
DIPLOMA THESIS

A Rigid Double Cone Microphone Array Prototype

Revised Version

Florian Pausch

Graz, June 2013

University of Music and Performing Arts Graz
Institute of Electronic Music and Acoustics

Graz University of Technology

Advisors:

DI Hannes Pomberger
DI Dr. rer. nat. Franz Zotter

Assessor:

o.Univ.-Prof. Mag. art. DI Dr. techn. Robert Höldrich



institute of electronic music and acoustics



Abstract

Alternatively to AB, XY or equivalence-stereophony-based surround recordings, spherical arrays have been investigated thoroughly in recent years as they permit recordings with high-resolution imaging of spatial sound. Nevertheless, this type of surround sound imaging is still a challenging research area. The particular challenge is that we have to trade hardware effort off against signal-to-interference ratio or resolution of the three-dimensional sound field capture.

The prototype presented in this thesis tries to escape this narrow trade-off by a variation of the spherical geometry. All involved microphones are distributed on a spherical segment that is confined by two rigid cones to exclude a defined, unneeded range of spatial directions. Compared to a full spherical array, this permits an increased density of microphones for the desired panoramic recording angle without increased electronic hardware effort.

The rigid cones are considered when solving the wave equation in terms of suitable angular solutions which are represented by modified spherical basis functions, the so-called *spherical segment harmonics* (SSH). Practical decomposition into SSH using a microphone array requires an array layout that provides a minimal condition number for the decomposition. An optimal layout is obtained by an algorithm that selects optimal positions from a dense grid of pre-defined nodes.

The double cone microphone array prototype is evaluated by employing its acoustic measurement data in the analytic beamformer model.

Zusammenfassung

Als Gegenstück zu paarweiser Intensitäts-, Laufzeit- oder Äquivalenzstereophonie für Surroundaufnahmen wurden unlängst kugelförmige Mikrofonanordnungen eingehend beforscht, weil sie räumlichen Klang hochauflösend abbilden. Dennoch ist diese Art der räumlichen Klangabbildung immer noch ein herausforderndes Forschungsfeld. Insbesondere sind dreidimensionale Schallfeldaufnahmen stets ein Kompromiss von Ressourcenaufwand einerseits und Signal-Stör-Verhältnis bzw. Richtungsauflösung andererseits.

Der in dieser Diplomarbeit vorgestellte Prototyp versucht, durch Modifikation der kugelförmigen Bauweise, diese Kompromissfindung zu durchbrechen. Dabei nichtverwendete Richtungsgebiete werden ausgeblendet, indem alle Mikrofone auf einem Kugelsegment sitzen, das oben und unten durch die Flächen eines schallharten Doppelkegels begrenzt ist. Im Vergleich zu einem Kugelmikrofonarray erlaubt dieser Ansatz eine dichtere Mikrofonverteilung und führt ohne steigenden Hardwareaufwand zu einer besseren Auflösung im interessanten Aufnahmebereich.

Bei der Lösung der Wellengleichung werden die beiden Kegel unter geeigneten Randbedingungen miteinbezogen und durch modifizierte Kugelflächenfunktionen, die sogenannten *Kugelsegmentharmonischen*, repräsentiert. Ein numerischer Algorithmus, der ein gegebenes Raster an Punkten verwendet, bestimmt die Mikrofonpositionen. Die gefundene Anordnung ermöglicht eine wenig störanfällige Zerlegung der Schalldrücke an den Mikrofonen in Kugelsegmentharmonische, d.h. mit minimaler Konditionszahl.

Abschließende akustische Messungen evaluieren das aufgebaute Doppelkegelmikrofonarray und bestätigen die theoretischen Annahmen.

Acknowledgements

First of all, I would like to thank my advisors Hannes Pomberger and Franz Zotter for their guidance and help throughout the whole thesis. Many discussions and sessions helped me to improve my knowledge and to meet the targets.

Thank you, Thomas Musil, for the dedication to measure the array hardware. My compliments to Hermann Deutscher who did excellent work when manufacturing the cabinet of the microphone array. Additional thanks go to Alois Sontacchi for his support in extracurricular activities and Thomas Kößler for his prior work which saved me a lot of time.

My friends have to be mentioned separately. I really had a great time with you in Graz. A big thank you goes to Thomas Kubelka for the commitment and enthusiasm during the photo session.

I would like to express my gratitude to my family for their love. This work is dedicated to my parents who enabled my academic education and always supported me in so many ways.

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Graz,
.....
(Signature)

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am
.....
(Unterschrift)

Contents

Abstract	iii
Zusammenfassung	v
Acknowledgements	vii
Statutory Declaration / Eidesstattliche Erklärung	ix
1 Introduction	1
1.1 Motivation	2
1.2 Potential Application Areas	2
1.3 Thesis Outline	5
2 Theoretical Background	7
2.1 The Wave Equation	7
2.2 Solutions of the Wave Equation	9
2.2.1 Angular Solutions	9
2.2.2 Radial Solutions	13
2.3 Spherical Segment Basis Solutions	16
2.3.1 A Suitable Order Truncation	16
3 Spherical Segment Harmonics	19
3.1 Spherical Domain Transforms	20
3.1.1 The Continuous SSH Transform (SSHT)	20
3.1.2 Angular Band Limitation	21

4	Sampling the Spherical Segment	23
4.1	Vectorization of the SSH	23
4.1.1	Sound Field Decomposition	24
4.1.2	Spatial Aliasing	25
4.2	The Condition Number of the SSH Matrix	26
4.3	Sampling Strategies	27
4.3.1	Optimal Microphone Layout on a Fixed Raster	27
4.4	Open and Rigid Array Configurations	31
4.4.1	Open Array Configuration	31
4.4.2	Rigid Array Configuration	32
4.5	Holographic Filtering	33
4.5.1	Causal Holographic Filters	33
4.5.2	Digital Filter Design	35
4.6	Spherical Segment Signal Processing	36
5	The Prototype	37
5.1	Designing the Cabinet	37
5.1.1	Geometrical Properties	37
5.1.2	Functional Aspects	38
5.2	The Array Hardware	39
5.2.1	VDAM™ Units in the Model	40
5.3	Manufactured Cabinet	44
5.3.1	Assembling the Cabinet	44
6	Acoustic Measurements	47
6.1	Calibration of the Array Microphones	47
6.2	Preliminary Measurements	48
6.2.1	Measurement Setup A	48
6.2.2	Obtaining the Impulse Responses	50
6.2.3	Measured Impulse Responses	51
6.2.4	Level Deviations Among Gain Stages	51

6.2.5	Conclusions	53
6.3	Measuring the Prototype Characteristics	54
6.3.1	Measurement Setup B	54
6.3.2	Overall System Description	56
6.3.3	Equalization of the Measurement Loudspeaker Array	57
6.3.4	Spatial Impulse Responses of the Array Prototype	58
6.4	Array Microphone Directivity Patterns	61
7	Modal Beamforming in the Spherical Segment Domain	63
7.1	System Description	63
7.1.1	Analytic Beams	64
7.1.2	Beams Based on Measurement Data	65
8	Conclusion and Outlook	69
A	Quantitative Results	71
A.1	Related Values to the Double Cone Boundary Condition	71
A.2	Microphone Coordinates	74
B	Results of the Acoustic Measurements	77
B.1	Preliminary Measurements	77
C	User Instructions	85
C.1	Driver Installation	85
C.2	Recording Audio	85
C.3	Manipulations	86
C.3.1	Maintenance	86
C.3.2	Unmounting	87
	References	89

Chapter 1

Introduction

Today, our biggest expectations in innovating surround recordings by making them more plactical and well-resolved lie in spherical microphone arrays such as the 'Eigenmike' [EM03]. Arrays of this type also have a variety of other applications than Ambisonic surround recording such as *acoustic scene analysis*, i.e. information retrieval of a surrounding sound field (sound pressure/velocity/intensity) and subsequent extraction of source parameters (spherical source strength distribution, direction of arrival).

In general, spherical array signal processing techniques exploit the spatial diversity of a distribution of microphones on a rigid spherical surface. The required filters and mathematical procedures are well-defined and they emphasize the signal differences between the microphones to obtain a high-definition image of uniform spatial resolution for all directions. In order to avoid boosting noise in the recordings, the resolution at low frequencies has to be limited in some ways, e.g. [BPF11, MDB06], as usual for any super-directional microphone.



Figure 1.1: Comparison of the ornated ceiling of the *Sala de Dos Hermanas* in the famous *Alhambra* of Granada and the unpopulated spherical segment of the microphone array prototype.

This thesis presents a new prototype for three-dimensional sound field capture with a variation of the spherical geometry. More precisely, the array microphones are distributed on a spherical segment which is confined by two rigid cones. The two cones change the theory for signal processing as the underlying equation has to be solved under new boundary con-

ditions. In particular, angular and radial solutions change, and, instead of the spherical harmonics, the so-called *spherical segment harmonics* (SSH) are employed for decomposing a surrounding sound field.

1.1 Motivation

Spherical arrays without restriction of the angular pick-up range identify sounds from all directions. Consequently, in order to leave away sound from unwanted directions requires the capability to identify what is being removed. On the other hand, if one designs arrays based on assumptions such as: there cannot be any sound arriving from the ceiling, one might simplify the array hardware, but get a lot of trouble and a significant error if the assumption does not hold in a realistic room, cf. [ME08, Pom13].

The same procedure seems to be necessary to obtain non-uniform resolution as convex arrays tend to have spherically uniform resolution at low frequencies [Kai12]. To obtain a non-uniform resolution, first of all, sound has to be identified with uniform resolution before the resolution can be made non-uniform.

The new design considers the additional introduction of theoretically infinite long rigid cones as diffraction objects as investigated in [Pom13, Köß11] as well as in other areas of research [HC97]. The main advantage of this step is the *physical* exclusion of unwanted directions instead of a removal through identification. Moreover, this is expected to improve the noise behavior and the resolution on the limited angular range.

1.2 Potential Application Areas

Similar to other spherical microphone arrays, the realized prototype has a wide range of application areas. The following illustrations exemplify a few of them.

Higher-order Ambisonic recording / playback. Auditory scenes can be captured with the assembled 64-channel microphone hardware, encoded into Ambisonics format and played-back after decoding on a three-dimensional loudspeaker arrangement.

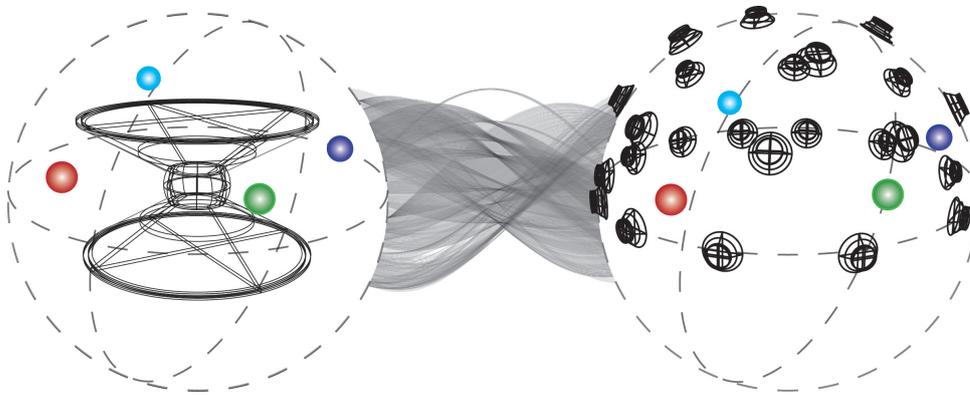


Figure 1.2: Sound field capture for a playback on a three-dimensional loudspeaker arrangement (Ambisonic).

Conference beamforming system. Excellent beam characteristics are achieved within the desired recording angle which leads to an accurate source separation in combination with a remarkable flexibility regarding the possible steering directions.

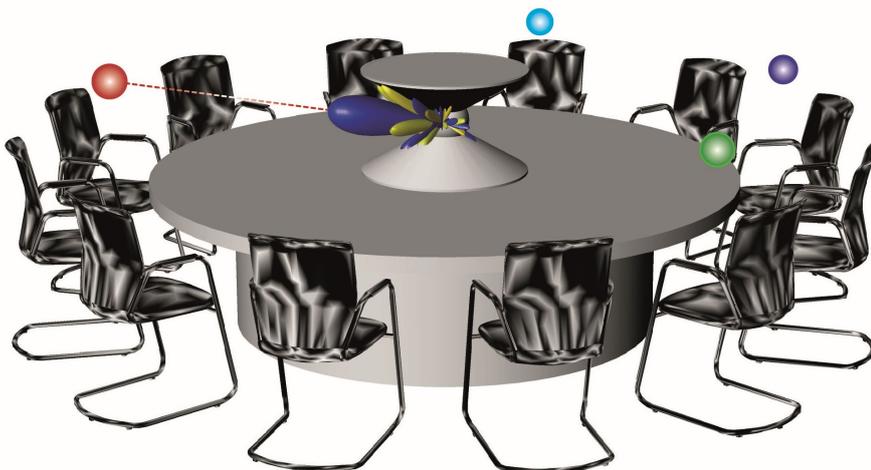


Figure 1.3: Teleconference system with controllable beams in form and direction.

Directional analysis of room impulse responses. With the help of modal beamforming, the prototype upgrades the information content of measurements in room acoustics by analyzing spatial directions in room impulse responses.

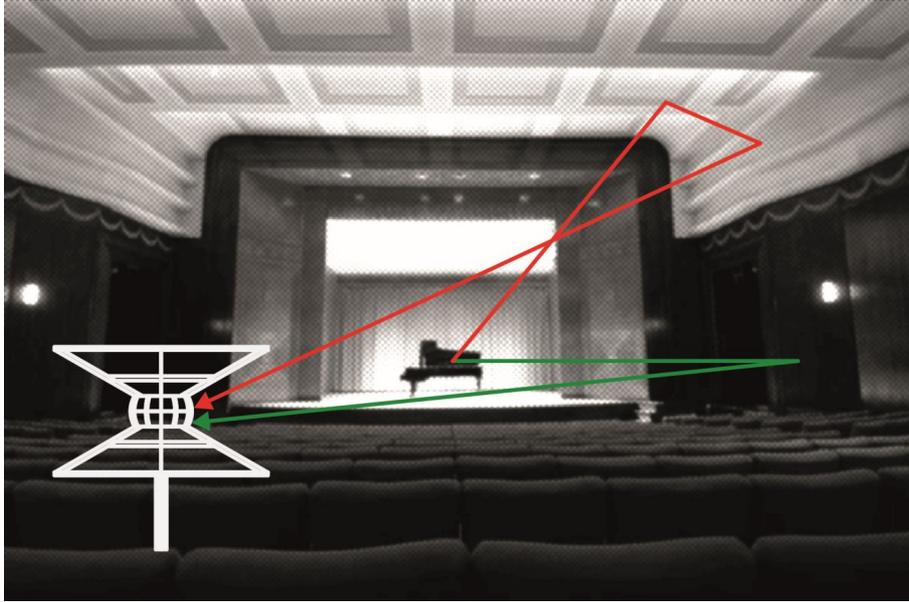


Figure 1.4: Upgrade of the information content when measuring room acoustics.

Sound mark recordings. Immersive environments with natural sounds, industry sounds and sounds produced by humans can be recorded and reproduced in a different location to convey local "acoustic fingerprints".

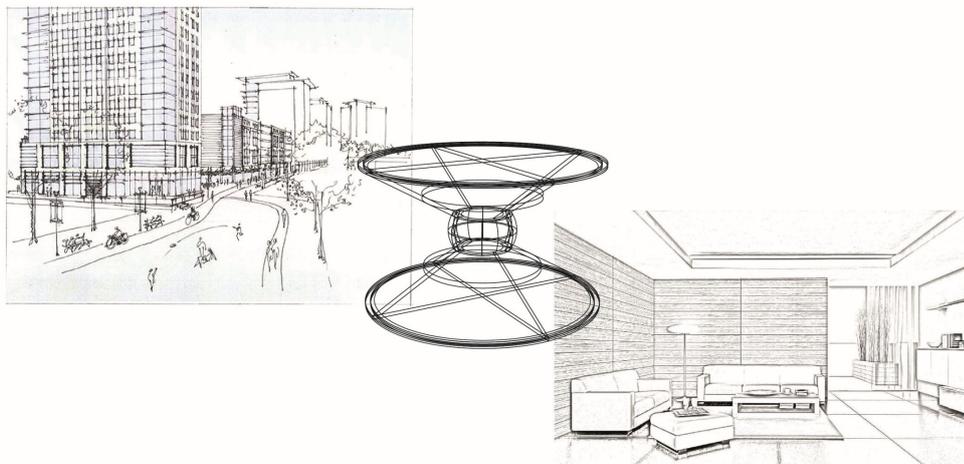


Figure 1.5: Reproduction of natural environments to convey local acoustic atmospheres.

1.3 Thesis Outline

This document is divided into the following chapters:

Chapter 2 explains the fundamentals of the underlying signal processing theory. The modified angular and radial solutions due to the introduction of rigid cones are presented in detail.

Chapter 3 describes the spherical segment basis functions that are necessary to decompose a surrounding soundfield into its modal components through a spatial Fourier transform.

Chapter 4 deals with strategies to sample the spherical segment and presents an algorithm to determine an optimal microphone layout. Different array scenarios are simulated to show the influence on the modal responses. The resulting analytic holographic filters for radial extrapolation of the captured soundfield are realized as finite impulse response filters.

In **chapter 5**, the design process of the prototype, its manufacturing and construction details are presented.

Chapter 6 gives information about the directional microphone characteristics after measuring the prototype in the acoustic domain. The damping effects of the rigid double cone are analyzed through transfer functions.

One of the main purposes of the prototype is its application as a beamforming system. **Chapter 7** provides the theory to implement such a system in the modal domain. For this purpose, not only microphone signals generated by an analytic model but also measurement data is employed in the holographic model.

Appendix A presents quantitative results in connection with the double cone boundary condition. Moreover, the exact coordinates of the optimized microphone distribution on the spherical segment are given.

In **appendix B**, the entire results of the preliminary acoustic measurements are shown by means of transfer functions.

Appendix C provides detailed user instructions to use the array hardware properly.

Chapter 2

Theoretical Background

The following sections explain the mathematical fundamentals underlying the acoustical array application. It is necessary to make the descriptions rather detailed because the solutions under conical boundary conditions are not as common as the spherical solutions without angular boundary conditions.

2.1 The Wave Equation

The Helmholtz equation describes a sound pressure variation $p(\mathbf{r}, \omega)$ in a homogeneous fluid with no viscosity [Wil99]:

$$(\Delta + k^2)p(\mathbf{r}, \omega) = 0 \quad (2.1)$$

with the Laplacian Δ and the wave number $k = \omega/c$ that contains $\omega = 2\pi f$, the angular frequency.

The vector $\mathbf{r} \in \mathbb{R}^3$ is expressed by

$$\mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = r \begin{pmatrix} \sin(\vartheta) \cos(\varphi) \\ \sin(\vartheta) \sin(\varphi) \\ \cos(\vartheta) \end{pmatrix} = r\boldsymbol{\theta},$$

where r is the radius, $\varphi \in [0, 2\pi)$ is the azimuth angle and $\vartheta \in [0, \pi]$ is the zenith angle. The inverse relation between spherical and Cartesian coordinates is given by

$$\begin{pmatrix} r \\ \varphi \\ \vartheta \end{pmatrix} = \begin{pmatrix} \sqrt{x^2 + y^2 + z^2} \\ \arctan\left(\frac{y}{x}\right) \\ \arctan\left(\frac{z}{r}\right) \end{pmatrix}.$$

The coordinate system is shown in fig. 2.1.

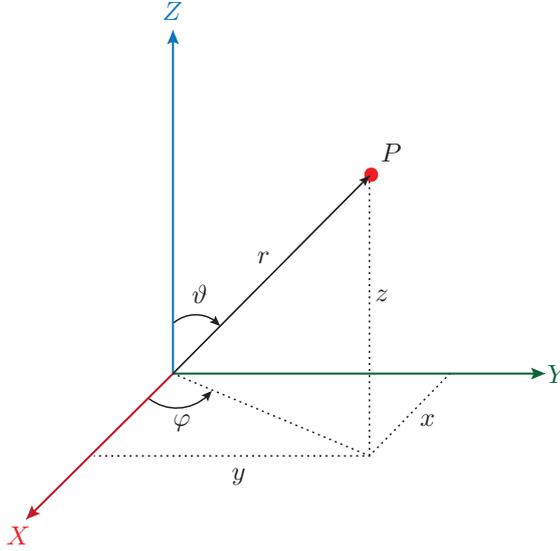


Figure 2.1: Definition of a point P relative to Cartesian (x, y, z) or spherical coordinates (r, φ, ϑ) .

The Laplacian in spherical coordinates for a function f is written as [Zot09a]

$$\begin{aligned} \Delta f &= \Delta_r f + \Delta_\vartheta f + \Delta_\varphi f = \\ &= \frac{1}{r^2} \frac{\partial^2 r^2 f}{\partial r^2} + \frac{1}{r^2 \sin^2(\vartheta)} \frac{\partial^2 f}{\partial \varphi^2} + \frac{1}{r^2 \sin \vartheta} \frac{\partial}{\partial \vartheta} \left(\sin(\vartheta) \frac{\partial f}{\partial \vartheta} \right). \end{aligned} \quad (2.2)$$

A detailed derivation can be found in [Dav07].

Inserting eq. (2.2) in eq. (2.1) yields the Helmholtz equation in spherical coordinates, a homogeneous second order partial differential equation:

$$\frac{1}{r^2} \frac{\partial^2 r^2 p(\mathbf{r}, \omega)}{\partial r^2} + \frac{1}{r^2 \sin^2(\vartheta)} \frac{\partial^2 p(\mathbf{r}, \omega)}{\partial \varphi^2} + \frac{1}{r^2 \sin \vartheta} \frac{\partial}{\partial \vartheta} \left(\sin(\vartheta) \frac{\partial p(\mathbf{r}, \omega)}{\partial \vartheta} \right) + k^2 p(\mathbf{r}, \omega) = 0. \quad (2.3)$$

2.2 Solutions of the Wave Equation

The Helmholtz equation in spherical coordinates, eq. (2.3), is solved through a product ansatz

$$p(kr, \boldsymbol{\theta}) = R(r)\Phi(\varphi)\Theta(\vartheta) \quad (2.4)$$

and separates into the following three ordinary differential equations (ODE) [Wil99]:

$$\frac{\partial^2 \Phi}{\partial \varphi^2} + m^2 \Phi = 0, \quad (2.5)$$

$$\frac{1}{\sin(\vartheta)} \frac{\partial}{\partial \vartheta} \left(\sin(\vartheta) \frac{\partial \Theta}{\partial \vartheta} \right) + \left[l(l+1) - \frac{m^2}{\sin^2(\vartheta)} \right] \Theta = 0, \quad (2.6)$$

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) + k^2 R - \frac{l(l+1)}{r^2} R = 0. \quad (2.7)$$

Each of the above ODE can be solved separately. Considering suitable boundary conditions ensures physically feasible results.

2.2.1 Angular Solutions

Physical azimuth solutions. The solution of eq. (2.5) is [Wil99]

$$\Phi(\varphi) = \Phi_1 \cos(\mu\varphi) + \Phi_2 \sin(\mu\varphi).$$

with Φ_1 and Φ_2 being complex valued constants. Without other azimuthal boundary conditions, the sound pressure $p(kr, \boldsymbol{\theta})$ in a homogeneous field must be continuous on a sphere and hence periodic with a period of 2π , i.e.

$$\Phi(\varphi) = \Phi(\varphi + 2\pi).$$

Thus, $\mu = m \in \mathbb{Z}$ has to be an integer value [Wil99]:

$$\Phi(\varphi) = \Phi_1 \cos(m\varphi) + \Phi_2 \sin(m\varphi). \quad (2.8)$$

Physical zenith solutions. For real-valued arguments $\cos(\vartheta)$ on the interval $0 \leq \vartheta \leq \pi$ the so-called *associated Legendre differential equation*, eq. (2.6), has the solution

$$\Theta(\vartheta) = \Theta_1 P_\nu^{|m|}(\cos(\vartheta)) + \Theta_2 Q_\nu^{|m|}(\cos(\vartheta)) \quad (2.9)$$

which contains the *associated Legendre functions* of first and second kind $P_\nu^{[m]}$ and $Q_\nu^{[m]}$ with the *order* ν and the *degree* m [Pom10]. To facilitate notation we define

$$\eta = \cos(\vartheta).$$

For $\nu \in \mathbb{N}_0$, $P_\nu^{[m]}(\eta)$ has a singularity at the south pole and $Q_\nu^{[m]}(\eta)$ has singularities at the north and south pole [Wil99], i.e.

$$\begin{aligned} \lim_{\eta \rightarrow -1} P_\nu^{[m]}(\eta) &= \pm\infty, \\ \lim_{\eta \rightarrow \pm 1} Q_\nu^{[m]}(\eta) &= \pm\infty. \end{aligned} \quad (2.10)$$

Our goal is to describe homogeneous sound fields within rigid angular boundaries which leads to boundary conditions when solving eq. (2.9). As long as these boundary conditions are fulfilled and exclude the singularities, the singularities themselves are unproblematic.

Let us consider two different scenarios for physical boundaries as depicted in fig. 2.2:

1. a semi-infinite rigid cone with a cone angle of $\vartheta_{c,2}$ ($\eta_{c,2} = \cos(\vartheta_{c,2})$) and
2. an infinite rigid double cone with the cone angles $\vartheta_{c,1}$ ($\eta_{c,1} = \cos(\vartheta_{c,1})$) and $\vartheta_{c,2}$ ($\eta_{c,2} = \cos(\vartheta_{c,2})$).

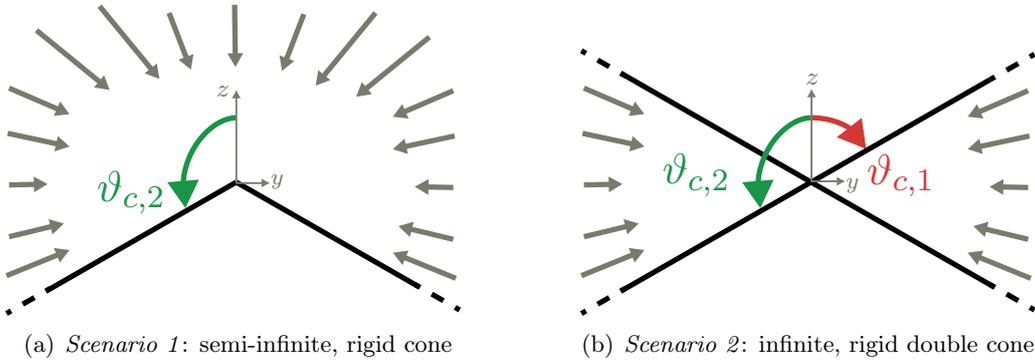


Figure 2.2: Introduction of physical boundaries.

Scenario 1:

A semi-infinite, rigid cone (with rotational symmetry around the z axis) restricts possible directions of arrival in ϑ to a spherical cap, i.e. $\vartheta \in [0, \vartheta_{c,2}]$, fig. 2.2(a). From the two solutions $P_\nu^{[m]}(\eta)$ and $Q_\nu^{[m]}(\eta)$ in eq. (2.9), $Q_\nu^{[m]}(\eta)$ is unphysical because of the singularities in the considered domain, cf. eq. (2.10). As the particle velocity vanishes in normal direction to the conical boundary and is proportional to the derivative of $p(kr, \boldsymbol{\theta})$, the boundary condition is expressed by rewriting eq. (2.9) to

$$f(\nu) = \left. \frac{\partial P_{\nu_l(|m|)}^{(|m|)}(\eta)}{\partial \vartheta} \right|_{\eta=\eta_{c,2}} = 0. \quad (2.11)$$

Solving eq. (2.11) through a root-finding algorithm results in an infinite, discrete number of *non-integer* orders $\nu_l(|m|)$ with the assignment index $l = 1, 2, \dots, \infty$ and the related degree m [Pom13]. Olver and Smith present an algorithm in [OS83] to calculate the associated Legendre functions and their derivatives for non-integer orders.

Figure 2.3 shows different $\nu_l(|m|)$ which fulfill this particular boundary condition for $\vartheta_{c,2} = 120^\circ$.

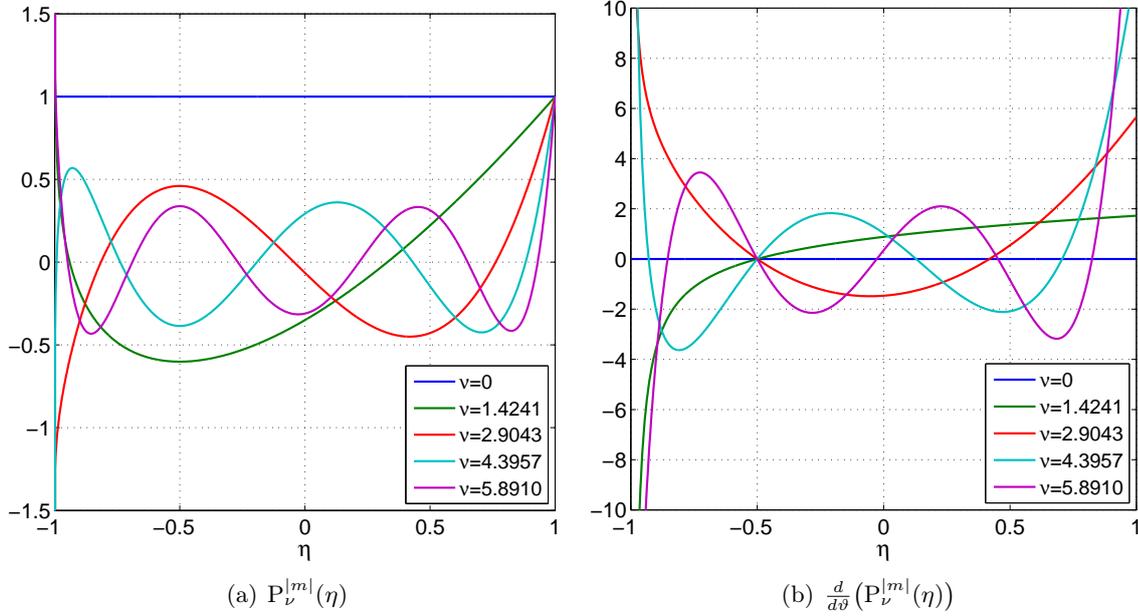


Figure 2.3: Associated Legendre functions of first kind and their derivatives with respect to ϑ , $\nu \in \mathbb{R}_0^+$ fulfilling the boundary condition in *scenario 1* for $\vartheta_{c,2} = 120^\circ$.

Scenario 2:

If we introduce an infinite, rigid double cone (with rotational symmetry around the z axis), the angular range in ϑ is restricted to $\vartheta \in [60^\circ, 120^\circ]$, fig. 2.2(b). This allows to use both types of associated Legendre functions $P_{\nu}^{(|m|)}(\eta)$ and $Q_{\nu}^{(|m|)}(\eta)$ in eq. (2.9) as none of the singularities are within the considered domain.

To combine the two functions, a weighting factor $\alpha_l(|m|)$ is introduced [Pom13]:

$$\Theta_{\nu_l(|m|)}^{(|m|)} = P_{\nu_l(|m|)}^{(|m|)}(\eta) - \alpha_l(|m|)Q_{\nu_l(|m|)}^{(|m|)}(\eta). \quad (2.12)$$

Similar to *scenario 1*, the two boundary conditions for the cone angles $\vartheta_{c,1}$ and $\vartheta_{c,2}$ are formulated by constraining the derivative of eq. (2.12) to zero:

$$\left. \frac{\partial \Theta_{\nu_l(|m|)}^{(|m|)}(\vartheta)}{\partial \vartheta} \right|_{\vartheta=\vartheta_{c,1}, \vartheta_{c,2}} = \left[\frac{\partial P_{\nu_l(|m|)}^{(|m|)}(\eta)}{\partial \vartheta} - \alpha_l(|m|) \frac{\partial Q_{\nu_l(|m|)}^{(|m|)}(\eta)}{\partial \vartheta} \right]_{\eta=\eta_{c,1}, \eta_{c,2}} = 0. \quad (2.13)$$

Inserting the two cone angles into eq. (2.13) yields

$$f(\nu) = \left[\frac{\partial P_{\nu_l(|m|)}^{(|m|)}(m)(\eta_{c,1})}{\partial \vartheta} \frac{\partial Q_{\nu_l(|m|)}^{(|m|)}(m)(\eta_{c,2})}{\partial \vartheta} + \frac{\partial P_{\nu_l(|m|)}^{(|m|)}(m)(\eta_{c,2})}{\partial \vartheta} \frac{\partial Q_{\nu_l(|m|)}^{(|m|)}(m)(\eta_{c,1})}{\partial \vartheta} \right] = 0, \quad (2.14)$$

with eliminated factor $\alpha_l(|m|)$. The non-integer orders $\nu_l(|m|)$ are determined by a root-finding algorithm which results in an infinite, discrete set of solutions for $l = 1, 2, \dots, \infty$. Finally, we are able to calculate $\alpha_l(|m|)$ using eq. (2.13) [Pom13].

The resulting functions $\Theta_{\nu_l(|m|)}^{(|m|)}$, eq. (2.12), are plotted in fig. 2.4 for different $\nu_l(|m|)$. It is easy to see that their derivatives at the defined cone angles ($\vartheta_{c,1} = 60^\circ$ and $\vartheta_{c,2} = 120^\circ$) are zero.

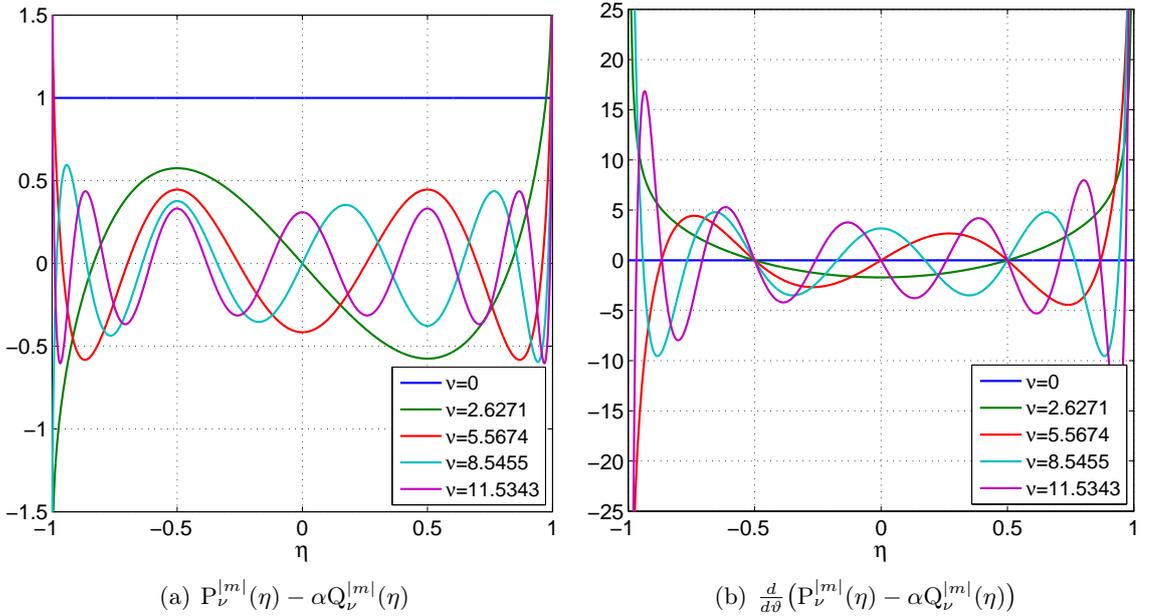


Figure 2.4: A combination of the associated Legendre functions and their derivatives with respect to ϑ , $\nu \in \mathbb{R}_0^+$ fulfilling the boundary conditions in *Scenario 2* for $\vartheta_{c,1} = 60^\circ$ and $\vartheta_{c,2} = 120^\circ$.

The functions with the determined non-integer $\nu_l(|m|)$ and $\alpha_l(|m|)$ form an orthogonal set of solutions

$$\int_{\tilde{\vartheta}_{c,1}}^{\tilde{\vartheta}_{c,2}} \Theta_{\nu_l(m)}^{(|m|)}(\vartheta) \Theta_{\nu_{l'}(m)}^{(|m|)}(\vartheta) d\vartheta = \begin{cases} 0 & \text{for } \nu_l(m) \neq \nu_{l'}(m) \\ \sigma_{\nu_l(|m|)}^2 & \text{for } \nu_l(m) = \nu_{l'}(m), \end{cases}$$

with the squared energy $\sigma_{\nu_l(|m|)}^2$ on the intervall $\vartheta_{c,1} \leq \vartheta \leq \vartheta_{c,2}$ [Pom10].

2.2.2 Radial Solutions

The radial differential equation, eq. (2.7), has the solution [Wil99]

$$R(r) = R_1 j_{\nu_l(|m|)}(kr) + R_2 h_{\nu_l(|m|)}(kr) \quad (2.15)$$

where R_1 and R_2 are complex valued constants and $j_{\nu_l(|m|)}(kr)$ and $h_{\nu_l(|m|)}(kr)$ are the *spherical Bessel functions* of first kind and the *spherical Hankel functions* of second kind, respectively.

To determine the physically feasible radial solutions the problem has to be split up into an *interior* and an *exterior* problem [Zot09a].

Physical Radial Solutions – Interior Problem. In case of an interior problem all sources are located outside a defined radius $r > r_a$.

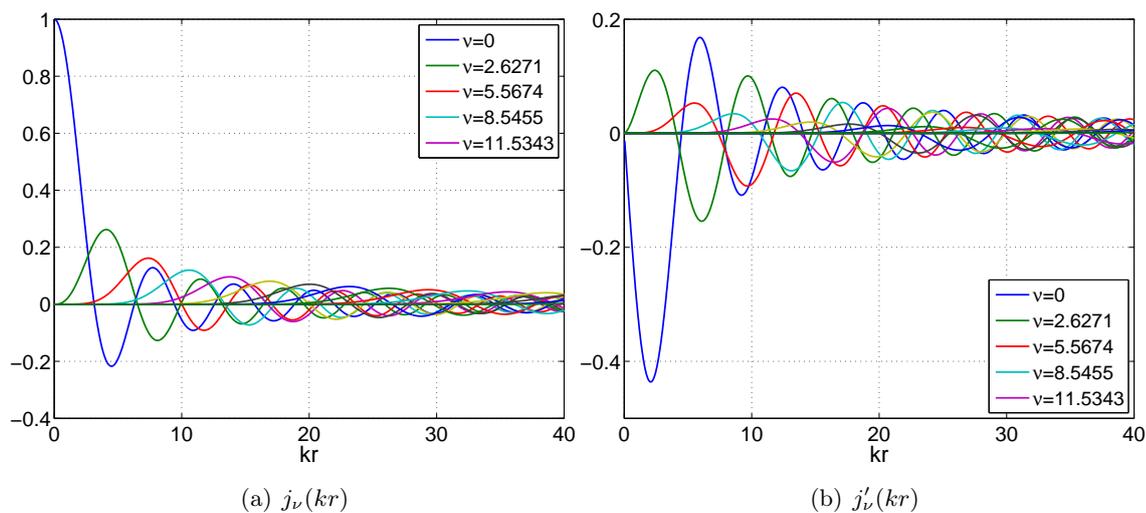


Figure 2.5: Spherical Bessel functions $j_\nu(kr)$ of first kind and their derivatives $j'_\nu(kr)$ with respect to kr for non-integer orders ν .

Setting R_2 to zero (cf. [Wil99, Zot09a]), eq. (2.15) reduces to

$$R(r) = R_1 j_{\nu_l(|m|)}(kr) \quad (2.16)$$

and describes incident sound fields within the range $0 \leq r \leq r_a$.

The derivatives $j'_{\nu_l(|m|)}$ in fig. 2.5(b) can be calculated by the recurrence relationship [Dlm27]

$$R'_{\nu_l(|m|)}(kr) = R_{\nu_l(|m|)-1}(kr) - \frac{\nu_l(|m|) + 1}{kr} R_{\nu_l(|m|)-1}(kr). \quad (2.17)$$

Physical Radial Solutions – Exterior Problem. In case of an exterior problem all sources lie inside a defined radius $r < r_a$. The sound field can be described within the range $r_a \leq r < \infty$ by solutions that fulfill the *Sommerfeld radiation condition*

$$\lim_{r \rightarrow \infty} r \left(\frac{\partial}{\partial r} p(kr, \boldsymbol{\theta}) + ikp(kr, \boldsymbol{\theta}) \right) = 0$$

which is only met by

$$R(r) = R_2 h_{\nu_l(|m|)}(kr), \quad (2.18)$$

the spherical Hankel functions of second kind [Zot09a, Sch92, Wil99]. The magnitude of these functions as shown in fig. 2.6(a) clearly indicate the near- and far-field region. Equation (2.17) is used to calculate the derivatives in fig. 2.6(b).

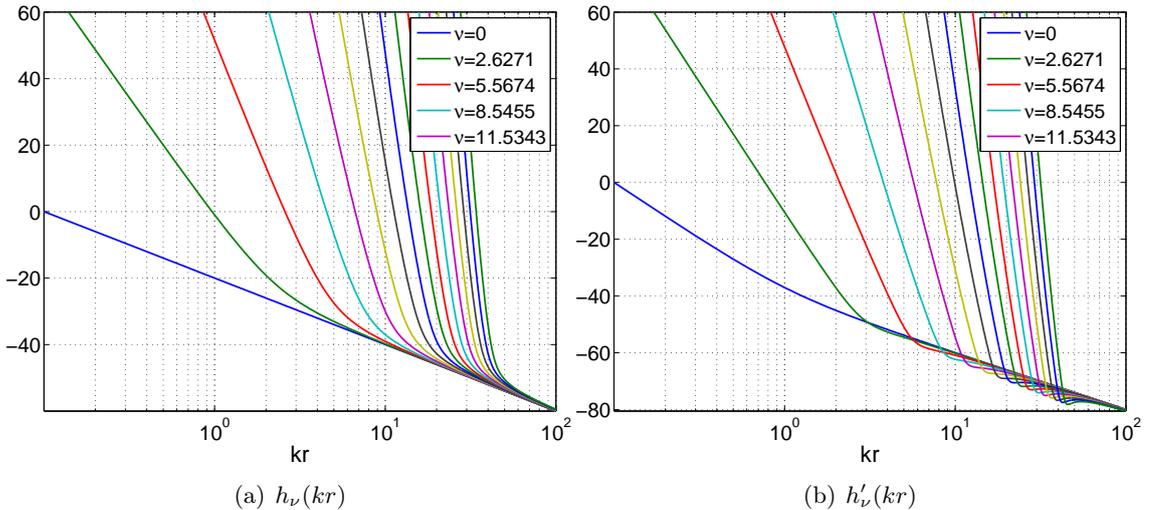


Figure 2.6: Magnitude of the spherical Hankel functions of second kind $h_\nu(kr)$ and their derivatives $h'_\nu(kr)$ with respect to kr for non-integer orders ν .

The spherical Hankel functions $h_{\nu_l(|m|)}(kr)$ are approximated by [Zot09a]

$$|h_{\nu_l(|m|)}(kr)| \propto \begin{cases} (2\nu_l(|m|) + 1)!! / (kr)^{\nu_l(|m|)+1}, & \text{near-field } kr \ll 2 \\ 1/(kr), & \text{far-field } kr \gg \frac{\nu_l(|m|)(\nu_l(|m|)+1)}{2} \end{cases} \quad (2.19)$$

where $(\cdot)!!$ denotes the double factorial. As we are dealing with non-integer, positive arguments z , the Gamma function [MWG22]

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$$

is applied and yields [MWD22]

$$z!! = \sqrt{\frac{2^{z+1}}{\pi}} \Gamma\left(\frac{z}{2} + 1\right).$$

For $z = 0$, the common double factorial as defined in [Wil99]

$$z!! = \frac{z!}{2^z z!}$$

with $z! = \prod_{k=1}^z k$ and $0! = 1$ is used.

Figure 2.7 visualizes the far-field approximation of eq. (2.19) by plotting the minimum source radius r_s in dependency of frequency. Above this radius, sources can be simplified to emitting plane waves.

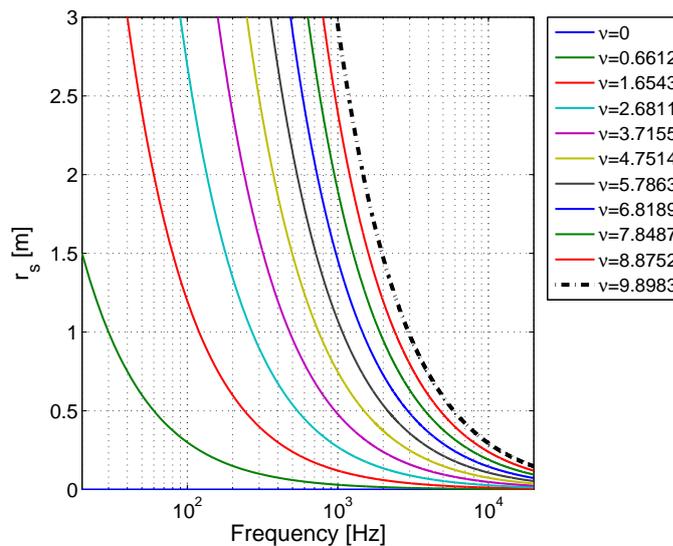


Figure 2.7: Minimum source radius r_s to simplify sources to emitting plane waves.

2.3 Spherical Segment Basis Solutions

The entire solution of eq. (2.4) describes the homogeneous sound field and is given by the physical solutions, also called *spherical segment basis solutions*, for the *incident* sound field (eq. (2.16), (2.8), (2.9))

$$\begin{aligned} p_i(kr, \boldsymbol{\theta}) &= R(r)\Phi(\varphi)\Theta(\vartheta) = \\ &= R_{1,l}j_{\nu_l(|m|)}(kr) \left[\Phi_{1,m} \cos(m\varphi) + \Phi_{2,m} \sin(m\varphi) \right] \cdot \\ &\quad \cdot \left[\Theta_{1,l,m} P_{\nu_l(|m|)}^{|m|}(\eta) + \Theta_{2,l,m} Q_{\nu_l(|m|)}^{|m|}(\eta) \right] \end{aligned}$$

and the *radiating* or *scattered* sound field (eq. (2.18), (2.8), (2.9))

$$\begin{aligned} p_r(kr, \boldsymbol{\theta}) &= R_{2,l}h_{\nu_l(|m|)}(kr) \left[\Phi_{1,m} \cos(m\varphi) + \Phi_{2,m} \sin(m\varphi) \right] \cdot \\ &\quad \cdot \left[\Theta_{1,l,m} P_{\nu_l(|m|)}^{|m|}(\eta) + \Theta_{2,l,m} Q_{\nu_l(|m|)}^{|m|}(\eta) \right] \end{aligned}$$

which yields [Kö&11]

$$\begin{aligned} p(kr, \boldsymbol{\theta}) &= p_i(kr, \boldsymbol{\theta}) + p_r(kr, \boldsymbol{\theta}) = \\ &= \left[R_{1,l}j_{\nu_l(|m|)}(kr) + R_{2,l}h_{\nu_l(|m|)}(kr) \right] \left[\Phi_{1,m} \cos(m\varphi) + \Phi_{2,m} \sin(m\varphi) \right] \cdot \\ &\quad \cdot \left[\Theta_{1,l,m} P_{\nu_l(|m|)}^{|m|}(\eta) + \Theta_{2,l,m} Q_{\nu_l(|m|)}^{|m|}(\eta) \right]. \end{aligned}$$

2.3.1 A Suitable Order Truncation

In general, there exists an infinite number of $\nu_l(|m|)$ that fulfills the double conical boundary condition described in sec. 2.2.1. However, practical limitation demand an appropriate limitation. First, let us introduce a linear index $q := l^2 + l + m + 1 = 1, 2, \dots, Q$. For notational convenience, we define $\nu_q = \nu_l(|m|)$ and $\alpha_q = \alpha_l(|m|)$.

Figure 2.8 depicts two limited sets of solutions $\nu_l(|m|)$ both with 64 entries but truncated differently. More precisely, the first set is truncated with respect to the index l which results in $(l+1)^2$ solutions whereas the second set is limited by a maximal order N and yields Q solutions. The aim is to find a reasonable truncation of the set of solutions motivated by acoustical considerations.

Examining the spherical Bessel functions in fig. 2.9, which describe the incident sound field, cf. sec. 2.2.2, for a low and high order ν and a radius $r_a = 0.1$ m provides a straight answer which is formulated in the next paragraph.

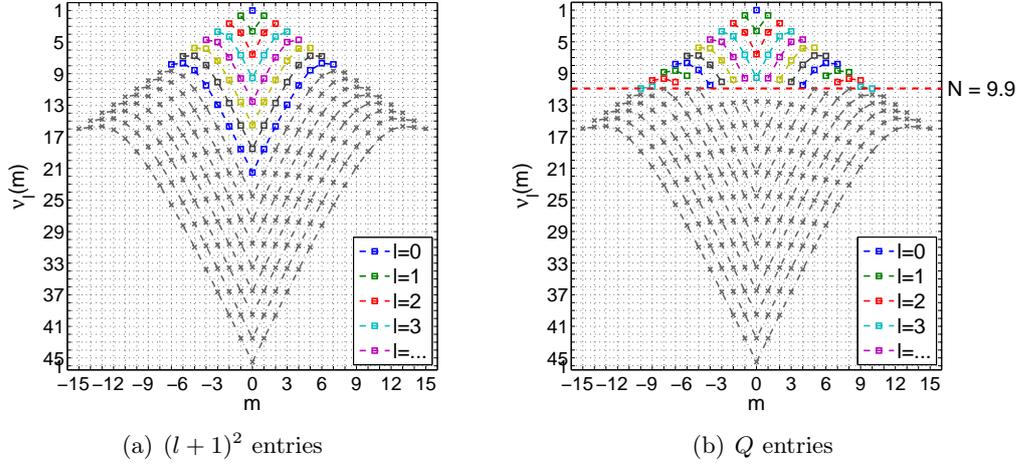


Figure 2.8: Two differently truncated sets of solutions ν that fulfill the double conical boundary condition for $\vartheta_{c,1} = 60^\circ$, $\vartheta_{c,2} = 120^\circ$. Both sets contain 64 entries.

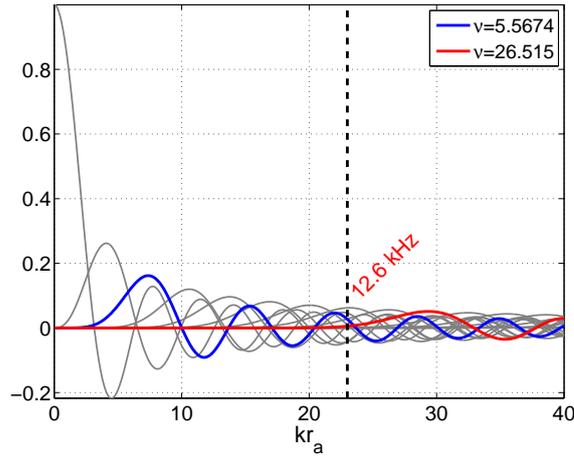


Figure 2.9: Spherical Bessel functions $j_\nu(kr_a)$ for a low (blue line) and a high (red line) order ν , $r_a = 0.1$ m.

Upper spatial frequency limit. In practice, the upper spatial frequency f_u , up to which we are able to decompose a surrounding soundfield, is directly linked to the largest order ν_Q and the aperture radius r_a , i.e. $kr_a < \nu_Q$, which yields

$$f_u \approx \frac{1}{2\pi} \frac{\nu_Q}{r_a} c.$$

Thence, higher-order modes are reflected into lower ones and make harmonic components indistinguishable. This effect is called *spatial aliasing*.

In our special case we are able to control $Q = 64$ sound field components to decompose a surrounding sound field up to $f_u \approx 5.4$ kHz ($r_a = 0.1$ m, $c = 343$ m/s). The spherical

Bessel function $j_\nu(kr_a)$ in fig. 2.9 evaluated for a high order results in a function that is zero up to $f \approx 12.6$ kHz, so large ν in fig. 2.8 can be neglected.

For this reason, the set of orders ν_q is truncated wrt. a maximal order N (fig. 2.8(b)):

$$\begin{aligned}\nu_q &: 0 \leq \nu_q < \nu_Q \leq N < \infty \\ q &= 1, 2, \dots, Q, \quad Q < \infty\end{aligned}$$

with pairwise appearance of $\nu_{q>0}$ (e.g. $\nu_2 = \nu_3$, cf. fig 2.8) [Pom13].

Chapter 3

Spherical Segment Harmonics

Similar to spherical array signal processing, the angular solutions in φ and ϑ can be summarized to spherical segment basis functions. As the array area is no longer spherical but restricted in ϑ , cf. sec. 2, these basis functions are called *spherical segment harmonics* (SSH) in the following. The normalized, real-valued SSH are defined by [Köf11]

$$Y_{\nu_q}(\boldsymbol{\theta}) = Y_{\nu_l(|m|)}^m(\boldsymbol{\theta}) = \Xi_{\nu_l(|m|)}^{|m|} \Theta_{\nu_l(|m|)}^{|m|}(\vartheta) \begin{cases} \sin(m\varphi), & m < 0, \\ \cos(m\varphi), & m \geq 0. \end{cases}$$

The sine/cosine terms can be written as to be selected by the sign of m [Zot09a]. $\Xi_{\nu_l(|m|)}^{|m|}$ symbolizes the real-valued normalization term which is identical for $\pm m$ and normalizes the set of basis functions in a way that they reduce to a product of *Kronecker deltas* after integration on the spherical segment,

$$\int_{\tilde{\varphi}=0}^{2\pi} \int_{\tilde{\vartheta}=\vartheta_{c,1}}^{\vartheta_{c,2}} Y_{\nu_l(|m|)}^m(\boldsymbol{\theta}) Y_{\nu_{l'}(|m'|)}^{m'}(\boldsymbol{\theta}) \sin(\vartheta) d\vartheta d\varphi = \delta_{ll'} \delta_{mm'} \quad (3.1)$$

with

$$\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

and thus become *orthonormal* [Pom13].

For non-integer values of $\nu_l(|m|)$, there is no simple analytical expression for the normalization terms $\Xi_{\nu_l(|m|)}^{|m|}$. Hence, they were determined numerically using the relation [HC97, Köf11]

$$\Xi_{\nu_l(|m|)}^{|m|} = \left[\tau(m) \int_{\tilde{\vartheta}_{c,1}}^{\tilde{\vartheta}_{c,2}} (\Theta_{\nu_l(|m|)}^{|m|}(\vartheta))^2 \sin(\vartheta) d\vartheta \right]^{-1/2} \quad (3.2)$$

with the term

$$\tau(m) = (2 - \delta_m) \pi.$$

Figure (3.1) exemplarily shows the normalized absolute values of the analytical, real-valued SSH $|Y_{\nu_q}(\boldsymbol{\theta})| = |Y_{\nu_l(|m|)}^m(\boldsymbol{\theta})|$ for $l = 0 \dots 3$ and $m = -l \dots l$ ($q = 1, 2, \dots, 9$) with color-

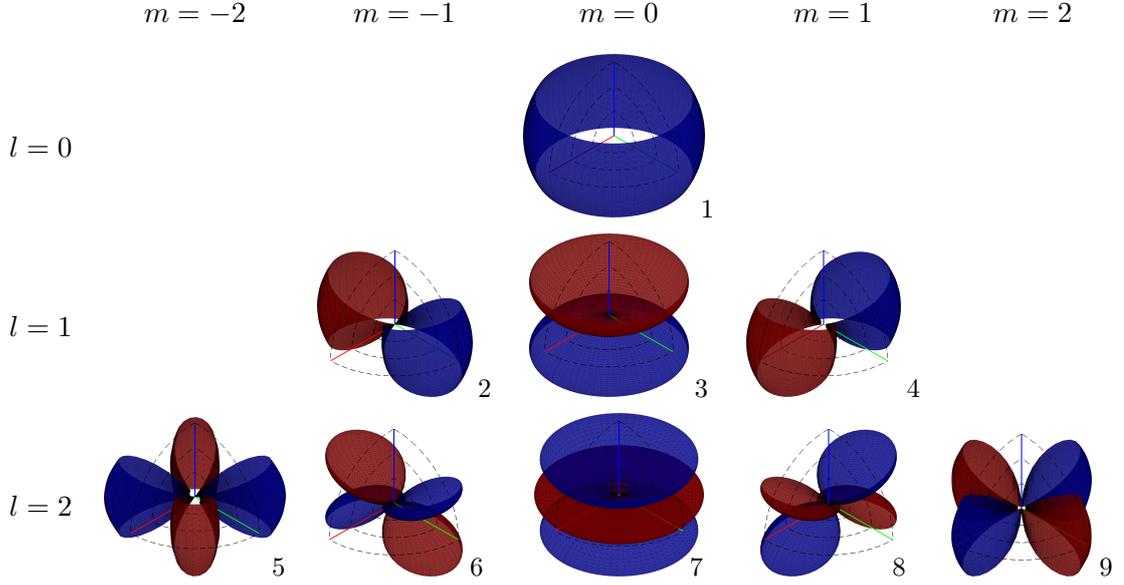


Figure 3.1: Normalized real-valued SSH $|Y_{\nu_q}(\theta)| = |Y_{\nu_l(|m|)}^m(\theta)|$ up to index $l = 2$, $m = -l \dots l$ ($q = 1, 2, \dots, 9$) evaluated within the cone angles $\vartheta_{c,1} = 60^\circ$, $\vartheta_{c,2} = 120^\circ$.

coded phase in $[rad]$ (blue: 0, red: π). The restricted interval in ϑ and a rotation of $\pi/(1 + |m|)$ is clearly visible. Additionally, they are symmetric (every $2(q - 1) + 1$ SSH) or skew-symmetric (every $2q$ SSH) with respect to the z axis. The number of nodal circles is defined by the index l whereas the degree $|m|$ counts how many nodal circles lie on vertical planes.

3.1 Spherical Domain Transforms

Similar to *time-frequency transforms* with the Fourier kernel $e^{\pm i\omega t}$, the following spherical transforms can be seen as *spatial Fourier transforms* where the indices l and m substitute the frequency index ω .

3.1.1 The Continuous SSH Transform (SSHT)

With the help of SSH, a continuous sound pressure $p(kr, \theta)$ on the radius r and the defined angular ranges $0 \leq \varphi < 2\pi$ and $\vartheta_{c,1} \leq \vartheta \leq \vartheta_{c,2}$ can be transformed into the SSH domain

using the *spherical segment harmonics transform* (SSHT) to obtain the *spherical wave spectrum* [Köb11]

$$\psi_q = \int_{\tilde{\varphi}=0}^{2\pi} \int_{\tilde{\vartheta}=\vartheta_{c,1}}^{\vartheta_{c,2}} p(kr, \boldsymbol{\theta}) Y_{\nu_q}(\boldsymbol{\theta}) \sin \vartheta d\vartheta d\varphi = \text{SSHT}\{p(kr, \boldsymbol{\theta})\}.$$

The *inverse spherical segment transform* (ISSHT) without sound field truncation, i.e. the sound field is analyzed up to $q \rightarrow \infty$, is complete and can be formulated as a weighted sum of SSH

$$\begin{aligned} p(kr, \boldsymbol{\theta}) &= \text{ISSHT}\{\psi_q\} = \sum_{q=1}^{\infty} \psi_q Y_{\nu_q}(\boldsymbol{\theta}) \\ &= \sum_{q=1}^{\infty} [b_q j_{\nu_q}(kr) + c_q h_{\nu_q}(kr)] Y_{\nu_q}(\boldsymbol{\theta}) \end{aligned} \quad (3.3)$$

where b_q and c_q are complex coefficients usually referred to as *wave spectra* for the incident and the scattered sound field [Zot09a, Teu07]. It perfectly reconstructs the original sound pressure, meaning that

$$\text{ISSHT}\{\psi_q\} = \text{ISSHT}\{\text{SSHT}\{p(kr, \boldsymbol{\theta})\}\} \equiv p(kr, \boldsymbol{\theta})$$

and fulfills the Parseval theorem, i.e.

$$\int_{\tilde{\varphi}=0}^{2\pi} \int_{\tilde{\vartheta}=\vartheta_{c,1}}^{\vartheta_{c,2}} |p(kr, \boldsymbol{\theta})|^2 \sin(\vartheta) d\vartheta d\varphi = \sum_{q=1}^{\infty} |\psi_q|^2.$$

3.1.2 Angular Band Limitation

For a complete analysis of the sound field, the index q in eq. (3.3) has to reach infinity, i.e. $q \rightarrow \infty$. However, practical limitations need to be considered thus it is common to limit the set of solutions with respect to a maximal order N as shown in sec. 2.3.1, which is equal to an angular band-limitation. This results in the band-limited sound field with Q components

$$p_N(kr, \boldsymbol{\theta}) = \sum_{q=1}^Q \psi_q Y_{\nu_q}(\boldsymbol{\theta}) \quad (3.4)$$

$$= \sum_{q=1}^Q [b_q j_{\nu_q}(kr) + c_q h_{\nu_q}(kr)] Y_{\nu_q}(\boldsymbol{\theta}). \quad (3.5)$$

Chapter 4

Sampling the Spherical Segment

4.1 Vectorization of the SSH

In reality, there is no 'continuous' pressure microphone to record a sound pressure distribution on a spherical segment with a radius r_a . This problem is overcome by sampling the considered area with S microphones at the positions $r_a \boldsymbol{\theta}_s$, $s = 1, 2, \dots, S$, leading to a concise vector / matrix notation described in [Zot09a] and extended to SSH.

In general, summations can be represented by scalar products or matrix multiplications. Equation (3.4) is reformulated which yields the *discrete SSH transform* (DSSHT):

$$\mathbf{p}_N = \mathbf{Y}_N \boldsymbol{\psi}_N = \quad (4.1)$$

$$= \mathbf{Y}_N \left[\mathbf{J}_N \mathbf{b}_N + \mathbf{H}_N \mathbf{c}_N \right] \quad (4.2)$$

with the truncated SSH matrix

$$\mathbf{Y}_N = \begin{pmatrix} \mathbf{y}_N^T(\boldsymbol{\theta}_1) \\ \mathbf{y}_N^T(\boldsymbol{\theta}_2) \\ \vdots \\ \mathbf{y}_N^T(\boldsymbol{\theta}_S) \end{pmatrix}, \quad \mathbf{y}_N(\boldsymbol{\theta}_s) = (Y_{\nu_1}(\boldsymbol{\theta}_s), Y_{\nu_2}(\boldsymbol{\theta}_s), \dots, Y_{\nu_Q}(\boldsymbol{\theta}_s))^T$$

up to order N with $S \times Q$ entries. The wave spectra are written into vectors of the form

$$\mathbf{b}_N := [b_1, b_2, \dots, b_Q]^T,$$

$$\mathbf{c}_N := [c_1, c_2, \dots, c_Q]^T.$$

The radial functions are employed in eq. (4.2) as diagonal matrices

$$\begin{aligned}\mathbf{j}_N &:= [j_{\nu_1}(kr), j_{\nu_2}(kr), \dots, j_{\nu_Q}(kr)]^T, \\ \mathbf{h}_N &:= [h_{\nu_1}(kr), h_{\nu_2}(kr), \dots, h_{\nu_Q}(kr)]^T, \\ \mathbf{J}_N &:= \text{diag}\{\mathbf{j}_N\}, \\ \mathbf{H}_N &:= \text{diag}\{\mathbf{h}_N\}.\end{aligned}$$

4.1.1 Sound Field Decomposition

Sound field decomposition within the spherical segment domain is strongly dependent on the SSH matrix and thus the microphone positions. The relation between the spherical wave spectrum $\boldsymbol{\psi}_N$ and the sound pressure \mathbf{p}_N , eq. (4.1), is a system of S equations with Q unknown variables. Its pseudoinverse solution depends on the number of sampling points.

As described in [Pom08, Zot09a], there are three scenarios:

1. $S = Q$, the inverse is $(\mathbf{Y}_N^T)^{-1}$: If it exists, the measured spherical wave spectrum $\boldsymbol{\psi}$ is equivalent to the given, order-limited spherical wave spectrum $\boldsymbol{\psi}_N$,
 $\boldsymbol{\psi} = (\mathbf{Y}_N^T)^{-1} \mathbf{Y}_N \boldsymbol{\psi}_N$, and the measured sample vector is equivalent to the reconstructed sample vector, $\mathbf{p}_N = \mathbf{Y}_N (\mathbf{Y}_N^T)^{-1} \mathbf{p}$,
2. $S > Q$ the pseudoinverse is left-inverse $(\mathbf{Y}_N^T \mathbf{Y}_N)^{-1} \mathbf{Y}_N^T$: If it exists, the measured spherical wave spectrum is equivalent to the given, order-limited spherical wave spectrum, $\boldsymbol{\psi} = (\mathbf{Y}_N^T \mathbf{Y}_N)^{-1} \mathbf{Y}_N^T \mathbf{Y}_N \boldsymbol{\psi}_N$, or
3. $S < Q$ the pseudoinverse is right-inverse $\mathbf{Y}_N (\mathbf{Y}_N \mathbf{Y}_N^T)^{-1}$: If it exists, the measured sample vector is equivalent to the reconstructed sample vector,
 $\mathbf{p}_N = \mathbf{Y}_N \mathbf{Y}_N^T (\mathbf{Y}_N \mathbf{Y}_N^T)^{-1} \mathbf{p}$.

In order to distinguish original from reconstructed signals, \mathbf{p}_N was replaced by \mathbf{p} and $\boldsymbol{\psi}_N$ by $\boldsymbol{\psi}$, respectively, in the model equation (4.1).

Scenario 1 and 2 are suitable to calculate acoustic holography which is subject to an order limitation in the spherical wave spectrum. Thus, it is necessary that $\boldsymbol{\psi}_N$ is uniquely determined as soon as the order limitation is fulfilled. Ideally, this yields

$$\|\mathbf{p}_N - \mathbf{Y}_N \boldsymbol{\psi}_N\|^2 = 0.$$

In this thesis, we choose $S = Q$, thus eq. (4.1) is inverted by

$$\boldsymbol{\psi}_N = (\mathbf{Y}_N^T)^{-1} \mathbf{p}_N. \quad (4.3)$$

4.1.2 Spatial Aliasing

Similar to band-limited time-domain sampling with respect to an upper frequency limit which is defined by half of the sampling frequency f_s (*Nyquist-Shannon* criterium [OSB10]) there are as well limitations in the spatial domain. *Spatial aliasing* occurs as soon as the limited bandwidth (in terms of harmonic orders) is exceeded and manifests in artifacts due to higher order modes which are mirrored into the lower order modes. This phenomenon leads to uncontrolled sidelobes in beampatterns and results in a low signal-to-interference ratio (SIR).

Spatial Aliasing Error. To show the phenomenon of spatial aliasing qualitatively, a matrix product can be formulated as described in [Ple09] by

$$\begin{aligned}\hat{\boldsymbol{\psi}}_N &= (\mathbf{Y}_N^T)^{-1} \mathbf{Y}_{\tilde{N}} \boldsymbol{\psi}_{\tilde{N}} \\ &= (\mathbf{Y}_N^T)^{-1} \begin{bmatrix} \mathbf{Y}_N & \mathbf{Y}_{\nu > N} \end{bmatrix} \boldsymbol{\psi}_{\tilde{N}} \\ &= \begin{bmatrix} \mathbf{I} & (\mathbf{Y}_N^T)^{-1} \mathbf{Y}_{\nu > N} \end{bmatrix} \boldsymbol{\psi}_{\tilde{N}}\end{aligned}$$

with a high simulation order $\tilde{N} = 25$. This results in the estimated spherical wave spectrum $\hat{\boldsymbol{\psi}}_N$ that is made up of a $Q \times Q$ identity matrix \mathbf{I} (cf. orthonormality of SSH, eq. (3.1)) if $\mathbf{Y}_{\tilde{N}}$ is evaluated up to the same order N as $(\mathbf{Y}_N^T)^{-1}$. Otherwise, $\hat{\boldsymbol{\psi}}_N$ shows spatial aliasing artefacts, caused by modal orders which lie outside the controllable bandwidth, in the right hand entries of the matrix product $(\mathbf{Y}_N^T)^{-1} \mathbf{Y}_{\tilde{N}}$, fig. 4.1.

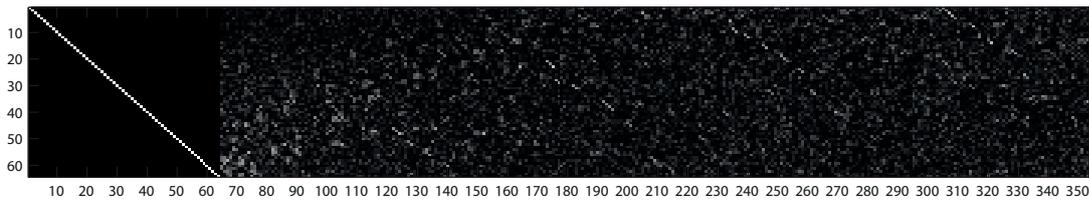


Figure 4.1: Matrix product $(\mathbf{Y}_N^T)^{-1} \mathbf{Y}_{\tilde{N}}$ to show spatial aliasing artefacts for orders $\nu > N$, $N = 9.9$, $\tilde{N} = 25$.

4.2 The Condition Number of the SSH Matrix

For a robust decomposition of a surrounding sound field into SSH it is important to find a microphone arrangement that guarantees a small *condition number* of the SSH matrix

$$1 \leq \kappa(\mathbf{Y}_N) \ll \infty.$$

The condition number is defined by the ratio of the largest and smallest singular value determined by a *singular value decomposition* [KL80].

Singular Value Decomposition (SVD). Assume an arbitrary rectangular, complex, unitary¹ matrix $\mathbf{A} \in \mathbb{C}^{[S \times Q]}$. This matrix can be rewritten as a product of three matrices with special properties:

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H. \quad (4.4)$$

$\mathbf{U} \in \mathbb{C}^{[S \times S]}$ and $\mathbf{V}^H \in \mathbb{C}^{[Q \times Q]}$ contain the left and right singular vectors² of \mathbf{A} columnwise and are unitary as well. $\mathbf{\Sigma} \in \mathbb{C}^{[S \times Q]}$ has the form

$$\mathbf{\Sigma} = \left[\begin{array}{ccc|ccc} & & & \vdots & & \\ & & & \vdots & & \\ & & \text{diag}\{\sigma_q\} & \vdots & \dots & 0 & \dots \\ & & & \vdots & & & \\ \hline & & & \vdots & & & \\ \dots & & 0 & \dots & \dots & 0 & \dots \\ & & & \vdots & & & \end{array} \right]$$

and contains a diagonal matrix made up by the singular values of \mathbf{A} in descending order, i.e.

$$\boldsymbol{\sigma}_q = [\sigma_1, \sigma_2 \dots \sigma_Q]^T, \text{ with } \sigma_1 \geq \sigma_2 \dots \geq \sigma_Q > 0.$$

To solve the matrix equation (4.4) with respect to $\mathbf{\Sigma}$ we multiply \mathbf{A} by \mathbf{U}^H and \mathbf{V} from left and right, respectively, i.e.

$$\mathbf{\Sigma} = \mathbf{U}^H \mathbf{A} \mathbf{V}.$$

Finally, the condition number of \mathbf{A} is defined by the ratio of the largest and smallest singular value:

$$\kappa(\mathbf{A}) = \frac{\max(\boldsymbol{\Sigma})}{\min(\boldsymbol{\Sigma})} = \frac{\sigma_{\max}(\mathbf{A})}{\sigma_{\min}(\mathbf{A})}.$$

¹ A matrix is unitary if and only if its columns are orthonormal, i.e. $\mathbf{A}^H \mathbf{A} = \mathbf{A} \mathbf{A}^H = \mathbf{I}$ (identity matrix).

² These are the orthonormal eigenvectors of $(\mathbf{A} \mathbf{A}^H)$ and $(\mathbf{A}^H \mathbf{A})$, respectively.

The condition number indicates the sensitivity of the system to errors. A matrix with a condition number close to 1 is classified as *well-conditioned* and can be inverted properly. An *ill-conditioned* matrix ($\kappa \gg 1$) is prone to produce strongly fluctuating results in case of slight input variations (noise, calibration errors, positioning errors, rounding errors in numerical evaluation systems) or one even fails in inverting the (singular) matrix ($\kappa \rightarrow \infty$). A matrix inversion is inevitable for decomposing the array microphone signals into SSH, cf. eq. (4.3), so a small condition number is an essential quality requisite for the prototype. As the condition number for the maximal order N of choice is determined by the sampling scheme, the following section discusses sampling strategies.

4.3 Sampling Strategies

There are various schemes to sample a sphere with microphones aiming at a small condition number κ and find a robust arrangement contra noise and positioning errors. Zotter lists some conventional and well studied sampling schemes:

- extremal points for hyperinterpolation [SW03]
- spiral points [RSZ95]
- equal area decomposition [SK97]
- Gauss-Legendre raster [Sne94]
- equiangular cylindric raster [Sne94]
- equiangular raster [DH94]
- hierarchical equal area isoLatitude pixelization (HEALPix) raster [GHB⁺05]

and analyzes some of them in [Zot09a, Zot09b].

In [Köf11] different approaches to optimize a microphone distribution on a spherical cap are presented. They either rely on a geometrical regularity (principle of repellant charges, minimizing the area spread of Voronoi cells and the iterative decomposition of Voronoi cell centroids, HEALPix grid) or try to minimize the condition number (direct approach, maximizing the smallest singular value).

4.3.1 Optimal Microphone Layout on a Fixed Raster

The most suitable and efficient algorithm to determine optimal sampling nodes for a spherical cap was selecting a subset of nodes from a pre-defined HEALPix grid to achieve a minimal condition number for \mathbf{Y}_N ($S = Q$) [Köf11].

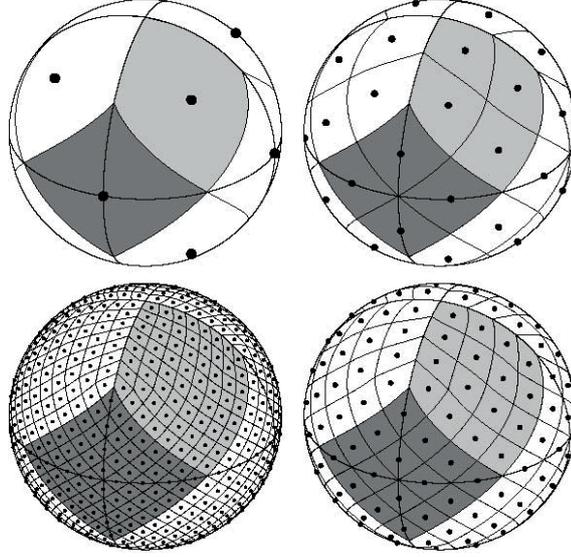


Figure 4.2: Moving clockwise: HEALPix raster on a sphere with $N_{side} = 1, 2, 4, 8$ which results in $N_{pix} = 12, 48, 192, 768$ pixels with center points on $N_{ring} = 3, 7, 15, 31$ rings of constant latitude [GHB⁺05].

HEALPix Tessellation. The basic HEALPix tessellation consists of 4 equatorial and 8 polar patches (4 north and 4 south) with equal area. This basis tessellation is subdivided into $N_{side} \in 2^a$ ($a \in \mathbb{Z}^+$) subpixels, which yields a resolution of $N_{pix} = 12N_{side}^2$ HEALPix points. Additionally, the pixel centers lie on $N_{ring} = 4N_{side} - 1$ rings of constant latitude [GHB⁺05].

The original optimization process has been designed for a spherical cap and is described in detail in [Köß11]. Kößler starts with an initial distribution where the first 64 HEALPix points located around the north pole on the spherical cap are occupied. In each *iteration*, the s -th microphone is moved by one node over all vacant grid positions, and the condition number $\kappa(\mathbf{Y}_N)$ of the resulting SSH matrix is calculated for each step. The best position is kept and the iteration of the next microphone starts. This procedure is repeated for all S microphones, which defines a *run*. If the resulting condition number is better than that of the run before the underlying distribution will be used as the new start distribution in the following run. The condition number approaches the minimum asymptotically until the process is aborted manually.

Adaptation of the Optimization Process.

- After some modifications to upgrade the performance of the algorithm, the optimization process is adapted for the spherical segment scenario, i.e. only a HEALPix raster $N_{pix,r}$ restricted in ϑ is permitted and the related functions for the double cone boundary condition are included.

- Because of construction constraints, an upper and a lower circular *taboo ring* as well as longitudinal *taboo belts* are introduced which diminish the arrangement possibilities (see fig. 4.3, black areas).

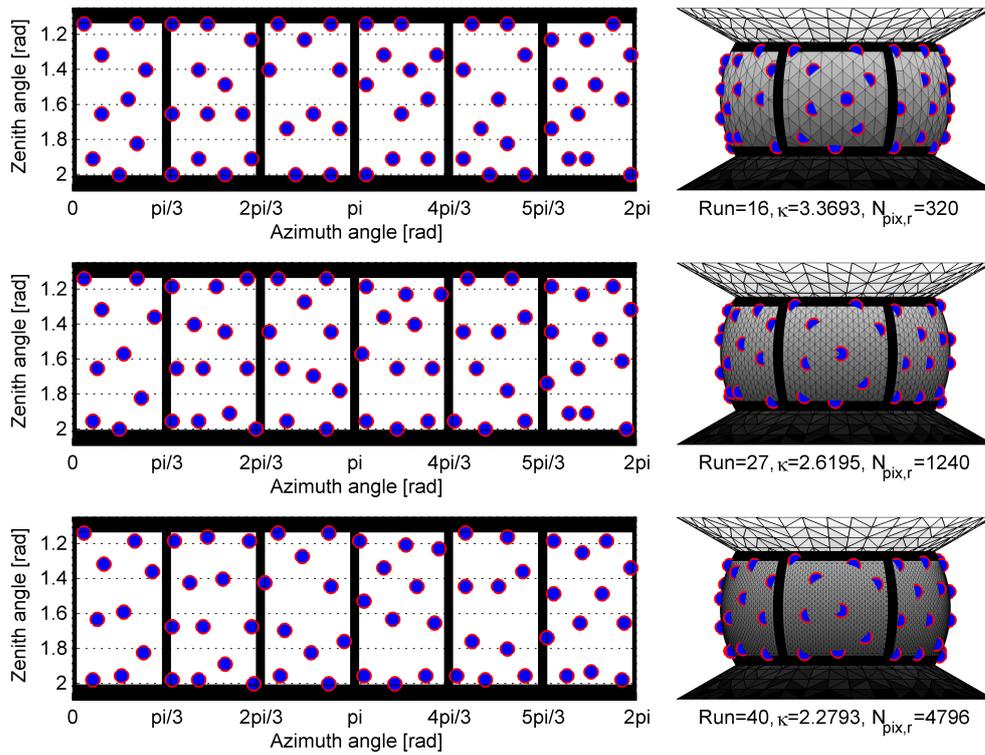


Figure 4.3: Optimal microphone distribution using the described HEALPix subset algorithm with a resulting condition number $\kappa(\mathbf{Y}_N) \approx 2.3$, $S = 64$, $Q = 64$ spherical segment basis functions, $N_{pix,r} = 320, 1240, 4796$ on a spherical segment with the cone angles $\vartheta_{c,1} = 60^\circ$, $\vartheta_{c,2} = 120^\circ$.

- We start with a random subset which normally has an a priori better condition number than a microphone distribution using the first S HEALPix nodes.
- The possibility to converge towards a local minimum is minimized by selecting the microphone positions in random order within each run.
- Starting with a relative small HEALPix grid to select from ($N_{pix,r} = 320$, fig. 4.3 above), the resolution of the HEALPix grid is increased ($N_{pix,r} = 1240$, fig. 4.3 middle) as soon as there is no improvement in consecutive runs, which is indicated by the first vertical dashed red line in fig. 4.4. This refinement obviously allows to improve the condition number. The algorithm is aborted whenever there is no improvement of κ in consecutive runs after increasing the grid resolution twice from $N_{pix,r} = 320$ via $N_{pix,r} = 1240$ to $N_{pix,r} = 4796$ (cf. fig. 4.3 at the bottom).

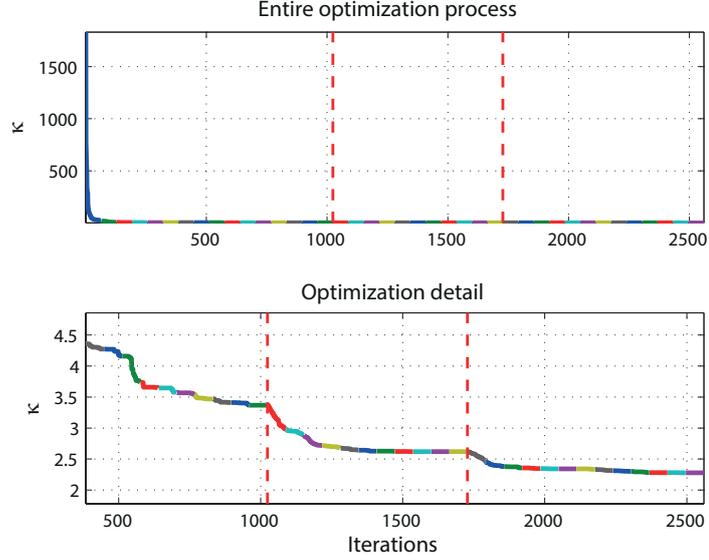


Figure 4.4: History of the optimization process with $S = 64$ microphones, $Q = 64$ spherical segment basis functions, $N_{pix,r} = 320, 1240, 4796$ grid points on a spherical segment with the cone angles $\vartheta_{c,1} = 60^\circ$, $\vartheta_{c,2} = 120^\circ$. The runs consist of S iterations and are depicted colored, the vertical red dashed lines indicate a raster refinement.

Due to the intrinsic random execution, the algorithm was repeated for several times and different overall executions yield a different final result. Finally, the selected subset of $S = 64$ points for $Q = 64$ SSH successfully achieved an acceptable condition number $\kappa(\mathbf{Y}_N) \approx 2.3$, using the microphone distribution depicted in fig. 4.3 (at the bottom). A list of all microphone coordinates $r_a \boldsymbol{\theta}_s$ can be found in app. A.2.

The upcoming section investigates the possibilities how to arrange the determined microphone distribution on an aperture radius r_a .

4.4 Open and Rigid Array Configurations

In order to analyze the frequency-dependent behavior of spherical modes, different array configurations and their impact on the modal transfer functions are simulated in the following.

4.4.1 Open Array Configuration

Although an open array configuration [AW02, GRS02] intuitively seems to be the optimal choice, pursuing the ideal of little sound field perturbations, numerical problems arise. The radius-dependent terms in eq. (3.5) are then [Teu07, Raf05]

$$\rho_{\nu_q}^{(h)}(kr) = j_{\nu_q}(kr)$$

as there is no outgoing sound field ($c_q \equiv 0$). In fig. 4.5(a), the magnitude of the modal responses in open apertures are depicted for omnidirectional pressure transducers.

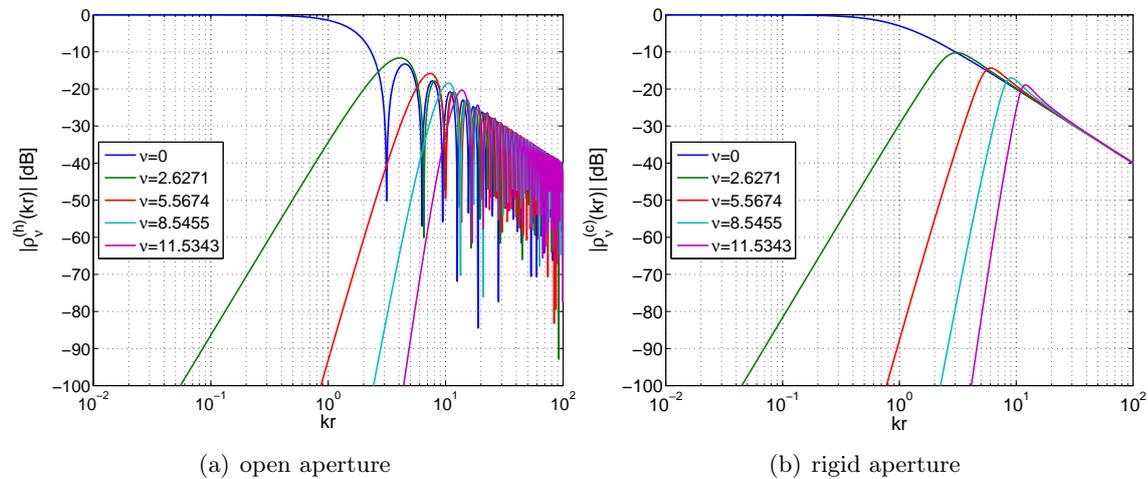


Figure 4.5: Magnitude of the modal responses $\rho_{\nu}(kr)$.

The problem gets obvious when regarding the resulting notches above $kr > \nu_q$ which coincide with the zeros of the spherical Bessel functions [Zau12]. These notches are prone to lead to numerical instabilities when inverting the modal response curves which is necessary for a regularization to calculate *acoustic holography*. Although it was theoretically shown that radially oriented cardioid transducers solve this problem [AW02, GRS02], such transducers tend to be less accurate than omnidirectional pressure microphones.

Alternatively, the filter curves can be characterized as high-pass filters for $q > 0$ with a slope of $6\nu_q$ dB/octave within the range $kr < \nu_q$ [Köf11].

4.4.2 Rigid Array Configuration

A different approach to overcome the problem of singularities is assembling a rigid diffusor with omnidirectional pressure transducers. A rigid element demands a radially vanishing sound velocity component which is proportional to the sound pressure of the incident and the reflected sound field [Köck11] meaning that

$$\frac{\partial}{\partial r} \left(p_i(kr_a, r_a \boldsymbol{\theta}_s) + p_r(kr_a, r_a \boldsymbol{\theta}_s) \right) = 0.$$

Under this assumption, eq. (3.4) can be rewritten to

$$\sum_{q=1}^Q b_q j'_{\nu_q}(kr_a) Y_{\nu_q}(\boldsymbol{\theta}_s) = - \sum_{q=1}^Q c_q h'_{\nu_q}(kr_a) Y_{\nu_q}(\boldsymbol{\theta}_s).$$

Exploiting the orthogonality of SSH, we obtain the scattered wave spectrum from the incident one

$$c_q = -b_q \frac{j'_{\nu_q}(kr_a)}{h'_{\nu_q}(kr_a)}. \quad (4.5)$$

This results in the combination of radius-dependent expansions for the incident and the scattered sound field, i.e. [Teu07, Raf05]

$$\rho_{\nu_q}^{(c)}(kr) = j_{\nu_q}(kr) - \frac{j'_{\nu_q}(kr_a)}{h'_{\nu_q}(kr_a)} h_{\nu_q}(kr)$$

with the main advantage of being well-conditioned except for very low frequencies and high orders ν (fig. 4.5(b)). The inversion to realize *acoustic holography* is no longer problematic as the radial terms for the incident field do not share common zeros with the ones for the scattered sound field [Zau12].

Inserting the wave spectrum [Zot09a]

$$b_q = -ikh_{\nu_q}(kr_s) \phi_q$$

of a spherical source strength distribution ϕ_q situated at r_s into eq. (4.5) and using the *Wronskian relation* for spherical Bessel functions [Wil99]

$$j_{\nu_q}(kr_a) h'_{\nu_q}(kr_a) - j'_{\nu_q}(kr_a) h_{\nu_q}(kr_a) = \frac{1}{i(kr_a)^2}$$

we obtain a more compact term for the spherical wave spectrum:

$$\psi_q = - \underbrace{\frac{h_{\nu_q}(kr_s)}{kr_a^2 h'_{\nu_q}(kr_a)}}_{\mathcal{G}_{\nu_q}(k, r_a, r_s)} \phi_q. \quad (4.6)$$

4.5 Holographic Filtering

In general, the term *acoustic holography* refers to the process of evaluating sound field quantities (pressure, velocity, intensity) situated at a radius r_s by recording the resulting soundfield at an aperture radius r_a and subsequent radial focusing. Holographic filters are necessary for a compensation of the frequency dependent radiation terms to obtain flattened frequency responses. This filtering is usually referred to as *spherical convolution* [Zot09a], which is similar, however not entirely compatible, to the spherical segment case.

4.5.1 Causal Holographic Filters

The *causal* version of the holographic filters $\mathcal{H}_{\nu_q}(k, r_a, r_s) = \mathcal{G}_{\nu_q}^{-1}(k, r_a, r_s)$, cf. eq. (4.6), without the linear phase term, which results due to the speed-limited sound propagation, is written as [Kö11]

$$\mathcal{H}_{\nu_q}^c(k, r_a, r_s) = -\frac{kr_a^2 h'_{\nu_q}(kr_a)}{h_{\nu_q}(kr_s)} \frac{1}{e^{ik(r_s-r_a)}}.$$

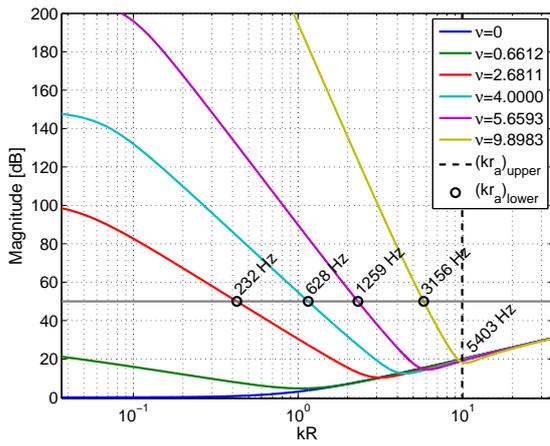
Finally, this filtering procedure yields the desired spherical source strength distribution

$$\phi_q = \mathcal{H}_{\nu_q}^c(k, r_a, r_s)\psi_q.$$

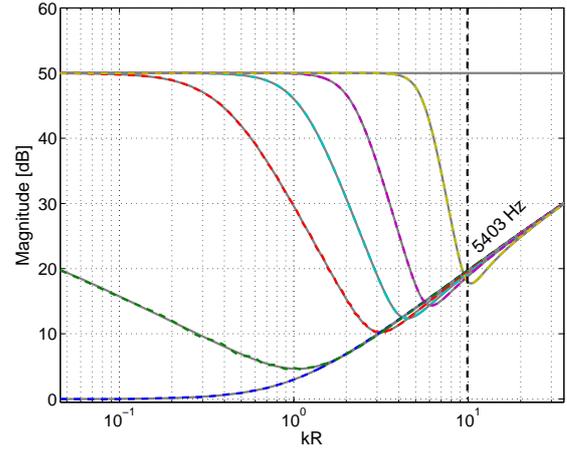
Figure 4.6(a) and 4.6(c) depicts the magnitude and phase characteristics of the analytic, causal *holographic filters* $\mathcal{H}_{\nu_q}^c(k, r_a, r_s)$.

The resulting transfer functions show upper knees at $kr_a = \nu_q$. Additionally, $\nu_q > 0$ feature lower knees which are determined by the source radius r_s and exhibit a slope of $-6\nu_q$ dB/octave (cf. magnitude of spherical Hankel functions $|h_{\nu_q}(kr)|$, fig. 2.6). Consequently, they exceed the feasible dynamic range and have to be adapted to the dynamic range of the used microphones (50 dB are considered). This restriction determines the lower bounds of the valid frequency ranges for each order ν_q (black circles), whereas the upper bound is set by the highest order ν_Q (dashed black line at $kr_a = \nu_Q$) [Kö11].

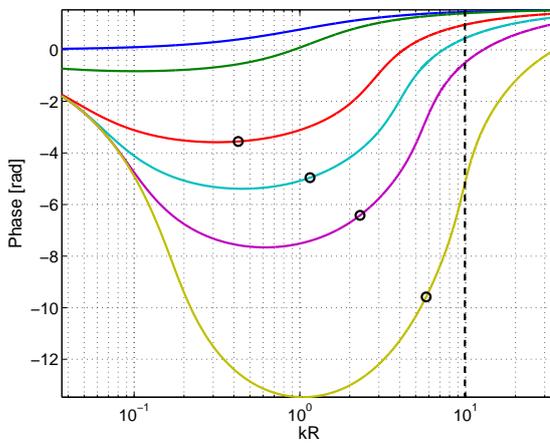
With this knowledge we can limit the array's valid frequency range reasonably with respect to the orders ν_q .



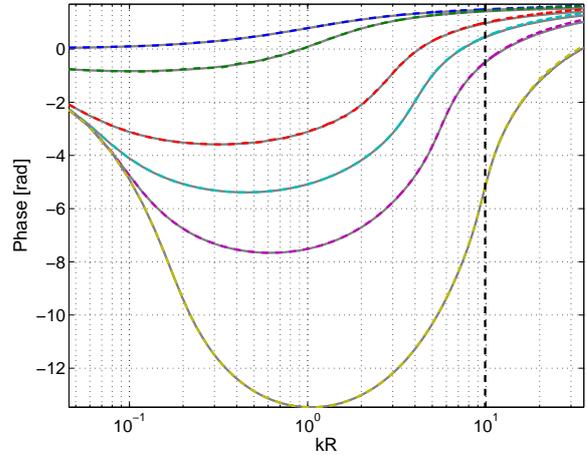
(a) magnitude responses (analytic filters, not regularized)



(b) magnitude responses (regularized FIR filters)



(c) phase responses (analytic filters, not regularized)



(d) phase responses (regularized FIR filters)

Figure 4.6: Analytic holographic filters $\mathcal{H}_{\nu_q}^c$ (not regularized) and the regularized counterparts $\hat{\mathcal{H}}_{\nu_q}^c$ realized as FIR filters ($B_{\#} = 4096$ coefficients) to compensate the frequency dependent mode behaviour, $r_a = 0.1$ m, $r_s = 5$ m, maximal dynamic range: 50 dB.

4.5.2 Digital Filter Design

IIR filter design. Kößler uses a curve fitting algorithm which delivers the *infinite impulse response* (IIR) filter coefficients to obtain an error bounded magnitude and phase response in the frequency area of interest [Köß11]. Unlike in the case of spherical arrays for which the IIR filter order is limited by $\nu + 1$, Kößler shows required IIR filter orders for the filters with non-integer ν .

FIR filter design. The preferred filter design in this work was a *finite impulse response* (FIR) design. This is realized through the *frequency sampling method* [PB87] which is obtained by sampling the analytical transfer functions³ $H_{\nu_q}^c$ at a number $\hat{N} = 2^a$ ($a \in \mathbb{N}$) of equidistantly spaced frequency bins on the interval $[0, 2\pi)$. As a result, we obtain FIR filters which are valid at the considered bins with the disadvantage of an uncontrolled behaviour between the bins, due to the intrinsic sinc-interpolation of the Fourier transform.

The following steps outline the design procedure:

1. To obtain filters with a dynamic range of $d_{yn} = 50$ dB (cf. sec. 4.5) some sort of regularization has to be done. In terms of acoustics, a soft transition ("soft knee") due to dynamic limits is preferred. Thus, we apply a regularization term on $|\mathcal{G}_{\nu_q}^c|$ while preserving the phase characteristics similar to [BPSW11]:

$$\bar{\mathcal{G}}_{\nu_q}^c = \left(|\mathcal{G}_{\nu_q}^c| + \max\{|\mathcal{G}_{\{\nu\}}^c|\} \cdot 10^{-d_{yn}/20} \right) \cdot e^{j \cdot \arg\{\mathcal{G}_{\nu_q}^c\}}. \quad (4.7)$$

2. Consequently, an inverse Fourier transform with \hat{N} bins is applied on⁴ $\bar{\mathcal{H}}_{\nu_q}^c = (\bar{\mathcal{G}}_{\nu_q}^c)^{-1}$, i.e.

$$\hat{h}_{\nu_q}(\hat{n}) = \frac{1}{\hat{N}} \sum_{\hat{k}=0}^{\hat{N}-1} \bar{\mathcal{H}}_{\nu_q}^c \left(\frac{2\pi}{\hat{N}} \hat{k} \right) e^{j2\pi \hat{k} \hat{n} / \hat{N}}. \quad (4.8)$$

3. The obtained, approximated impulse responses / FIR filter coefficients $\hat{h}_{\nu_q}(\hat{n})$ have to be shifted circularly to get causal versions.

The resulting FIR filters $\hat{\mathcal{H}}_{\nu_q}^c$ which approximate the analytical counterparts are shown in fig. 4.6(b) and 4.6(d) as dashed lines. To get reasonable results, we choose a filter length of $\hat{N}/2 = B_{\#} = 4096$ coefficients.

³ The dependency on k, r_a, r_s is omitted for better readability.

⁴ Note that the magnitude and phase spectra of $\bar{\mathcal{H}}_{\nu_q}^c$ have to be strictly conjugate symmetric to get a real-valued impulse response $\hat{h}_{\nu_q}(\hat{n})$.

4.6 Spherical Segment Signal Processing

To sum up this chapter, a concise synopsis is given by the following flowchart, which depicts the signal processing steps in the context of the acoustic array application.

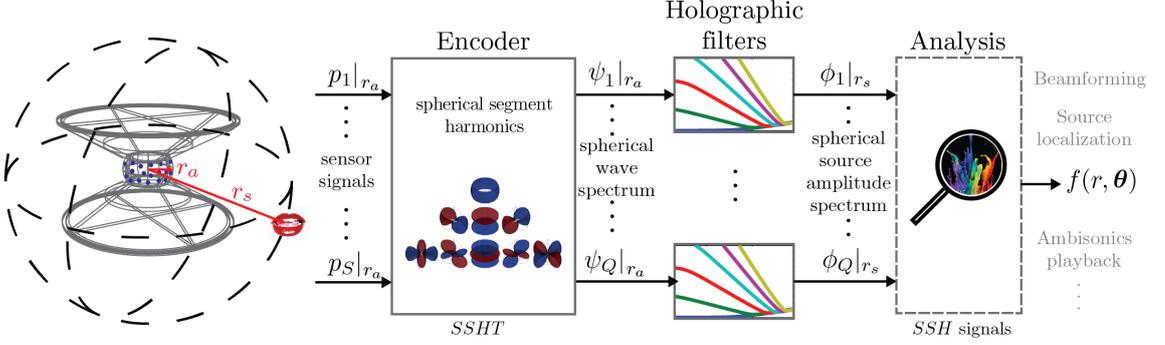


Figure 4.7: Signal processing steps to obtain a spherical source amplitude spectrum through a *spherical segment harmonic transform* (SSHT) and subsequent radial filtering.

Figure 4.7 shows the processing structure to attain *spherical segment harmonics* (SSH) signals. In a first step, the signals from S omnidirectional pressure transducers $p_s|_{r_a}$, $s = 1, 2, \dots, S$ at the sampling positions $r_a \theta_s$ on the aperture radius r_a are encoded through applying a frequency-independent *SSH transform* (cf. sec. 3.1 and sec. 4) to obtain the *spherical wave spectrum* $\psi_q|_{r_a}$, $q = 1, 2, \dots, Q$. To extrapolate the sound field to a source radius r_s we apply *holographic filters* / *radial filters* (cf. sec. 4.5) and obtain the *spherical source amplitude spectrum* $\phi_q|_{r_s}$. These signals can be decoded for a playback on an Ambisonics loudspeaker arrangement [ZPN10]. Further applications are source localization [STMK11, SH11] and beamforming [ME08, Pet04].

Chapter 5

The Prototype

5.1 Designing the Cabinet

The cabinet design of the prototype was accomplished using the software "Rhinceros 4.0", a NURBS¹-based 3-D modelling software by Robert McNeel & Associates. The graphical user interface of Rhino allows diverse manipulations methods of the 3D object model and the pre-defined nodes (sec. 4.3.1) as well as other object definitions (sec. 5.2.1) were imported through *.txt files.

5.1.1 Geometrical Properties

Spherical segment. As the aperture radius r_a directly influences the upper frequency limit (sec. 4.1.2), it is important to keep the dimensions of the spherical segment as small as possible. Whereas this would actually reduce the low frequency resolution of the array, the practical challenge lies in miniaturization. The critically limited space demands a careful design with respect to practicability and is outlined in sec. 5.2.1. Finally, a minimum radius of $r_a = 100$ mm has been determined.

Rigid cones. The influence of cones with finite length on the decomposition of a surrounding sound field into SSH is a question of current research [Pom13]. To guarantee a sufficient attenuation for sources that lie outside the angular range of interest, and because of practical limitations, we set the length of the rigid cones to $L_c = 3r_a = 300$ mm (measured from the array surface). This leads to an outer cone diameter of $D_c = 690$ mm.

The exact dimensions of the designed prototype are shown in fig. 5.1.

¹ *Non-uniform rational basis spline* (NURBS) is a mathematical model to represent curves and surfaces and is used frequently in computer aided design and engineering [Wik07].

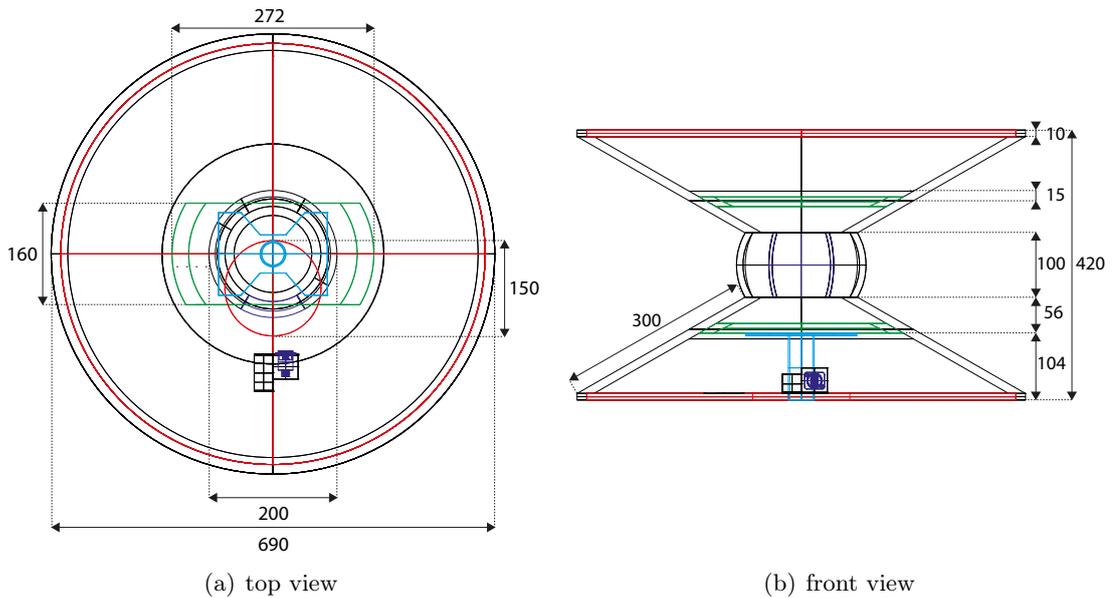


Figure 5.1: Dimensions of the designed prototype in [mm].

5.1.2 Functional Aspects

The prototype (fig. 5.2) mainly consists of two geometrical elements: the two symmetric cones and the spherical segment which is made up by six staves. To guarantee mechanical decomposability as well as stability these parts are stuck together via timber connectors and interlocked by a threaded M10 bar which goes vertically through the center of the spherical segment being screwed on two robust carrier plates. The array can be mounted on a standard speaker stand (pole diameter: 36 mm) with the help of a mounting adapter. For an application in a conference scenario it is also possible to place the prototype on a table (fig. 5.8).

To minimize the influence of disturbing noise, i.e. sound impinging from $\vartheta \in [0, 60^\circ]$ and $\vartheta \in (120^\circ, 180^\circ]$, and because of aesthetic reasons the cabinet is closed with a cover plate and a baseplate, respectively. The baseplate comprises an excentric 150 mm hole which enables the adjustment of the mounting adapter. Furthermore, the outsourced USB adapter and the barrel connector for power supply can be reached without demounting the base plate.

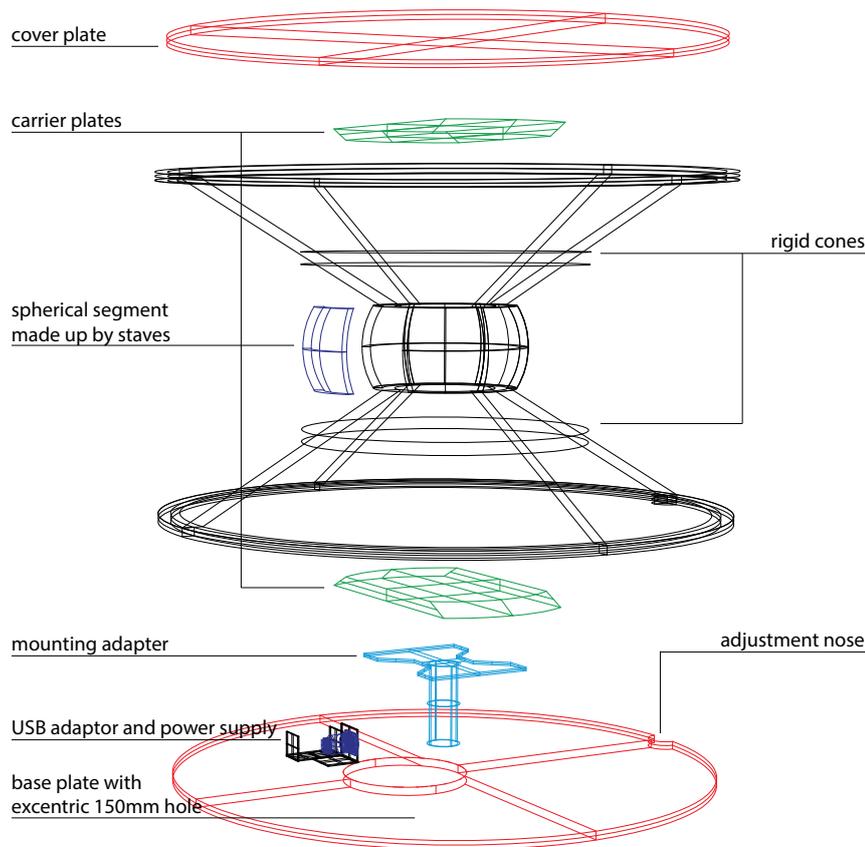


Figure 5.2: Exploded assembly drawing of the designed prototype with a description of the main elements.

5.2 The Array Hardware

The cabinet is assembled with 64 'VisiSonics' Digital Array Microphones (VDAM™)². Each VDAM™ unit contains an electret microphone, amplifies the analog input signal and converts it into a digital one at a fixed sampling rate of 44.1 kHz. The multiplexed, digitized data is sent to a serialization controller using flexible flat cables (FFC). Synchronization is realized via U.FL connectors and 1.32 mm double shielded cables (fig. 5.3). External synchronization was not possible during the experimental validation as the provided clock signal was not yet conform to word clock standards.

All modules are daisy chained, and the 4 busses, each containing 16 channels, are fed into the VDAM™ array controller that is linked to an 'Opal Kelly XEM3010' FPGA-powered integration board. Onboard, a simple USB 2.0 interface is available as well as a barrel connector for power supply, cf. fig. 5.4 [Vis07].

² VisiSonics Corporation: <http://www.visisonics.com/>

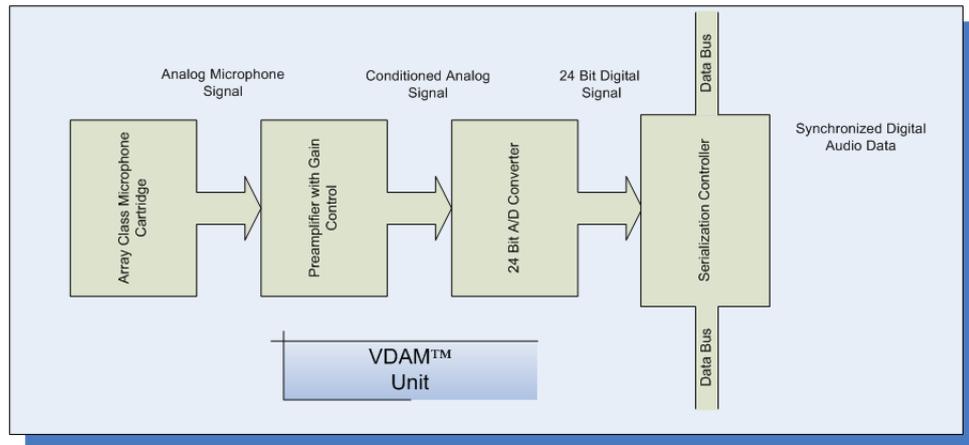


Figure 5.3: Signal chain within a VDAM™ unit depicting the microphone, the preamp, the A/D converter and the serialization controller [Vis07].

VisiSonics additionally provides ASIO³ drivers to facilitate the setup and integration in real-time recording systems (for detailed user instructions the reader is referred to app. C).

5.2.1 VDAM™ Units in the Model

To check the practical feasibility of the determined optimal microphone distribution, a simplified scale model of the array microphones has been designed (fig. 5.5). These units are sequentially build with the help of a RHINO command file which was generated beforehand in MATLAB. Consequently, each microphone model is rotated and moved to the corresponding microphone position (lst. 5.1).

Listing 5.1: Generated *.txt file in RHINO command file format to build and position a simplified array microphone (exemplarily shows the process for the 12th microphone model).

```

1  _ Polyline 0.35,0.3,0 r0.8,0.85,0 r3.07,0,0 r0,-2.3,0 r-3.07,0,0 r-0.8,0.85,0
   r0,0.3,0 0.35,0.3,0 % generate base line
2  _SelLast
3  _SetObjectName Baseline
4  _ExtrudeCrv R 0 0,0,1 0.49 % extrude base line to build platine
5  _SelLast
6  _SetObjectName Plate12
7  _SelNone
8  _SelName Baseline
9  _Delete
10 _SelName Plate12

```

³ *Audio Stream Input/Output* is a low-latency driver protocol for sound cards developed by 'Steinberg' for multichannel digital audio transfer and software/hardware communication [Wik25].

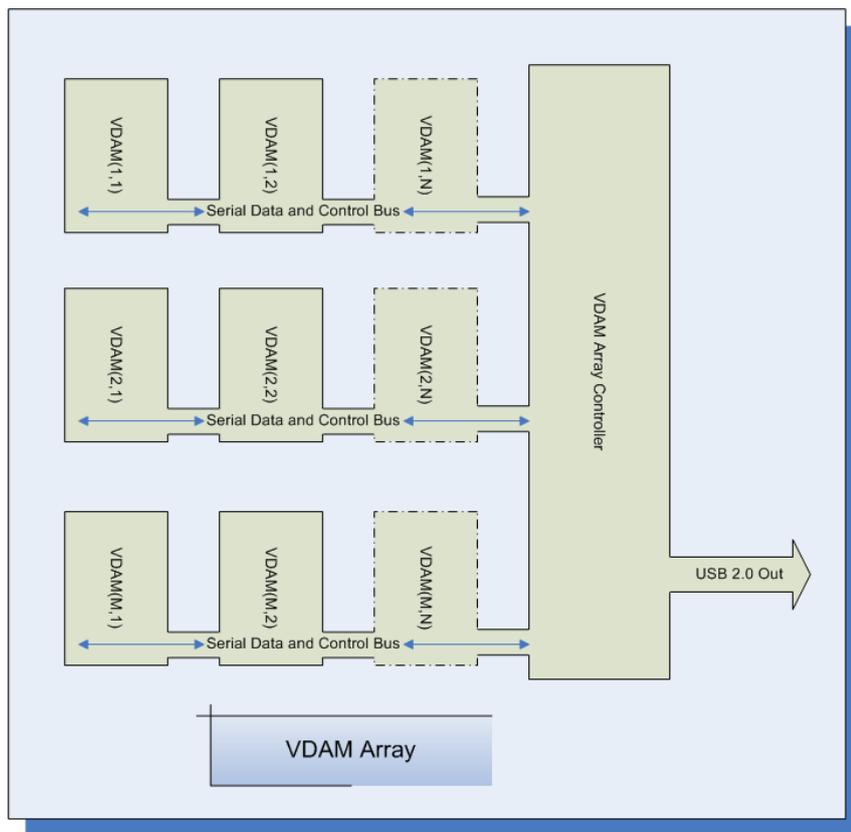


Figure 5.4: Schematic description of the VDAMTM array, $N = 16$, $M = 4$ (here, N and M stands for the number of modules per chain and the number of busses, respectively) [Vis07].

```

11 _Move 0.3,0,0 0.3,0,-0.3
12 _Cylinder R V 0.35,0,0 0.3 0.35 % generate capsule
13 _SelLast
14 _SetObjectName capsule12
15 _SelName capsule12
16 _Rotate3D 0.35,0,0 0.35,1,0 -90
17 _SelNone
18 _Box 1.15,-0.85,0.19 1.82,-0.18,0.19 0.47 % generate calibration elements
19 _SelLast
20 _SetObjectName CalElement12x
21 _SelNone
22 _SelName CalElement12x
23 _Copy 1.15,-0.85,0.19 2.232,0.2,0.19 _Enter
24 _SelNone
25 _SelLast
26 _SetObjectName CalElement12y
27 _SelNone
28 _SelName CalElement12x
29 _SelName CalElement12y
30 _SelName Plate12

```

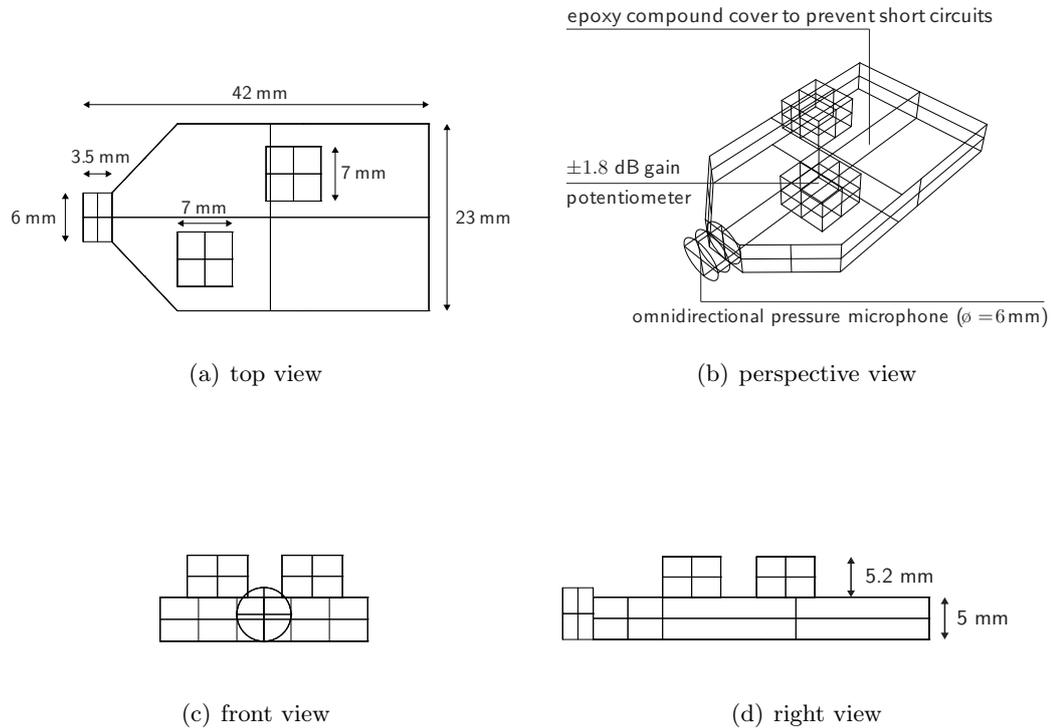


Figure 5.5: Simplified scale model of the assembled VDAM™ units.

```

31 _SelName capsule12
32 _Group
33 _SelLast
34 _Rotate3D 0 0,1,0 170.4059 % rotate microphone vertically (theta [deg])
35 _Rotate3D 0 0,0,1 336.0938 % rotate microphone horizontally (phi [deg])
36 _Move 0 9.0142,-3.9957,1.6667 % move microphone to corresponding position
37 _Ungroup
38 _SelNone

```

As we are dealing with a relative small radius of the spherical segment ($r_a = 100$ mm) which contains the 64 microphones, the space is critically limited and an intersection of array microphones has to be avoided, naturally. This issue can be solved by carefully rotating the elements radially in the model. For this purpose, another command script has been written which allows the user to rotate each element by a desired amount of degrees and crosscheck the result visually (lst. 5.2).

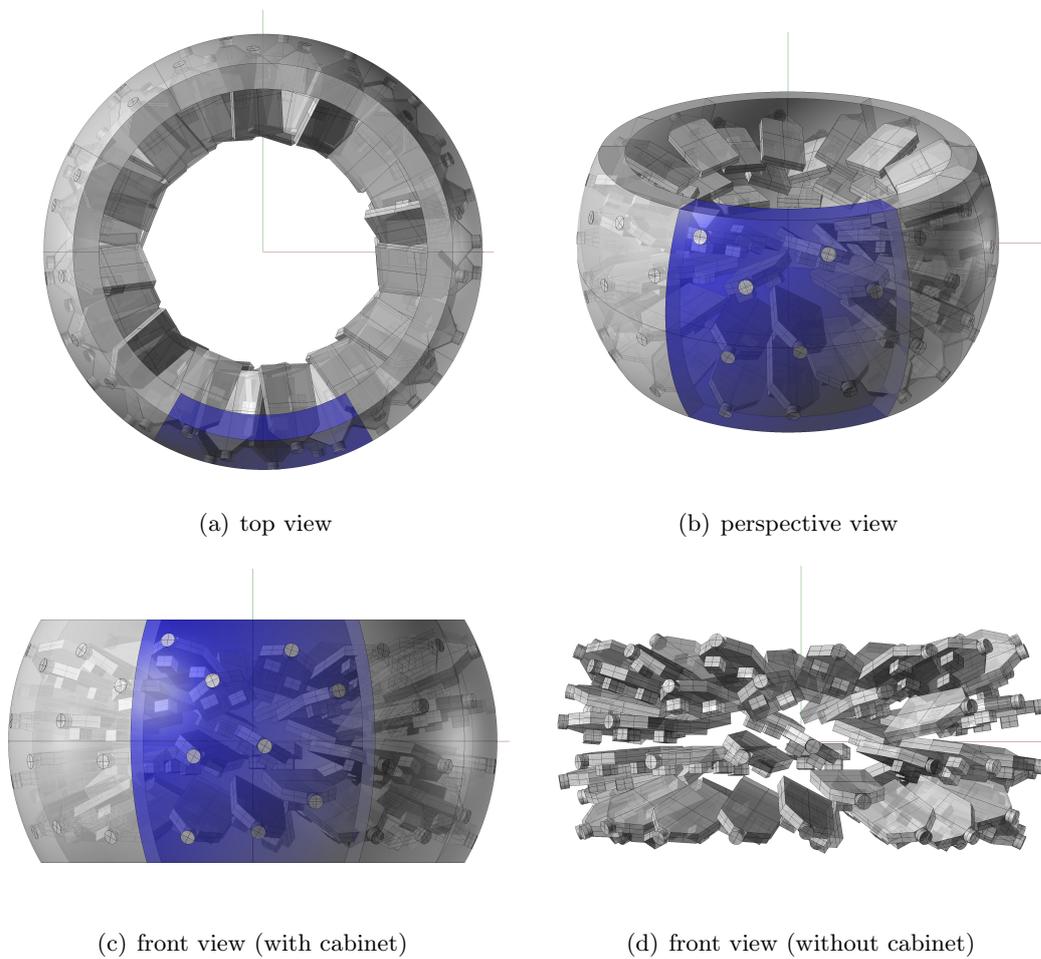


Figure 5.6: Spherical segment of the designed prototype assembled with 64 simplified scale models of the array microphones.

Listing 5.2: Generated *.txt file in RHINO command file format for a user-defined radial rotation of microphones (exemplarily shows the process for the 12th microphone model).

```

1  _Rotate3D
2  _SelName CalElement12x
3  _SelName CalElement12y
4  _SelName Plate12
5  _SelName capsule12
6  _Enter
7  9.0142,-3.9957,1.6667 6.31,-2.797,1.1667 % define axis of rotation
8  _Pause % enter the desired angle of rotation [deg]

```

The final microphone distribution with rotated elements is shown in fig. 5.6.

5.3 Manufactured Cabinet

The cabinet of the prototype has been manufactured using a *computerized numeric control* (CNC) machine by Hermann Deutscher⁴, a carpenter located in Hart/Pischelsdorf near Graz. At this point, I want to thank him for his excellent joinery. To get an impression of the different steps during the manufacturing process the reader is referred to fig. 5.7. The final workpiece deviates from the designed one with respect to the carrier plates and in some minor details.

Used materials. As the advantages of *medium density fibreboard* (MDF) lie in a homogeneous resonance behaviour in and off the fiber-direction and in cheap acquisition costs it was considered to be the ideal material [Ker10]. To avoid frazzled boreholes the carpenter made the suggestion to produce the spherical segment from polyurethane plastic plates. Finally, the individual components were glued together and finished in black.

5.3.1 Assembling the Cabinet

Before mounting the array microphones the ribbon cables were glued to the FFC connector with the aid of two-component epoxy compound for long-term stability. This step—even though it complicates maintenance—was necessary to guarantee permanent plug contact as a rapid heat increase can be diagnosed in case of partial loosening which potentially causes melting of the involved ribbon cable and/or the VDAM[™] element.

⁴ *deutscher – 2D and 3D CNC for wood and plastic material*: <http://www.deutscher.at/>



(a) milling the cone's bottom



(b) milling the cone's top



(c) raw unpainted cone



(d) drilling one of the six staves



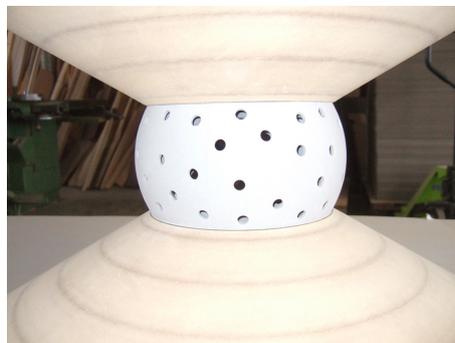
(e) finished staves (back side)



(f) finished staves (front side)



(g) unvarnished cabinet (w/o cover plate)



(h) unvarnished spherical segment

Figure 5.7: Visual documentation of the manufacturing process.

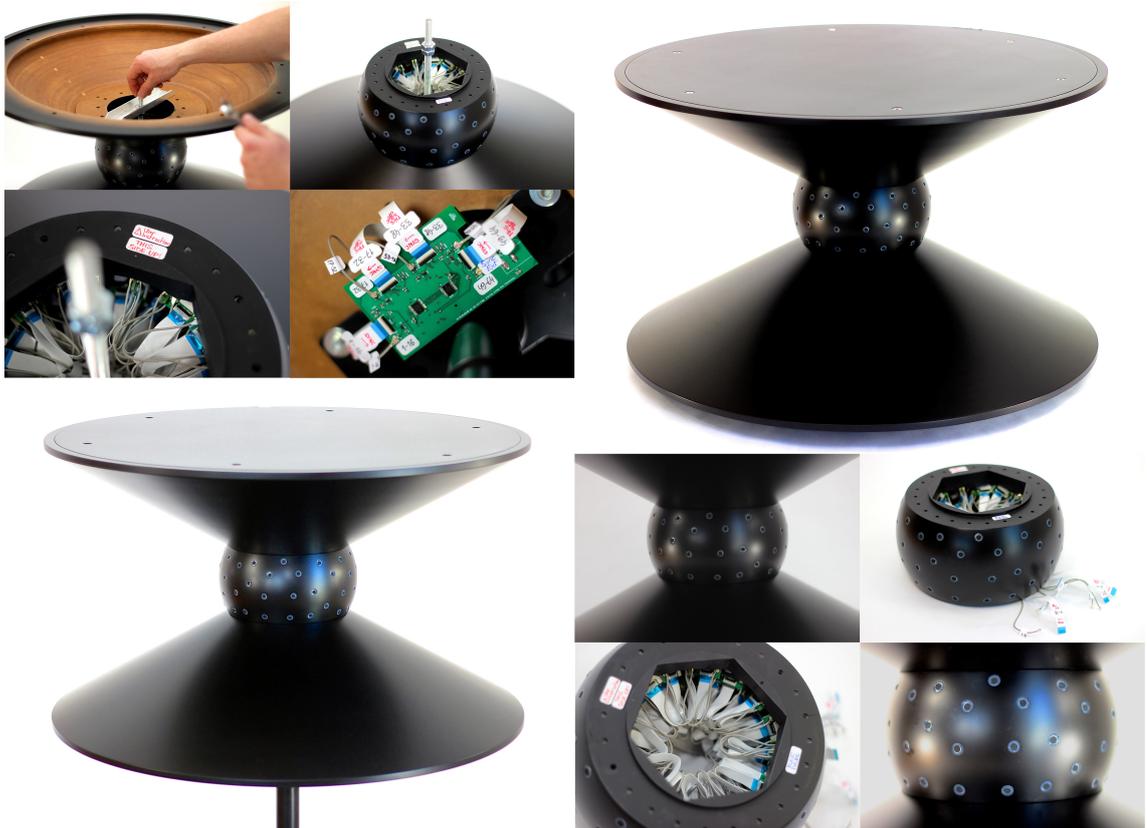


Figure 5.8: Finished prototype - the large pictures depict the array mounted on a speaker stand and placed on a table, respectively. The smaller ones show some construction details and the serialization controller with the four signal busses and the synchronization cables.

Chapter 6

Acoustic Measurements

This chapter describes the procedure of measuring the prototype in the acoustic domain. In a first step, the microphones are calibrated. Subsequently, preliminary measurements were carried out to examine the precision of the different gain stages. In the second part of this chapter, the directional impulse responses of the microphones are measured. The damping effects of the rigid cones are analyzed by means of transfer functions. Finally, the array microphone directivity patterns are visualized.

In chapter 7, the acoustic measurement data is applied in the analytic beamformer model.

6.1 Calibration of the Array Microphones

The ASIO drivers provided by VisiSonics offer three different *.dll files to set the gain of the array microphones (cf. app. C.2). Table 6.1 lists the three gain settings—by default, the file `vsAsioDriver_g1.dll` is used.

Table 6.1: ASIO gain settings.

*.dll file	gain
<code>vsAsioDriver_g1.dll</code>	-20 dB
<code>vsAsioDriver_g10.dll</code>	0 dB
<code>vsAsioDriver_g100.dll</code>	+20 dB

After experiments with gain settings, the 0 dB operating level was considered to be the most suitable level for practical music or speech recordings as it administers a compromise between sensitivity and headroom. This was the main reason why the calibration was done at this gain stage.

As reference source, a calibrator pistonphone [Kja] is used and sequentially attached onto the array microphone capsules. Level differences were balanced by a gain potentiometer with a setting range of $\approx \pm 1.8$ dB (cf. fig. 5.5, the potentiometer which is located nearer to the capsule) inspecting the received level on the DAW (precision: ± 0.05 dB).

6.2 Preliminary Measurements

6.2.1 Measurement Setup A

To get a rough idea of the individual array microphone impulse responses (IR) and the related transfer functions, the microphones are measured mounted in foam (fig. 6.1).

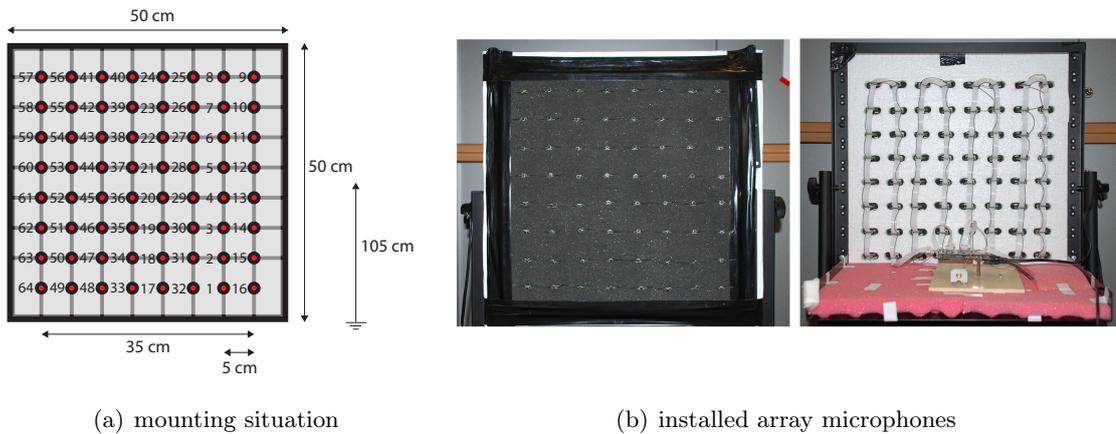


Figure 6.1: Array microphones mounted in foam.

Measurement chain. The signal chain of the measurement setup is shown in fig. (6.2).

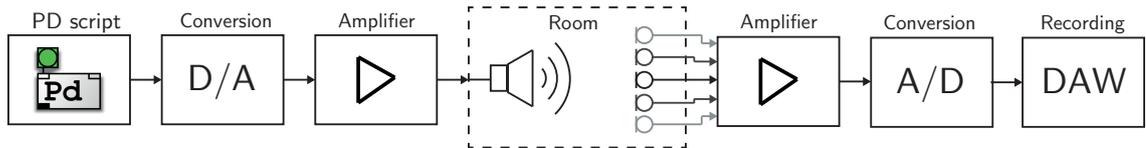


Figure 6.2: Flowchart to measure the sweep responses of the array microphones.

A PD¹-patch is used to send an exponential sweep $\tilde{s}(t)$ which is converted by a D/A converter [RME13b] and played back by an active loudspeaker [Gen05]. After analog-to-digital conversion the sweep responses are recorded on a DAW.

As we are dealing with a relative analysis of impulse responses, an equalization of the transducer frequency response can be omitted; A/D and D/A conversion as well as the amplifier are assumed to be ideal.

¹ PD stands for *Pure Data*, an open-source visual programming language for real-time applications in audio, video and graphical processing created by Miller Puckette in the 1990s (<http://puredata.info>).

Excitation signal. Mathematically, an exponential sweep can be described as [DF00]

$$\tilde{s}(t) = \sin[f(t)] = \sin\left(\zeta \cdot (e^{t/\xi} - 1)\right)$$

in time domain with the two variables

$$\zeta = \frac{\omega_1 \cdot T}{\ln\left(\frac{\omega_2}{\omega_1}\right)} \quad \text{and} \quad \xi = \frac{T}{\ln\left(\frac{\omega_2}{\omega_1}\right)}$$

where ω_1 and ω_2 represent the start and end frequencies, respectively, and T stands for the total duration of the sweep (fig. 6.3). The instantaneous frequency $\omega(t)$ can be calculated as the derivative of $f(t)$ and yields

$$\frac{d[f(t)]}{dt} = \frac{\zeta}{\xi} \cdot e^{\frac{t}{\xi}}.$$

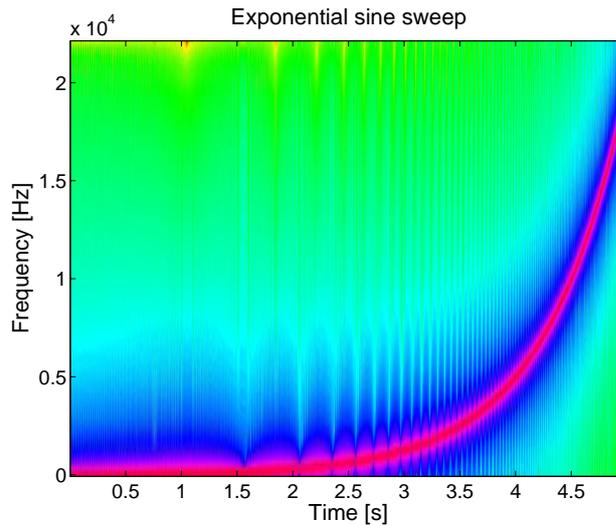


Figure 6.3: Spectrogram of an exponential sine sweep which covers the audible frequency range, $T = 5$ s, $f_s = 44.1$ kHz.

Measurement Room. The measurements were taken in the office of diploma and Ph.D. students where a measurement room, which is clearly not anechoic, is set up (fig. 6.4). Nevertheless, by recording the decaying impulse responses long enough and cutting away the acoustic reflections, a fair impression of the frequency responses is accessible. The acoustic center of the speaker [Gen05] points at the array center (height: 1.05 m, fig. 6.1). During the measurements the room temperature was constantly at 21°C ($c \approx 344$ m/s).

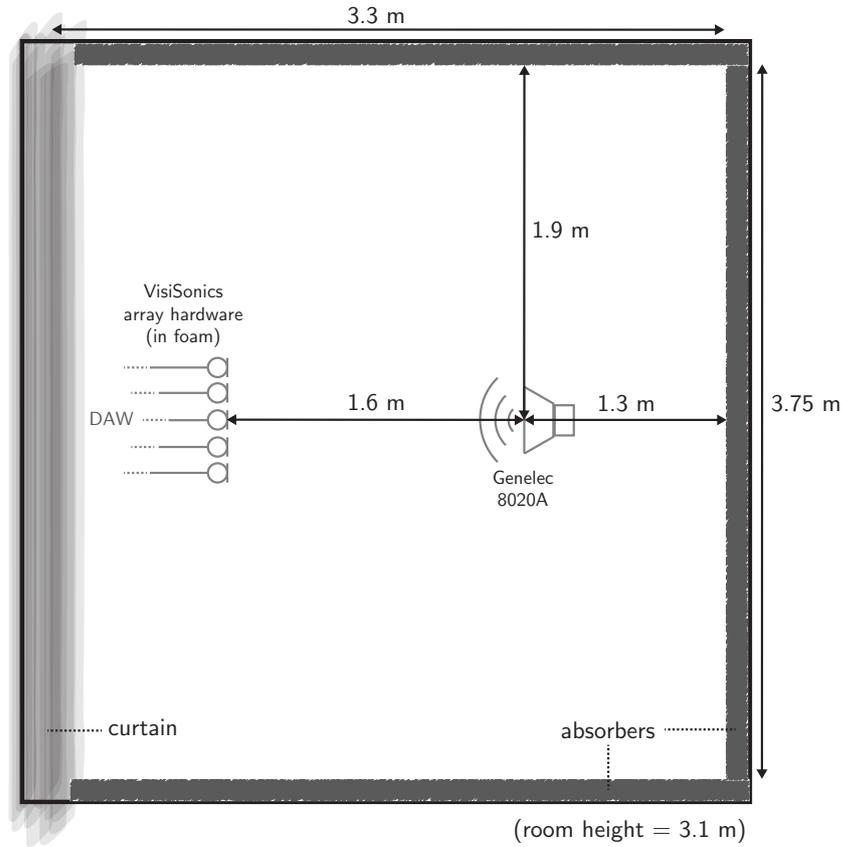


Figure 6.4: Setup and arrangement for the preliminary measurements of the array microphones.

6.2.2 Obtaining the Impulse Responses

For each gain setting (cf. tab. 6.1) the sweep responses of all array microphones are recorded ($f_s = 44.1$ kHz, 24 Bit); the excitation level is set in a way that the microphone levels at the +20 dB stage do not exceed 0 dB FS.

The impulse response $h_s[n]$ of the s -th array microphone ($s = 1, 2, \dots, 64$) can be calculated in the frequency domain through a simple division of spectra followed by an inverse DFT transform, i.e.

$$h_s[n] = IDFT \left[\frac{DFT(y_s[n])}{DFT(\tilde{s}[n])} \right] \quad (6.1)$$

where $y_s[n]$ is the measured s -th sweep response. To avoid artefacts due to cyclic deconvolution, the DFT length is chosen to be twice long as the longer of the two signals in eq. (6.1) [Zau12].

6.2.3 Measured Impulse Responses

In order to exclude reflections of the surrounding walls or reflecting objects the measured impulse responses were windowed (early reflections from the floor are expected after ≈ 3 ms wrt. the direct sound) with the first and the second half of a Hann window (cf. fig. 6.5).

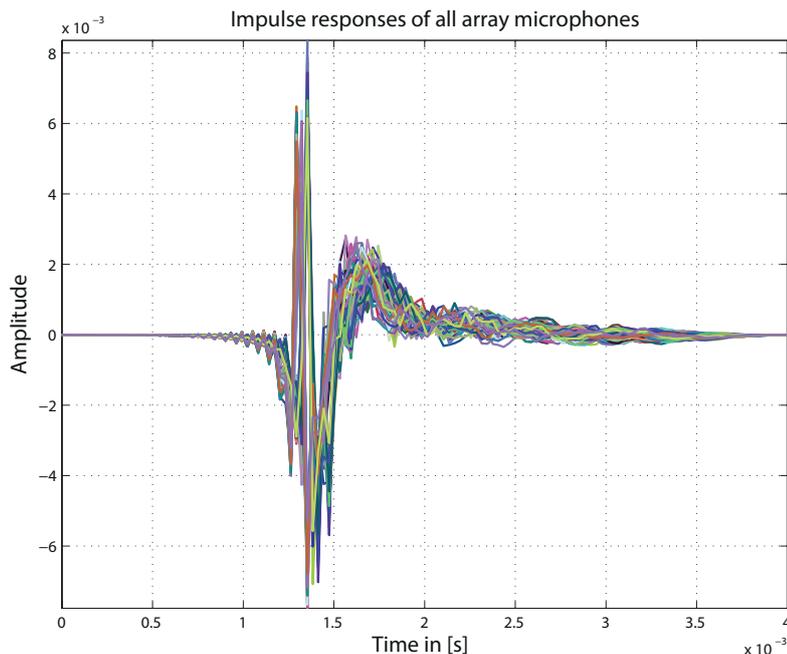


Figure 6.5: Windowed impulse responses of all array microphones ($f_s = 44.1$ kHz, *_g10.d11).

Figure 6.6 exemplarily shows the corresponding frequency domain transfer functions for array microphone 20 (which lies in the center of the foam) using different gains. The complete measurement results are presented in app. B.1. Aside from an attenuated response in the lower frequency range there are ripples from about 2 kHz upwards, which are mainly caused by diffraction because of the mounting situation.

6.2.4 Level Deviations Among Gain Stages

Local array microphone deviations. In practical applications, the user might set an application-dependent sensitivity level, thus a comparison of array microphone responses Λ_{s,i° [dB] among different gain stages has to be drawn. Thereto, i° octave band-related deviations ($i^\circ = 1, 2, \dots, 7$) of $S = 64$ array microphones ($s = 1, 2, \dots, S$) are analyzed

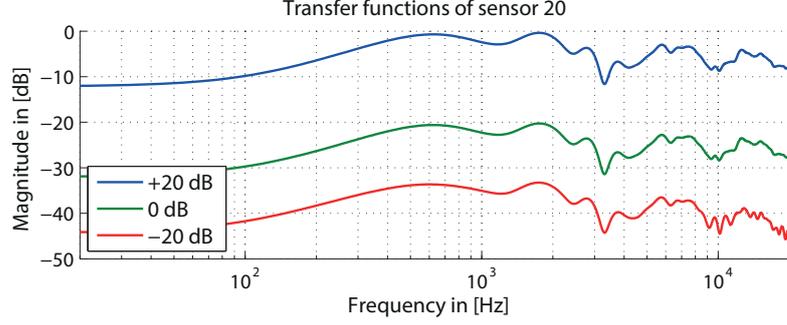


Figure 6.6: Magnitude responses of array microphone 20 (middle of foam) using different gain levels, normalized to the maximum of all microphone frequency responses in +20 dB stage (*_g100.d11), $N_{\text{FFT}} = 4096$ points.

after subtracting the mean of attenuation per octave band (eq. (6.2), eq. (6.3), tab. 6.2). The deviations are defined in the following way:

$$\begin{aligned}\Delta\Lambda_{1,s,\iota^\circ} &= \Lambda_{-20\text{dB},s,\iota^\circ} - \Lambda_{0\text{dB},s,\iota^\circ} - \text{mean}_{1,\iota^\circ}, \\ \Delta\Lambda_{2,s,\iota^\circ} &= \Lambda_{20\text{dB},s,\iota^\circ} - \Lambda_{0\text{dB},s,\iota^\circ} - \text{mean}_{2,\iota^\circ}\end{aligned}$$

with the octave band-related means

$$\text{mean}_{1,\iota^\circ} = \frac{1}{S} \sum_{s=1}^S (\Lambda_{-20\text{dB},s,\iota^\circ} - \Lambda_{0\text{dB},s,\iota^\circ}), \quad (6.2)$$

$$\text{mean}_{2,\iota^\circ} = \frac{1}{S} \sum_{s=1}^S (\Lambda_{20\text{dB},s,\iota^\circ} - \Lambda_{0\text{dB},s,\iota^\circ}). \quad (6.3)$$

The results are plotted in fig. 6.7.

Table 6.2: Octave band-related means and global means of array microphone level attenuations [dB].

	[Hz]	125	250	500	1k	2k	4k	8k	global
mean _{1,ι°} [dB]		-11.5	-11.7	-12.5	-12.5	-12.7	-12.9	-14.0	
mean _{2,ι°} [dB]		19.9	19.9	19.9	19.9	19.9	19.8	19.7	
mean _{3,1:4} [dB]		→	→	→	→				-12.1
mean _{4,1:4} [dB]		→	→	→	→				19.9

Global array microphone deviations. For a more practical result, the evaluation is repeated for an expected global attenuation in [dB] in the whole frequency range. To be fair, the global means are only calculated using values of the first four octave bands (125 Hz-1 kHz) as the values for higher octave bands are not trustworthy enough due to the mounting situation.

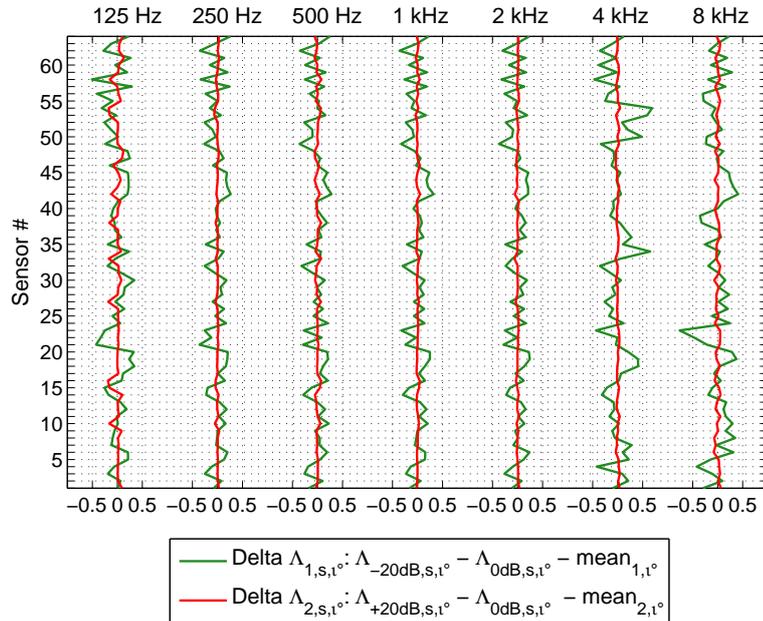


Figure 6.7: Zero mean array microphone deviations among different level stages per octave band when subtracting the local means in [dB].

Obviously, the absolute value of the level difference between the 0 dB and -20 dB stage is not 20 dB (fig. 6.6). Anyway, fig. 6.8 shows the array microphone deviations for the 'expected' -12 dB ($\text{mean}_{3,1:4} = -12.1$ dB) and 20 dB ($\text{mean}_{4,1:4} = 19.9$ dB) gains, respectively, among the gain stages.

6.2.5 Conclusions

Based on these results, we can conclude that the two stages with the higher gain levels (*_g100.d11 and *_g10.d11) merely have minor array microphone-related level deviations over all octave bands. So these two stages can be confidently used in practice with respect to errors due to gain mismatches. Reasons for the rather defective behaviour of the second deltas ($\Lambda_{1,s,t}^{\circ}, \Lambda_{3,s,t}^{\circ}$) can be most likely found in a small SNR when recording at the least sensitive gain stage (*_g1.d11) and the resulting influence of noise which leads to manipulated transfer functions.

Consequences of gain mismatches. Let us assume a gain mismatch factor of $g_{\varepsilon} = 1.01$ (≈ 0.09 dB) between two array microphones on the spherical segment. The deviation error for this scenario is $\varepsilon = 0.01$ (≈ -40 dB) so an acoustic pressure gradient between two array microphones in the order of this error cannot be detected correctly. This fact has a significant influence on the maximal dynamic of the holographic filters (sec. 4.5.1). If we equalize the sound field modes with a gain of 40 dB the resulting spherical source amplitude

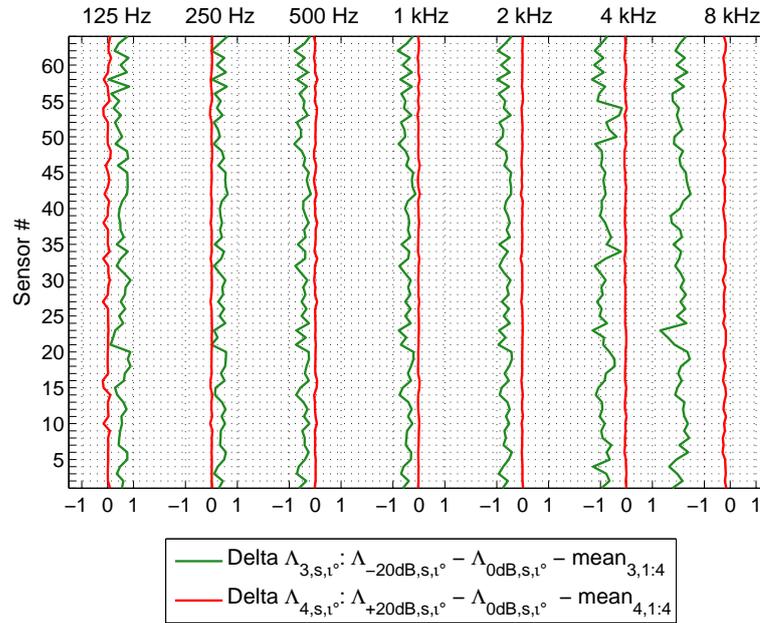


Figure 6.8: Zero-mean array microphone deviations among different level stages per octave band when subtracting the global means in [dB]

distribution, which is employed as beamformer input signal, will not be correct anymore as the resulting error is of the same magnitude as the signal.

6.3 Measuring the Prototype Characteristics

To assess the quality of the realized prototype the directional spatial impulse responses were measured and applied on the analytic beamforming algorithm (sec. 7).

6.3.1 Measurement Setup B

Measurement environment. The measurements were taken in the IEM CUBE² [ZRS03]. A loudspeaker semicircle (radius $r_s = 1.4$ m) with 16 loudspeakers which are displaced in vertical steps of $\Delta\vartheta = 11.25^\circ$ (initial offset: 5.625°) surrounds the centered prototype as shown in fig. 6.9. For a whole measurement run, a turntable rotates the array in 36 equidistant azimuthal steps ($\Delta\varphi = 10^\circ$) to obtain a spherical grid layout with $V = 576$ sampling points. The room temperature during the measurements was constantly at 22°C ($c \approx 345$ m/s).

² CUBE = ComputerUnterstützte BeschallungsEinheit (engl.: computer-aided sound system).

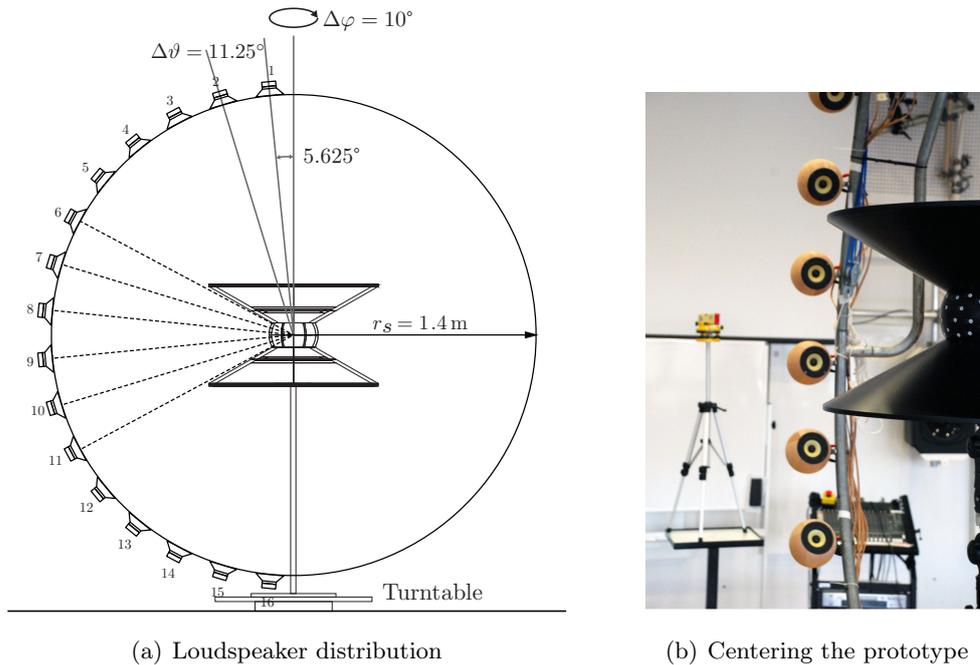


Figure 6.9: Grid layout to measure spatial impulse responses.

Measurement chain. Figure 6.10 depicts the whole measurement chain. Two PD instances with different drivers are set up (1st instance: 'VisiSonics' ASIO, 2nd instance: 'RME Hammerfall II Digiface' + turntable control patch). It is possible to send the playback trigger signal by packing an OSC³ message and transmitting it via UDP⁴ to the second instance. There, after unpacking the OSC signal, the digital USB interface [RME13a] is used to send the excitation signals to a D/A converter [RME13b] via ADAT. The analog signals are amplified ('Apart PA 4060', 16 x 60 W amplifiers) and played back by the 16 loudspeakers. Finally, the signals of the array microphones are amplified and converted on board and recorded by the first PD instance. The recordings were done at a sampling rate of $f_s = 44.1$ kHz and a resolution of 24 Bit.

Reference measurements. After matching the electric outputs of the loudspeaker amplifiers using a multimeter, reference measurements were taken in which a reference microphone is placed in the center of the surrounding loudspeakers instead of the prototype. The reference hardware (omnidirectional pressure transducer: 'Earthworks M55', preamp: 'Earthworks LAB 1') was initially calibrated with the help of a pistonphone [Kja] before recording one measurement cycle. The obtained signals are used in the measurement evaluation not only to gain-correct and equalize the impulse responses of the array micro-

³ *Open Sound Control* is a message-based communication protocol.

⁴ *User Datagram Protocol* is a wireless network protocol.

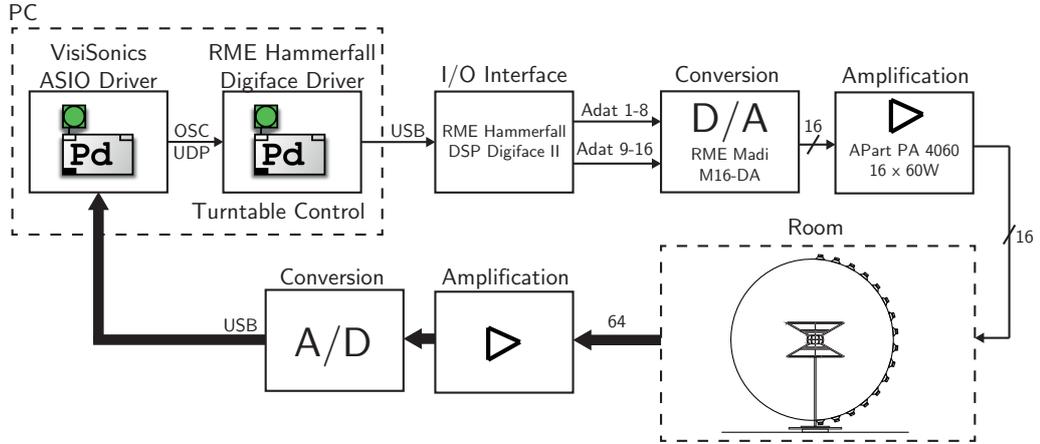


Figure 6.10: Setup to measure the prototype's characteristics.

phones but also to compensate for time discrepancies due to deviations of the loudspeaker semicircle from a perfect circle (cf. sec. 6.3.3).

Method. We used the exponentially-swept-sine technique by Farina [DF00] with a sweep length of 1.5 s (sweep fire rate: 1 s). For a more detailed description of the excitation signal the reader is referred to sec. 6.2.1.

Synchronization of the measurement cycles. As the array hardware does not provide a clock signal to synchronize external devices, a peak finder determined the start value of each measurement cycle (per azimuthal step) to keep the phase error small.

6.3.2 Overall System Description

The measurement setup can be described as a *multiple-input-multiple-output* (MIMO) system with $V = 576$ sources and $S = 64$ array microphones (fig. 6.11). The microphones receive signals which can be calculated through convoluting the loudspeaker signals $u_v(t)$ with the appendant transmission paths $w_{v,s}(t)$, for the v -th loudspeaker and the s -th array microphone, respectively.

In the frequency domain, this can be simply expressed by a multiplication, i.e.

$$\mathbf{x}(\omega) = \mathbf{W}(\omega) \cdot \mathbf{u}(\omega) \quad (6.4)$$

where $\mathbf{W}(\omega)$ is the $S \times V$ system matrix which contains all Fourier-transformed impulse responses, $\mathbf{x}(\omega)$ and $\mathbf{u}(\omega)$ are the array microphone signals and the loudspeaker signals, respectively. In the following, the frequency dependency will be omitted for better readability.

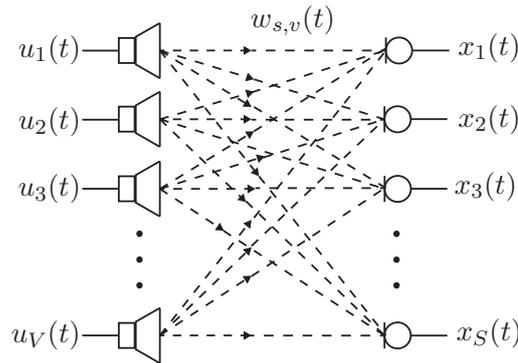
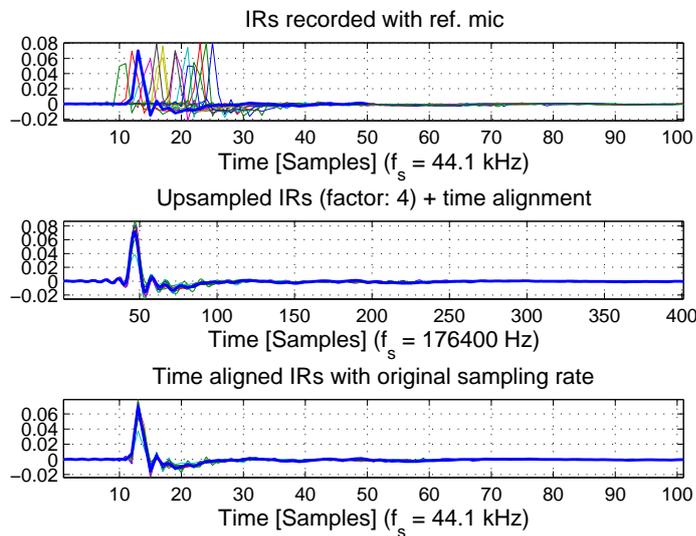


Figure 6.11: MIMO system description.

6.3.3 Equalization of the Measurement Loudspeaker Array

Reference data. To minimize phase errors due to geometric deviations of the loudspeaker semicircle from a perfect circle, we use the measured reference impulse responses $h_{ref,v}$, upsampled by a factor $\hat{U} = 4$, and calculate the resulting delays wrt. a reference loudspeaker (e.g. loudspeaker 8). Subsequently, these delays are used for time compensation of the measured array microphone impulse responses (fig. 6.12).

Figure 6.12: Time alignment of the reference impulse responses through upsampling by a factor $\hat{U} = 4$ (wrt. the 8th loudspeaker).

Additionally, the transfer functions between all V loudspeakers and the reference microphone are calculated. For gain-correction, the 1 kHz frequency bin of the reference loudspeaker is used to match the transducer gains. If we assume the conversion and amplification to be ideal we can delimit the influence of the loudspeaker characteristics through deter-

mined equalization filters. This measurement loudspeaker equalization is implemented as minimum-phase filters (fig. 6.13) using the real cepstrum [PL06, T.N01].

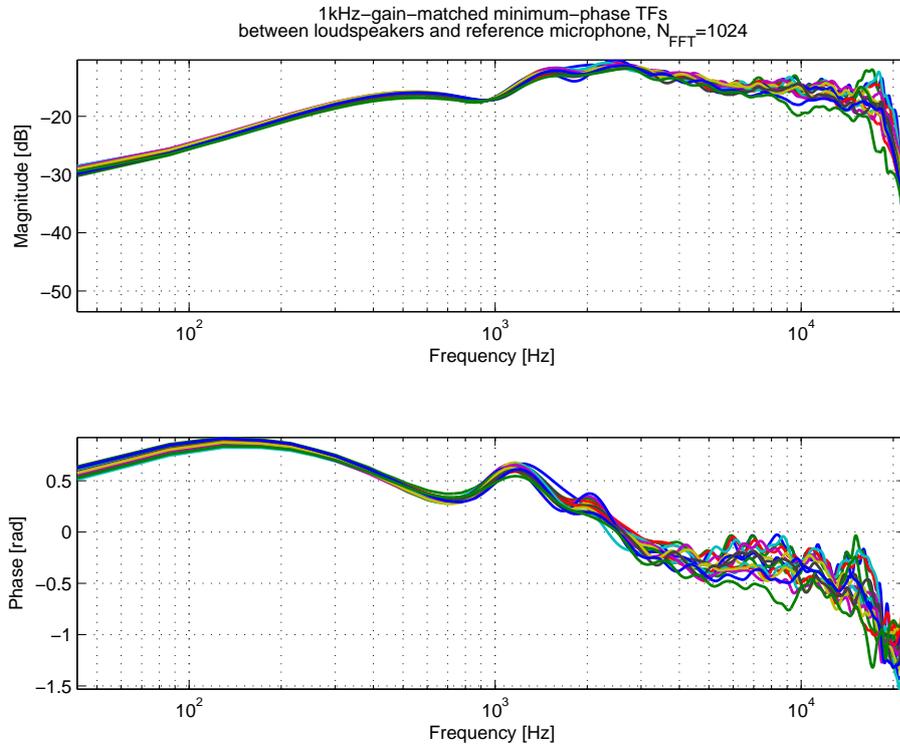


Figure 6.13: Minimum-phase equalization filters to compensate the transducer characteristics, $N_{\text{FFT}} = 1024$.

6.3.4 Spatial Impulse Responses of the Array Prototype

In a first step, the loudspeaker gain factors as well as the time delays are applied on the deconvolved (cf. eq. (6.1)) directional impulse responses of the array microphones. As the measurements were done in a non-anechoic measurement room, we apply a window of 128 samples (≈ 3 ms @ 44.1 kHz) so early reflections, with a difference of 1 m compared to the direct path, are suppressed (reflections and resulting distortions by the rigid cones are preserved). Particularly the loudspeakers in the lower half of the measurement hemisphere demand this short window size as they are placed closely to the reflecting floor (fig. 6.9).

To minimize Gibbs' phenomenon, the cut and equalized impulse responses are faded in and out using the first and second half of a Hann window.

Results. Figure 6.14 exemplarily shows impulse responses of all 64 array microphones for a particular source direction in time and frequency domain. To examine the effect of a varying source direction, the magnitude and phase spectra of two scenarios, either with

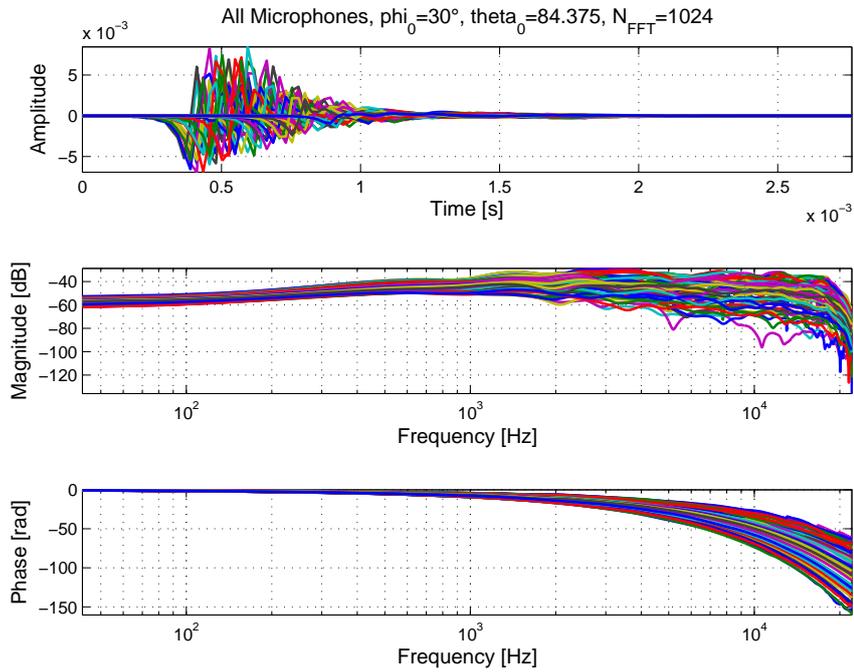


Figure 6.14: Impulse responses of all array microphones for a source direction of $\varphi_0 = 30^\circ$, $\vartheta_0 = 84.375^\circ$ in time and frequency domain ($N_{\text{FFT}} = 1024$).

fixed angle φ_0 or ϑ_0 , are shown in fig. 6.15 and fig. 6.16, respectively.

In the first scenario, we observe notches above $f \approx 1$ kHz in the magnitude spectra mainly due to reflections of the rigid cones which result in cancellations. When varying the source direction φ_0 (second scenario) the damping effect of the spherical segment becomes manifest in attenuated magnitude responses in the upper frequency range.

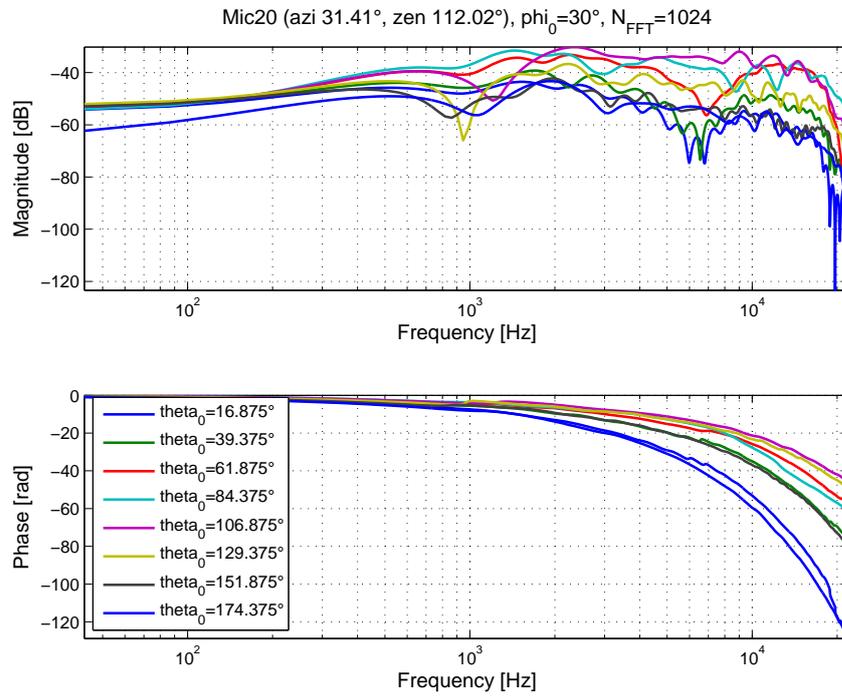


Figure 6.15: Frequency spectra of array microphone 20 for a varying source direction ϑ_0 , $\varphi_0 = 30^\circ$ ($N_{\text{FFT}} = 1024$).

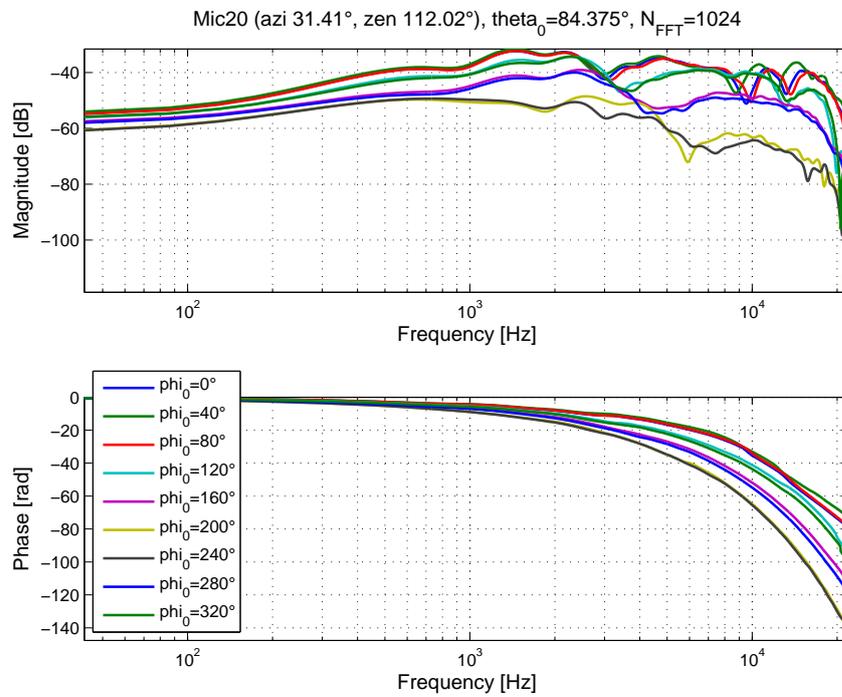


Figure 6.16: Frequency spectra of array microphone 20 for a varying source direction φ_0 , $\vartheta_0 = 84.375^\circ$ ($N_{\text{FFT}} = 1024$).

6.4 Array Microphone Directivity Patterns

To get a more general idea of the frequency-dependent array microphone characteristics, one row of the transfer matrix \mathbf{W} in eq. (6.4) can be selected and plotted which results in a directivity pattern of one array microphone.

Interpolated directivity patterns. With the help of the *spherical harmonics*⁵ (SH) *interpolation* the array microphone directivity patterns are evaluated at a denser grid $\{\boldsymbol{\theta}_k\}_+$ (fig. 6.17(a)). For the s -th directivity pattern we write⁶

$$\mathbf{d}_s(\{\boldsymbol{\theta}_k\}_+) = \mathring{\mathbf{Y}}_{\mathring{N}}^T(\{\boldsymbol{\theta}_k\}_+) \mathring{\boldsymbol{\psi}}_{\mathring{N}} = \mathring{\mathbf{Y}}_{\mathring{N}}^T(\{\boldsymbol{\theta}_k\}_+) \mathring{\mathbf{Y}}_{\mathring{N}}^\dagger(\{\boldsymbol{\theta}_v\}_{1s}) \mathbf{w}, \quad (6.5)$$

$$\mathring{\mathbf{Y}}_{\mathring{N}}(\{\boldsymbol{\theta}_v\}_{1s}) = \begin{pmatrix} \mathring{\mathbf{y}}_{\mathring{N}}^T(\boldsymbol{\theta}_1) \\ \mathring{\mathbf{y}}_{\mathring{N}}^T(\boldsymbol{\theta}_2) \\ \vdots \\ \mathring{\mathbf{y}}_{\mathring{N}}^T(\boldsymbol{\theta}_V) \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} w_{s,1} \\ w_{s,2} \\ \vdots \\ w_{s,V} \end{pmatrix}, \quad (6.6)$$

$$\mathring{\mathbf{y}}_{\mathring{N}}(\boldsymbol{\theta}_v) = \left(\mathring{Y}_0^0(\boldsymbol{\theta}_v), \mathring{Y}_{-1}^1(\boldsymbol{\theta}_v) \dots \underbrace{\mathring{Y}_{-l}^l(\boldsymbol{\theta}_v) \dots \mathring{Y}_l^l(\boldsymbol{\theta}_v)}_{2l+1} \dots \mathring{Y}_{\mathring{N}}^{\mathring{N}}(\boldsymbol{\theta}_v) \right)^T.$$

The maximal order \mathring{N} of $\mathring{\mathbf{Y}}_{\mathring{N}}(\{\boldsymbol{\theta}_v\}_{1s})$ in eq. (6.5) is determined by analyzing its condition number as the pseudoinverse has to be calculated. For a small error, a weighted pseudoinverse $\kappa(\mathring{\mathbf{Y}}_{\mathring{N}}^T(\{\boldsymbol{\theta}_v\}_{1s}) \text{diag}\{\mathbf{w}\} \mathring{\mathbf{Y}}_{\mathring{N}}(\{\boldsymbol{\theta}_v\}_{1s}))$ is introduced [Pom08]. The equivalent surface fraction weights \mathbf{w} emphasize sampling nodes located near the equator. Finally, the order is set to $\mathring{N} = 15$ (fig. 6.17(b)) and results in $(\mathring{N} + 1)^2 = 256$ columns and rows, respectively, in $\mathring{\mathbf{Y}}_{\mathring{N}}(\{\boldsymbol{\theta}_k\}_+)$ and $\mathring{\mathbf{Y}}_{\mathring{N}}^\dagger(\{\boldsymbol{\theta}_v\}_{1s})$.

Figure 6.18 shows the normalized, interpolated directivity patterns for different octave bands of one particular array microphone with color-coded phase. As expected, the directivity index approaches 1 (0 dB) towards lower frequencies; in other words we get an omnidirectional directivity. Cancellations due to reflections of the rigid cones as well as damping effects become visible at higher frequencies ($f \geq 1$ kHz).

⁵ A thorough introduction to *spherical harmonics* is given in [Wil99, Zot09a].

⁶ Note that the symbol (\cdot) indicates variables in the SH domain.

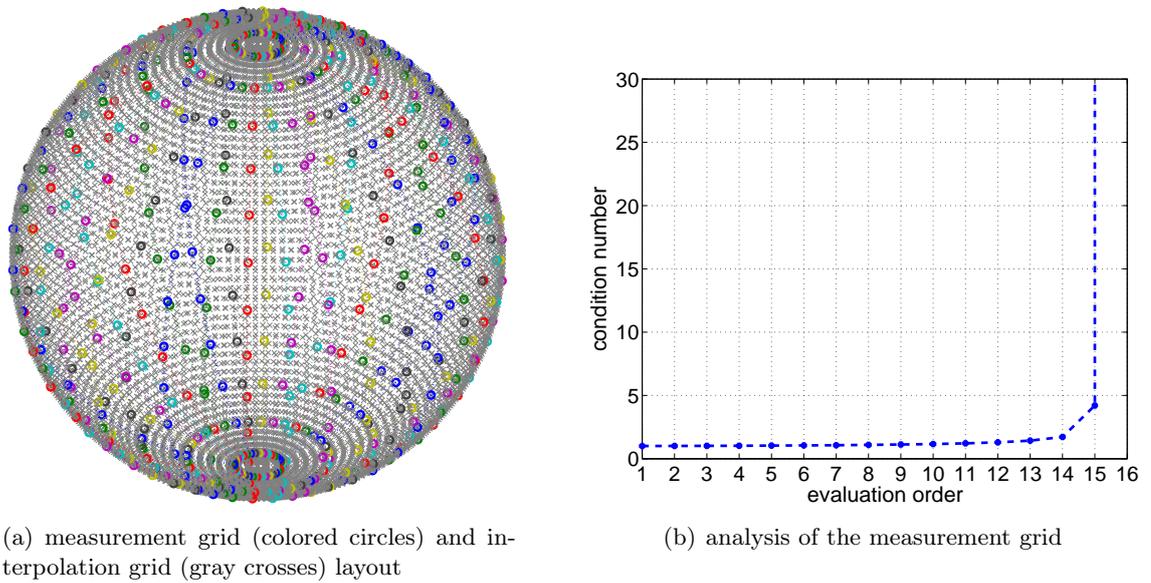


Figure 6.17: (a) Grid layout to obtain refined directivity patterns, (b) condition number $\kappa(\mathbf{Y}_N^T(\{\theta_v\}_{1s}) \text{diag}\{\mathbf{w}\} \mathbf{Y}_N(\{\theta_v\}_{1s}))$.

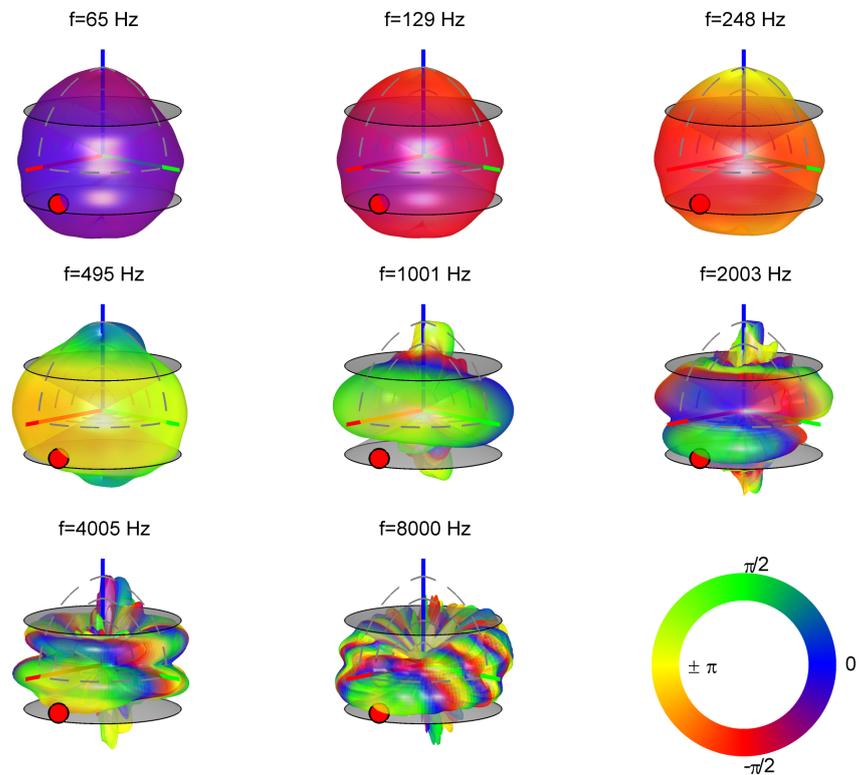


Figure 6.18: Normalized directivity patterns $\mathbf{d}_1(\{\theta_k\}_+)$ evaluated at octave bands for array microphone 1. The red dot indicates the microphone position.

Chapter 7

Modal Beamforming in the Spherical Segment Domain

The microphone array prototype has a potential application as beamforming system in a teleconference situation. Such an application requires a concept to gain control over the system in the spherical segment domain. More specific, a beamforming system with variable steering direction has to be implemented. To assess the quality of the prototype, measurement data is used as input to the holographic model.

In the following, the signal processing background to realize such a beamforming system in the spherical segment domain is provided.

7.1 System Description

The overall system to realize a beamformer is depicted in fig. 7.1 and consists of two main blocks, a *decomposition unit* and a *beamforming unit*. The decomposition unit accomplishes a band-limited decomposition of a surrounding soundfield, which is sampled at the optimal microphone positions $r_a \boldsymbol{\theta}_s$. This is achieved by encoding the received signals $\tilde{\mathbf{p}}$ at the array radius r_a in terms of SSH as shown in sec. 4.1. The coefficients of the excitation pattern ϕ_N (spherical source strength distribution) are calculated through a projection of the spherical wave spectrum ψ_N onto a desired source radius r_s via holographic filters (sec. 4.5). The beamformer unit finally weights and sums these excitation coefficients to obtain a beam $d_b(\boldsymbol{\theta}_{\rightarrow}, \boldsymbol{\theta}_i)$ for a source direction $\boldsymbol{\theta}_i$ and the desired look direction $\boldsymbol{\theta}_{\rightarrow}$.

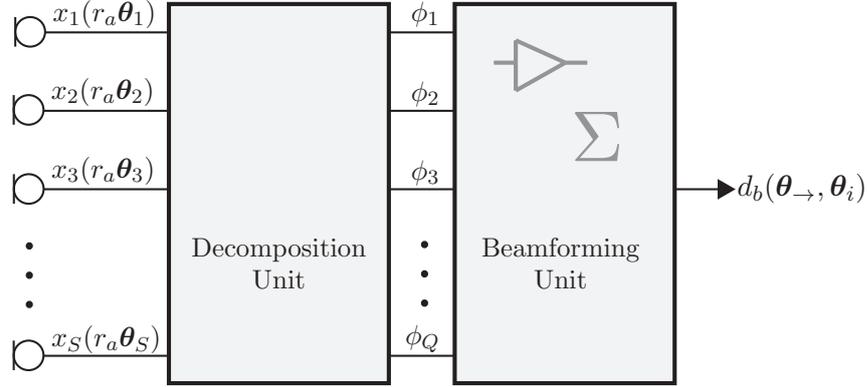


Figure 7.1: Schematic description of the concept for modal beamforming.

7.1.1 Analytic Beams

To generate analytic beams we simulate the sound pressure at the microphone positions on the array surface. The analytic *array model* is based on the wave spectrum¹ of a plane wave impinging from θ_i [Zot09a]

$$\begin{aligned}\tilde{\mathbf{b}} &= \tilde{\mathbf{F}} \cdot \mathbf{y}_{\tilde{N}}(\theta_i), \\ \tilde{\mathbf{F}} &= \text{diag}\left\{4\pi e^{i\frac{\pi}{2}\nu_1}, 4\pi e^{i\frac{\pi}{2}\nu_2}, \dots, 4\pi e^{i\frac{\pi}{2}\nu_{\tilde{Q}}}\right\}^T.\end{aligned}$$

This SSH wave spectrum is simulated with a high order \tilde{N} which results in \tilde{Q} components. To obtain the spherical wave spectrum $\tilde{\psi}_{\tilde{N}}$, we project $\tilde{\mathbf{b}}$ onto the array radius r_a (cf. sec. 4.5), i.e.

$$\begin{aligned}\tilde{\psi}_{\tilde{N}} &= \tilde{\mathbf{G}} \cdot \tilde{\mathbf{b}}, \\ \tilde{\mathbf{G}} &= \text{diag}\left\{(\mathcal{G}_{\nu_1}^c, \mathcal{G}_{\nu_2}^c \dots \mathcal{G}_{\nu_{\tilde{Q}}}^c)^T\right\}.\end{aligned}$$

The analytic microphone signals in terms of \tilde{Q} SSH are written as

$$\tilde{\mathbf{p}} = \mathbf{Y}_{\tilde{N}}(\{\theta_s\}_{\text{mic}}) \cdot \tilde{\psi}_{\tilde{N}}. \quad (7.1)$$

In the *holographic model*, the microphone signals are decomposed with a limited order N into Q spherical segment wave spectra²

$$\psi_N = (\mathbf{Y}_N^T(\{\theta_s\}_{\text{mic}}))^{-1} \tilde{\mathbf{p}}$$

¹ Note that the symbol $(\tilde{\cdot})$ symbolizes a high simulation order \tilde{N} and indicates variables in the analytic array model.

² In the holographic model, the number of microphones is equal to the number of SSH, i.e. $S = Q$.

with the SSH encoder matrix $(\mathbf{Y}_N^T(\{\boldsymbol{\theta}_s\}_{\text{mic}}))^{-1}$. Sound field extrapolation is achieved through regularized holographic filters $\bar{\mathbf{H}} = \text{diag}\{(\bar{\mathcal{H}}_{\nu_1}^c, \bar{\mathcal{H}}_{\nu_2}^c \dots \bar{\mathcal{H}}_{\nu_Q}^c)^T\}$, i.e.

$$\boldsymbol{\phi}_N = \mathbf{F}^{-1} \cdot \bar{\mathbf{H}} \cdot \boldsymbol{\psi}_N \quad (7.2)$$

with the inverse wave term

$$\mathbf{F}^{-1} = \text{diag}\{(4\pi e^{i\frac{\pi}{2}\nu_1}, 4\pi e^{i\frac{\pi}{2}\nu_2}, \dots, 4\pi e^{i\frac{\pi}{2}\nu_Q})^T\}^{-1}$$

containing Q elements on its diagonal.

To obtain a beam pattern with the desired look direction $\boldsymbol{\theta}_{\rightarrow}$, the source amplitude spectrum is weighted with the beam coefficients $\mathbf{y}_N^T(\boldsymbol{\theta}_{\rightarrow})$ and $\boldsymbol{\theta}_i$ is varied through expansion to $\{\boldsymbol{\theta}_k\}_+$ ($\varphi_+ \in [0, 2\pi)$, $\vartheta_+ \in [\vartheta_{c,1}, \vartheta_{c,2}]$)

$$\begin{aligned} \mathbf{d}_b(\boldsymbol{\theta}_{\rightarrow}, \{\boldsymbol{\theta}_k\}_+) &= \mathbf{y}_N^T(\boldsymbol{\theta}_{\rightarrow}) \cdot \underbrace{[\mathbf{F}^{-1} \cdot \bar{\mathbf{H}} \cdot (\mathbf{Y}_N^T(\{\boldsymbol{\theta}_s\}_{\text{mic}}))^{-1}]}_{\text{holographic model}} \\ &\quad \cdot \underbrace{[\mathbf{Y}_{\tilde{N}}^T(\{\boldsymbol{\theta}_s\}_{\text{mic}}) \cdot \tilde{\mathbf{G}} \cdot \tilde{\mathbf{F}}]}_{\text{analytic array model}} \cdot \mathbf{Y}_{\tilde{N}}^T(\{\boldsymbol{\theta}_k\}_+). \end{aligned}$$

Results. Figure 7.2 depicts the resulting beam patterns for different frequencies and a steering direction of $\boldsymbol{\theta}_{\rightarrow} = (0^\circ, 75^\circ)$. The polar plots show a cut through the 3D balloon plot in horizontal direction (solid black line) and vertical direction (dashed black line). For the vertical cut, the double cone is indicated through dashed grey lines in the polar plot.

7.1.2 Beams Based on Measurement Data

To assess the performance of the prototype, the measured transfer functions \mathbf{W} (for one specific frequency) from eq. (6.4) are employed in the holographic model instead of microphone signals generated by the analytic model.

With a source located at $\boldsymbol{\theta}_i$, eq. (7.1) is rewritten in terms of $(\overset{\circ}{N} + 1)^2$ spherical harmonics to

$$\tilde{\mathbf{p}} = \mathbf{W} \cdot (\overset{\circ}{\mathbf{Y}}_N^T(\{\boldsymbol{\theta}_v\}_{\text{ls}}))^\dagger \cdot \overset{\circ}{\mathbf{y}}_N(\boldsymbol{\theta}_i), \quad (7.3)$$

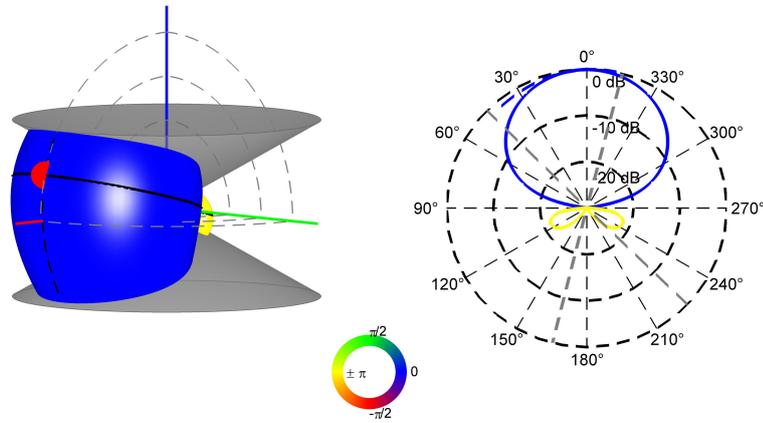
with the Ambisonics decoder matrix $(\overset{\circ}{\mathbf{Y}}_N^T(\{\boldsymbol{\theta}_v\}_{\text{ls}}))^\dagger$ as defined in eq. (6.6).

Substituting eq. (7.3) into eq. (7.2) yields the spherical source amplitude spectrum based on measured microphone signals which is weighted with beamforming coefficients and eval-

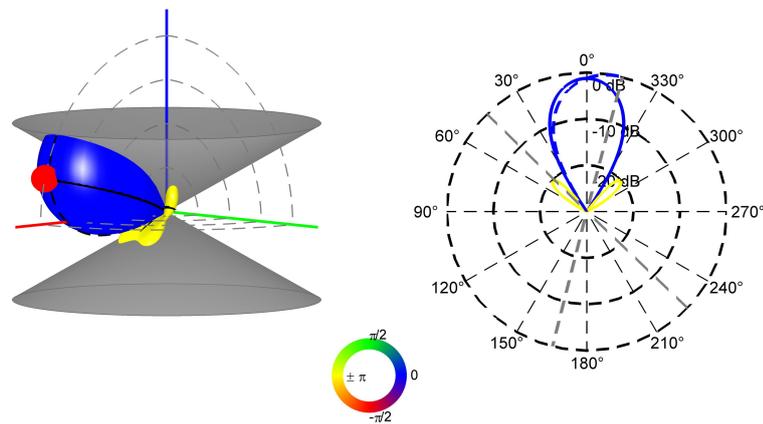
uated at a dense grid of possible source directions $\boldsymbol{\theta}_i = \{\boldsymbol{\theta}_k\}_+$. This yields a beam pattern with the desired look direction $\boldsymbol{\theta}_{\rightarrow}$:

$$\mathbf{d}_b(\boldsymbol{\theta}_{\rightarrow}, \{\boldsymbol{\theta}_k\}_+) = \underbrace{[\mathbf{F}^{-1} \cdot \bar{\mathbf{H}} \cdot (\mathbf{Y}_N^T(\{\boldsymbol{\theta}_s\}_{\text{mic}}))^{-1}]}_{\text{holographic model}} \cdot \underbrace{[\mathbf{W} \cdot (\mathbf{Y}_N^T(\{\boldsymbol{\theta}_v\}_{\text{ls}}))^\dagger]}_{\text{measured microphone signals}} \cdot \mathbf{Y}_N^T(\{\boldsymbol{\theta}_k\}_+).$$

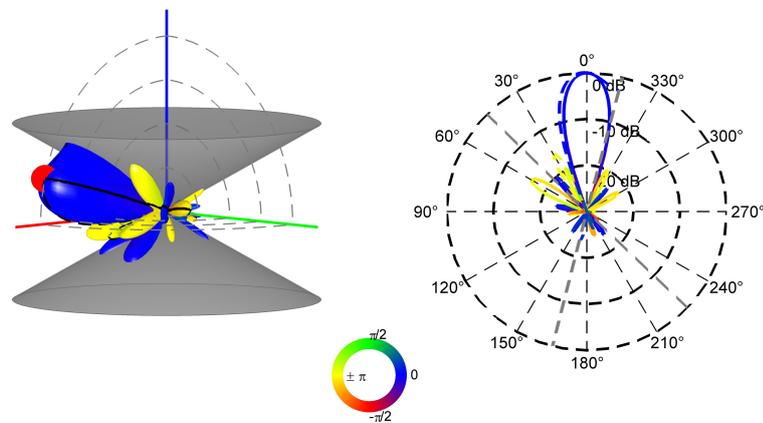
Results. The generated beampatterns based on measurement data are exemplarily shown for low, mid and high frequencies in fig. 7.3. Especially for high frequencies, the influence of the double cone with finite length is negligible in comparison to the analytic beams (fig. 7.2). Examining the form of the main lobe and the side lobes, the resulting beam patterns promise a satisfying source separation.



(a) $f = 250$ Hz



(b) $f = 1.5$ kHz



(c) $f = 4$ kHz

Figure 7.2: Generated beam patterns for low, mid and high frequencies using simulated microphone signals and regularized holographic filters with a dynamic range of 30 dB. The beam is steered in direction $\theta_{\rightarrow} = (0^{\circ}, 75^{\circ})$.

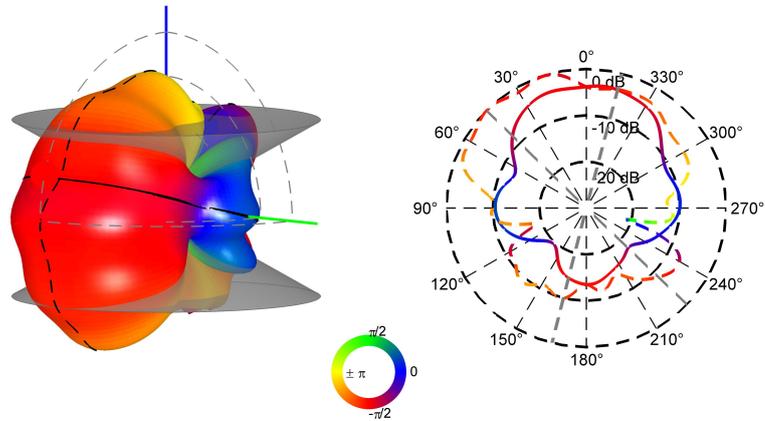
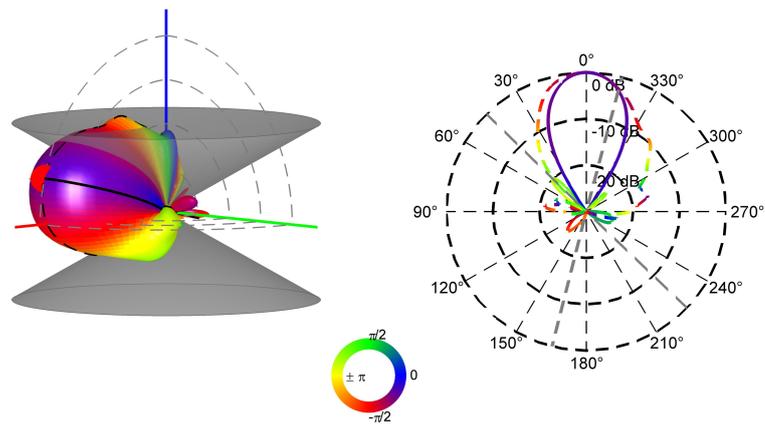
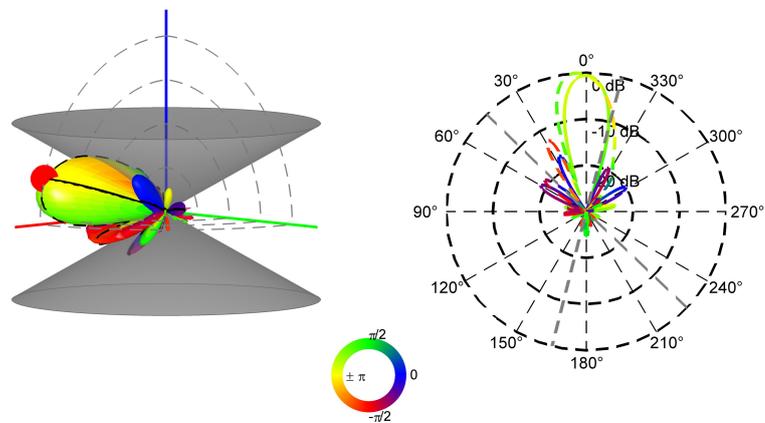
(a) $f = 250$ Hz(b) $f = 1.5$ kHz(c) $f = 4$ kHz

Figure 7.3: Generated beam patterns for low, mid and high frequencies using measured microphone signals and regularized holographic filters with a dynamic range of 30 dB. The beam is steered in direction $\theta_{\rightarrow} = (0^\circ, 75^\circ)$.

Chapter 8

Conclusion and Outlook

In this thesis, a rigid double cone microphone array prototype has been designed and implemented.

The signal processing theory presented in chapter 2 and 3 allows to decompose a surrounding soundfield on a restricted angular range through applying spherical segment harmonics (SSH). As a robust decomposition requires a carefully designed array layout, an algorithm to achieve a small condition number of the decomposition matrix was presented in chapter 4. The spherical source amplitude spectrum is obtained by utilizing holographic filters which were realized as FIR filters.

Chapter 5 outlined the prototype design and presented the assembled array hardware.

Acoustic measurements, as described in chapter 6, were carried out to obtain the spatial impulse responses of the array microphones. After calibrating the array microphones, preliminary measurements examined the microphone characteristics decoupled from the cabinet and analyzed the microphone-related deviations comparing the different gain stages. Thereafter, the directional impulse responses were measured and presented by directivity patterns.

The prototype application as modal beamforming system has been analyzed in chapter 7. For this purpose, the analytic beamformer was evaluated using microphone signals of an analytic simulation as well as measured data. The resulting pick-up patterns of the measured prototype are quite promising in the context of source separation.

Outlook. As the prototype can be used as a flexible beamforming system, the implementation of real-time algorithms for source localization would be a challenging task. In this context, additional measurements that include a table as reflecting object have to be made.

To suppress side lobes in beampatterns and to shape the form of the main lobe, order-related weights need to be derived for the spherical segment scenario. However, this is not assumed to improve the performance significantly.

Comparative studies including listening tests are necessary to assess the performance of the prototype in comparison to other microphone arrays with respect to resolution and perception.

To analyze array imperfections, a detailed error analysis would be informative. First of all, the presumption of an infinitely long cone cannot be realized in practice and needs to be investigated by means of an analytic model. Several types of non-idealistic issues are additionally introduced in practice: positioning errors, microphone gain mismatches as well as spatial aliasing have to be considered and can be summarized by defining a *holographic error*. Nevertheless, an open question of how to calibrate the array microphones mounted in the spherical segment should be practically solved.

Finally, a miniaturized version of the prototype with less sensors fulfilling industry standards could be a promising product.

Appendix A

Quantitative Results

A.1 Related Values to the Double Cone Boundary Condition

Tables A.1 - A.3 list all orders $\nu_l(|m|)$, normalization factors $\Xi_{\nu_l(|m|)}^{|m|}$, $n_p = 1000$ (number of supporting points to evaluate the integral in eq. (3.2) numerically) as well as the weighting factors $\alpha_l(|m|)$ of the associated Legendre functions for the double cone boundary condition (cf. sec. 2.2.1) up to the order $N = 9.9$ which results in $Q = 64$ SSH (cf. fig. A.1(a)).

Table A.1: Summary of orders $\nu_l(|m|)$ related to the double cone boundary condition up to $N = 9.9$.

$l \setminus m $	0	1	2	3	4	5	6	7	8	9	10
0	0										
1	2.627	0.661									
2	5.567	2.808	1.654								
3	8.545	5.659	3.299	2.681							
4		8.607	5.927	4.000	3.716						
5			8.788	6.351	4.828	4.751					
6				9.083	6.905	5.731	5.786				
7					9.482	7.562	6.680	6.819			
8							8.302	7.660	7.849		
9								9.104	8.658	8.875	
10										9.669	9.898

Table A.2: Summary of normalization factors $\Xi_{\nu_l(|m|)}^{|m|}$ related to the double cone boundary condition up to order $N = 9.9$.

$l \setminus m $	0	1	2	3	4	5	6	7	8	9	10
0	0.398										
1	0.999	0.547									
2	1.316	0.154	0.207								
3	1.566	0.199	0.114	0.045							
4		0.187	0.059	0.024	0.007						
5			0.033	0.004	0.003	7.89e-04					
6				4.6e-4	1.5e-4	3.2e-4	7.4e-5				
7					2.4e-4	6.3e-5	2.7e-5	5.8e-6			
8							7.8e-6	1.9e-6	3.9e-7		
9								6.5e-7	1.2e-7	2.3e-8	
10										6.6e-9	1.3e-9

Table A.3: Summary of weighting factors $\alpha_l(|m|)$ of the associated Legendre functions related to the double cone boundary condition up to order $N = 9.9$

$l \setminus m $	0	1	2	3	4	5	6	7	8	9	10
0	0										
1	-0.422	-0.375									
2	-0.514	2.046	-0.384								
3	-0.552	1.074	0.324	-0.349							
4		0.896	-0.073	-3.8e-16	-0.305						
5			-0.220	-1.035	-0.177	-0.262					
6				-4.839	4.228	-0.286	-0.222				
7					0.601	0.776	-0.349	-0.186			
8							0.326	-0.376	-0.154		
9								0.105	-0.379	-0.126	
10										-0.364	-0.103

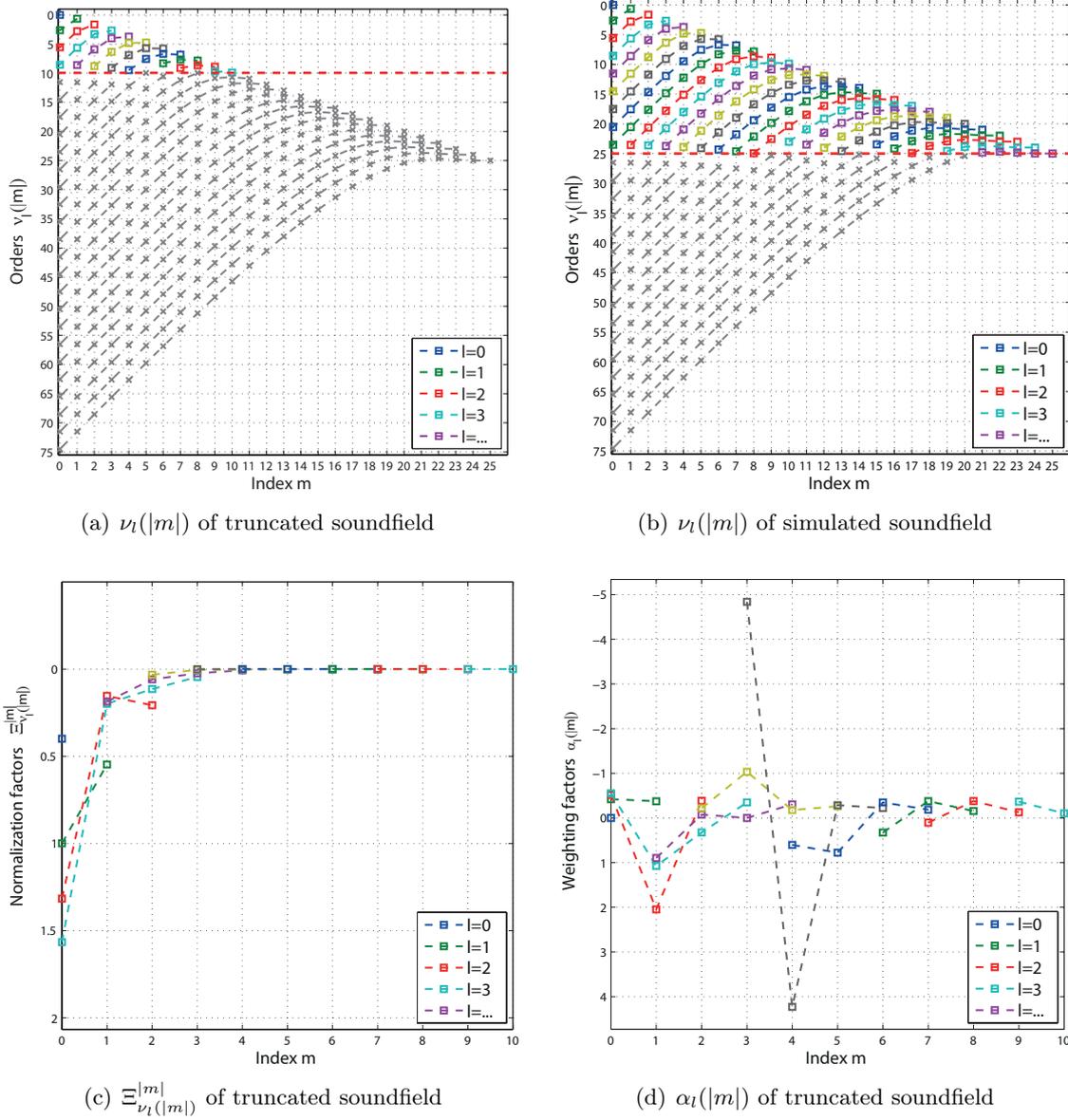


Figure A.1: Illustration of orders $\nu_l(|m|)$, normalization factors $\Xi_{\nu_l(|m|)}^{|m|}$ and weighting factors $\alpha_l(|m|)$ related to the double cone boundary condition up to (a) $N = 9.9$, (b) $\tilde{N} = 25$ (horizontal red dashed line), $\vartheta_{c,1} = 60^\circ$, $\vartheta_{c,2} = 120^\circ$ (associated values related to the index l are linked via a dashed line).

A.2 Microphone Coordinates

The microphone distribution with channel numbers is shown in fig. (A.2).

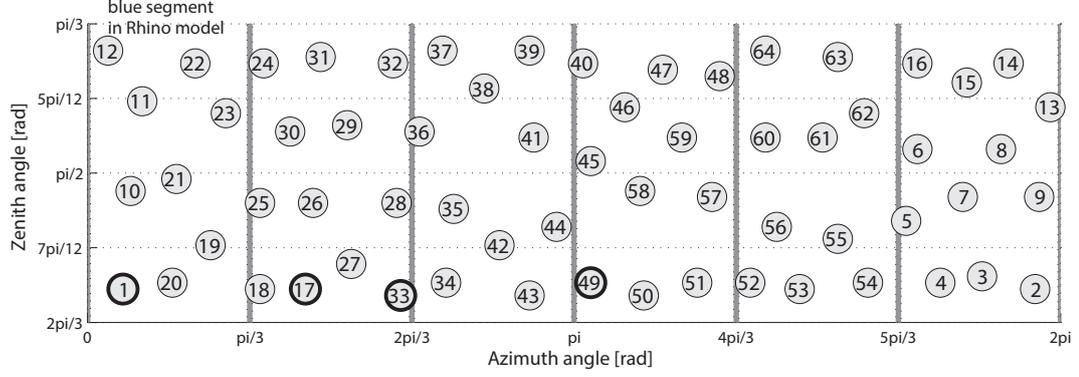


Figure A.2: Microphone distribution on the spherical segment of the prototype with channel numbers. Channels 1, 17, 33, 49, respectively, are the first channels of the four busses.

Table A.4 depicts the resulting positions of all $S = 64$ microphones in spherical and Cartesian coordinates after the optimisation process described in section 4.3.1.

Table A.4: Microphone coordinates for the $S = 64$ omnidirectional pressure transducers used in the prototype, dimensions: r_a [mm], φ [deg], ϑ [deg], Cartesian coordinates (x, y, z) in [mm]

microphone #	spherical coordinates $(r_a, \varphi, \vartheta)$	Cartesian coordinates (x, y, z)
1	100,13.13,113.32	18.56,11.60,97.58
2	100,350.62,113.32	7.24,-20.65,97.58
3	100,330.94,110.74	33.98,62.11,-70.62
4	100,315.47,112.02	-22.70,-84.88,47.75
5	100,302.81,99.59	-27.72,-75.65,59.23
6	100,307.03,85.22	-25.66,28.90,-92.23
7	100,323.91,94.78	-48.15,-16.07,86.16
8	100,337.97,85.22	-9.47,37.47,-92.23
9	100,352.03,94.78	50.00,8.73,86.16
10	100,15.94,93.58	60.05,14.03,78.73
11	100,20.16,75.52	3.24,11.96,99.23
12	100,7.50,65.38	19.51,52.79,-82.66

Continued on next page

Table A.4 – continued from previous page

microphone #	spherical coordinates $(r_a, \varphi, \vartheta)$	Cartesian coordinates (x, y, z)
13	100,356.25,76.75	-30.79,-92.67,21.53
14	100,340.78,67.98	-7.40,-90.53,41.82
15	100,325.31,71.79	7.06,-44.43,-89.31
16	100,307.03,67.98	-60.30,67.93,41.82
17	100,80.63,113.32	10.77,-19.05,97.58
18	100,63.75,113.32	13.29,17.39,97.58
19	100,45.47,104.48	-6.07,-71.82,-69.31
20	100,31.41,112.02	-87.86,0.85,47.75
21	100,32.81,91.19	-1.52,-8.61,-99.62
22	100,39.84,67.98	49.31,-76.29,41.82
23	100,51.09,77.98	36.17,39.42,-84.48
24	100,65.16,67.98	62.16,-66.24,41.82
25	100,63.75,95.98	59.95,78.43,-15.99
26	100,83.44,95.98	-18.19,97.02,-15.99
27	100,97.50,108.21	-97.87,-10.87,17.41
28	100,114.38,95.98	28.52,94.50,-15.99
29	100,96.09,80.41	26.00,-92.07,29.10
30	100,75.00,81.61	-6.17,2.59,99.78
31	100,86.25,66.68	9.33,64.41,-75.93
32	100,112.97,67.98	-90.09,11.65,41.82
33	100,115.78,114.62	-89.62,44.14,4.38
34	100,132.66,112.02	-66.67,-57.23,47.75
35	100,135.47,97.18	-19.23,-7.69,-97.83
36	100,122.81,81.61	6.41,1.92,99.78
37	100,131.25,65.38	43.16,-36.12,-82.66
38	100,146.72,73.04	41.94,-56.95,-70.70
39	100,163.59,65.38	54.79,12.88,-82.66
40	100,183.28,67.98	-43.70,-79.63,41.82
41	100,165.00,82.82	-6.02,90.57,41.96
42	100,152.34,104.48	-1.69,-72.06,-69.31
43	100,163.59,114.62	97.25,22.87,4.38
44	100,173.44,100.81	-21.69,-16.49,96.22
45	100,186.09,87.61	25.51,23.28,93.85
46	100,198.75,76.75	-65.93,-72.04,21.53
47	100,212.81,69.26	9.74,-10.35,98.99
48	100,233.91,70.53	14.02,97.77,15.64
49	100,186.09,112.02	64.90,59.23,47.75
50	100,205.78,114.62	0.69,-99.90,4.38
51	100,225.47,112.02	-65.71,58.33,47.75
52	100,245.16,112.02	-87.31,-9.82,47.75
53	100,263.44,113.32	19.64,-9.64,97.58
54	100,288.75,112.02	-84.53,23.99,47.75
55	100,277.50,103.25	20.86,35.52,-91.12
56	100,255.00,100.81	-23.50,-13.80,96.22

Continued on next page

Table A.4 – continued from previous page

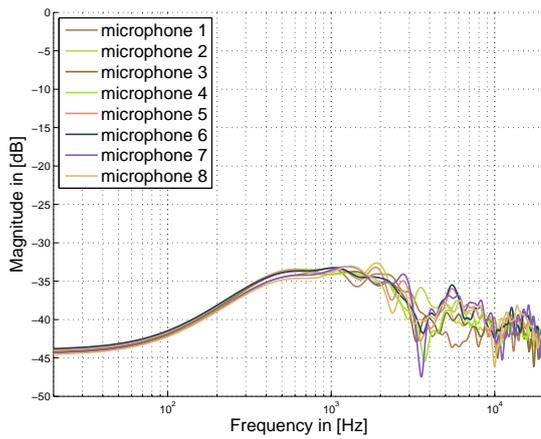
microphone #	spherical coordinates $(r_a, \varphi, \vartheta)$	Cartesian coordinates (x, y, z)
57	100,231.09,94.78	9.42,-49.88,86.16
58	100,204.38,93.58	60.76,10.52,78.73
59	100,219.84,82.82	90.56,-6.14,41.96
60	100,250.78,82.82	77.57,-47.15,41.96
61	100,271.88,82.82	-11.52,90.04,41.96
62	100,287.34,77.98	-5.98,-53.17,-84.48
63	100,277.50,66.68	-32.95,-56.12,-75.93
64	100,250.78,65.38	48.09,-29.23,-82.66

Appendix B

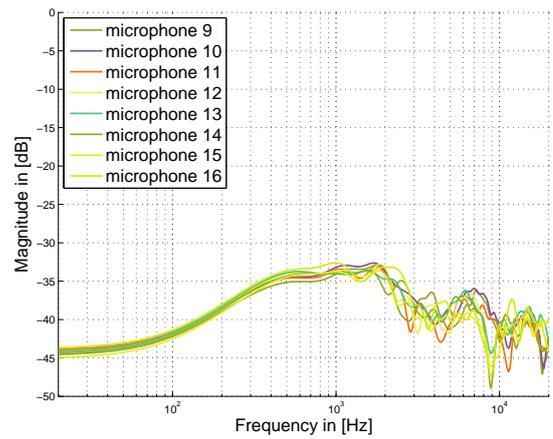
Results of the Acoustic Measurements

B.1 Preliminary Measurements

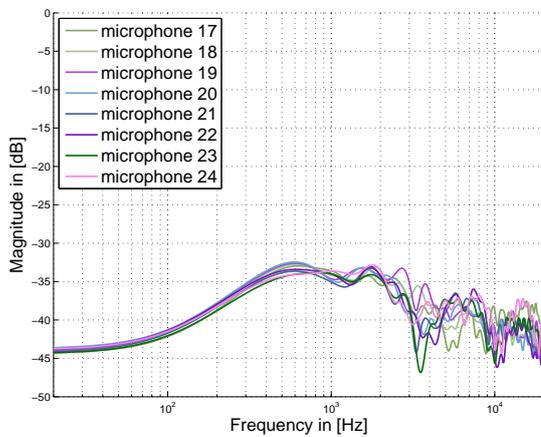
Figures B.1 - B.6 show the results of the measurements described in sec. 6.2.



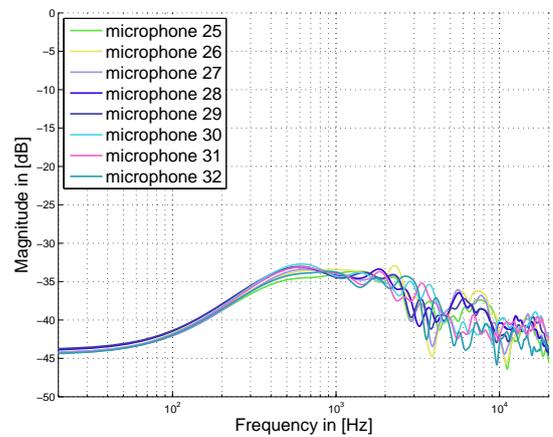
(a) microphones 1-8



(b) microphones 9-16

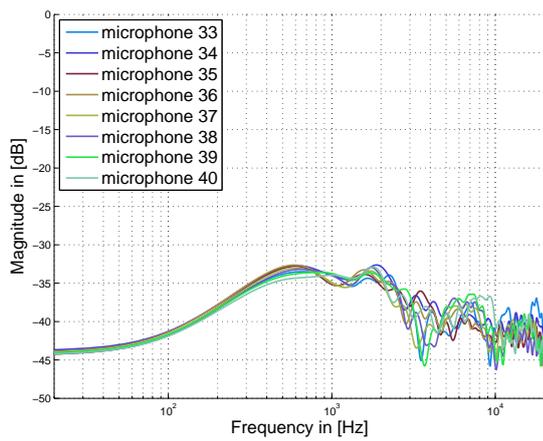


(c) microphones 17-24

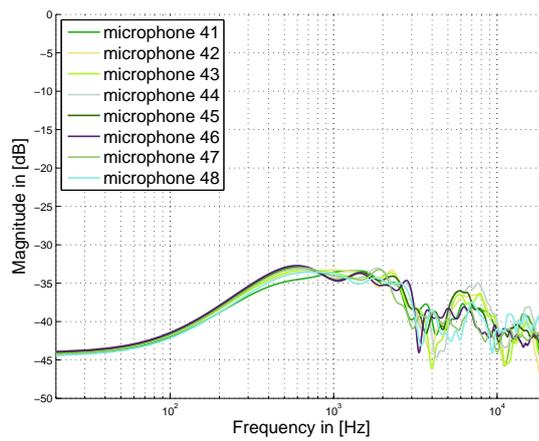


(d) microphones 25-32

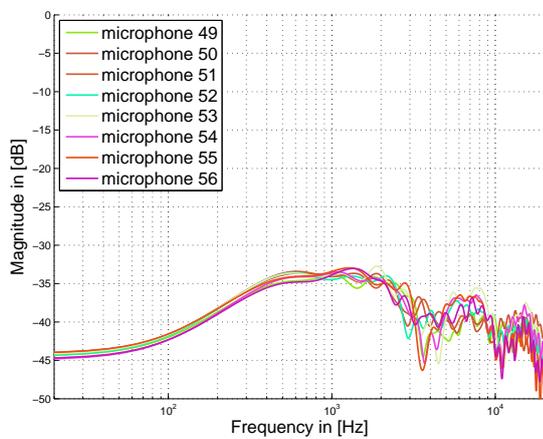
Figure B.1: Transfer functions of the array microphones 1-32, -20 dB gain setting (*_g1.d11).



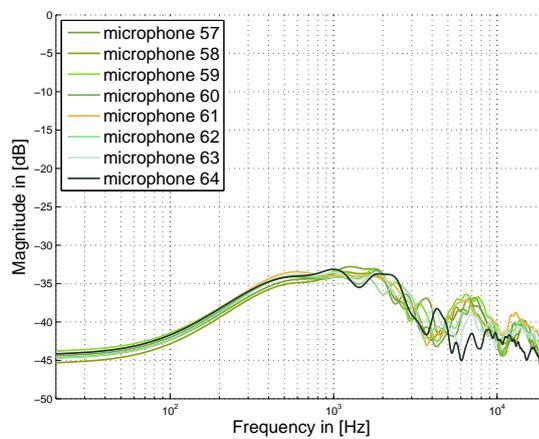
(a) microphones 33-40



(b) microphones 41-48

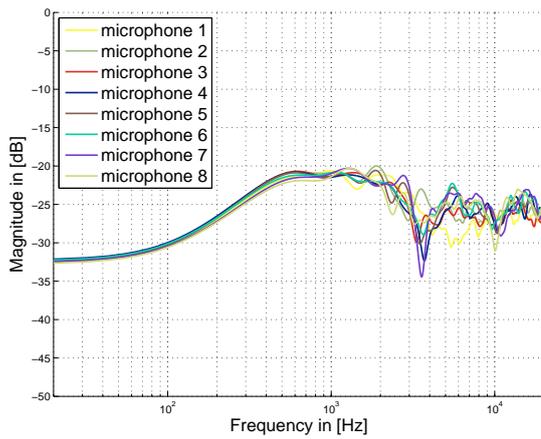


(c) microphones 49-56

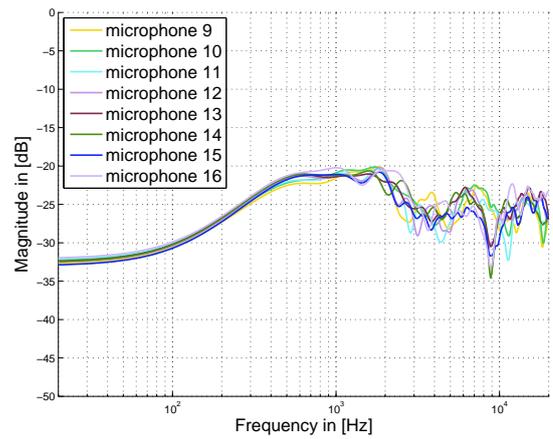


(d) microphones 57-64

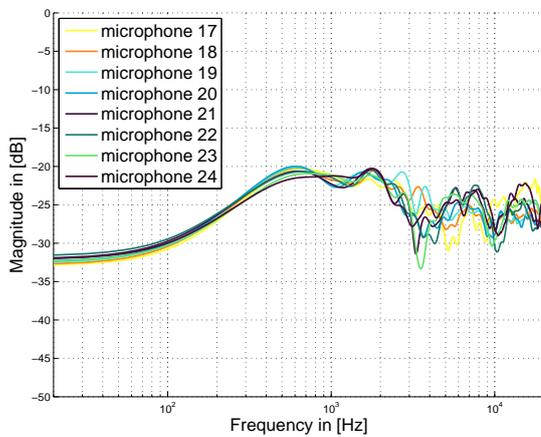
Figure B.2: Transfer functions of the array microphones 33-64, -20 dB gain setting (*_g1.d11).



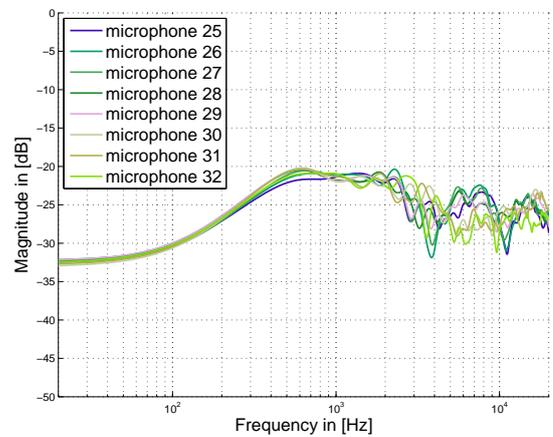
(a) microphones 1-8



(b) microphones 9-16

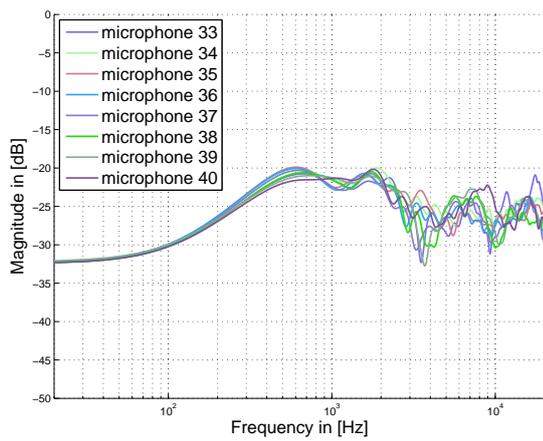


(c) microphones 17-24

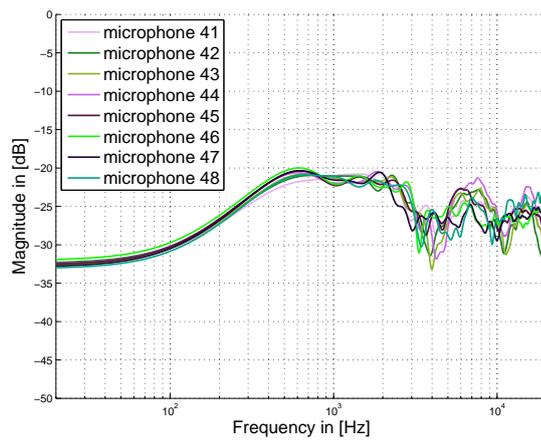


(d) microphones 25-32

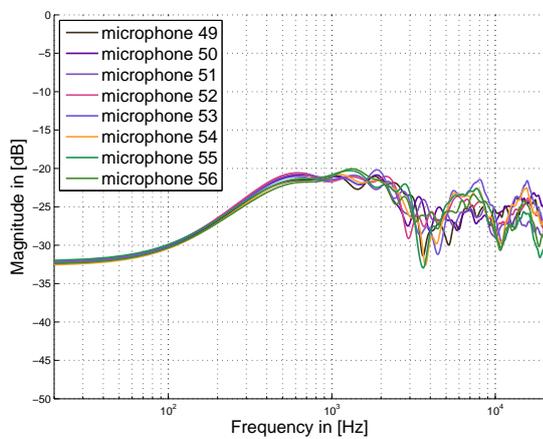
Figure B.3: Transfer functions of the array microphones 1-32, 0 dB gain setting (*_g10.d11).



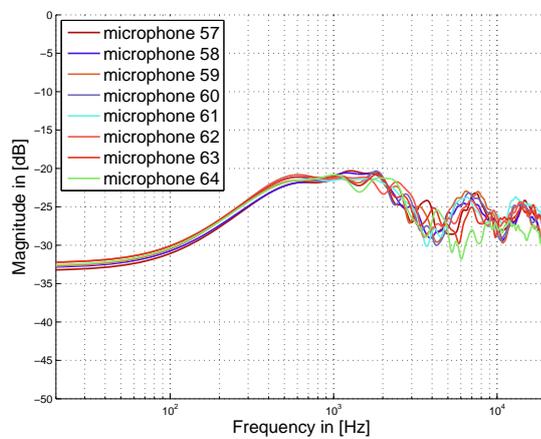
(a) microphones 33-40



(b) microphones 41-48

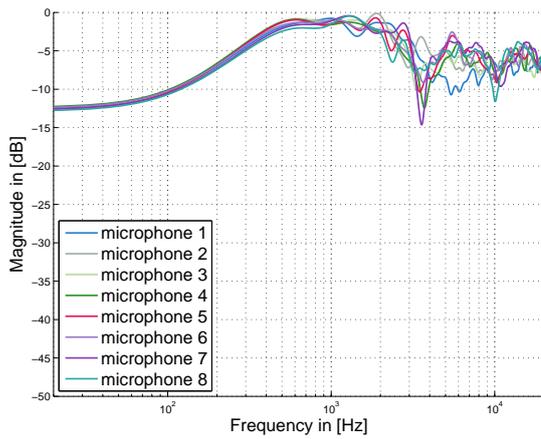


(c) microphones 49-56

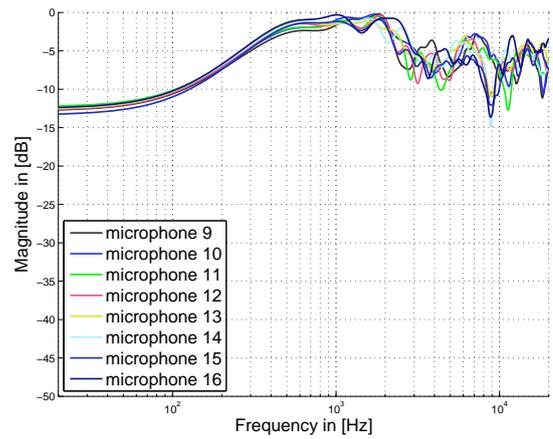


(d) microphones 57-64

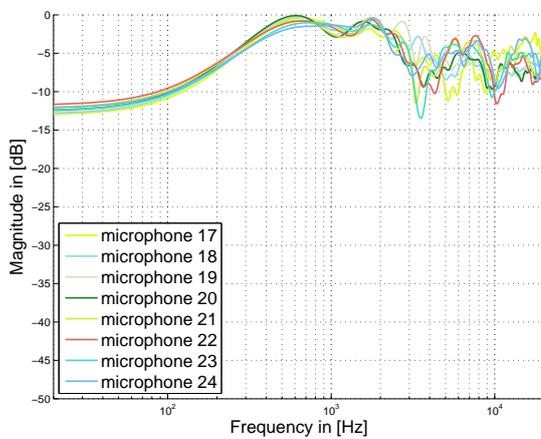
Figure B.4: Transfer functions of the array microphones 33-64, 0 dB gain setting (*_g10.dll).



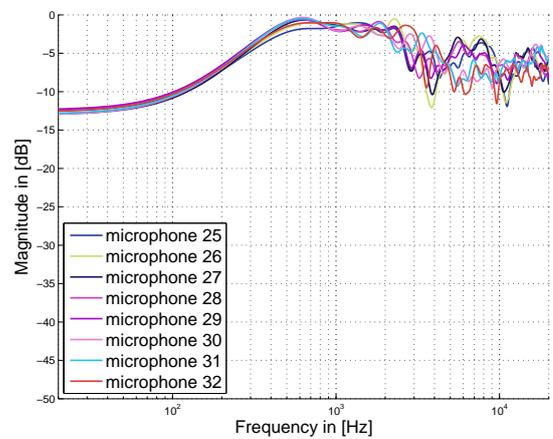
(a) microphones 1-8



(b) microphones 9-16

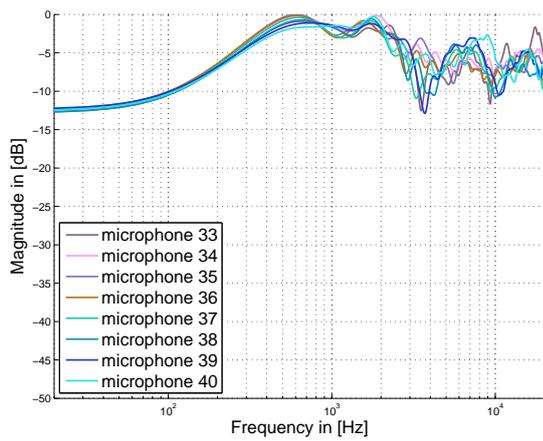


(c) microphones 17-24

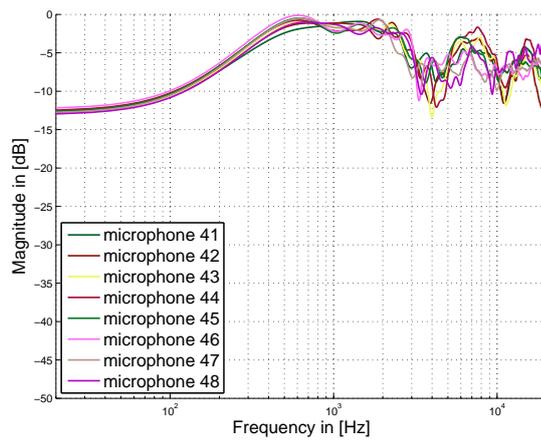


(d) microphones 25-32

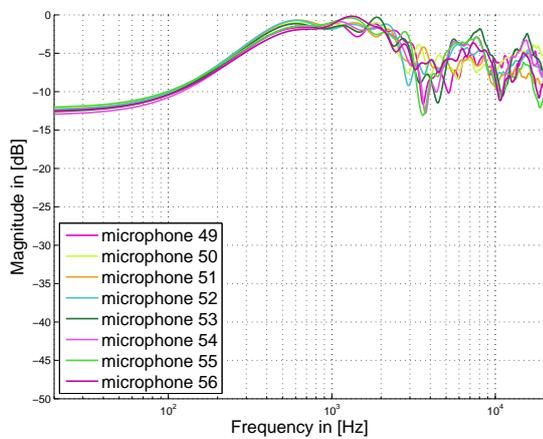
Figure B.5: Transfer functions of the array microphones 1-32, +20 dB gain setting (*_g100.d11).



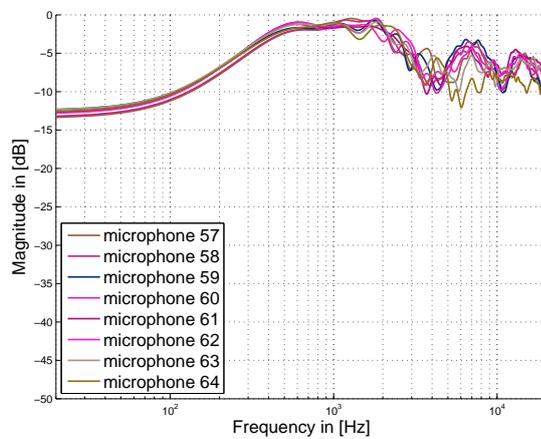
(a) microphones 33-40



(b) microphones 41-48



(c) microphones 49-56



(d) microphones 57-64

Figure B.6: Transfer functions of the array microphones 33-64, +20 dB gain setting (*_g100.d11).

Appendix C

User Instructions

Please read the following instructions BEFORE using or even manipulating the VisiSonics array hardware.

C.1 Driver Installation

First-time installation of the 'Opal Kelly XEM3010' front panel USB driver and the VisiSonics ASIO driver:

- Do not plug in the array hub board into the PC's USB port until both driver software installation procedures have been completed.

Run

`FrontPanelUSB-DriverOnly-Win-x64-4.0.8.exe`

to install the integration board drivers.

- Follow the installation instructions of the `ReadMe.txt` file (located in the folder "2013-03-07_v4_ASIO_Drivers") for a successful ASIO driver installation! Always safely remove hardware before disconnecting the hub board—if not needed on long-term view additionally disconnect the power supply.

C.2 Recording Audio

- On a Win7 system change the priority of your DAW to 'High' to avoid dropouts due to priority issues.
- The sampling rate of $f_s = 44.1$ kHz is fixed by VisiSonics and cannot be changed by the user. External synchronization was not yet possible during the experimental validation.
- To change the ASIO buffer size simply modify the value in the file `vsBufferSize.txt` (in the installation directory, $value \stackrel{!}{=} 2^{\tilde{k}} \geq 4096, \tilde{k} \in \mathbb{Z}$), save and close the file. Use

a buffer size of at least 4096 to avoid buffer dropouts during recordings. Restart the DAW to apply the updated value.

- Gain settings can be changed via the provided batch files (`register_g#.bat`, cf. tab. 6.1) which will unregister all registered *.dll files and register the selected gain stage. The batch files have to be executed in the driver installation directory via a command window run in administrator mode (Win7). For a successful registration it is necessary to restart the DAW. Run `regedit` and check the key `HKEY_LOCAL_MACHINE\Software\Wow6432Node\ASIO\VisiSonics RealSpace... 5/64 Audio Camera`—the key description contains the entry `VisiSonics Audio... Camera - 32G# (#...1, 10, 100)` and confirms the selected gain level.
- To find the correct orientation of the array, a red marker dot indicates 0° azimuth direction on the edge of the lower cone.

C.3 Manipulations

- Carefully read the instruction manual (`V DAM_array_connection_instructions.pdf`) by VisiSonics!
- Never disconnect or connect microphones in the chain while the hub board is powered or connected via USB (this could destroy microphones or other electronic elements)!
- It is necessary to unplug the USB cable as well as the power plug from the outsourced connectors BEFORE removing the base plate, otherwise the hub board gets seriously damaged.
- The FFC connector can be unlocked by deliberately folding up the black lever on the back side of the connector. Note that the *blue side* of the ribbon cable (labelled "*This side up*") faces *upwards* with respect to the hub board—ignoring this will destroy the first microphone in the chain!
- Each chain of 16 channels has an appendant U.FL synchronization connector which is always on the *right* side of the FFC connector (see labels).
- Channels are only working if bus ribbon cables are connected in ascending order, i.e. to receive input signals e.g. on channels 33-48, the two preceding busses (ch. 1-16 *and* ch. 17-32) have to be integrated.

C.3.1 Maintenance

- In case of failures it could be necessary to replace a microphone in the chain. When embedding the new microphone pay attention to the right direction of installation ("*TOHUB*", "*TOSTR*", cf. VisiSonics connection instructions for further details).

Failing in doing so will destroy the first incorrectly connected microphone due to polarity reversal!

C.3.2 Unmounting

NEVER load the prototype collaterally as the weight should lie vertically on the mounting adapter and the spherical segment, respectively. Lateral forces may destroy the spherical segment! Only lift the cabinet at the lower cone (fig. C.1)!

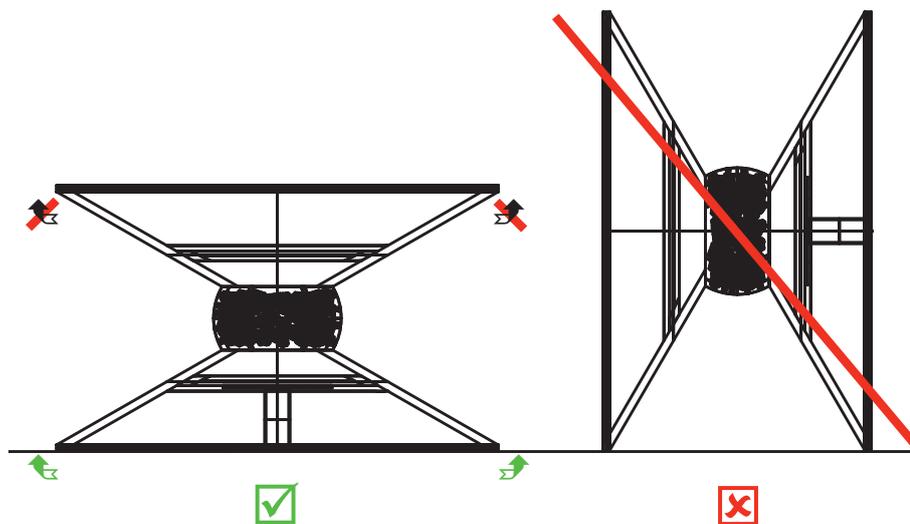


Figure C.1: Instructions for placement and lifting of the prototype to avoid damages. Lateral forces may destroy the spherical segment!

Unmounting steps:

- (1) Disconnect the power plug and the USB cable on the backside of the outsourced connectors (inside the lower cone).
- (2) Unscrew the six screws on the bottom and demount the base plate.
- (3) Carefully unlock the FFC connectors, remove the four ribbon cables and the appendant U.FL coaxial wires from the hub board.
- (4) Unscrew the six screws on the top, demount the cover plate and remove the aluminium bracket.
- (5) Do not twist the spherical segment as it is kept in position by cropped bolts in the lower cone—only take it off vertically!
- (6) When reassembling the array note the green and red marker dots which define the azimuthal position of the spherical segment. Do not overbolt the nut of the threaded bar; the resulting forces may damage the spherical segment. The same goes for the screws in the cover plate and the base plate, respectively, as the threaded sleeves are glued.

References

- [AW02] T. D. Abhayapala and D. B. Ward, “Theory and design of high order sound field microphones using spherical microphone array.” in *ICASSP*. IEEE, 2002, pp. 1949–1952.
- [BPF11] R. Baumgartner, H. Pomberger, and M. Frank, “Practical Implementation of Radial Filters for Ambisonic Recordings,” in *Proc. Int. Conf. Spatial Audio, Detmold*, 2011.
- [BPSW11] B. Bernschütz, C. Pörschmann, S. Spors, and S. Weinzierl, “Soft-Limiting der modalen Amplitudenverstärkung bei sphärischen Mikrofonarrays im Plane Wave Decomposition Verfahren,” in *Fortschritte der Akustik - DAGA 2011*. Dt. Ges. für Akustik, 2011.
- [Dav07] C. W. David, “The Laplacian in Spherical Polar Coordinates,” Chemistry Education Materials, Paper 34, 2007.
- [DF00] A. F. Dipartimento and A. Farina, “Simultaneous Measurement of Impulse Response and Distortion With a Swept-Sine Technique,” in *108th AES Convention*, 2000, pp. 18–22.
- [DH94] J. Driscoll and D. Healy, “Computing Fourier Transforms and Convolutions on the 2-Sphere,” *Advances in Applied Mathematics*, vol. 15, no. 2, pp. 202 – 250, 1994.
- [Dlm27] “NIST Digital Library of Mathematical Functions,” <http://dlmf.nist.gov/14.3>, Accessed: 2012-11-27.
- [EM03] G. Elko and J. Meyer, “Audio system based on at least second-order eigenbeams,” *WO2003 061336, world-wide Patent*, 2003.
- [Gen05] Genelec, “Data Sheet - Genelec 8020A, Active Monitoring System,” Olvitie 5 FIN-74100, Iisalmi Finland, Tech. Rep. BBA0034001, 2005.
- [GHB⁺05] K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, “HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere,” *The Astrophysical Journal*, vol. 622, no. 2, p. 759, 2005.

- [GRS02] B. N. Gover, J. G. Ryan, and M. R. Stinson, "Microphone array measurement system for analysis of directional and spatial variations of sound fields." *The Journal of the Acoustical Society of America*, vol. 112, no. 5 Pt 1, pp. 1980–1991, Nov. 2002.
- [HC97] C. Hwang and S.-K. Chen, "Fully normalized spherical cap harmonics: application to the analysis of sea-level data from TOPEX/POSEIDON and ERS-1," *Geophysical Journal International*, vol. 129, no. 2, pp. 450–460, 1997.
- [Kai12] F. Kaiser, "Modal Analysis of Rigid Microphone Arrays using Boundary Elements," Master's thesis, University of Music and Performing Arts Graz, 2012.
- [Ker10] M. Kerscher, *Compact Spherical Loudspeaker Array: Implementation of a System for Variable Sound-Radiation*. VDM Publishing, 2010.
- [Kja] B. & Kjaer, "Product data sound level calibrator – type 4231," Headquarters: DK-2850 Nærum - Denmark, Tech. Rep. BP1311–14. [Online]. Available: <http://www.signalysis.com/pdf/BK4231MicrophoneCalibrator.pdf>
- [KL80] V. Klema and A. Laub, "The singular value decomposition: Its computation and some applications," *IEEE Transactions on Automatic Control*, vol. 25, no. 2, pp. 164–176, 1980.
- [Köß11] T. Kößler, "Microphone Array on a Spherical Cap for three dimensional Recordings of the Wave Field," Master's thesis, University of Music and Performing Arts Graz, 2011.
- [MDB06] S. Moreau, J. Daniel, and S. Bertet, "3D Sound Field Recording with Higher Order Ambisonics – Objective Measurements and Validation of a 4th Order Spherical Microphone," in *120th AES Conf., Paris, May, 2006*.
- [ME08] J. Meyer and G. Elko, "Spherical harmonic modal beamforming for an augmented circular microphone array," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, 31 2008–april 4 2008, pp. 5280–5283.
- [MWD22] "Double factorial," <http://mathworld.wolfram.com/DoubleFactorial.html>, Accessed: 2013-05-22.
- [MWG22] "Gamma function," <http://mathworld.wolfram.com/GammaFunction.html>, Accessed: 2013-05-22.
- [OS83] F. Olver and J. Smith, "Associated legendre functions on the cut," *Journal of Computational Physics*, vol. 51, no. 3, pp. 502 – 518, 1983.
- [OSB10] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, *Discrete-Time Signal Processing (Prentice-Hall Signal Processing Series)*, 3rd ed. Prentice Hall, 2010.

- [PB87] T. W. Parks and C. S. Burrus, *Digital filter design*. New York, NY, USA: Wiley-Interscience, 1987.
- [Pet04] S. O. Petersen, “Localization of sound sources using 3D microphone array,” Master’s thesis, University of Southern Denmark, September 2004.
- [PL06] S.-C. Pei and H.-S. Lin, “Minimum-Phase FIR Filter Design Using Real Cepstrum,” *Circuits and Systems II: Express Briefs, IEEE Transactions on*, vol. 53, no. 10, pp. 1113–1117, 2006.
- [Ple09] P. Plessas, “Rigid Sphere Microphone Arrays for Spatial Recording and Holography,” Master’s thesis, University of Music and Performing Arts Graz, 2009.
- [Pom08] H. Pomberger, “Angular and Radial Directivity Control for Spherical Loudspeaker Arrays,” Master’s thesis, University of Music and Performing Arts Graz, 2008.
- [Pom10] H. Pomberger, “Microphone Arrays Utilizing Rigid Cones for Sound Field Analysis,” 2010, University of Music and Performing Arts Graz, unpublished.
- [Pom13] —, “Modal Sound Field Decomposition Applicable for a Limited Range of Directions,” in *Fortschritte der Akustik, AIA-DAGA, March 2013*, March 2013.
- [Raf05] B. Rafaely, “Analysis and design of spherical microphone arrays,” *Speech and Audio Processing, IEEE Transactions on*, vol. 13, no. 1, pp. 135–143, Jan. 2005.
- [RME13a] RME, “Data Sheet - Hammerfall Digiface DSP,” Audio AG, Am Pfanderling 60, D-85778 Haimhausen, Germany, Tech. Rep., 2013. [Online]. Available: http://www.rme-audio.de/download/sheets/hdsp_digiface_e.pdf
- [RME13b] —, “Data Sheet - RME M-16, DA,” Audio AG, Am Pfanderling 60, D-85778 Haimhausen, Germany, Tech. Rep., 2013. [Online]. Available: http://www.rme-audio.de/download/sheets/m32da_m16da.pdf
- [RSZ95] E. A. Rakhmanov, E. B. Saff, and Y. M. Zhou, “Electrons on the sphere,” *Computational Methods and Function Theory*, 1995.
- [Sch92] S. Schot, “Eighty years of Sommerfeld’s radiation condition,” *Historia Mathematica*, vol. 19, no. 4, pp. 385–401, Nov. 1992.
- [SH11] T. Song and Q. Huang, “Localization of multiple narrowband acoustic sources in spherical harmonic domain,” in *Image and Signal Processing (CISP), 2011 4th International Congress on*, vol. 5, oct. 2011, pp. 2418–2421.
- [SK97] E. Saff and A. Kuijlaars, “Distributing many points on a sphere,” *The Mathematical Intelligencer*, vol. 19, pp. 5–11, 1997.
- [Sne94] N. Sneeuw, “Global spherical harmonic analysis by least-squares and numerical quadrature methods in historical perspective,” *Geophysical Journal International*, vol. 118, no. 3, pp. 707–716, 1994.

- [STMK11] H. Sun, H. Teutsch, E. Mabande, and W. Kellermann, “Robust localization of multiple sources in reverberant environments using EB-ESPRIT with spherical microphone arrays,” in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, may 2011, pp. 117–120.
- [SW03] I. H. Sloan and R. S. Womersley, “Extremal systems of points and numerical integration on the sphere,” *Advances in Computational Mathematics*, vol. 21, pp. 107–125, 2003.
- [Teu07] H. Teutsch, *Modal Array Signal Processing: Principles and Applications of Acoustic Wavefield Decomposition*, 1st ed., ser. Modal Array Signal Processing: Principles and Applications of Acoustic Wavefield Decomposition, M. M. Thoma, Ed. Berlin: Springer, 2007.
- [T.N01] T. Nagarajan, V. Kamakshi Prasad and Hema A. Murthy, “The minimum phase signal derived from the magnitude spectrum and its application to speech segmentation,” *Sixth Biennial Conference on Signal Processing and Communications, Bangalore, India*, 2001.
- [Vis07] “VisiSonics Digital Array Microphones,” <http://www.visisonics.com/products-realspace/digital-array-microphones>, Accessed: 2012-12-07.
- [Wik07] “Non-uniform rational B-spline,” http://en.wikipedia.org/wiki/Non-uniform_rational_B-spline, Accessed: 2012-12-07.
- [Wik25] “Audio stream input/output,” http://en.wikipedia.org/wiki/Audio_Stream_Input/Output, Accessed: 2013-02-25.
- [Wil99] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. Academic Press, 1999.
- [Zau12] M. Zaunschirm, “Modal beamforming using planar circular microphone arrays,” Master’s thesis, University of Music and Performing Arts Graz, 2012.
- [Zot09a] F. Zotter, “Analysis and Synthesis of Sound Radiation with Spherical Arrays,” Ph.D. dissertation, University of Music and Performing Arts Graz, 2009.
- [Zot09b] —, “Sampling Strategies for Acoustic Holography/Holophony on the Sphere,” *NAG-DAGA*, 2009.
- [ZPN10] F. Zotter, H. Pomberger, and M. Noisternig, “Ambisonic Decoding with and without Mode-Matching: A Case-Study Using the Hemisphere,” in *Proc. of the 2nd International Symposium on Ambisonics and Spherical Acoustics, Paris, France*, May 6-7 2010.
- [ZRS03] J. M. Zmolnig, W. Ritsch, and A. Sontacchi, “The IEM CUBE,” in *Proceedings of the 9th International Conference on Auditory Display (ICAD2003)*, E. Brazil and B. Shinn-Cunningham, Eds., 2003.