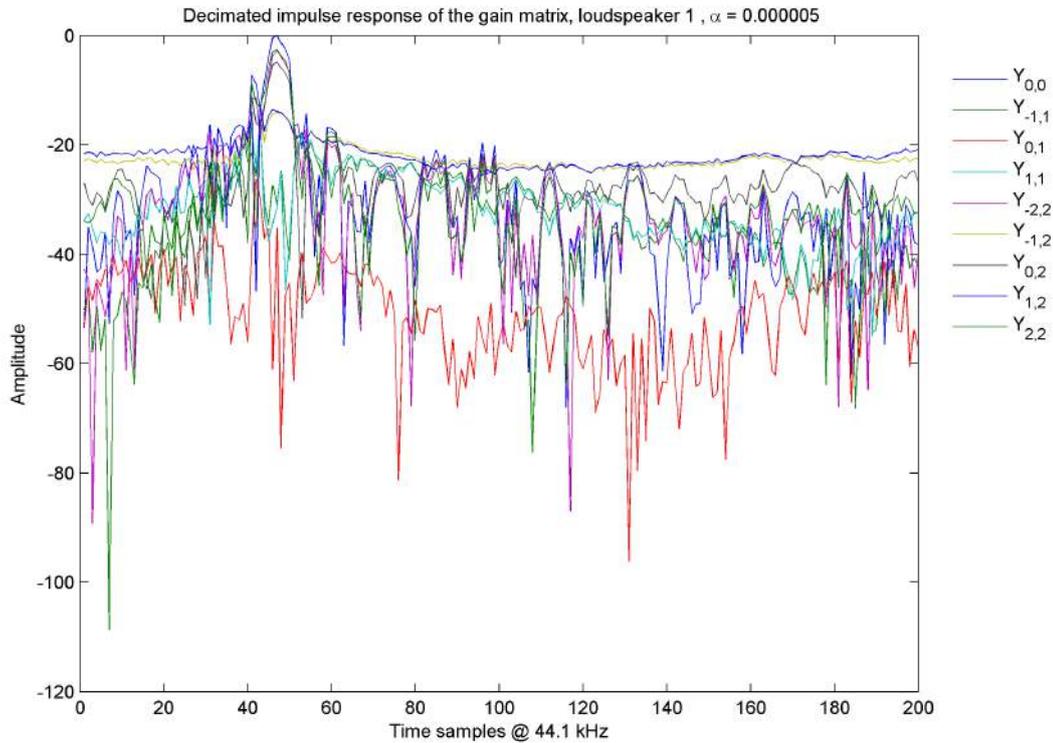
The measured radiation pattern of the physical transducers has been till now analysed in the frequency domain. Since the weighting vector \mathbf{g}_T represents the impulse response of an holographic filter, it is sensible to study the effect of the scaling factor α on the performance of the holographic filter. For that purpose, a Dirac impulse was defined as $\hat{\mathbf{d}}_T$. This implies, that every spherical harmonic function has to be perfectly controllable. The order was set to $N=3$. The calculation was performed in the frequency domain by defining:

$$\mathbf{I} = \hat{\mathbf{D}}_{\text{rot}} \cdot \mathbf{G} \quad (5.11)$$

$$\mathbf{G} = (\hat{\mathbf{D}}_{\text{rot}}^H \cdot \hat{\mathbf{D}}_{\text{rot}} + \alpha \cdot \mathbf{I})^{-1} \cdot \hat{\mathbf{D}}_{\text{rot}} \quad (5.12)$$

In order to overcome computational restrictions, a script was implemented to decimate the impulse responses by an arbitrary factor in the time domain. The decimation factor chosen to study the impulse response was 2, which means that only half of the time samples have been considered. Firstly, lets consider single transducers such as transducers 1, 11 and 22:

It is evident that the synthesis of an arbitrary three dimensional directivity can not be performed with single transducers. Nevertheless, figures 5.1 and 5.2 show clearly the trade off the matrix regularization poses: The regularization term has to be held as little as possible to keep the regularization error low, yet yielding a satisfactory impulse response. By increasing the α the

Figure 5.1: transducer 1, $\alpha = 0.005$ Figure 5.2: Transducer 1, $\alpha = 5 \cdot 10^{-6}$

impulse response may look better and better, but the resulting matrix diverts more and more from the original matrix. Equivalently to filters designed in the time domain, the impulse response of an holographic filter should be as close to an impulse as possible. The same analysis was performed using transducers 11 and 22 and from these observations it was concluded, that the impulse response envelope of the three transducer types behave similarly in terms of sensitivity towards the regularization parameter α . Consequently, to show the impulse response envelope of the whole MIMO system, it is sensible to compute the quadratic impulse responses in dB for different regularization factors. This way, it can be decided which value for α is optimal for any arbitrary amount of loudspeakers being described in the MIMO system. The impulse responses yielded by an arbitrary number of transducers have been evaluated using $\varphi_{\text{rot}} = 24^\circ$, $\varphi_{\text{rot}} = 18^\circ$ and $\varphi_{\text{rot}} = 15^\circ$. Based on the results presented in figures 5.2 to 5.1, a scaling factor $\alpha = 0.005$ has been chosen.

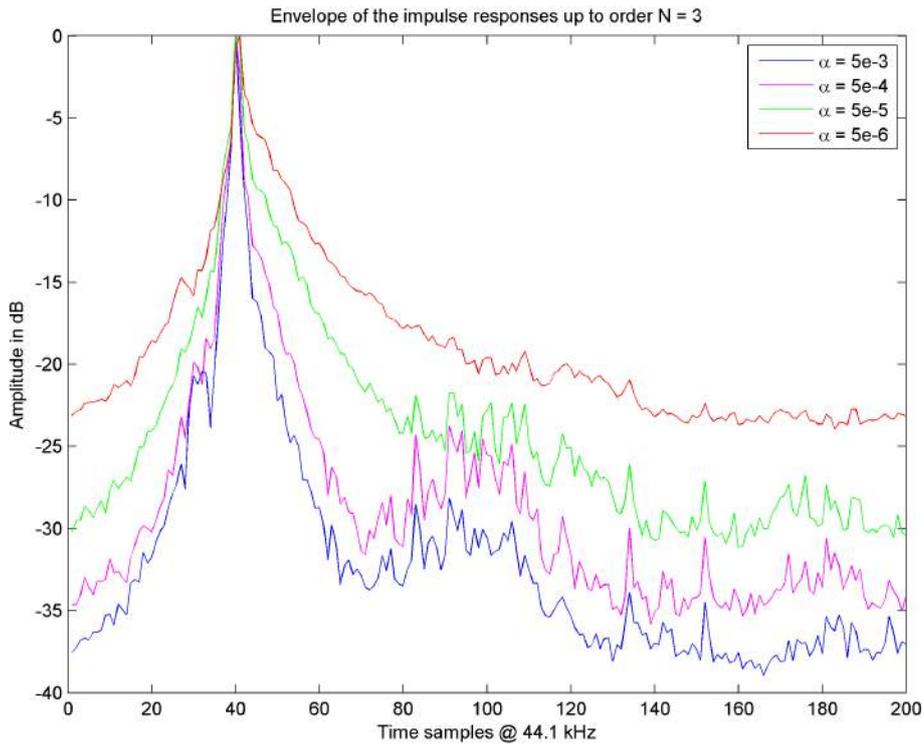


Figure 5.3: Impulse response envelope using 28 transducers

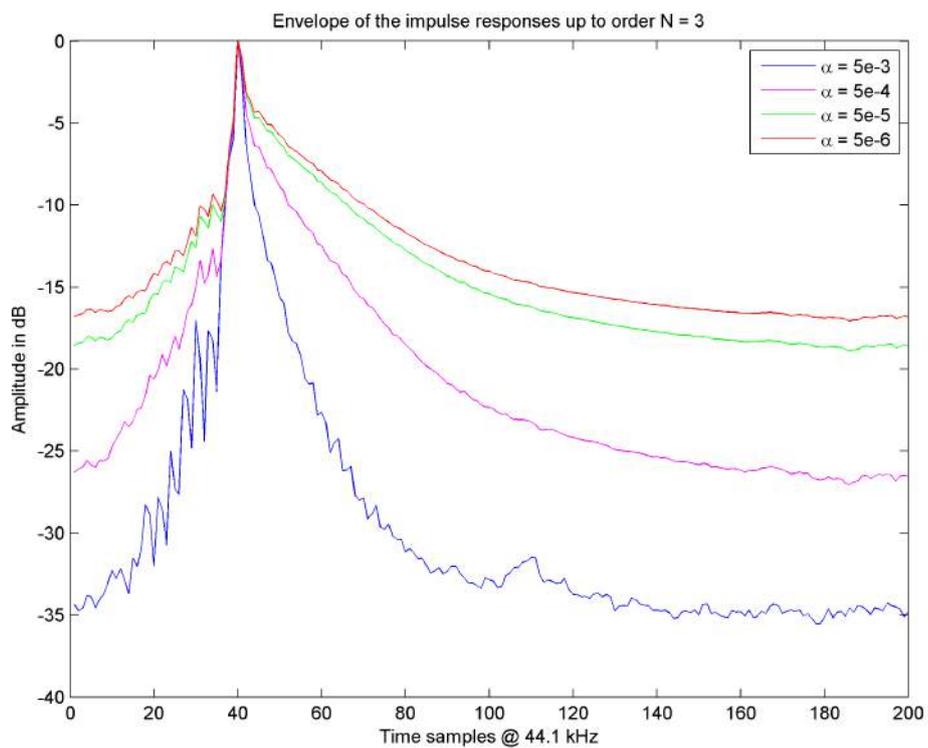


Figure 5.4: Impulse response envelope using 420 transducers

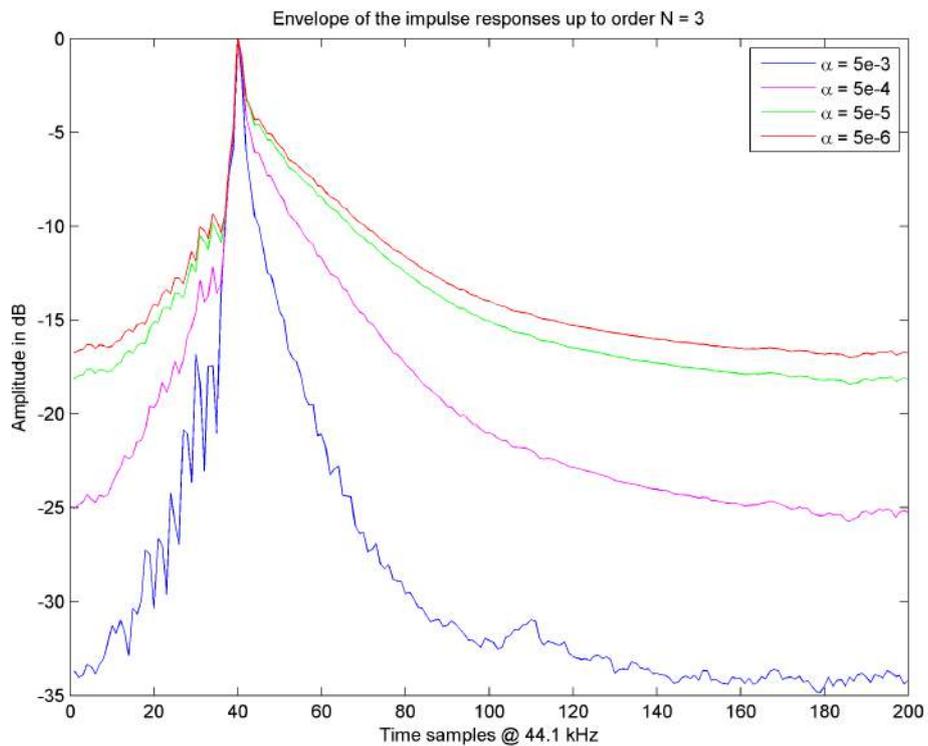


Figure 5.5: Impulse response envelope using 560 transducers

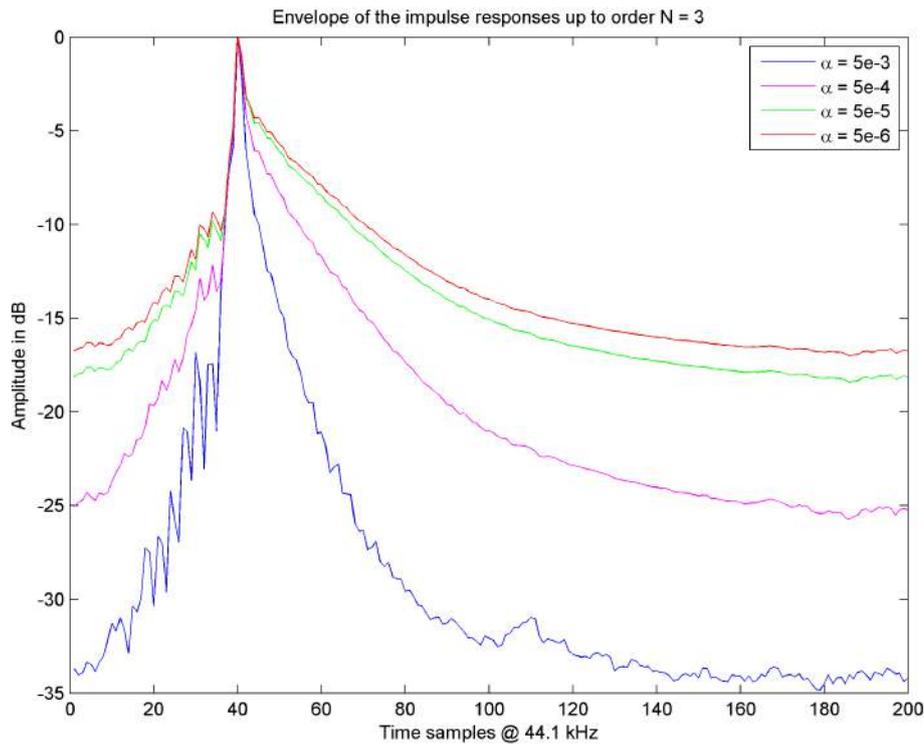


Figure 5.6: Impulse response envelope using 672 transducers

Greater values for α make the decay of the impulse responses more steep. It can be observed that the impulse responses for different regularization parameters do not vary largely if the number of virtual transducers is set to 420, 560 or 672. The reason for this fact lies within the fact, that low orders can be correctly synthesized using fewer loudspeakers. To obtain the minimal amount of loudspeakers to synthesize an arbitrary order for any given frequency is a topic of further research. Based on the results presented in figures 5.3 to 5.6, a scaling factor $\alpha = 0.005$ has been chosen for the further calculations. Since synthesis up to order $N = 11$ is theoretically achievable, the performance of the source has been studied regarding $N = 11$. To understand the results of this analysis, the concept of *spatial aliasing* needs to be clarified.

5.3 Spatial aliasing

In this section, spatial aliasing is shown based on the discrete inverse spherical harmonic transformation. Recalling the Parseval's theorem in eq. 2.6, the total amount of energy needs to be held within a finite order N . Analogue to the Parseval's theorem is to require that the function \mathbf{f} , when defined in spherical coordinates, does not hold angular structures which can not be represented by the superposition of spherical harmonic coefficients. If $\hat{\mathbf{f}}$ does hold such angular structures, *spatial aliasing* will be noticeable.

The following example illustrates the effect of spatial sampling on the discrete spherical harmonic transformation. Consider a Gaussian sampling of order $N = 10$. As explained in 2.4, a Gaussian sampling of order N enables generates $2 \cdot (N + 1)^2 = 242$ nodes on a spherical surface. Therefore, the spherical harmonic decomposition of a function f in $(N + 1)^2 = (10 + 1)^2 = 121$ spherical harmonic

coefficients can be correctly performed. As \mathbf{f} the \mathbf{Y} matrix as defined in eq. 2.9 was chosen. Hence, eq. 2.15 can be rewritten as:

$$\hat{\mathbf{f}} = \overline{\mathbf{Y}} \cdot \text{diag}\{\mathbf{w}\} \cdot \mathbf{Y} \quad (5.13)$$

In eq. 5.13, the matrix \mathbf{Y} can be perfectly reconstructed in the spherical harmonic domain. Since $\overline{\mathbf{Y}}$ is just the hermitian transpose of \mathbf{Y} , the absolute value of their coefficients remains the same. This fact can be seen by studying $\hat{\mathbf{f}}$:

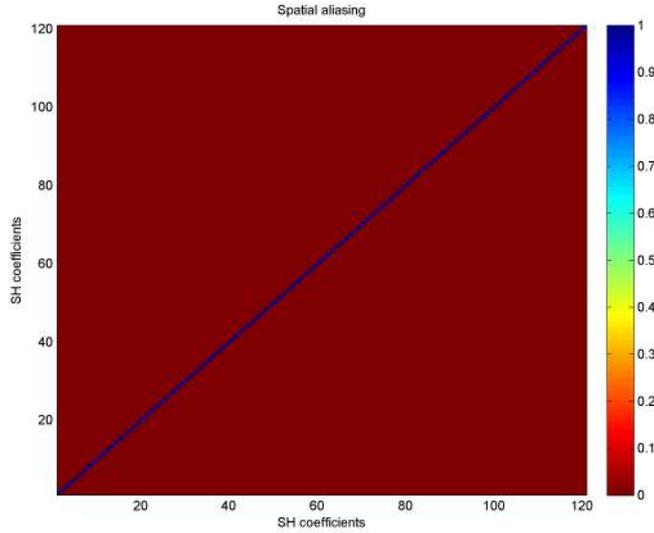


Figure 5.7: Aliasing

Now consider that the matrix \mathbf{Y} holds all spherical harmonic functions up to order 15 while the node number is kept equal:

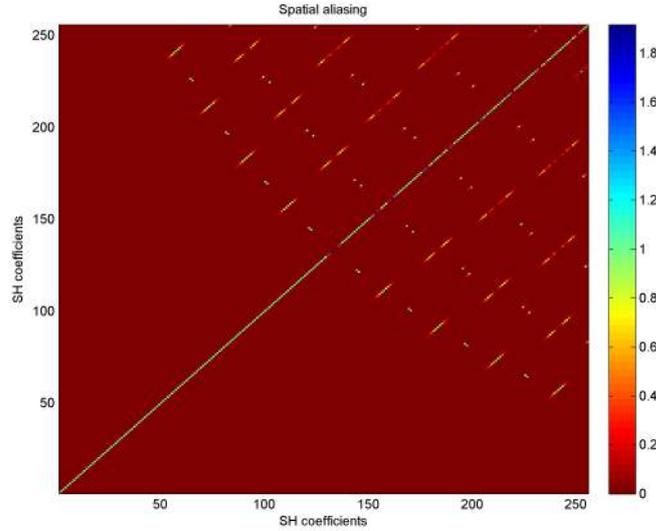


Figure 5.8: Aliasing

Since the number of nodes remains the same, by transforming \mathbf{Y} into the spherical harmonic domain, spherical harmonic functions of orders higher than 10 are mirrored into orders lower than $N = 10$. Since the energy within spherical harmonic orders needs to remain constant (see 2.2.2), the energy within the orders is spread throughout the mirrored and the non mirrored coefficients. This case has been illustrated in figure 5.8.

The *DSHT* is required to reproduce sound after having defined the radiation pattern so that the short example in section 5.3 shows the entire chain in order to define an arbitrary radiation pattern: The measured transducer directivities are expanded into spherical harmonic functions. After being conveniently combined, the resulting radiation pattern needs to be transferred into the spatial domain, facing the problem of spatial aliasing. In 3.1.3 and 3.5, the results of the directivity measurement were first discussed regarding spherical harmonic orders and afterwards by means of the spherical harmonic functions within the orders. In analogy to this procedure, the analysis of the spatial aliasing and its link to the regularization parameter α will be done in the same way. To illustrate the maximal order that can be synthesized correctly, a similar approach has been employed. Firstly, the number of spherical harmonics in matrix $\hat{\mathbf{D}}$ was limited to 961 ($n_{\text{synthesis}} = 30$) due to computational restrictions. In [12], a theoretical approach was presented to quantify the spatial aliasing error in which the transformation matrix \mathbf{Y} holds values of order $N = \infty$. Consequently, $\hat{\mathbf{D}}$ should, theoretically, hold spherical harmonic functions up to an infinite order N . Virtual loudspeakers were created by using $\varphi_{\text{rot}} = 15^\circ$, $\varphi_{\text{rot}} = 18^\circ$ and $\varphi_{\text{rot}} = 24^\circ$. Having created the matrix $\hat{\mathbf{D}}_{\text{rot}}$, the synthesis error is computed by defining a matrix \mathbf{E} [12]

$$\mathbf{IO} = \begin{pmatrix} \mathbf{I} \\ (n_{\text{synthesis}}+1)^2, (n_{\text{synthesis}}+1)^2 \\ \mathbf{O} \\ (N - (n_{\text{synthesis}}+1)^2)^2, (n_{\text{synthesis}}+1)^2 \end{pmatrix} \quad (5.14)$$

$$\mathbf{E} = \hat{\mathbf{D}}_{\text{rot},N} \oplus \hat{\mathbf{D}}_{\text{rot},Q} - \mathbf{IO} \quad (5.15)$$

An ideal error matrix should be a matrix which only holds zeros: The spherical harmonic functions within the orders up to $n_{\text{synthesis}}$ can be perfectly reconstructed and equal unity, so the subtraction

with matrix \mathbf{I} yields a zero matrix \mathbf{O} of dimensions $(n_{\text{synthesis}} + 1)^2, (n_{\text{synthesis}} + 1)^2$. Spherical harmonics belonging to orders higher than $n_{\text{synthesis}}$ are not mirrored into lower orders, which equals a zero when reconstructed. Although the regularization parameter $\alpha = 0.005$ is rather high comparing it to other works ($\alpha = 1 \cdot 10^{-6}$ in 18), it shows the best behaviour in the time domain. The synthesis error (= spatial aliasing) is the variable that shows in a direct way if the chosen scaling factor is suitable or not. Pictures 5.9 to 5.11 show the synthesis error for three different virtual loudspeaker arrays with 420, 560 and 672 virtual loudspeakers respectively. The synthesis error has been computed for every frequency using equation 5.15. Pictures 5.9 to 5.11 show the synthesis error, summed of for all spherical harmonic functions up to $N = 11$. The virtual array

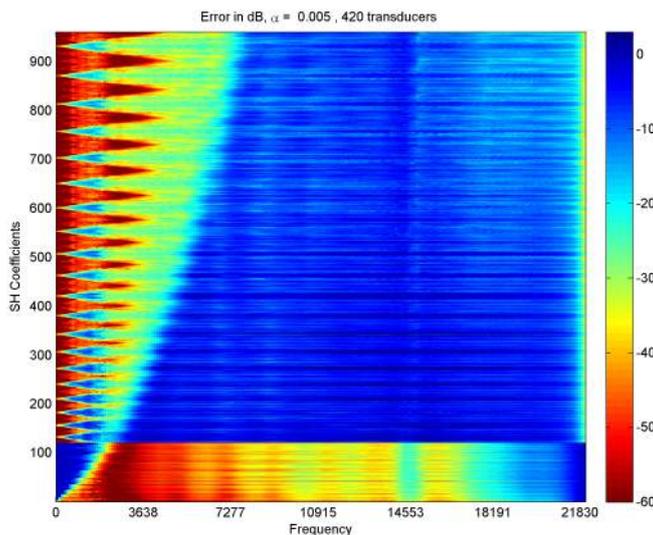


Figure 5.9: Synthesis error using 420 virtual loudspeakers

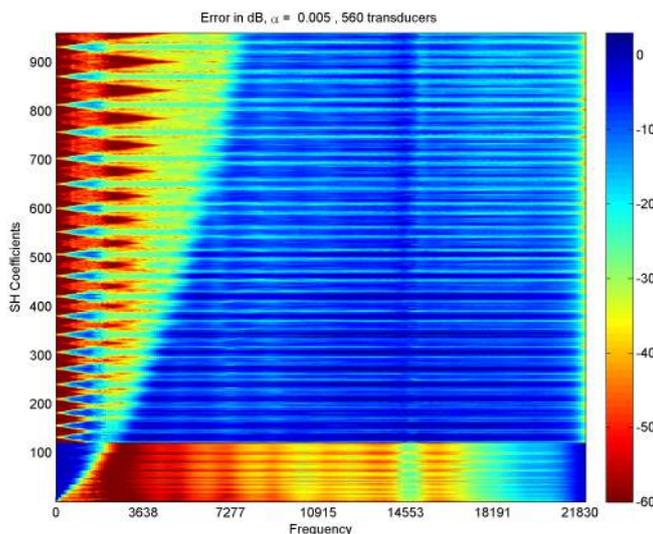


Figure 5.10: Synthesis error for 560 transducers

with 420 loudspeakers allows an almost error free synthesis over a wide frequency range while enlarging the number of virtual loudspeakers results in a wider frequency range in which the error

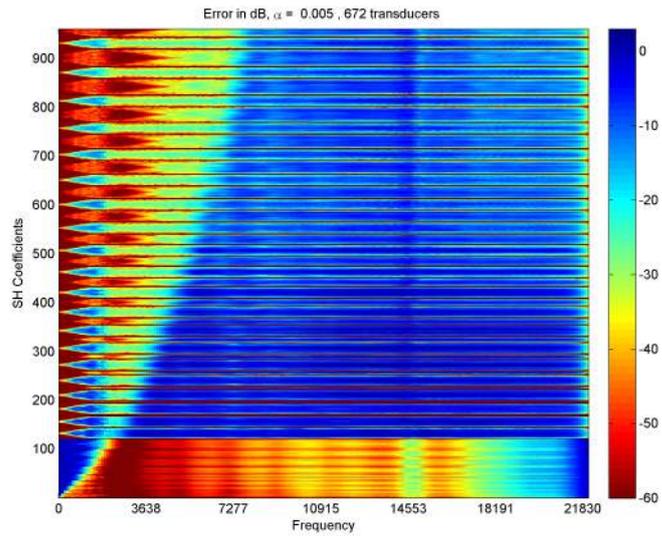


Figure 5.11: Synthesis error for 672 transducers

value under -30 dB. The error in the frequency range below 1 kHz is mainly dependent on the regularization error rather than on the number of transducers, which is evident since spherical harmonic functions held in low orders can be correctly radiated with a small number of transducers. Recalling the introduction in chapter 5, virtual transducer positions are crucial to represent complex spatial patterns, present in higher frequencies. Comparing a virtual array of 420 loudspeakers and $\alpha = 0.00005$ and $\alpha = 0.005$ respectively, this fact becomes clear:

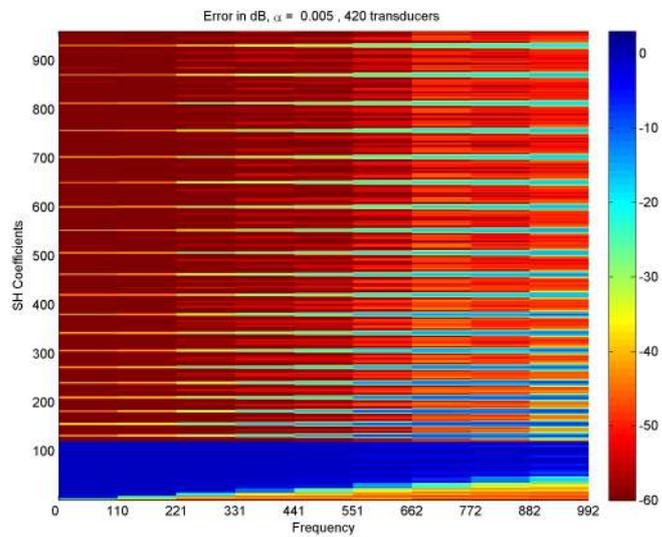


Figure 5.12: Synthesis error for 420 transducers, $\alpha = 0.005$ up to 1 kHz

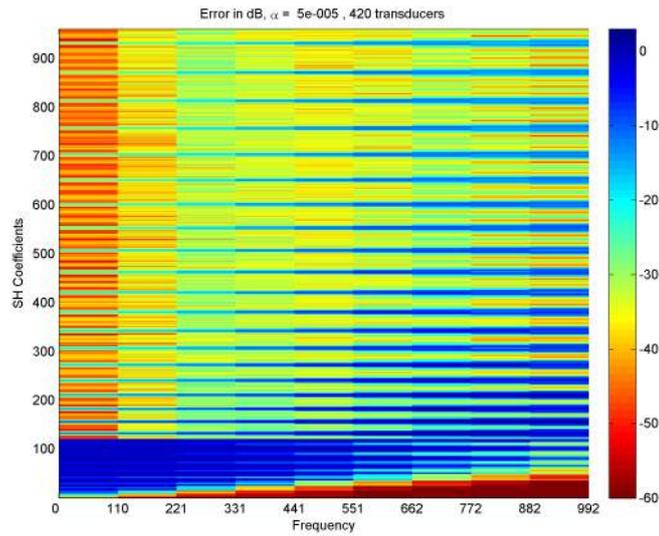


Figure 5.13: Synthesis error for 420 transducers, $\alpha = 0.00005$ up to 1 kHz

To illustrate the risk of enlarging or shrinking α without taking into account the synthesis error, figure shows the synthesis error for a virtual array of 420 loudspeakers and an $\alpha = 0.00005$ and $\alpha = 0.1$ respectively.

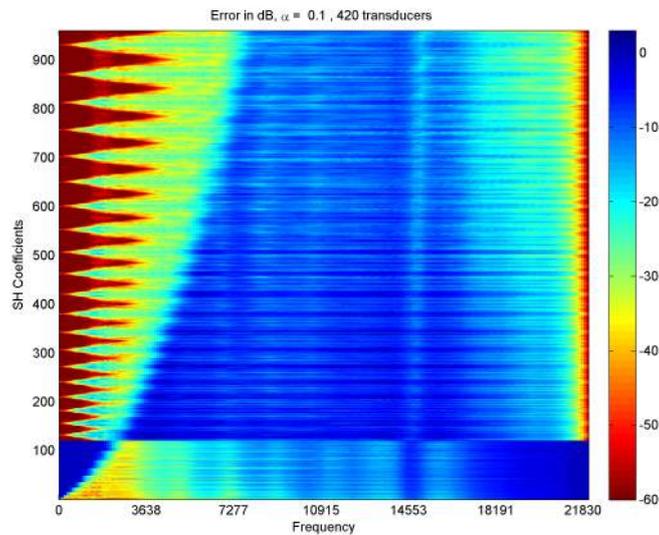


Figure 5.14: Synthesis error for 420 transducers, $\alpha = 0.1$

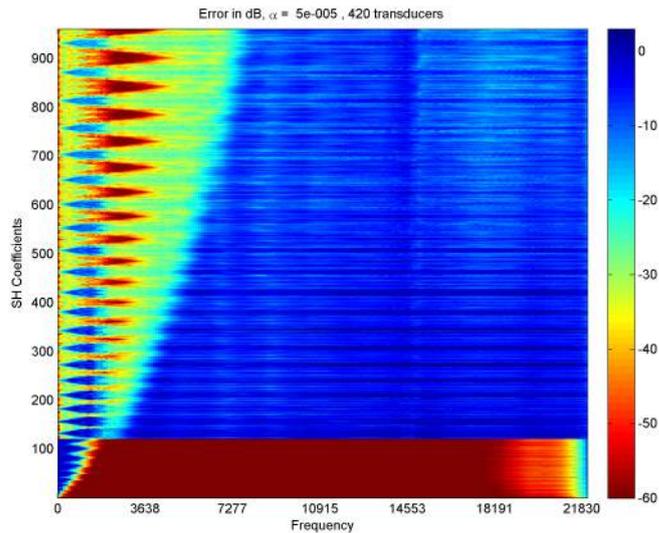


Figure 5.15: Synthesis error for 420 transducers, $\alpha = 0.00005$

These examples show the importance of taking into account the time domain and the frequency domain properties of holographic filters: The time domain shows the necessity to choose a scaling factor high enough to ensure a good impulse response and its practical implementation can be illustrated in the frequency domain.

5.4 Synthesis of a Dirac delta function

In [30] and [31], the possibility and feasibility of radiation control of dodecahedron loudspeakers was studied by trying to reproduce a Dirac delta- distribution depending on φ_T and ϑ_T :

$$\delta(\varphi - \varphi_T) \cdot \delta(\vartheta - \vartheta_T) = \begin{cases} \infty & \text{at the point}(\varphi_T, \vartheta_T) \\ 0 & \text{otherwise} \end{cases} \quad (5.16)$$

The expansion of a Dirac delta- distribution on a sphere can be easily calculated since

$$\delta_{nm} = Y_n^{m*}(\vartheta_T, \varphi_T)$$

The direction of the beam resulting from the expansion of the Dirac delta-distribution has been arbitrarily set to $(\vartheta_T = 90^\circ, \varphi_T = 245^\circ)$. As it has been pointed out in [30], the beam can be either expanded into a large number of spherical harmonic coefficients (*real beam*) or into a beam just holding spherical harmonic functions up to an order $N = \sqrt{L} - 1$ which can be correctly radiated by the array. Since the synthesis using this measurement source has to be possible for arbitrary sources, a real beam has been defined by expanding the Dirac delta- distribution into spherical harmonic functions up to order $N = 20$. A real beam is strictly speaking not real since every function on the sphere has to be expanded in an infinite set of spherical harmonic functions. In practice, most sources present an angular unbounded radiation pattern, so that expanding a beam using a high order is closer to a reality than expanding it with a lower order. The target directivity function is depicted in figure 5.16 The synthesis was performed using a regularization parameter $\alpha = 0.005$ and



Figure 5.16: Target directivity

672 virtual transducers, the maximal number of transducers which can be generated using the first rotation mode, as explained at the beginning of this chapter. The synthesis was conducted for the frequencies 220 Hz, 1 kHz, 2 kHz, 5 kHz, 8 kHz and 13 kHz.



Figure 5.17: 220 Hz



Figure 5.18: 1 kHz



Figure 5.19: 2 kHz



Figure 5.20: 5 kHz



Figure 5.21: 8 kHz



Figure 5.22: 13 kHz

As shown in figure 5.11, even the virtual array with 672 transducers can not match the target directivity by any means. For higher frequencies though, the shape of the synthetic beam gets closer and closer to the target directivity. Approximately around 13 kHz, the error due to spatial aliasing starts to enlarge, which can be seen in the growing side lobes, which for 13 kHz are not very severe, but, as it can be deduced out of figure 5.11, will keep on growing for higher frequencies. Nevertheless, the synthesis of the target directivity works remarkably well along a wide frequency range, which shows the benefits of using this measurement source for the synthesis of directivity pattern and their associated room impulse responses instead of a dodecahedron loudspeaker.

6 Outlook

This work is an analysis of the capabilities of the measurement source SLAYER regarding the synthesis of arbitrary radiation patterns. Since the measured transducer directivity functions have proved to be suitable, the creation of virtual loudspeaker arrays involving the tilted and the untilted measurements would be the next logical step. During my stay at RWTH Aachen, several room impulse response measurements with the measurement source SLAYER were performed in different locations on the institute. One of these measurements involved a *reciprocal* acoustic measurement. The impulse response of the acoustic path between a dodecahedron loudspeaker and a dummy head

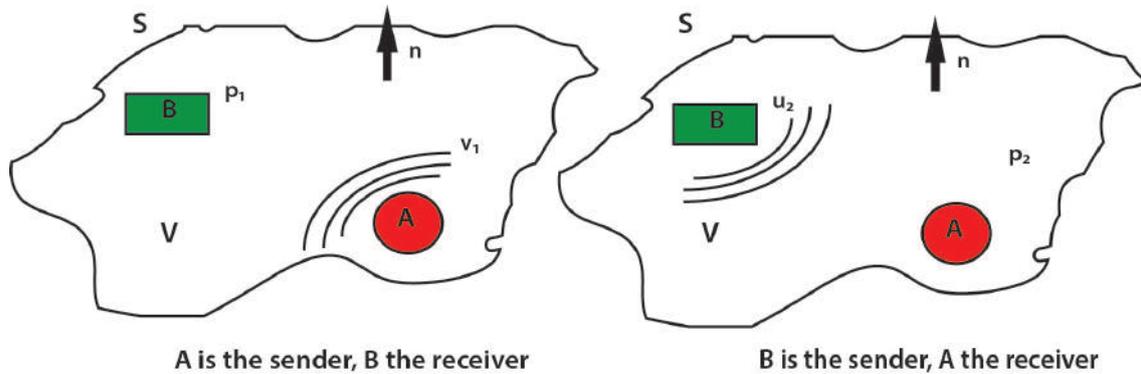


Figure 6.1: Reciprocity

was measured in one of the enclosures of the institute. Then, the measurement source SLAYER was set at the same place as the dodecahedron and the dummy head was replaced by a microphone with spherical recording pattern. The idea behind these two measurements is to recreate the recording pattern of the dummy head (right or left ear) using the measurement source SLAYER and analyse the differences in the room impulse response. Since the measured room impulse responses and the radiation pattern measurement signals have different lengths, it is necessary to apply a zero-padding to the transducer directivity measurement signals. Otherwise, the frequency bins will not be equal. Since the amount of room impulse responses is rather long, the experience gained by the author advises to decimate the room impulse responses as much as possible.

7 Anhang

7.1 Coordinate Systems

Throughout this work two right handed coordinate systems will be used: The Cartesian coordinate system and the Spherical coordinate system. The Cartesian coordinate system is characterized by its 3 axes x, y, z . x and y span the horizontal plane while the z axis allows the representation of vertical planes. The Cartesian coordinate system used throughout this work is depicted in figure 7.1:

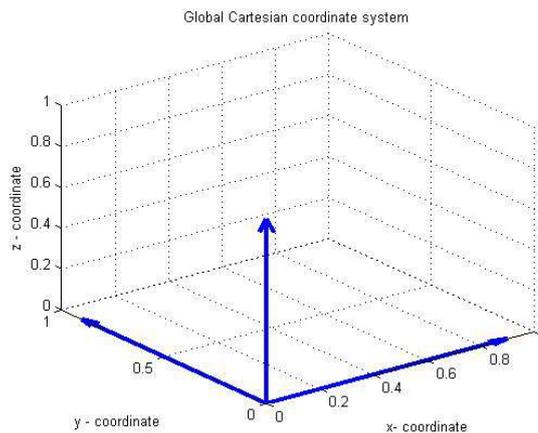


Figure 7.1: 3 dimensional Cartesian coordinate system

The spherical coordinate system also allows three dimensional representations and is characterized by r , which is the radial component, the azimuthal angle φ and the elevation angle ϑ . The relationship between both coordinate systems is given by [5]

$$x = r \sin \vartheta \cos \varphi \qquad r = \sqrt{(x^2 + y^2 + z^2)} \qquad (7.1)$$

$$y = r \sin \vartheta \sin \varphi \qquad \vartheta = \arccos \frac{z}{r} \qquad (7.2)$$

$$z = r \cos \vartheta \qquad \varphi = \arctan \frac{y}{x} \qquad (7.3)$$

The definition of the angles φ and ϑ can be derived from figure 7.2

7.2 Spherical harmonic expanded data

7.3 Radiation pattern control

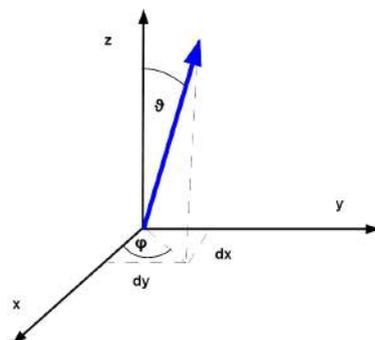


Figure 7.2: 3 dimensional spherical coordinate system

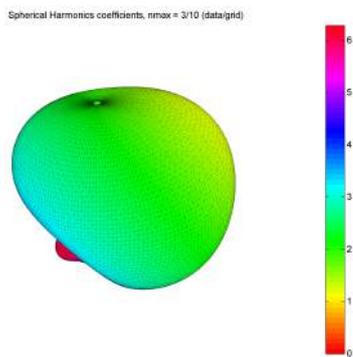


Figure 7.3: Transducer 1 for $f = 500$ Hz

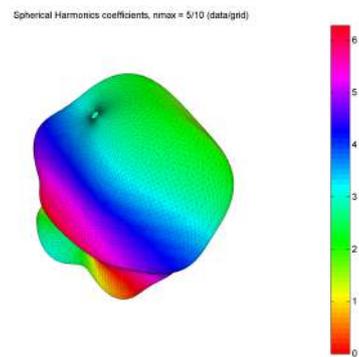


Figure 7.4: Transducer 1 for $f = 1$ kHz, SH

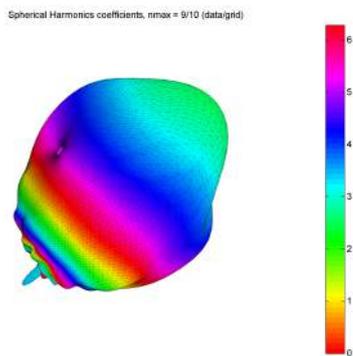


Figure 7.5: Transducer 11 for $f = 2$ kHz

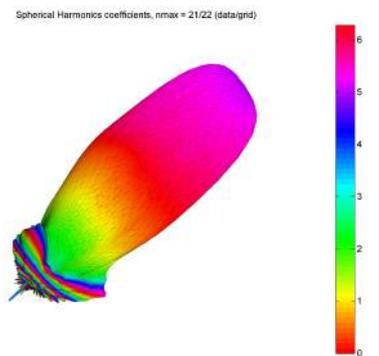
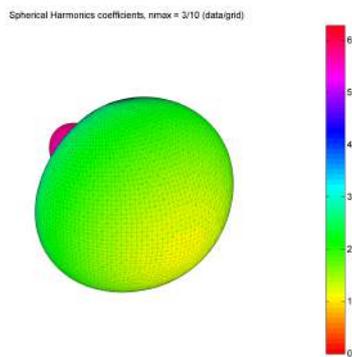
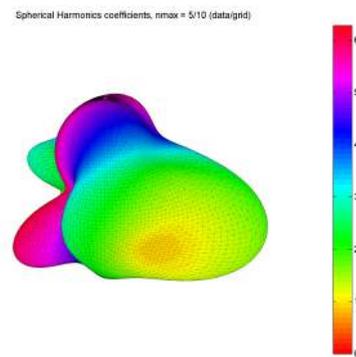
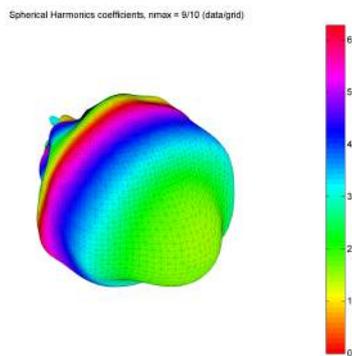
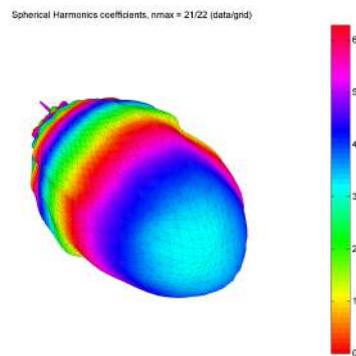


Figure 7.6: Transducer 11 for $f = 5500$ Hz, SH

Figure 7.7: Transducer 28 for $f = 500$ HzFigure 7.8: Transducer 28 for $f = 1$ kHz, SHFigure 7.9: Transducer 1 for $f = 2$ kHzFigure 7.10: Transducer 1 for $f = 5500$ Hz, SH

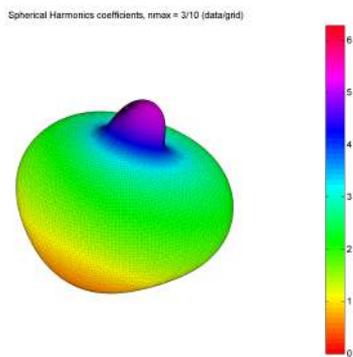


Figure 7.11: Transducer 28 for $f = 500$ Hz

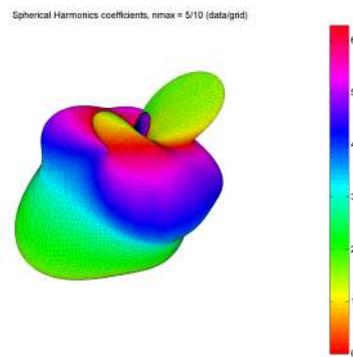


Figure 7.12: Transducer 28 for $f = 1$ kHz, SH

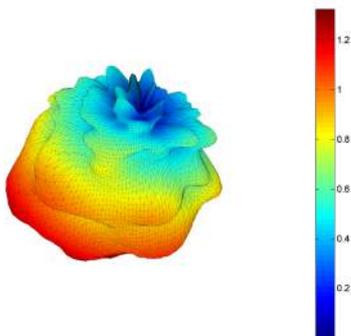


Figure 7.13: Transducer 28 for $f = 2$ kHz

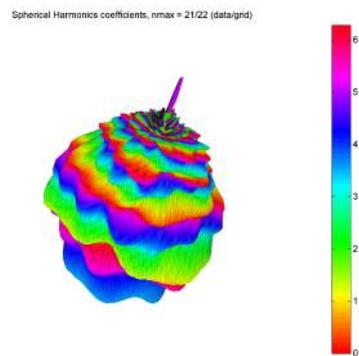


Figure 7.14: Transducer 28 for $f = 5500$ Hz, SH

Bibliography

- [1] Rafaely: *Analysis and Design of spherical microphone arrays*
Speech and Audio Processing, IEEE , Volume 13, Issue 1, p.135-143, 2005

- [2] Williams: *Fourier Acoustics and Nearfield Acoustical Holography*
Elsevier, 1999 Washington,

- [3] Pape: *Vergleich Robuster Mikrofonarrays*
Research report on the seminar *Seminar Aufnahmetechnik 3*, University of Music and Performing Arts Graz 2005

- [4] Plessas: *Rigid Sphere Microphone Arrays for Spatial Recording and Holography*
Diploma Thesis, University of Music and Performing Arts Graz 2009

- [5] Zotter: *Akustische Holophonie und Holographie*
Notes on the lecture *Akustische Holophonie und Holographie*, University of Music and Performing Arts Graz 2012

- [6] Noisternig: *Spherical Acoustics and Applications*, 2010 Graz

- [7] Klein, Pollow, Dietrich: *Optimized System for the Synthesis of Room Impulse Responses of Arbitrary Sound Sources*
38th German Annual Conference on Acoustics (DAGA) Meran 2012

- [8] Klein: *Optimization of a Method for the Synthesis of Transfer Functions of Variable Sound Source Directivities for Acoustical Measurements*
Diploma Thesis, RWTH Aachen 2012

- [9] Mueller, Massarani: *Transfer - Function Measurement with Sweeps*
Journal of the Audio Engineering Society, vol. 49, pp.44-471, 2001

- [10] Kunkemöller: *Synthesis of Room Impulse Responses for Variable Source Characteristics*
Diploma Thesis, RWTH Aachen 2011

- [11] Pollow : *Variable directivity of dodecahedron loudspeakers*
Diploma Thesis, RWTH Aachen 2007

- [12] Zotter : *Analysis and Synthesis of Sound Radiation with Spherical Arrays*
PhD thesis, University of Music and Performing Arts Graz 2009

- [13] Bronstein, Semendjajew, Musiol, Mühlig: *Taschenbuch der Mathematik*
Springer, 2012.

- [14] Rafaely, Weiss, Bachmat: *Spatial Aliasing in Spherical Microphone Arrays* Signal Processing,
IEEE Transactions on 55.3, p.1003-1010, 2007

- [15] Weseleak, Graber: *Elektroakustik*
Notes on the lecture *Elektroakustik*, Technical University Graz 2009

- [16] Rafaely: *Analysis and Design of Spherical Microphone Arrays*
Speech and Audio Processing, IEEE Transactions on 2005, 13. Jg., Nr. 1, S. 135-143

- [17] Oppenheim, Schaffer: *Zeitdiskrete Signalverarbeitung*
3.Edition, Oldenbourg, 1999

- [18] Müller-Trapet, Pollow, Vorländer: *Spherical Harmonics as a Basis for Quantifying Scattering and Diffusing Objects*
submitted for Forum Acusticum, Aalborg 2011

- [19] Weselak, Graber: *Raumakustik*
Notes on the lecture *Raumakustik*, Technical University Graz Version 5.2 Graz 2008

- [20] Farina: *Simultaneous measurement of impulse response and distortion with a swept-sine technique*
Audio Engineering Society Convention 108. Audio Engineering Society, Paris 2000.

- [21] Majdak, Balazs, Laback: *Multiple Exponential Sweep Method for Fast Measurement of Head - Related Transfer Functions*
J.Audio Eng. Soc. Am. vol. 113, no. 5, pp. 623-637 2007

- [22] Dietrich, Masiero: *On the Optimization of the Multiple Exponential Sweep Method*
J.Audio Eng. Soc. Am. vol. 61, no. 3, 2013

- [23] Kinsler, Frey, Coppens, Sanders: *Fundamentals of Acoustics*
Third Edition 1982

- [24] Mechel: *Formulas of Acoustics*
Springer, Second Edition 2008

- [25] ISO 3382: *Acoustics - Measurement of room acoustic parameters - Part 1: Performance*

spaces

ISO TC 43 / SC 2 , 2009

- [26] Weinzierl: *Handbuch der Audiotechnik*
Springer 2008

- [27] Zotter, Bank: *Geometric Error Estimation and Compensation in Compact Spherical Loudspeaker Array Calibration*
I2MTC, IEEE, Graz, April, 2012

- [28] Strang: *Lineare Algebra*
Springer 2003

- [29] Novak: *Identification of non linear models in acoustics*
PhD Thesis, Université du Maine, Le Mans, 2009.

- [30] Kerscher: *Compact Spherical Loudspeaker Array for Variable Sound-Radiation*
Diploma Thesis, University of Music and Performing Arts Graz 2010

- [31] Pollow, Klein, Dietrich, Behler Vorländer: *Optimized Spherical Sound Source for Room Reflection Analysis* International Workshop on Acoustic Signal Enhancement 2012, 4-6 September 2012, Aachen

- [32] Duraiswami, Zotkin, Gumerov: *Interpolation and Range Extrapolation of HRTFs*
Applications of Signal Processing to Audio and Acoustics, 2009. IEEE Workshop, New York