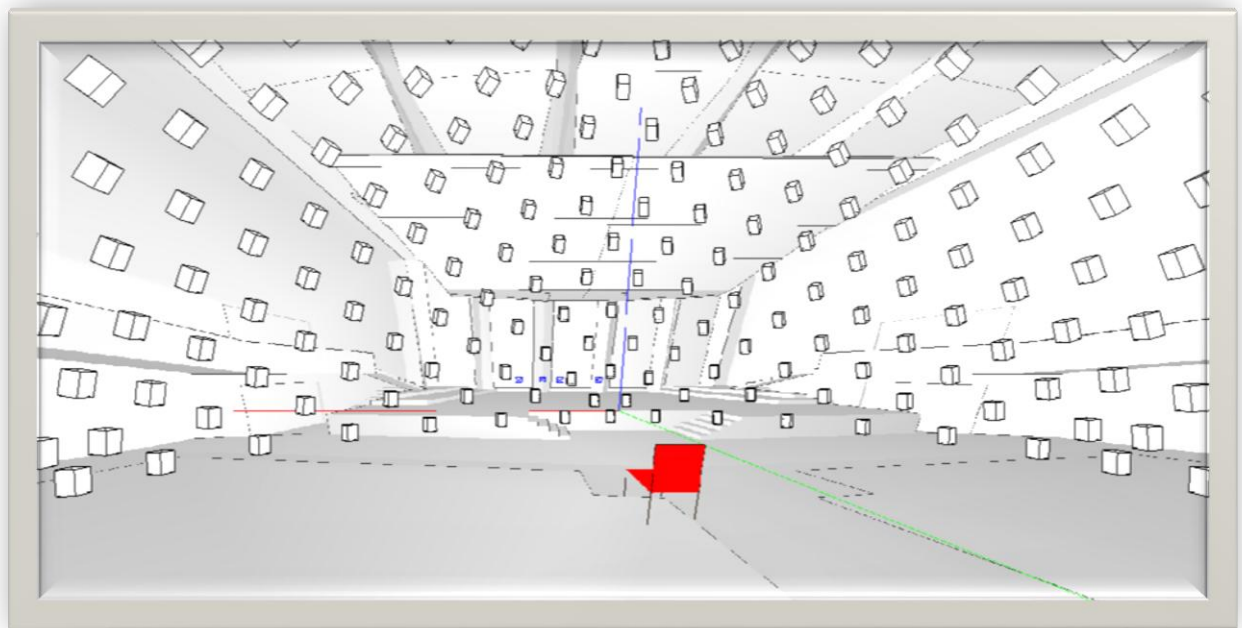# Simulation of array-based sound field synthesis methods

Simulation of a Sound Field Reproduction Approach for an
Innovative Musical Venue (III)

Simulation arraybasierter Schallfeldsyntheseverfahren

Simulation eines Beschallungskonzepts für eine
innovative musikalische Spielstätte (III)

Vorgelegt von:

Johnny Nahas

▄▄▄▄▄▄▄▄▄▄

Betreut von:

Prof. Dr. rer. nat. M. Vorländer, ITA an der RWTH-Aachen
Bruno Masiero, MSc

Prof. Dr. S. Weinzierl, Fachgebiet Audiokommunikation an der TU-Berlin
Alexander Lindau, M.A.
Frank Schultz, Dipl.-Ing. (FH)

Acknowledgments

First of all, I would like to thank Prof. Dr. Stefan Weinzierl for giving me the chance to participate in this project und to improve my knowledge and skills.

I would also like to specially thank Prof. Dr. Michael Vorländer from the RWTH Aachen for enabling me to carry out this thesis in Berlin, for his interest in my work and for reviewing this thesis.

My special thanks go to my supervisors both in Aachen and in Berlin; Bruno Masiero, Alexander Lindau and Frank Schultz, for their helpful advices and support till the last minute.

Further, I thank all members and students of the Audio communication group at the TU Berlin, and especially those who participated in the project, for helpful discussions.

Finally, my very special thanks go to my girlfriend Marga for her outstanding patience and support. And I would like to thank my family for all the support I get.

Table of Contents

Chapter 1

Introduction

Since several decades, the problem of physically recreating a given wave field has been addressed in the audio community. Different reproduction techniques have been developed with the intention of reconstructing sound fields over a finite region of space where a listener positioned in this region would experience the original sound field.

1.1 Project Background

The Klangakademie Berlin (Figure 1.1) is a joint project of the ADAM Audio GmbH and the Audio communication group at the Technical University in Berlin. The main goal of the overall project is to have controlled room acoustics in the planned, innovative musical venue. In order to achieve this, the reproduction of arbitrary room acoustics in the venue (quasi-anechoic) should be simulated *via an array based reproduction system*. This thesis deals with the evaluation of three suitable reproduction techniques: 1) Wave Field Synthesis (WFS), 2) Higher Order Ambisonics (HOA), 3) Vector Base Amplitude Panning (VBAP).

Figure 1.1: Rendered interior CAD-design of the Klangakademie Berlin.

Both the theoretical basis and practical implementation strategies for sound field reproduction using different loudspeaker arrays will be presented. The reproduction strategies of interest primarily concern Wave Field Synthesis (WFS) and Higher Order Ambisonics (HOA) techniques, which are based on the synthesis of the desired sound field to be reproduced. However, nowadays other classical periphonic reproduction techniques are reconsidered and optimized for different applications; as an example the Vector Base Amplitude Panning (VBAP) technique will be discussed.

The implementation part of this work will give an overview over the three different reproduction techniques for evaluation purposes. First of all, reproduction systems for real and virtual environments were designed that could be used for simulations of real world applications. Eventually, by taking the simulation a step further, the reproduction system was applied to reproduce desired room acoustics. The central question was to verify which (electro-acoustical) reproduction system was best able to reproduce the desired room acoustics (Figure 1.2). This final result is presented as a working demonstrator for all three sound reproduction methods.

More specifically, so-called driving functions (interpreted as filters) for different reproduction techniques were derived and the dynamic auralization of a simulated room acoustics were developed and demonstrated.
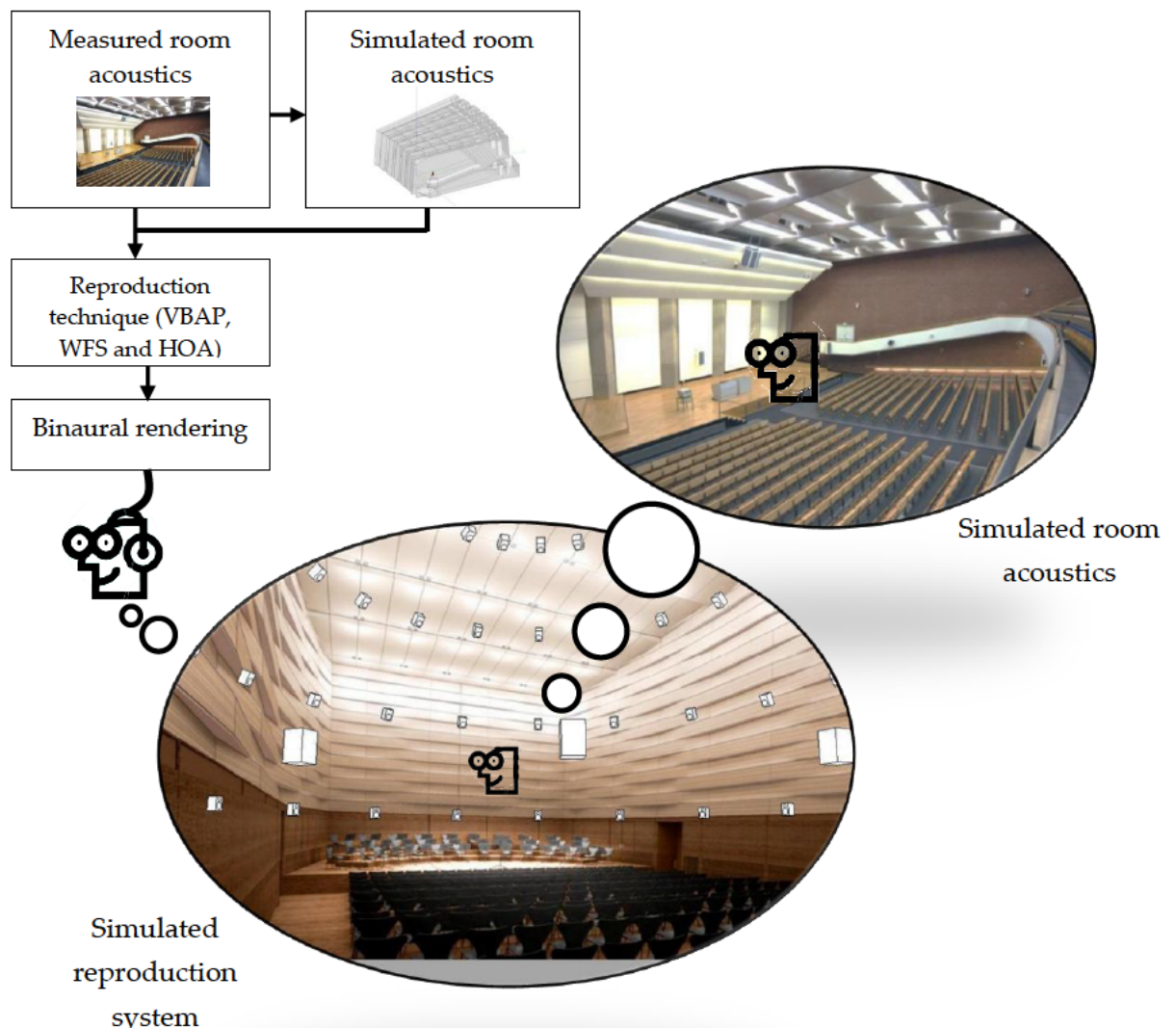
Figure 1.2: Graphical illustration of this thesis' content

1.2 Thesis Outline

After having clarified the goals of the thesis, the structure of the rest of this thesis will be presented here.

Chapter 2 introduces the theory and terminology of sound wave field propagation. This includes basically the derivation of the wave equation solutions in different coordinate systems for different problems. The wave field description (synthesis) based on Kirchhoff-Helmholtz integral and Spherical Harmonics will be presented.

Chapter 3 introduces three reproduction techniques considered in this thesis. It provides theoretical and mathematical formulation for derivation of driving function meant to be implemented.

Chapter 4 will introduce the geometrical basics for designing loudspeaker layouts meeting different criteria for the reproduction techniques presented in Chapter 3.

Chapter 5 introduces the concept of Auralization, which simulates the binaural listening experience at a given position in a virtual acoustic environment.

Chapter 6 introduces the practical implementation steps performed to build a software tool in MATLAB. This enables designing loudspeaker layouts (Chapter 4), calculating the driving functions (Chapter 2) for a given reproduction technique and generation of binaural dataset for real-time, offline dynamic auralization experience, which will be demonstrated for a modelled room acoustics and set of simulated reproduction systems.

Chapter 7 summarizes the conclusions from this thesis and suggests future improvements.

Chapter 2

Fundamentals of Sound Propagation

Introduction

In this Chapter the fundamentals of sound propagation will be introduced as basis for the discussion of wave field analysis and sound reproduction in the remainder of this work. The Chapter is outlined as follows: first the wave equation and its homogeneous solutions are derived, followed by a discussion of the inhomogeneous wave equation and reasonable choices for boundary conditions. Then the solution to the inhomogeneous wave equation with respect to arbitrary boundary conditions will be presented. The mathematical basics needed to understand the concepts presented in this thesis will be discussed[1].

2.1 Derivation of the Homogeneous Acoustic Wave Equation

The acoustic wave equation provides the mathematical description of sound propagation through fluids. A homogeneous acoustic wave equation describes the acoustic wave propagation for the free-field case (no boundaries are present) of a source free volume. The derivation is basically referred to [Vor08; Wil99; Spo05]. In order to derive the lossless acoustic wave equation for case of acoustic wave propagation in air we assume that the propagation medium is homogeneous; this implies that the parameters of the medium are position independent. Furthermore, the propagation medium is quiescent; this assures that the parameters do not vary in time and that there is no gross movement of the medium. In addition, the propagation medium can be characterized as an ideal gas, whose state changes be modelled as adiabatic processes; this means that there is no energy exchange in form of heat conduction with the medium (no propagation losses). Finally, the pressure and density perturbations due to wave propagation are small compared to the static pressure p_0 and the static density ρ_0. This allows the field variables and medium characteristics to be linearized around an operating point [Vor08; Spo05].

---

[1] It was recommended to introduce the related mathematical background and derivations sufficiently (Chapter 2 and 3).

Nomenclature:

For the mathematical derivations and equations presented in this thesis we define some conventions, as used in the referenced literature.

- For scalar variables lower case denotes the time domain (e.g. t, f), where upper case denotes the temporal or spatial frequency domain (e.g. $P(.)$). For scalar temporal in time or frequency domain variables we add Hat symbol above the variable (e.g. $\hat{p}(t), \hat{P}(\omega)$).
- Vectors are denoted by lower case bold face and could have index to refer to the Coordinate system considered or other useful declarations (e.g. $\boldsymbol{x}, \boldsymbol{x}_{index}$). A vector in space will be defined by its coordinates as (3×1) vector formed as its transpose vector $\boldsymbol{x} = [x_1 \ x_2 \ x_3]^T$.

The wave equation can be derived from the conservation of mass principle and the momentum equation [Wil99]. The first principle describes the mass balance in an infinitesimal volume element. Its mathematical formulation is given as follows:

$$\frac{\partial \rho}{\partial t} + \rho \nabla v(\boldsymbol{x}, t) = 0 \ , \tag{2.1}$$

where ρ denotes the density of the propagation medium, ∇ the *nabla operator* and $v(\boldsymbol{x}, t)$[2] the acoustic particle velocity at the position $\boldsymbol{x}$ and the time t. The time derivative of the density ρ in Eq. (2.1) can be expressed by the acoustic pressure $p(\boldsymbol{x}, t)$ can be found as:

$$\frac{\partial p(\boldsymbol{x}, t)}{\partial t} = c^2 \frac{\partial \rho}{\partial t} \ , \tag{2.2}$$

where c denotes the speed of sound. The speed of sound is dependent on the characteristics of the propagation medium. For air, it mainly depends on the temperature and the humidity. A value which reflects typical conditions (T =20 °C, 50% relative humidity) for wave propagation in air c = 343 [m/s]. Using Eq. (2.2) to eliminate the temporal derivative of the density ρ in Eq. (2.1) yields

$$-\frac{\partial p(\boldsymbol{x}, t)}{\partial t} = \rho_0 c^2 \nabla v(\boldsymbol{x}, t) \ , \tag{2.3}$$

where ρ_0 denotes the static density of air. The momentum equation (also called Newton's equation) relates the force applied to an infinitesimal volume (mass) element to the acceleration of that volume element as a result of the force. This results the *Euler's* equation [Wil99] as

$$\rho_0 \frac{\partial v(\boldsymbol{x}, t)}{\partial t} = -\nabla p(\boldsymbol{x}, t) \tag{2.4}$$

---

[2] We refer to [Vor08] for better understanding of the sound field physical characteristics.

Equation (2.3) can be used to eliminate the acoustic particle velocity $v(\boldsymbol{x}, t)$ from *Euler's* equation (2.4). The result is the well-known homogeneous acoustic wave equation

$$\nabla^2 p(\boldsymbol{x}, t) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}p(\boldsymbol{x}, t) = 0 \ , \tag{2.5}$$

where ∇^2 operator is referred as *Laplace operator* $\nabla^2 = \Delta$ [AW01]. The Laplace operator in the wave equation (2.5) is independent from the particular coordinate system used for the position vector $\boldsymbol{x}$ and has to be specified to the particular coordinate system used.

The wave equation can be described in frequency domain by applying a Fourier Transformation [OL10] of the acoustic pressure $p(\boldsymbol{x}, t)$ and the particle velocity $v(\boldsymbol{x}, t)$. The Fourier transform pair of the acoustic pressure with respect to the time is given as

$$P(\boldsymbol{x}, \omega) = \mathcal{F}_t\{p(\boldsymbol{x}, t)\} = \int_{-\infty}^{\infty} p(\boldsymbol{x}, t)e^{-j\omega t}dt \tag{2.6a}$$

$$p(\boldsymbol{x}, t) = \mathcal{F}_t^{-1}\{P(\boldsymbol{x}, \omega)\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} P(\boldsymbol{x}, \omega)e^{j\omega t}dt \ , \tag{2.6b}$$

where $\omega = 2\pi f$ denotes the temporal (radial) frequency and $\mathcal{F}_t\{.\}$ the Fourier transformation with respect to the time t. Introducing $P(\boldsymbol{x}, \omega)$ into the wave equation (2.5) and applying the differentiation theorem of the Fourier transformation [OL10] derives the wave equation formulated in the frequency domain

$$\nabla^2 P(\boldsymbol{x}, \omega) + \left(\frac{\omega}{c}\right)^2 P(\boldsymbol{x}, \omega) = 0 \ . \tag{2.7}$$

This form is known as the *Helmholtz* equation. In general, the term ω/c denotes the acoustic wavenumber k

$$k^2 = k^2(\omega) = \left(\frac{\omega}{c}\right)^2 \tag{2.8}$$

It will be assumed in the sequel, that the wavenumber can be expressed by the temporal frequency using Eq. (2.8) whenever needed. The wavelength λ, measured in m, is given by $\lambda = \frac{c}{f} = \frac{2\pi}{k}$.

2.2 Solutions of the Homogeneous Wave Equation in Cartesian Coordinates

This Section introduces the free-field solutions of the homogeneous wave equation (2.5) formulated in Cartesian coordinates. The vectors and functions evaluated in this coordinate system will be denoted by the index C attached to the respective variables. A well-known solution to the wave equation formulated in Cartesian coordinates is given as [Spo05]

$$p_c(\boldsymbol{x_c}, t) = f(ct - \boldsymbol{n_c^T} \boldsymbol{x_c}) \tag{2.9}$$

where $f(\cdot)$ denotes an arbitrary function and $\boldsymbol{n_c^T}$ a normal vector with $|\boldsymbol{n_c^T}| = 1$. Eq. (2.9) describes propagating wave fronts with the shape of $f(\cdot)$ that propagate with the speed of sound into the direction given by $\boldsymbol{n_c}$. This type of waves is termed as plane waves. Figure 2.1 illustrates an arbitrary shaped plane wave travelling in two-dimensional space. Performing a temporal Fourier transformation of Eq. (2.9) according to Eq. (2.6a) yields [Spo05]

$$P_c(\boldsymbol{x_c}, \omega) = \underbrace{\frac{1}{c} F\left(\frac{\omega}{c}\right)}_{\hat{P}(\omega)} e^{-i\frac{\omega}{c} \boldsymbol{n_c^T} \boldsymbol{x_c}} \quad, \tag{2.10}$$

where $\hat{P}(\omega)$ denotes the temporal spectrum (frequency dependent parts), as denoted above and $\boldsymbol{k_c^T} = \frac{\omega}{c} \boldsymbol{n_c^T}$, allows to derive a more compact form of the frequency-domain solution as

$$P_c(\boldsymbol{x_c}, \omega) = \hat{P}(\omega) e^{-i \boldsymbol{k_c^T} \boldsymbol{x_c}} \tag{2.11}$$
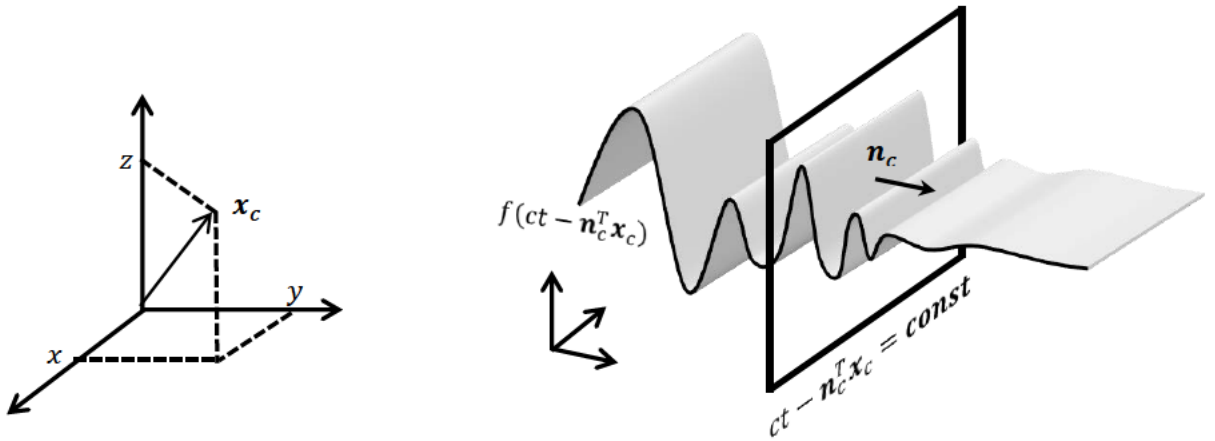
Figure 3.1: Left: Illustration of solution of the wave equation in Cartesian coordinates. Right: An arbitrary shaped plane wave $f(ct - \boldsymbol{n_c^T} \boldsymbol{x_c})$ traveling in two-dimensional space is shown for a fixed time t (right). The plane of constant value ($ct - \boldsymbol{n_c^T} \boldsymbol{x_c} = const$) moves with the speed of sound c into the direction given by $\boldsymbol{n_c}$.

Introducing Eq. (2.1) into the homogeneous Helmholtz equation (2.7) yields

$$k^2 = k_x^2 + k_y^2 + k_z^2 = |\boldsymbol{k}_c|^2 \qquad (2.12)$$

Equation (2.12) states that the acoustic wavenumber k is equal to the length of the vector $\boldsymbol{k}_c$, which will denote the wave vector of a plane wave. Each wave vector $\boldsymbol{k}_c$ belongs to a specific (temporal) frequency ω_0. A signal in the time domain of the form $e^{-i\omega_0 t}$ is called a monofrequent or monochromatic signal. The constant $\omega_0 = 2\pi f_0$ denotes the angular frequency, where f_0 is the number of cycles per second the signal exhibits. Due to these considerations, the term $e^{-i\,\boldsymbol{k}_c^T\,\boldsymbol{x}_c}$ will be denoted as monochromatic plane wave. The wave vector $\boldsymbol{k}_c$ of a plane wave can be interpreted as a vector consisting of the spatial frequencies $\boldsymbol{k}_c = [\,k_x\,k_y\,k_z]^T$, where each spatial frequency denotes 2π times the number of cycles per meter of the monochromatic plane wave in x, y and z-direction. Figure 2.2 shows the pressure field of a monochromatic plane wave traveling in the x-y plane.
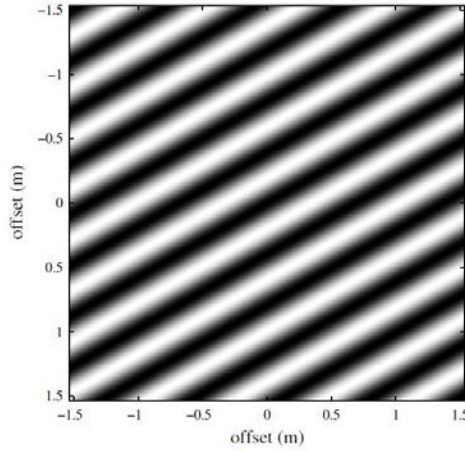
Figure 2.2: Pressure field of a traveling monochromatic plane wave the x-y-plane. The gray level denotes the amplitude.

2.3 Solutions of the Homogeneous Wave Equation in Cylindrical Coordinates

In this Section the free-field solutions of the homogeneous wave equation (2.5) formulated in cylindrical coordinates will be considered. We use the index **Y** for denoting all vectors and functions evaluated in Cylindrical Coordinate system. This Section is mainly based on the work of [Wil99] and [Spo05]. The wave equation, as given by Eq. (2.5), can be specified straightforwardly to the case of Cylindrical Coordinates

$$\nabla^2 p_{\mathrm{Y}}(\boldsymbol{x}_{\mathrm{Y}}, t) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\, p_{\mathrm{Y}}(\boldsymbol{x}, t) = 0 \quad , \qquad (2.13)$$

where the Laplace operator $\nabla^2 = \Delta$ is given as [Wil99]

$$\nabla^2 = \Delta = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial \alpha^2} + \frac{\partial^2}{\partial z^2} \qquad (2.14)$$

and $x_Y = [\alpha\ r\ z]^T$ denotes the position in Cylindrical Coordinate system (Figure 2.3).
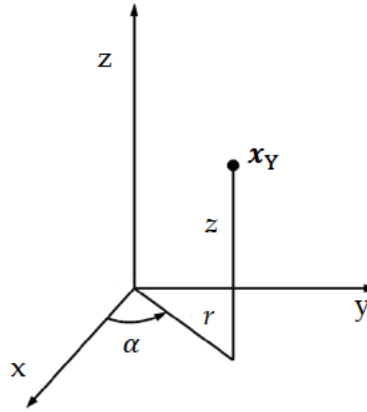
Figure 2.3: Illustration of the Cylindrical Coordinate system.

A standard technique often utilized to solve partial differential equations (PDE) of this type is the separation of variables [AW01]. The solution of a PDE can be written in terms of a product of functions which are only dependent on one variable. Applying this principle to the wave equation (2.15) states that the solution can be written as a product of functions which are only dependent from one of the three spatial variables α, r, z and the time t as

$$p_Y(\alpha, r, z, t) = p_\alpha(\alpha) \cdot p_r(r) \cdot p_z(z) \cdot p_t(t) \tag{2.15}$$

Introducing the solution (Eq. 2.15) into the wave equation (2.13) results in three ordinary differential equations of second order for $p_\alpha(\alpha)$, $p_z(z)$ and $p_t(t)$. The solution to the radial part $p_r(r)$ is given by Bessel's differential equation [Wil99]

$$\frac{d^2 p(r)}{dr^2} + \frac{dp_r(r)}{rdr} + \left(k_r^2 - \frac{v^2}{r^2}\right) p_r(r) = 0 \tag{2.16}$$

The solutions of Bessel's differential equation are given by the Bessel functions of first, second and third kind.

$$p_r(r) = R_1 J_n(k_r r) + R_2 Y_n(k_r r) \quad, \tag{2.17}$$

where R_1, R_2 are arbitrary constants. The Bessel functions of the first kind and non-negative integer degree n of r, denoted by $J_n(.)$ and are the solutions of Bessel's differential equation that are finite at the origin r=0. They can be defined by its Taylor series expansion around $r = 0$ as follows [Wil99]

$$J_n(r) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!\,\Gamma(m+n+1)} \left(\frac{r}{2}\right)^{2m+n} \tag{2.18}$$

where $\Gamma(.)$ is the gamma function, a generalization of the factorial function to non-integer values. Figure 2.4 shows the Bessel function behavior, which roughly looks like oscillating sine or cosine functions that decay proportionally to$1/\sqrt{x}$. For integer order, the following relationship is valid$J_{-n}(x) = (-1)^n J_n(x)$ [Wil99].

10
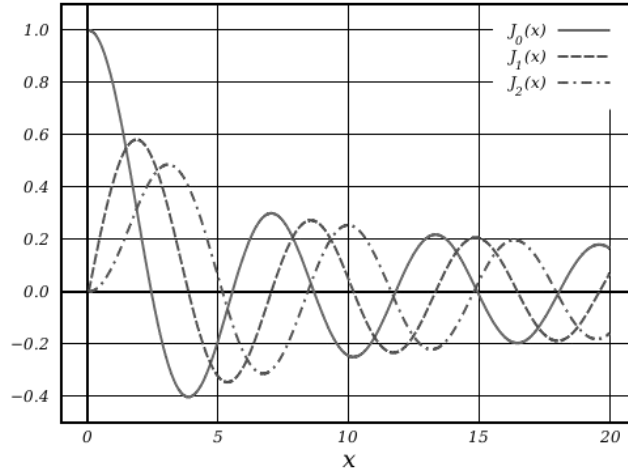
Figure 2.4: Bessel function of the first kind $J_n(x)$, for integer orders n= 0, 1, 2 [Wik1].

The Bessel functions of the second kind (Figure 2.5) also called the *Neumann* functions, denoted by $Y_n(x)$ are solutions of the Bessel differential equation. They have a singularity at the origin $(x = 0)$.
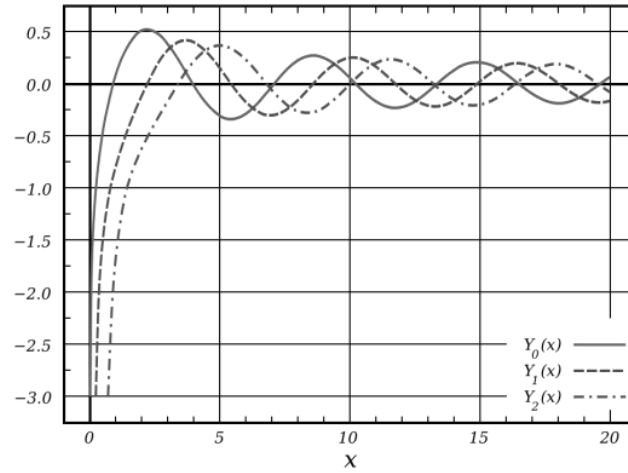
Figure 2.5: Bessel function of the second kind $Y_n(x)$, for integer orders n= 0, 1, 2 [Wik1].

A traveling wave solution for the Eq. (2.16) is given in terms of Hankel functions (Bessel functions of third kind) as follows [Spo05]

$$p_r(r) = R_1 H_n^{(1)}(k_r r) + R_2 H_n^{(2)}(k_r r) \tag{2.19}$$

where $H_n^{(1),(2)}(.)$ denotes the n-th order Hankel function of first/second kind, respectively and defined as follows

$$H_n^{(1)}(x) = J_n(x) + i\, Y_n(x) \tag{2.20}$$

$$H_n^{(2)}(x) = J_n(x) - i\, Y_n(x) \tag{2.21}$$

11

The general solution to the wave equation in cylindrical can be written to be proportional to [Wil99]

$$p_Y(x_Y, t) \propto e^{\pm in\alpha}.H_n^{(1),(2)}(k_r r).e^{\pm ik_z z}.e^{-i\omega t} \quad , \tag{2.22}$$

where n can be interpreted as the angular frequency, k_z as the wavenumber (or spacial frequency) in the z-direction and ω as the temporal frequency. The sign of the angular frequency n denotes the rotation direction, while the sign of k_z denotes the propagation direction of waves in the z-direction.

In order to find a similar interpretation for the radial part, the properties of *Hankel* functions have to be investigated. In the far-field ($k_r r \gg 1$) the *Hankel* functions can be approximated as follows [Wil99]

$$H_n^{(1)}(k_r r) \approx \sqrt{\frac{2}{\pi k_r r}} \; e^{i\left(k_r r - \frac{1}{2}n\pi - \frac{1}{4}\pi\right)} \tag{2.23a}$$

$$H_n^{(2)}(k_r r) \approx \sqrt{\frac{2}{\pi k_r r}} \; e^{-i\left(k_r r - \frac{1}{2}n\pi - \frac{1}{4}\pi\right)} \tag{2.23b}$$

These approximations of the Hankel functions can be used to express respectively the in- and outward-propagating cylindrical wave solutions of the cylindrical wave equation, respectively.

2.4 Solutions of the Homogeneous Wave Equation in Spherical Coordinates

In this Section the free-field solutions of the homogeneous wave equation (2.5) formulated in Spherical Coordinates will be considered. In this case all vectors and functions evaluated in this Coordinate system will be denoted by the index **S**. A point position in Cartesian coordinate system is given as $x = [x\ y\ z]^T$ is linked to the spherical coordinates $x_S = [\theta\ \varphi\ r]^T$ as follows $x = r\cos\theta\sin\varphi$, $y = r\sin\theta\cos\varphi$, and $z = r\cos\varphi$. The mathematical derivation done in this Section mainly referred to [Wil99]. Thus, the wave equation in Spherical Coordinates can be formulated as

$$\nabla^2 p_S(x_S,t) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\,p_S(x_S,t) = 0 \qquad , \qquad (2.24)$$

where the Laplace operator $\nabla^2 = \Delta$ is given in spherical Coordinate system as [Wil99]

$$\nabla^2 = \Delta = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\varphi}\frac{\partial}{\partial\varphi}\left(\sin\varphi\frac{\partial}{\partial\varphi}\right) + \frac{1}{r^2\sin^2\varphi}\frac{\partial^2}{\partial\theta^2}$$

and $x_S = [\theta\ \varphi\ r]^T$ denotes the position in spherical Coordinate system (Figure 2.6).
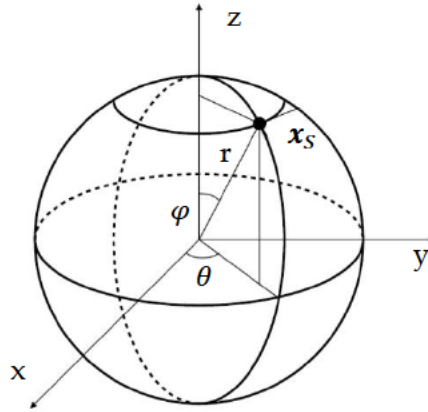
Figure 2.6: Illustration of the spherical Coordinate system.

Again the solution wave equation can be written as a product of functions which are only dependent from one of the three spatial variables θ, φ, r and the time t [Wil99]

$$p_S(\theta,\varphi,r,t) = p_\theta(\theta)\cdot p_\varphi(\varphi)\cdot p_r(r)\cdot p_t(t) \qquad (2.25)$$

To separate the Equation 2.25 into four differential equations in θ, φ, r and t the terms dependent on the respective other three quantities are replaced by a constant, which leads to four differential equations (Eq. 2.26 -2.29). The separated solutions will define important concepts for sound field description and reproduction technique, thus the derivation of the single solutions and the related mathematical basics will be introduced.

The separation of the variables in the Eq. (2.25) leads to the next four differential equations [Wil99]:

$$\frac{\partial^2 p_\varphi(\varphi)}{\partial \varphi^2} + m^2 p_\varphi(\varphi) = 0 \tag{2.26}$$

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial p_\theta(\theta)}{\partial\theta}\right) + \left(n(n+1) - \frac{m^2}{\sin^2\theta}\right)p_\theta(\theta) = 0 \tag{2.27}$$

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial p_r(r)}{\partial r}\right) + k^2\, p_r(r) - \frac{n(n+1)}{r^2}\, p_r(r) = 0 \tag{2.28}$$

$$\frac{1}{c^2}\left(r^2\frac{\partial p_t(t)}{\partial t^2}\right) + k^2\, p_t(t) = 0 \tag{2.29}$$

The solution of Eq. (2.29) with $k = \omega/c$ reads

$$p_t(t) = T_1 e^{-i\omega t} + T_2 e^{+i\omega t}$$

The solution of Eq. (2.26) is given as

$$p_\varphi(\varphi) = \Phi_1(m)e^{im\varphi} + \Phi_2(m)e^{-im\varphi}$$

where T_1, T_2, Φ_1, Φ_2 denote arbitrary constants. The solution of Eq. (2.27) is found using a transformation of variable. Let $\eta = cos\theta$, where $(-1 \leq \eta \leq 1)$ so that the differential equation (Legendre equation) for $p_\theta(\theta)$ becomes [Wil99]

$$\frac{\partial}{\partial\eta}\left((1-\eta^2)\frac{\partial p_\theta(\theta)}{\partial\eta}\right) + \left(n(n+1) - \frac{m^2}{1-\eta^2}\right)p_\theta(\theta) = 0 \tag{2.30}$$

The solution is given by *Legendre* functions of the first and second kinds, respectively,

$$p_\theta(\theta) = \Theta_1 P_n^m(\cos\theta) + \Theta_2 Q_n^m(\cos\theta),$$

where Θ_1, Θ_2 denote arbitrary constants .The Legendre functions of the second kind Q_n^m are finite at the poles where $\eta = \pm 1$ so this solution is discarded ($\Theta_2 = 0$). Further Information about Legendre functions are discussed in Chapter 2. The radial differential equation Eq. (2.28) can be rewritten as

$$\left(\frac{\partial^2}{\partial r} + \frac{2}{r}\frac{\partial}{\partial r} + k^2 - \frac{n(n+1)}{r^2}\right)p_r(r) = 0 \quad, \tag{2.31}$$

which would be Bessel's equation (2.16), except for the coefficient of $2/r$ instead of $1/r$. However, with suitable substitution, Eq. (2.31) can be transformed to Bessel's equation as follows [Wil99]

$$\text{Substitution: } p_r(r) = \frac{1}{r^{1/2}}\, u_r(r) \tag{2.32}$$

$$\left(\frac{\partial^2}{\partial r} + \frac{1}{r}\frac{\partial}{\partial r} + k^2 - \frac{n(n+1/2)^2}{r^2}\right)u_r(r) = 0 \tag{2.33}$$

14

Similarly to the solution of the Bessel's differential equation (Eq. 2.16) yields [Wil99]

$$p_r(r) = \frac{A_n}{r^{1/2}} J_{n+1/2}(kr) + \frac{B_n}{r^{1/2}} Y_{n+1/2}(kr) \ ,$$

where $J_{n+1/2}(.)$ and $Y_{n+1/2}(.)$ are Bessel functions of the first and second kind (see Section 2.3), respectively, and A_n and B_n denote arbitrary constants. Thus, a new form of Bessel functions, so-called spherical Bessel functions of the first and second kind can be defined as follows, respectively

$$j_n(x) = \left(\frac{\pi}{2x}\right)^{\frac{1}{2}} J_{n+\frac{1}{2}}(x) \tag{2.34}$$

$$y_n(x) = \left(\frac{\pi}{2x}\right)^{\frac{1}{2}} Y_{n+\frac{1}{2}}(x) \tag{2.35}$$

The spherical Hankel functions of the first and second kind (see Figure 2.7) in analogy to the Equations (2.20) and 2.21) are defined in terms of these solutions

$$h^{(1)}(x) = j_n(x) + iy_n(x) \tag{2.36}$$

$$h_n^{(2)}(x) = j_n(x) - iy_n(x) \tag{2.37}$$
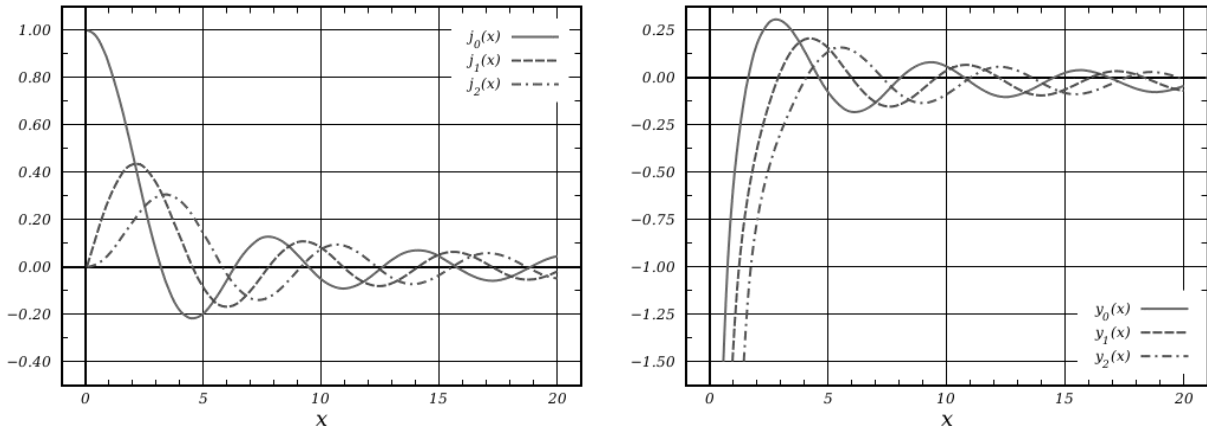
Figure 2.7: the spherical Bessel functions of the first (left) and second (right) kind for integer orders n= 0, 1, 2 [Wik1].

The angle functions are combined into a single function called a spherical harmonic Y_n^m defined by

$$Y_n^m(\theta, \varphi) = \sqrt{\frac{(2n+1)}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos \theta) e^{im\varphi} \tag{2.38}$$

Thus, any solution as outgoing traveling wave ($e^{-i\omega t}$) of Eq. (2.24) can be formulated as [Wil99]

$$P_S(\theta, \varphi, r, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left(A_{mn} j_n(kr) + B_{mn} y_n(kr)\right) Y_n^m(\theta, \varphi) \tag{2.39}$$

15

2.5 Solutions of the Inhomogeneous Wave Equation

The homogeneous acoustic wave equation (2.5) describes the acoustic wave propagation for the free-field case (no boundaries are present) of a source free volume. In order to calculate the wave field for generic scenarios, further information about the boundaries and the sources is required. Thus, the exact solution of the wave equation depends on [Spo05]

1. the initial conditions
2. the boundary conditions
3. the acoustic sources

The discussion of initial conditions lies beyond the scope of this thesis. It will be assumed that the acoustic pressure and velocity can be set to zero as initial condition. The boundary conditions will be briefly discussed in next Sections. This Section will introduce solutions of the inhomogeneous wave equation

$$\nabla^2 p(\boldsymbol{x}, t) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} p(\boldsymbol{x}, t) = -q(\boldsymbol{x}, t) \tag{2.40}$$

Or in Frequency domain (Helmholtz equation) as

$$\nabla^2 P(\boldsymbol{x}, \omega) + k^2 P(\boldsymbol{x}, \omega) = -Q(\boldsymbol{x}, \omega) \tag{2.41}$$

These solutions compromise acoustic sources (denoted by excitation $q(\boldsymbol{x}, t)$) as will be shown in the following Sections.

2.5.1 Point Source

The basis for the model of a point source is a radially oscillating (breathing) sphere [Vor08], which generates an outgoing and angle independent wave field (omni-directional). The model of a point source is derived when considering the limiting case for which the radius of the sphere becomes very small. The sphere will then degenerate to a single point in space. However, the same principle can be applied to nearly any arbitrary shaped source (with oscillating mass of fluid) if the dimensions of the source are small compared to the considered wavelength and the wave field is observed at a large distance compared to the source dimensions [Spo05]. The point source is a frequently used, ideal model for acoustic sources. Due to the omni-directional nature of the radiated pressure field it is convenient to use a spherical coordinate system to describe the wave field of a point source. The acoustic pressure field $P_S(\boldsymbol{x}_S, \omega)$ of a monochromatic point source placed at the origin is given as follows [Pie91]

$$P_S(\boldsymbol{x}_S, \omega) = P_S(r, \omega) = \hat{P}(\omega)\frac{1}{r}e^{-ikr} \quad, \tag{2.42}$$

where k denotes the wavenumber, r the radius and $\hat{P}(\omega)$ a frequency dependent pressure amplitude.

Transforming Eq. (2.42) back into the time-domain using the inverse Fourier transformation (2.6b) and its properties [OL10] yields

$$p_S(\boldsymbol{x}_S, t) = \frac{1}{2\pi r} \hat{p}\left(t - \frac{r}{c}\right)$$

(2.43)

This result states that Eq. (2.43) describes an outgoing spherical wave. The shape of the spherical wave in radial direction is given by $\hat{p}(t)$. The amplitude of a point source exhibits a $1/r$ decay and has a pole for $r = 0$. The point source is therefore also termed as acoustic monopole. This pole at $r = 0$ does not fit well into physical reality. However, the point source model provides a reasonable approximation well outside of this pole.

Evaluating the Helmholtz equation (2.7) for the point source yields

$$\nabla^2 P(\boldsymbol{x}, \omega) + k^2 P(\boldsymbol{x}, \omega) = -4\pi \hat{P}(\omega)\delta(\boldsymbol{x})$$

(2.44)

where $\delta(\boldsymbol{x})$ denotes a spatial Dirac pulse at the origin. Equation (2.44) states that the left hand side of the wave equation is equal to a spatial Dirac pulse multiplied by a frequency dependent factor. Thus, the excitation for the point source model (Eq. 2.42) is an infinitesimal small point in space. Figure 2.8 illustrates a snapshot of the pressure field a monopole source.
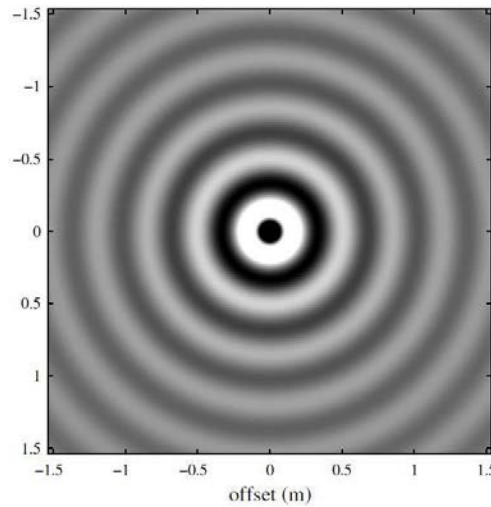
Figure 2.8: Illustration of snapshot of the pressure field of a 1 kHz monopole source. The plane wave is traveling in the x-y-plane (z=0).

2.5.2 Green's Functions

Green's functions provide a convenient way to compute arbitrary free field solutions of the inhomogeneous wave equation for arbitrary excitation $Q(x, \omega)$. The following Section reviews a background to the Green's functions referred to [Wil99].

The solution of the Eq. (2.44) can be generalized to the case of a point source placed at an arbitrary point x_0. The result is a shifted pressure field of a point source, which can be derived from Eq. (2.42) with $r = |x - x_0|$ as [Spo05]

$$G_{0,3D}(x|x_0, \omega) = \frac{1}{4\pi} \frac{e^{-i\frac{\omega}{c}|x - x_0|}}{|x - x_0|} \tag{2.45}$$

Introducing Eq. (2.45) into the Helmholtz equation (2.7) yields

$$\nabla^2 G_{0,3D}(x|x_0, \omega) + k^2 G_{0,3D}(x|x_0, \omega) = -\delta(x - x_0) \tag{2.46}$$

This special solution of the inhomogeneous wave equation is known as the free-field Green's function $G_{0,3D}(x|x_0, \omega)$ (or simply $G(x|x_0, \omega)$). Green's functions can be interpreted as the spatial-temporal impulse response [OL10] of the inhomogeneous wave equation.

Multiplying both sides of Eq. (2.46) by $Q(x, \omega)$ and integrating over the volume V' -which is identical to V except that it is associated with x' - yields, since ∇^2 does not depend on r' [Wil99]:

$$P(x, \omega) = \int_V Q(x', \omega) G_{0,3D}(x|x', \omega) dV' \tag{2.47}$$

Equation (2.47) indicates that once the Green function is known (the solution of Eq. (2.46)), then solutions to the general inhomogeneous wave equation, Eq. (2.41), are easily obtained by integration over the Green function. Figure 2.9 illustrates the geometry used in Eq. 2.47.
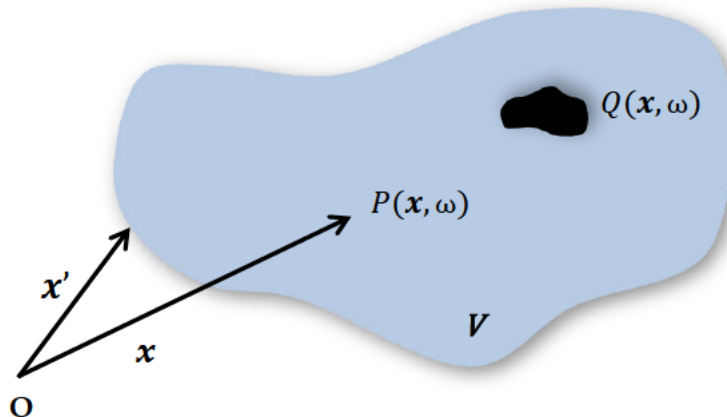
Figure 2.9: Illustration of the geometry used to describe general solution of the inhomogeneous wave equation for a defined region V. $Q(x, \omega)$ presents the source distribution and $P(x, \omega)$ its the pressure field.

2.5.3   Two-dimensional Free Field Green Function (Line Source)

Two dimensional acoustic fields have no dependence on the z coordinate (e. g. $P_Y(\alpha, r, z, \omega) = P_Y(\alpha, r, \omega)$). The analogon in this two-dimensional space to a point source in a three-dimensional space is the line source. This will present a line with infinite length in the z-direction and time-varying mass of fluid in radial direction. The concept of a line source is equal to the one of a two-dimensional point source for truly two-dimensional wave propagation. The wave field of a line source can be derived by calculating the field of an infinitely long radially oscillating cylinder [Spo05].

The inhomogeneous wave equation for a line source placed at the origin whose axis is perpendicular to the xy-plane is given as follows [Wil99; Spo05]

$$\nabla^2 P(\boldsymbol{x_c}, \omega) + k^2 P(\boldsymbol{x_c}, \omega) = -\underbrace{4\pi \hat{P}(\omega)\delta(x)\delta(y)}_{Q(\boldsymbol{x_c}, \omega)} \tag{2.48}$$

The solution to Eq. (2.48) is given in terms of the outgoing free-field Green's function by Eq. (2.47). Introduction of the line source excitation $Q(\boldsymbol{x_c}, \omega) = 4\pi \hat{P}(\omega)\delta(x)\delta(y)$ into (2.47) and exploitation of the sifting property of the Dirac-function [OL10] yields

$$P(\boldsymbol{x}, \omega) = \hat{P}(\omega) \int_V \delta(x')\delta(y')\, G_{0,3D}(\boldsymbol{x_c}|\boldsymbol{x_c'}, \omega).\, dV' \tag{2.49}$$

$$= \hat{P}(\omega) \int_{-\infty}^{\infty} \frac{e^{-jk\sqrt{x^2+y^2+(z-z')^2}}}{4\pi\sqrt{x^2+y^2+(z-z')^2}}.\, dz' \quad,$$

where $dV' = dx'\, dy'\, dz'$ denotes the volume element used for the first integral. The second integral can be solved using an integral definition of the Hankel function [AW01]. Due to symmetry of the problem it is useful to use a cylindrical coordinate system. Thus, the outgoing pressure field of a line source is given as follows

$$P_Y(\boldsymbol{x_Y}, \omega) = P(r, \omega) = \frac{i}{4}\hat{P}(\omega)H_0^{(2)}(kr) \quad, \tag{2.50}$$

where $H_0^{(2)}(.)$ is the zeroth Hankel function of the second kind. In order to derive a similar result for the amplitude decay ($1/r$ law) of the point source the Hankel function can be approximated as given by Equation 2.23b with $n = 0$ as

$$H_0^{(2)}(kr) \approx \sqrt{\frac{2}{\pi kr}}\, e^{-i\left(kr - \frac{1}{4}\pi\right)} \tag{2.51}$$

This approximation states, that the amplitude decay of a line source in the far field ($kr \gg 1$) will follow the behavior of $1/\sqrt{r}$. In the near-field the amplitude decay no simple conclusion can be drawn because of strong dependency on the small argument properties of the Hankel function [Spo05].

2.6 Boundary Value Problems

In the previous Sections only free-field propagation of acoustic waves was considered. In order to consider the wave propagation inside rooms (enclosures) the overall solution of the wave equation inside an enclosure must meet the acoustic conditions at the boundaries. This presents the Boundary Value Problems. Boundary value problems can be split up into three main types of problems, listed below [Zot09]. For simplicity the boundary values of the source-free field are in all cases known on one or two spheres, concentric to the origin. However the boundary problem classification can be introduced to arbitrary shaped boundaries (enclosures).

2.6.1 Interior Problem

The sound pressure distribution $p(\theta, \varphi, r_0)$ (Dirichlet problem) or particle velocity distribution $v(\theta, \varphi, r_0)$ (Neumann problem) due to sources outside r_0 is given. The interior free-field for $r \leq r_0$ is fully described mathematically and can be evaluated at every point.

This case represents the main interest in this work, e. g. sound field reproduction inside enclosed areas and will be mainly considered. Figure 2.10 illustrates the interior problem.
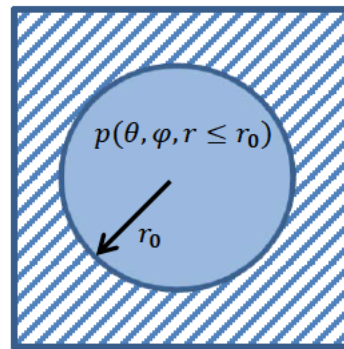
Figure 2.10: The pressure field inside the solid area (sphere) is fully described

2.6.2 Exterior Problem

A Dirichlet boundary value condition $p(\theta, \varphi, r_0)$ or a Neumann boundary value condition $v(\theta, \varphi, r_0)$ due to sources inside r_0 is given from measurements. The exterior free-field for $r \geq r_0$ is fully determined (Fig. 2.11).
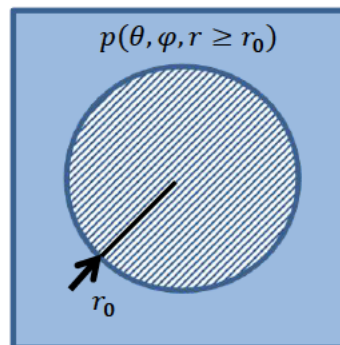
Figure 2.11: The pressure field outside the dashed area (sphere) is fully described.

2.6.3 Mixed Problems

a) Two Dirichlet boundary conditions defined by $p(\theta, \varphi, r_1)$ and $p(\theta, \varphi, r_2)$, or two Neumann boundary conditions $v(\theta, \varphi, r_2)$ and $v(\theta, \varphi, r_2)$ due to sources both inside r_2 and outside r_1 are known at two concentric spheres $0 < r_1 < r_2$. The free-field enclosed between the spheres is fully determined as Figure 2.11 illustrates.
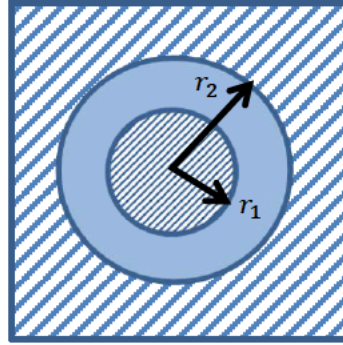
Figure 2.11: The pressure field enclosed between the spheres is fully described.

b) Two Dirichlet boundary conditions $p(\theta, \varphi, r_1)$ and $p(\theta, \varphi, r_2)$, or two Neumann boundary conditions $v(\theta, \varphi, r_1)$ and $v(\theta, \varphi, r_2)$ due to sources between r_1 and r_2 are known at two concentric spheres $0 < r_1 < r_2$. The two free-fields, one enclosed by the smaller sphere, the other one outside the larger sphere, are fully determined as Figure 2.12 illustrates.
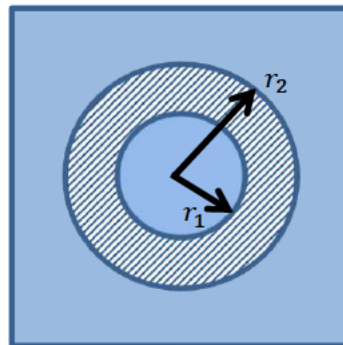
Figure 2.12: The pressure field inside the solid area is fully described

Notes:

a) Interior problems, or problems with irradiating fields, exhibit some exceptions. These are due to the zeros of the spherical Bessel function $j_n(kr)$ or its derivative $j_n'(kr)$ at specific frequencies. At those frequencies, the sound pressure or the sound particle velocity, respectively, cannot produce n-th order components of a field, alone. These exceptions will be discussed when the sound field reproduction techniques will be presented in Chapter 2.

b) The boundary conditions discussed in Section 2.6 can be classified also into two basic classes: homogeneous boundary conditions (stationary boundaries) and inhomogeneous boundary conditions (vibrating boundaries) [Sp005; Ahr10]. The total wave field inside the enclosure split into two components [Spo05]

$$P(\boldsymbol{x}, \omega) = P_s(\boldsymbol{x}, \omega) + P_b(\boldsymbol{x}, \omega) \quad , \tag{2.52}$$

where $P_s(\boldsymbol{x}, \omega)$ denotes the acoustic pressure wave field generated by sources present inside the enclosure and $P_b(\boldsymbol{x}, \omega)$ the wave field generated by the boundaries. Due to the main aim of the thesis, which considers the sound reproduction systems for enclosed listening area, it will focus on the wave field generated by the boundaries; without sources inside the enclosure (inhomogeneous boundary conditions).

2.7 The Kirchhoff-Helmholtz Integral (HIE)

The Kirchhoff-Helmholtz integral or Helmholtz Integral Equation (HIE(states that at any point within the source-free region V the sound pressure $P(\boldsymbol{x}, \omega)$ can be calculated if both the sound pressure $P(\boldsymbol{x_0}, \omega)$ and its directional derivative $\frac{\partial}{\partial \boldsymbol{n}} P(\boldsymbol{x_0}, \omega)$ are known on the boundary ∂V enclosing the volume V (Figure 2.13a). It is mathematically formulated as follows [Spo05]

$$P(\boldsymbol{x}, \omega) = -\oint_{\partial V} \left(G(\boldsymbol{x}|\boldsymbol{x_0}, \omega) \frac{\partial}{\partial \boldsymbol{n}} P(\boldsymbol{x_0}, \omega) - P(\boldsymbol{x_0}, \omega) \frac{\partial}{\partial \boldsymbol{n}} G(\boldsymbol{x}|\boldsymbol{x_0}, \omega) \right) dS_0 \tag{2.53}$$

HIE represents solutions to the homogeneous Helmholtz equation (2.7) with inhomogeneous boundary conditions (Note (b) p. 18) and it used to calculate a sound field emitted by a vibrating surface into a region or to calculate the a sound field inside a finite region produced by a source (also handled as virtual) outside the volume from measurements on the surface (Figure 2.13b).

Figure 2.13: (a) Geometry used for describing the Kirchhoff-Helmholtz integral. (b) Illustration of sound field reproduction (within V) of virtual source (outside V) based on Kirchhoff-Helmholtz integral formulation.

2.7.1 Three-dimensional Free-Space Kirchhoff-Helmholtz Integral

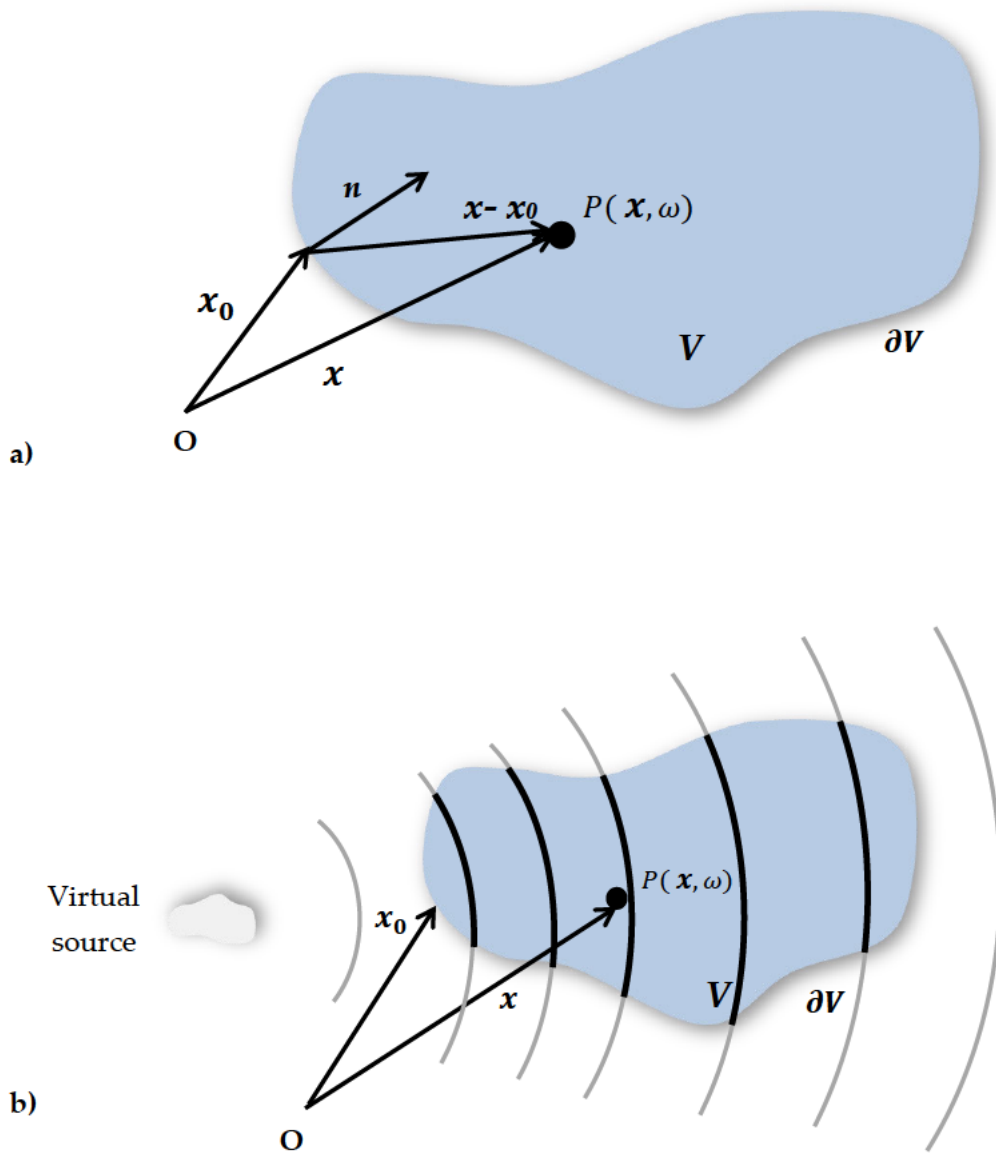The Green's function for a point source in free-space was derived in Section 2.5.2
Introduction of Green's function as given by Eq. (2.45) into the Kirchhoff-Helmholtz integral (2.53) yields

$$P(\boldsymbol{x},\omega) = -\frac{1}{4\pi}\oint_{\partial V}\left(\frac{\partial}{\partial \boldsymbol{n}}P(\boldsymbol{x_0},\omega)\frac{e^{-\mathrm{i}k|\boldsymbol{x}-\boldsymbol{x_0}|}}{|\boldsymbol{x}-\boldsymbol{x_0}|} - P(\boldsymbol{x_0},\omega)\frac{\partial}{\partial \boldsymbol{n}}\frac{e^{-\mathrm{i}k|\boldsymbol{x}-\boldsymbol{x_0}|}}{|\boldsymbol{x}-\boldsymbol{x_0}|}\right)dS_0 \,, \tag{2.54}$$

Equation 2.54 will be denoted by three-dimensional free-field Kirchhoff-Helmholtz integral in the following. Figure 2.13a illustrates the parameters used. The term $\frac{\partial}{\partial \boldsymbol{n}}P(\boldsymbol{x_0},\omega)$ can be interpreted as the particle velocity $V_n(\boldsymbol{x},\omega)$ on the surface ∂V in direction of the inward-pointing surface normal $\boldsymbol{n}$. It can be evaluated using Euler's equation (2.4) as the inner product [AW01] between the directional derivative of pressure and the normal $\boldsymbol{n}$ as follows

$$\frac{\partial}{\partial \boldsymbol{n}}P(\boldsymbol{x},\omega) = \langle \nabla P(\boldsymbol{x},\omega), \boldsymbol{n}\rangle = \langle -\rho_0\frac{\partial v(\boldsymbol{x},\omega)}{\partial t}, \boldsymbol{n}\rangle = -\mathrm{i}\omega\rho_0 V_n(\boldsymbol{x},\omega) \tag{2.55}$$

The term $\frac{e^{-\mathrm{i}k|\boldsymbol{x}-\boldsymbol{x_0}|}}{|\boldsymbol{x}-\boldsymbol{x_0}|}$ represents a monopole source distribution on the surface ∂V, where $P(\boldsymbol{x_0},\omega)$ the acoustic pressure on the surface ∂V.

The last term $\frac{\partial}{\partial \boldsymbol{n}}\frac{e^{-\mathrm{i}k|\boldsymbol{x}-\boldsymbol{x_0}|}}{|\boldsymbol{x}-\boldsymbol{x_0}|}$ constitutes the directional gradient of a monopole source. The gradient is given as follows [Spo05]

$$\frac{\partial}{\partial \boldsymbol{n}}\frac{e^{-\mathrm{i}k|\boldsymbol{x}-\boldsymbol{x_0}|}}{|\boldsymbol{x}-\boldsymbol{x_0}|} = -\frac{1+\mathrm{i}kr}{r}cos\varphi\frac{\partial}{\partial \boldsymbol{n}}\frac{e^{-\mathrm{i}k|\boldsymbol{x}-\boldsymbol{x_0}|}}{|\boldsymbol{x}-\boldsymbol{x_0}|} \tag{2.56}$$

It can be interpreted as the field of a dipole source placed on the surface, whose dipole axis lies in the direction of the surface normal $\boldsymbol{n}$ [Pie91].

2.7.2 Two-dimensional Free-Space Kirchhoff-Helmholtz Integral

This Section will introduce the two-dimensional wave fields as a special case of the three-dimensional Kirchhoff-Helmholtz integral (Eq. 2.54). Assuming that the wave field is independent of the z-coordinate the Kirchhoff-Helmholtz the surface ∂V will degenerate to a closed contour ∂S and the volume V to a surface S and a closed contour ∂S will surround the surface S (see Figure 2.1.4). The specialization of the three-dimensional free-space Kirchhoff-Helmholtz to two-dimensional geometry is therefore given as [Spo05]

$$P(\boldsymbol{x},\omega) = -\oint_{\partial S}\left(G_{0,2D}(\boldsymbol{x}|\boldsymbol{x_0},\boldsymbol{\omega})\frac{\partial}{\partial \boldsymbol{n}}P(\boldsymbol{x_0},\omega) - P(\boldsymbol{x_0},\omega)\frac{\partial}{\partial \boldsymbol{n}}G_{0,2D}(\boldsymbol{x}|\boldsymbol{x_0},\omega)\right)dL_0 \tag{2.57}$$

where dL_0 denotes a line element on ∂S. Figure 2.14 illustrates the geometry used for the two-dimensional, free-space Kirchhoff-Helmholtz integral.

The two-dimensional case of Kirchhoff-Helmholtz integral requires an appropriate two-dimensional Green's function. The two-dimensional free-space Green's function is a line source (Section 2.5.3) and can be generalized to an arbitrary source position x_0 and observation position x is given as [Spo05]

$$G_{0,2D}(x|x_0, \omega) = \frac{i}{4} H_0^{(2)}(k|x - x_0|) \tag{2.58}$$

The directional derivative of Green's function can be expressed as

$$\frac{\partial}{\partial n} G_{0,2D}(x|x_0, \omega) = \langle \nabla G_{0,2D}(x|x_0, \omega), n \rangle = \frac{ik}{4} H_1^{(2)}(k|x - x_0|) \cos\varphi \tag{2.59}$$

where φ denotes the angle between the inward pointing normal vector n of the closed contour ∂S and the vector $x - x_0$. Equation (2.59) can be interpreted as the field of a dipole line source whose axis lies parallel to the normal vector n [Spo05].
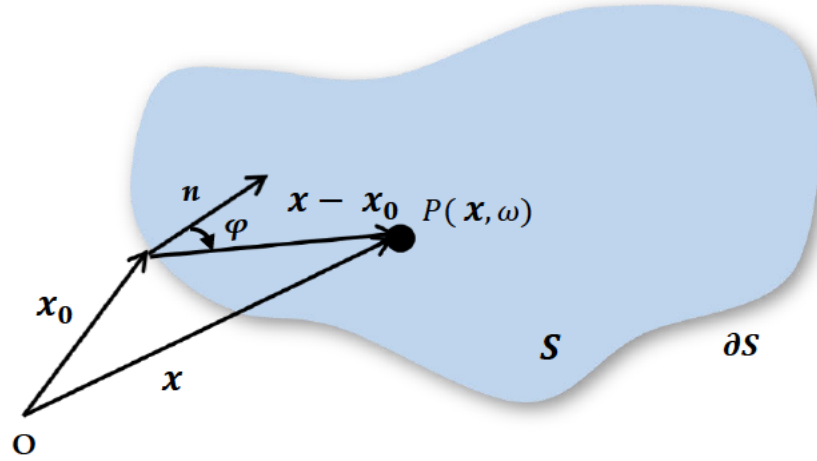
Figure 2.14: Geometry used for describing two-dimensional free-space Kirchhoff-Helmholtz integral

Chapter 3

Theory of Sound Field Synthesis

Introduction

Surround sound systems may be used for sound field reproduction without requiring information about the listener. The sound field is reconstructed over a finite region of space, and a listener positioned in this region may in principle experience the original sound field.

The theory of two-dimensional (2D) sound systems has received much attention, and a common implementation is the Ambisonics system [Ger85]. For regularly spaced, circular arrays, well-known panning driving functions may be derived by assuming an incident plane wave and applying a mode-matching procedure to determine the required loudspeaker weights [Pol05].

The theory of three-dimensional (3D) sound field reproduction has received increasing attention due to the potential it offers for more accurate reconstruction. There are several approaches to the problem. The Kirchhoff–Helmholtz integral shows that reproduction is possible inside a region if the pressure and normal velocity are known on the surface of the region. This is the basis for the wave field synthesis (WFS) approach. In practice simplifications can be made (WFS 2.5). A second approach is higher order Ambisonics (HOA) approach, which is based on a spherical harmonic decomposition of the sound field. The third approach is the inverse method, in which an inverse matrix is derived for a given geometry of loudspeakers, which allows the creation of the required sound pressure at a set of discrete points (e.g. VBAP).

In this Chapter we introduce approaches for sound reproduction based on the spherical harmonic/Bessel and Kirchhoff–Helmholtz descriptions of sound fields for bounded regions (interior problem). These methods are primarily concerned with wave field synthesis (WFS and WFS 2.5) and Higher Order Ambisonics (HOA).WFS and HOA present the sound field synthesis based reproduction methods. Additionally for ease of demonstration a panning based periphonic reproduction system (VBAP) will be also discussed.

3.1 Description of the Synthesized Sound Field

A 3D sound field of a point in space may be described in several ways. The Kirchhoff–Helmholtz integral is the mathematical form of Huygens' Principle. Alternatively, the sound field may be expanded in a Taylor series or a series of cylindrical or spherical coordinate eigenfunctions with corresponding cylindrical or spherical Bessel radial functions [Pol05].

3.1.1 Kirchhoff-Helmholtz Description

The Kirchhoff- Helmholtz integral (Section 2.7) can be illustrated with the principle of Huygens [Wik2]. Huygens stated that any point of a propagating wave front at any time - instant conforms to the envelope of spherical waves emanating from every point on the wave front at the prior instant. This principle can be used to synthesize acoustic wave fronts of arbitrary shape. By placing the loudspeakers on an arbitrary, fixed curve and by filtering the driving signals, an acoustic wave front of a virtual source can be synthesized with a loudspeaker array. Figure 3.1 illustrates this principle.
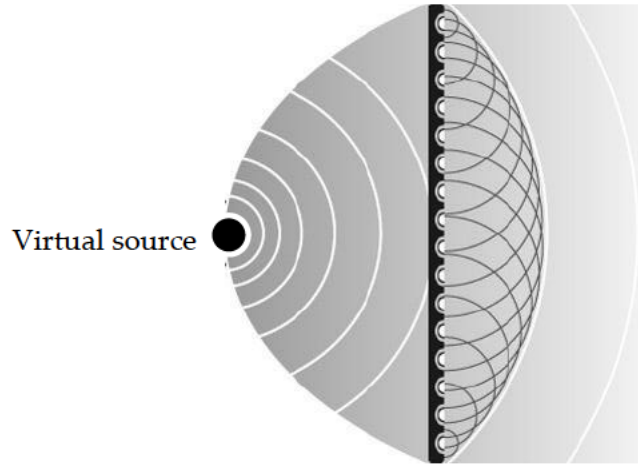
Figure 3.1: Illustration of Huygens principle to sound reproduction.

The solution of the homogeneous wave equation subject to inhomogeneous boundary conditions, as discussed in Section 2.7, is given by the Kirchhoff-Helmholtz integral (Eq. 2.53).

$$P(\boldsymbol{x}, \omega) = -\oint_{\partial V} \left(G(\boldsymbol{x}|\boldsymbol{x_0}, \omega) \frac{\partial}{\partial \boldsymbol{n}} P(\boldsymbol{x_0}, \omega) - P(\boldsymbol{x_0}, \omega) \frac{\partial}{\partial \boldsymbol{n}} G(\boldsymbol{x}|\boldsymbol{x_0}, \omega) \right) dS_0 \qquad (3.1)$$

Again this basic principle states that the sound pressure within a source-free bounded region V is fully determined by the inhomogeneous boundary conditions imposed on the closed surface ∂V surrounding V. The boundary conditions as introduced in Section 2.6 are given by a combination of the pressure itself and the directional derivative of the pressure of the wave field produced by the virtual source on ∂V [Spo08; Ahr10].

3.1.2 Spherical Harmonic/ Spherical Bessel Description

In spherical coordinates $\boldsymbol{x}_S = [\theta\ \varphi\ r]^T$ the solution to the wave equation may be written in terms of spherical Bessel functions and spherical harmonics. For the interior case (Section 2.6.1), where all sources lie outside the region of interest, the spatial variation of the sound field at a radial frequency ω, with c being the speed of sound, may be expressed as [Wil99]

$$P_S(\boldsymbol{x}_S, \omega) = P(\theta, \varphi, r, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} A_n^m(\omega) j_n\left(\frac{\omega}{c} r\right) Y_n^m(\theta, \varphi) \quad , \tag{3.2}$$

where $j_n(r)$ is the n-th spherical Bessel function of the first kind, $A_n^m(\omega)$ are so-called spherical harmonics expansion coefficients of the function $P_S(\theta, \varphi, r, \omega)$ and $Y_n^m(\theta, \varphi)$ are the Spherical Harmonics defined as

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

where $P_n^m(.)$ denotes the m-th order associated Legendre polynomial of the n-th degree [Wil99]. The mathematical basics of the Spherical Harmonics will be discussed in more detail in Section 3.5.1.

3.2 Sound Field Synthesis Based Reproduction

In this thesis synthesized sound field reproduction is introduced for interior domain (Interior problem) by the means of loudspeaker arrays based on following two approaches

1. Kirchhoff-Helmholtz Integral
2. Single layer Potential [Ahr10; Faz09]

The Kirchhoff-Helmholtz Integral approach aims to reproduce the synthesized wave field using simple monopole distribution surrounding the listening area. This can be done either by modification or approximation of Green's function in HIE. This approach will introduce WFS 3D and WFS 2.5 (Sections 3.3.1 and 3.3.2). The reproduced synthesized wave field $P(\boldsymbol{x}, \omega)$ inside the listening area V by the secondary source driving functions $D(\boldsymbol{x_0}, \omega)$ and the Green's functions $G(\boldsymbol{x}|\boldsymbol{x_0}, \omega)$ of the monopoles at the boundary ∂V is formulated as follows

$$P(\boldsymbol{x}, \omega) = \oint_{\partial V} D(\boldsymbol{x_0}, \omega) G_0(\boldsymbol{x}|\boldsymbol{x_0}, \omega)\, dS_0 \tag{3.3}$$

Single layer Potential approach formulates the synthesized wave field as *single layer potential* [Ahr10]. This will be achieved by calculating the driving function in Eq. 3.3 using the Spherical Harmonics expansion presentation of the desired synthesized wave field. Single layer potential approach will introduce basically the concepts of near field compensated higher order Ambisonics for three dimensional case (NFC-HOA) and two-dimensional case (NFC-HOA 2.5). The secondary source distributions used will be basically spherical [Ahr10; Faz09].

3.3 Wave Field Synthesis

Wave field synthesis (WFS) is a sound reproduction technique, which is based on the Kirchhoff-Helmholtz integral. However, in order to realize the system several theoretical and practical problems have to be solved. These constitute the concept of WFS, which was initially developed by the Technical University of Delft and has been developed further during the past two decades. One impressive practical example is the WFS-system built in the Technical University of Berlin.

In this thesis the two-dimensional WFS-systems will be basically considered. It was stated in Section 2.7.2 that a two-dimensional sound reproduction system can be realized by appropriately driving a monopole line source distribution surrounding the listening area. In theory, these secondary line sources would have to be of infinite length, which is impractical. Therefore, utilizing a distribution of monopole point sources, which can be approximated by loudspeakers, is reasonable for the realization of such a system. This approximation will introduce a modification of the original formulation of Kirchhoff-Helmholtz integral. The derivation of the driving function is done according to [Spo08].

3.3.1 Derivation of WFS (3D) –Driving Function

The solution of the homogeneous wave equation for a bounded region V with respect to inhomogeneous boundary conditions is given by the Kirchhoff-Helmholtz integral (Eq. 3.53)

$$P(\boldsymbol{x},\omega) = -\oint_{\partial V} \left(G(\boldsymbol{x}|\boldsymbol{x_0},\omega)\frac{\partial}{\partial \boldsymbol{n}} P(\boldsymbol{x_0},\omega) - P(\boldsymbol{x_0},\omega)\frac{\partial}{\partial \boldsymbol{n}} G(\boldsymbol{x}|\boldsymbol{x_0},\omega) \right) dS_0 \tag{3.4}$$

In order to use the monopole point source approximation as secondary source the second term in the Kirchhoff-Helmholtz integral, which presents a dipole source, must be eliminated. This can be achieved by modifying the Green's function. The modified Green's function $G_N(\boldsymbol{x}|\boldsymbol{x_0},\omega)$ has to obey the following condition

$$\frac{\partial}{\partial \boldsymbol{n}} G_N(\boldsymbol{x}|\boldsymbol{x_0},\omega)|_{x_0 \in \partial V} = 0 \tag{3.5}$$

Condition (3.5) formulates a homogeneous Neumann boundary condition imposed on the boundary ∂V. The modified Green's function is typically termed Neumann Green's function. A suitable Neumann Green's function for a planar/linear boundary ∂V is given by source term constituting a line source by the one constituting a monopole point source

$$G_N(\boldsymbol{x}|\boldsymbol{x_0},\omega) = G_{0,3D}(\boldsymbol{x}|\boldsymbol{x_0},\omega) + G_{0,3D}(\boldsymbol{x_m}(\boldsymbol{x})|\boldsymbol{x_0},\omega) \tag{3.6}$$

A solution fulfilling Eq. (3.6) is given by choosing the receiver point $x_m(x)$ as the point x mirrored at the planar boundary ∂V at the position x_0. Due to the specialized geometry, where $|x - x_0| = |x_m - x_0|$ and therefore $G_{0,3D}(x|x_0, \omega) = G_{0,3D}(\underbrace{x_m(x)}_{-x} |x_0)$ yields

$$G_N(x|x_0, \omega) = 2G_{0,3D}(x|x_0, \omega)$$

In this special case $G_N(x|x_0, \omega)$ is equal to a point source with double strength. Thus, the Equation (3.4) can be now take the general formulation

$$P(x, \omega) = -\oint_{\partial V} D(x_0, \omega) 2G_{0,3D}(x|x_0, \omega) \, dS_0 \;, \tag{3.7}$$

where $D(x_0, \omega)$ the driving function of the point source $G(x|x_0, \omega)$ at the position (x_0). The Equation (3.7) will introduce *Rayleigh* Integral [Wil99].
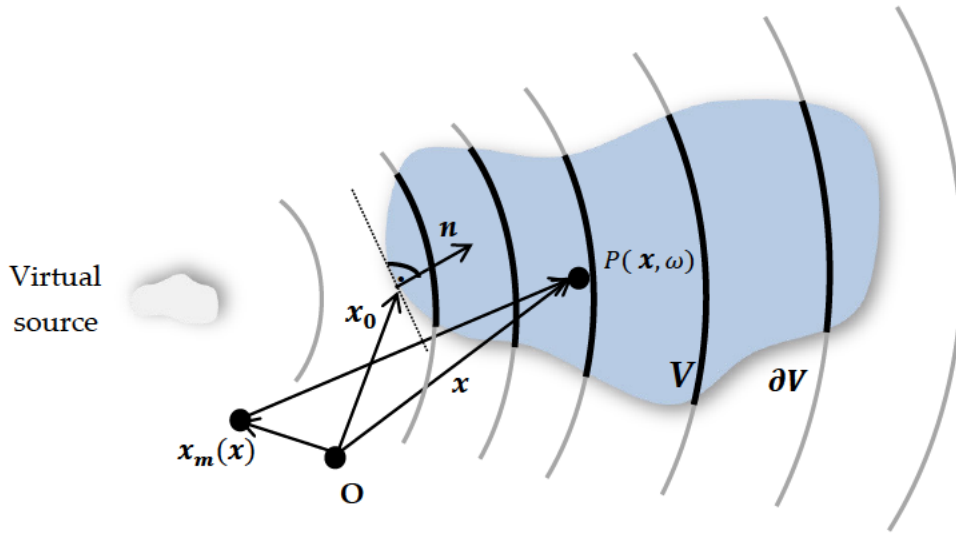
Figure3.2a: WFS 3D- Reproduction of sound field of a virtual source using monopole point source as secondary source.

It is assumed that Eq. (3.6) holds also approximately for other geometries [Spo05]. In this case the receiver point x_m is chosen as the point x mirrored at the tangent to the boundary ∂V at the position x_0 (Figure 3.2a). The elimination of the secondary dipole sources for an arbitrary secondary source contour ∂V will introduce a non-zero value of wave field outside of V and mismatch with desired virtual source field within V, which is a consequence of approximating the Neumann Green's function, which can be appropriate for a particular geometries. Using Neumann Green's function (Eq. 3.6) will produce undesired reflections secondary sources where the local propagation direction of the virtual wave field does not coincide with the normal vector n of the secondary source [Spo05].

To reduce the undesired reflections a windowing function could be applied, which will mute those secondary sources reproducing the undesired reflections. Thus, the reproduced wave field for a virtual source $S(x, \omega)$ reads [Spo05]

$$P(x, \omega) = -\oint_{\partial V} \underbrace{2a(x_0) \frac{\partial}{\partial n} S(x_0, \omega)}_{D_{3D}(x_0, \omega)} G_{0,3D}(x|x_0, \omega) dS_0 \quad , \tag{3.8}$$

where $a(x_0)$ denotes a suitable window function. It was proposed in [Spo07] to formulate this condition analytically on basis of the acoustic intensity vector. The Equation (3.8) provides a reasonable approximation for sound reproduction purposes.

3.3.2 Derivation of WFS (2.5) –Driving Function

For practical realization of WFS linear distribution of sources in a plane will be considered as two-dimensional WFS. This leads to generation of the wave field by a distribution of monopole line sources on the closed curve ∂S [Spo05]

$$P(x, \omega) = -\oint_{\partial S} D_{2D}(x_0, \omega) G_{0,2D}(x|x_0, \omega) dL_0 \tag{3.9}$$

Equivalent to

$$P(x, \omega) = -\frac{i}{4} \oint_{\partial S} D_{2D}(x_0, \omega) H_0^{(2)}\left(\frac{\omega}{c}|x - x_0|\right) dL_0 \tag{3.10}$$

Due to the practical application of two-dimensional WFS where loudspeakers approximated as monopoles are meant to be used it is useful to approximate the two-dimensional Green's function as three- dimensional green function. This can be performed considering the asymptotic expansion of the Hankel functions for large arguments [Spo08] is used to approximate the two-dimensional Green's function $G_{0,2D}(x|x_0, \omega)$ (Eq. 2.50) as follows

$$G_{0,2D}(x|x_0, \omega) \approx \sqrt{\frac{2\pi|x - x_0|}{i\frac{\omega}{c}}} \underbrace{\frac{1}{4\pi} \frac{e^{-i\frac{\omega}{c}|x-x_0|}}{|x - x_0|}}_{G_{0,3D}(x|x_0, \omega)} \tag{3.11}$$

Equation (3.11) states that $G_{0,2D}(x|x_0, \omega)$ is equivalent to $G_{0,3D}(x|x_0, \omega)$ when applying a spectral and amplitude correction. Introducing the approximation (3.11) of the two-dimensional free-field Green's function into Eq. (3.9) and modified three-dimensional free-field Green's function yields

$$P(x, \omega) = -\oint_{\partial S} \sqrt{\frac{2\pi|x - x_0|}{i\frac{\omega}{c}}} D_{2D}(x_0, \omega) G_{0,3D}(x|x_0, \omega) dL_0 \tag{3.12}$$

The required spectral correction denoted by $\sqrt{\frac{1}{i\frac{\omega}{c}}}$ is independent from the receiver position x. However, the amplitude correction $\sqrt{2\pi|x - x_0|}$ depends on the receiver position. As a consequence, the amplitude can only be corrected for one receiver position in the listening area, which will be denoted as a reference position x_{ref}. Thus, the corrected driving function is then given as

$$D_{2.5D}(\boldsymbol{x_0}, \omega) = \sqrt{\frac{2\pi|\boldsymbol{x_{ref}} - \boldsymbol{x_0}|}{i\frac{\omega}{c}}} D_{2D}(\boldsymbol{x_0}, \omega) \quad , \tag{3.11}$$

where the notation 2.5D denotes the combining of two dimensional reproduction using point sources as secondary sources.

 The reproduced two-dimensional wave field for a virtual source $S(\boldsymbol{x}, \omega)$ created by a distribution of monopole point sources on the closed curve ∂S reads

$$P(\boldsymbol{x}, \omega) = -\oint_{\partial S} D_{2.5D}(\boldsymbol{x_0}, \omega) \, G_{0,3D}(\boldsymbol{x}|\boldsymbol{x_0}, \omega) dL_0 \quad , \tag{3.12}$$

where

$$D_{2.5D}(\boldsymbol{x_0}, \omega) = 2a(\boldsymbol{x_0}) \sqrt{\frac{2\pi|\boldsymbol{x_{ref}} - \boldsymbol{x_0}|}{i\frac{\omega}{c}}} \frac{\partial}{\partial \boldsymbol{n}} S(\boldsymbol{x_0}, \omega) \tag{3.13}$$

Similarly to Geometry used in Figure 3.14 we illustrate again the reproduction of two dimensional wave field of a virtual source $S(\boldsymbol{x}, \omega)$ in Figure 3.2b.
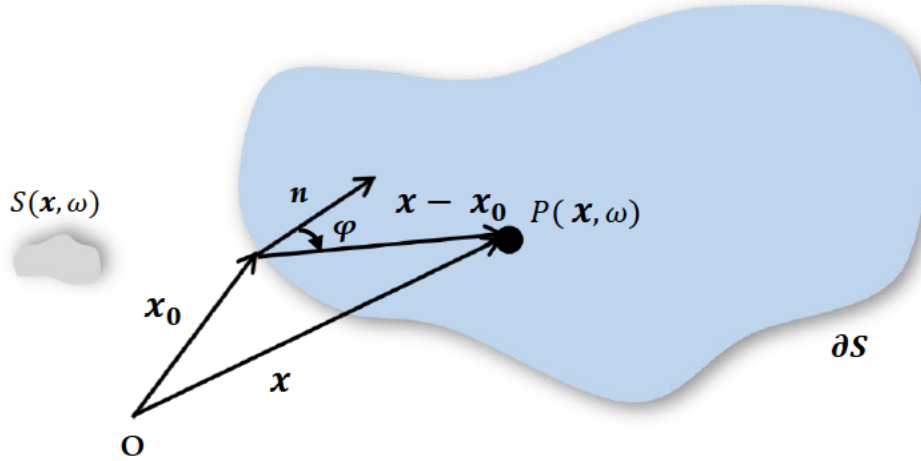

Figure3.2b: WFS 2.5D- Reproduction of sound field of a virtual source using monopole point source as secondary source.

3.3.3 Virtual Source Models

For the reproduction of two-dimensional wave field two virtual sources will be modeled either as plane or spherical waves [Spo08].

1. Plane Wave as virtual source

The wave field of a propagating plane wave in the direction $\boldsymbol{n_{pw}}$ is given as

$$S_{pw}(\boldsymbol{x}, \omega) = \hat{S}_{pw}(\omega) e^{-i\frac{\omega}{c}\boldsymbol{n}_{pw}^T \boldsymbol{x}} \quad , \tag{3.14}$$

where $\hat{S}_{pw}(\omega)$ is the temporal spectrum of the virtual plane wave.

The window-function $a_{pw}(\boldsymbol{x_0})$ for a virtual plane wave can be derived as [Spo07]

$$a_{pw}(\boldsymbol{x_0}) = \begin{cases} 1 & , if \langle \boldsymbol{n_{pw}}, n(\boldsymbol{x_0}) \rangle > 0 \\ 0 & , otherwise \end{cases} \tag{3.15}$$

The driving function for a virtual plane wave can be derived by introducing Eq. (3.14) and Eq. (3.15) into Eq. (3.13) as

$$D_{pw,2.5D}(\boldsymbol{x_0}, \omega) = 2a_{pw}(\boldsymbol{x_0}) \sqrt{\frac{2\pi|\boldsymbol{x_{ref}} - \boldsymbol{x_0}|}{i\frac{\omega}{c}}} (-i\frac{\omega}{c}\boldsymbol{n}_{pw}^T \boldsymbol{n}(\boldsymbol{x_0})) \hat{S}_{pw}(\omega) e^{-i\frac{\omega}{c}\boldsymbol{n}_{pw}^T \boldsymbol{x_0}}$$

$$= \underbrace{-2a_{pw}(\boldsymbol{x_0}) \sqrt{2\pi|\boldsymbol{x_{ref}} - \boldsymbol{x_0}|} \, \boldsymbol{n}_{pw}^T \boldsymbol{n}(\boldsymbol{x_0})}_{W_{pw}} \underbrace{\sqrt{i\frac{\omega}{c}} \, \hat{S}_{pw}(\omega) \, e^{-j\frac{\omega}{c}\boldsymbol{n}_{pw}^T \boldsymbol{x_0}}}_{F_{pw}(\omega)} \tag{3.16}$$

The time-domain driving function is given by applying the inverse Fourier transformation [OL10] of Eq. (3.16) is given as

$$d_{pw,2.5D}(\boldsymbol{x_0}, t) = w_{pw}(t)\delta\left(t - \frac{\boldsymbol{n}_{pw}^T \boldsymbol{x_0}}{c}\right) * \left(f_{pw}(t) * \hat{s}_{pw}(t)\right) \tag{3.17}$$

2. Spherical wave as virtual source

The wave field of a spherical wave with source position $\boldsymbol{x_S}$ is given as

$$S_{sw}(\boldsymbol{x}, \omega) = \hat{S}_{sw}(\omega) \frac{e^{-i\frac{\omega}{c}|\boldsymbol{x} - \boldsymbol{x_S}|}}{|\boldsymbol{x} - \boldsymbol{x_S}|} \quad , \tag{3.18}$$

where $\hat{S}_{sw}(\omega)$ is temporal spectrum of the virtual plane wave. The window-function $a_{sw}(\boldsymbol{x_0})$ for a virtual spherical wave can be derived as [Spo07]

$$a_{sw}(\boldsymbol{x_0}) = \begin{cases} 1 & , if \langle \boldsymbol{x} - \boldsymbol{x_S}, n(\boldsymbol{x_0}) \rangle > 0 \\ 0 & , otherwise \end{cases} \tag{3.19}$$

The driving function for a virtual spherical wave can be derived by introducing Eq. (3.18) and Eq. (3.19) into Eq. (3.13) as

$$D_{sw,2.5D}(\boldsymbol{x_0}, \omega) = -2a_{sw}(\boldsymbol{x_0}) \frac{(\boldsymbol{x_0} - \boldsymbol{x_S})^T \boldsymbol{n}(\boldsymbol{x_0})}{|\boldsymbol{x_0} - \boldsymbol{x_S}|} \sqrt{2\pi|\boldsymbol{x_{ref}} - \boldsymbol{x_0}|} \, \times$$

$$\times \underbrace{\left(\frac{1}{\sqrt{j\frac{\omega}{c}}|\boldsymbol{x_0} - \boldsymbol{x_S}|} + \sqrt{j\frac{\omega}{c}}\right)}_{F_{sw}(\omega)} \hat{S}_{sw}(\omega) \frac{e^{-j\frac{\omega}{c}|\boldsymbol{x_0} - \boldsymbol{x_S}|}}{|\boldsymbol{x_0} - \boldsymbol{x_S}|} \tag{3.20}$$

34

The time-domain broadband driving function of Eq. 3.20 is given by applying the inverse Fourier transformation of Eq. (3.20) is given as

$$d_{sw,2.5D}(\boldsymbol{x_0}, t) = w_{sw}(t)\delta\left(t - \frac{|\boldsymbol{x_0} - \boldsymbol{x_S}|}{c}\right) * \left(f_{sw}(t) * \hat{s}_{sw}(t)\right) \quad , \tag{3.21}$$

3.3.4 Spatial Sampling of the Secondary Monopole Source Distribution

The theory presented so far assumes a spatially continuous distribution of secondary sources. Practical implementations of WFS will consist of finite number of secondary sources that are placed at spatially discrete positions. This spatial sampling of the continuous distribution may lead to spatial aliasing artifacts in the reproduced wave field. Spatial aliasing constitutes a disturbance of the spatial structure of the reproduced wave field. For an arbitrary geometry of the secondary source distribution contour: the three main considerations are:

1. The spatial aliasing increases with the bandwidth of the virtual source signal and
2. The spatial aliasing artifacts depend on the listener position.
3. The distance between loudspeaker should be reduced (typically $\Delta x = 10...30$ cm).

To reduce the spatial aliasing artifact a bandwidth limitation of the reproduction signal can be implemented. For example the cut-off frequency for a virtual plane wave related to spacing between two sources Δx will be approximately given as

$$f = \frac{c}{2\Delta x} \tag{3.22}$$

The pre-filtering of the virtual source signal in WFS driving functions derived in 3.3.1 should be performed below the spatial aliasing frequency. Figure 3.3 illustrates the aliasing effect introduced to the wave fields of band-limited plane wave.
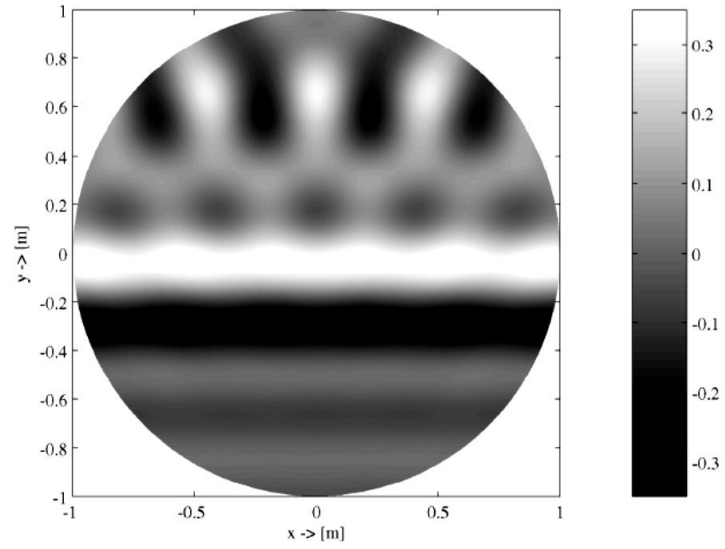
(a) reproduced wave field $P_{P,S}(x_P, \omega)$

Figure 3.3: Reproduction of a band-limited plane wave with $\theta_{pw} = 270°$ on a loudspeaker array with L = 48 loudspeakers and a radius of R = 1.5 m. The plane wave has a bandwidth of 1 kHz. The upper plot shows the reproduced wave field $P_{pw}(\boldsymbol{x}, \omega)$, the bottom one its aliasing contributions [Spo06].

3.4 Ambisonics/Higher Order Ambisonics (HOA)

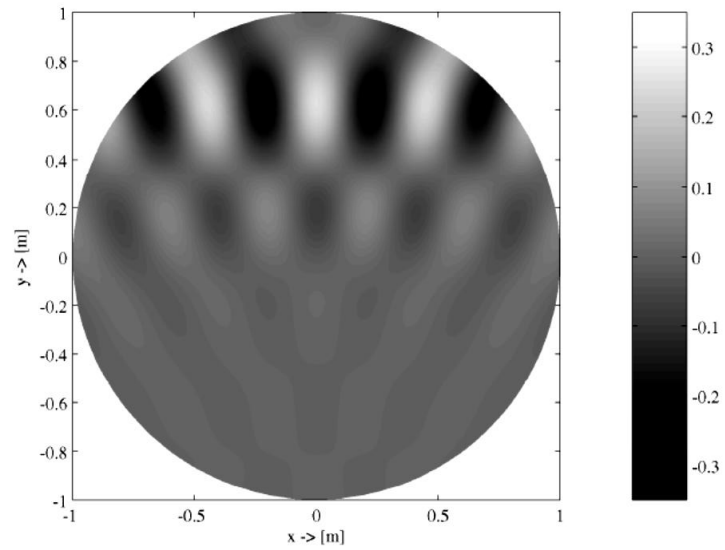Ambisonics is a spatial sound reproduction technology developed mostly by Michael Gerzon [Ger85] in the early1970s. The term Ambisonics traditionally refers to first order Ambisonics, involving a tetrahedron of loudspeakers. This setup is restricted to the reproduction of only zeroth -and first order spatial harmonics. The typical *Ambisonics* approach is based on Mode-Matching approach (Section 3.2), which involves a finite number of discrete loudspeakers whose emitted wave fields superpose matching an approximation of the desired one. Recent extensions have included a compensation for the properties of the sound field of a finite distance source for both virtual sources and loudspeakers. These extensions are typically termed near field corrections and the resulting system consequently near field corrected HOA (NFC-HOA) [Dan03].

Higher order Ambisonics (HOA) is a sound reproduction technique that utilizes a large number of loudspeakers to physically recreate a wave field in a specific listening area. Similar to WFS method, the reproduction by HOA can be formulated based on the Simple-Source approach [Pol05], which assumes a simple monopole distribution for the secondary sources. To illustrate Ambisonics it is typically assumed that the secondary sources are located on a sphere around the listener residing in the center, since a closed solution for the loudspeaker driving functions can be derived for this setup. In the following Sections Ambisonics/HOA approaches will be introduced. Practical consideration for realization of the system will be taken into account.

3.5 Mode-Matching based Ambisonics/Amplitude Panning Ambisonics (APA)

As introduced in Section 3.1.2 the sound field at a radial frequency ω, can be expressed as

$$P_S(\boldsymbol{x_S}, \omega) = P(\theta, \varphi, r, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} A_n^m(\omega) j_n\left(\frac{\omega}{c} r\right) Y_n^m(\theta, \varphi) \;, \tag{3.23}$$

where c sound speed, $j_n(.)$ is the spherical Bessel function of the first kind, $A_n^m(\omega)$ spherical harmonics expansion coefficients of the function $p(\theta, \varphi, r, \omega)$ and $Y_n^m(\theta, \varphi)$ are the spherical harmonics.

3.5.1 Spherical Harmonics

Again the spherical Harmonics $Y_n^m(\theta, \varphi)$ is defined as follows

$$Y_n^m(\theta, \varphi) = \sqrt{\frac{(2n+1)}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos\varphi) e^{im\theta} \;, \tag{3.24}$$

where $P_n^m(.)$ denotes the m-th order associated Legendre polynomial of n-th degree.

Any arbitrary function on a sphere $f(\theta, \varphi)$ can be expanded in terms of the spherical harmonics [Wil99]

$$f(\theta, \varphi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} A_n^m Y_n^m(\theta, \varphi) \tag{3.25}$$

The spherical harmonics are orthonormal, and expressed as follows [Wil99]

$$\int_0^{\pi} \int_0^{2\pi} Y_n^m(\theta, \varphi) Y_{n'}^{m'}(\theta, \varphi)^* \sin(\theta) \, d\theta d\varphi = \delta_{nn'} \delta_{mm'} \ , \tag{3.26}$$

where $\delta_{nn'}$ is 0 for $n = n'$ and 1 for $n \neq n'$.

Because of the orthonormality of the spherical harmonics the coefficients A_n^m in Eq. (3.25) are to be found as

$$A_n^m = \int_0^{\pi} \int_0^{2\pi} f(\theta, \varphi) Y_n^m(\theta, \varphi)^* \sin(\theta) \, d\theta d\varphi \tag{3.27}$$

Figure 3.4 illustrates the real part of a selection of spherical harmonics

Associated Legendre Functions

The associated Legendre functions (see Eq. 3.24) are solutions of the Legendre equation, which was presented in Section 2.3 as follows [Wil99]

$$\frac{\partial}{\partial \eta} \left((1 - \eta^2) \frac{\partial p_\theta(\theta)}{\partial \eta} \right) + \left(n(n+1) - \frac{m^2}{1 - \eta^2} \right) p_\theta(\theta) = 0 \tag{3.28}$$

For $m > 0$ t associated Legendre functions (Figure 3.6) are related to the Legendre polynomials $P_n(x)$ by the formula [Wil99]

$$P_n^m(x) = (-1)^m (1 - x^2)^{\frac{m}{2}} \frac{d^m}{dx^m} P_n(x) \ , \tag{3.29}$$

where $P_n(x)$ is the special case of $P_n^m(x)$ for $m = 0$. For $m < 0$ the associated Legendre functions are given as

$$P_n^{-m}(x) = (-1)^m \frac{(n-m)!}{(n+m)!} P_n^m(x), \quad \text{where m>0} \tag{3.30}$$

Figure 3.5 illustrates some associated Legendre's functions for positive and negative degrees.

(a) $n = 0$, $m = 0$ (b) $n = 1$, $m = 1$
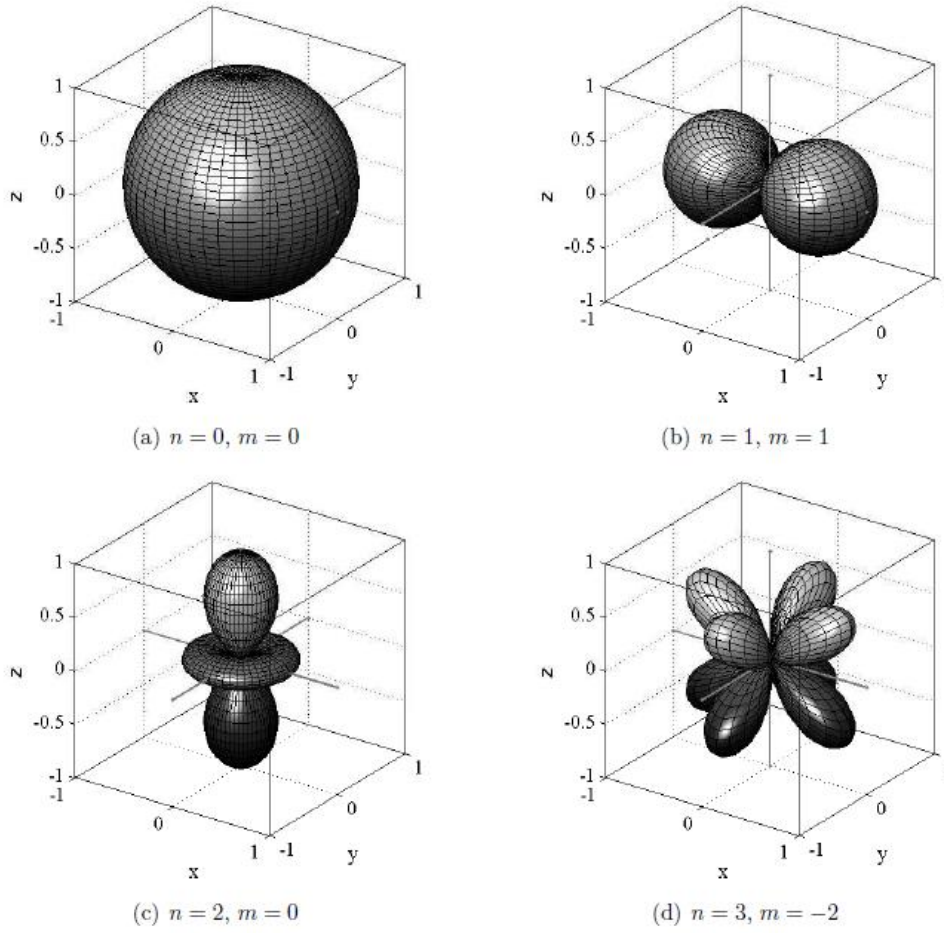
(c) $n = 2$, $m = 0$ (d) $n = 3$, $m = -2$

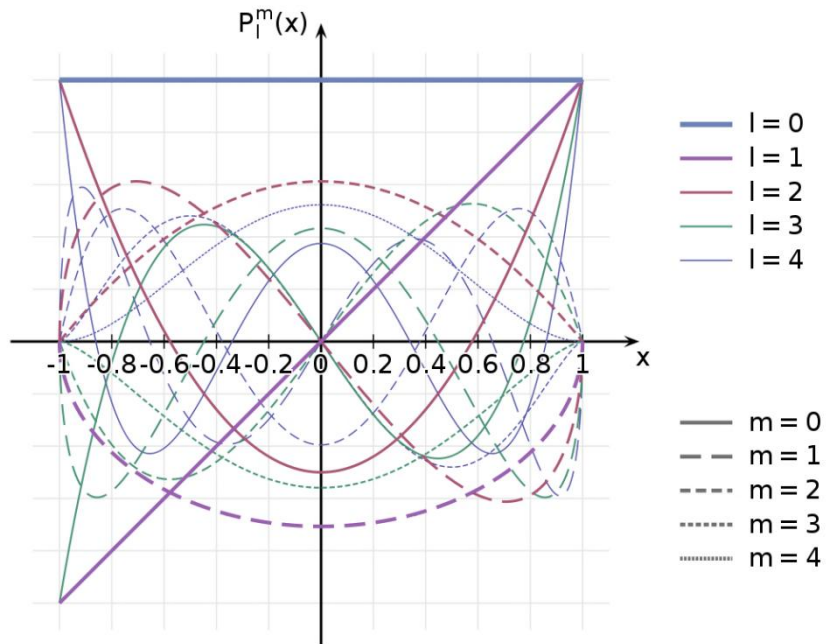**Figure 3.4**: Real part of the some spherical harmonics $Y_n^m(\theta, \varphi)$ [Ahr10].

Figure 3.5: Illustration of the associated Legendre's function $P_l^m(x)$ for order $l = 0...5$ and degree m = 0...4 [Wik3].

3.5.2 Plane and Spherical Wave Expansions

The spherical harmonics expansion of an outgoing plane wave of incident angles $(\theta_{pw}, \varphi_{pw})$ is given as follows [Ahr08].

$$S_{pw}(\boldsymbol{x}_S, \omega) = \hat{S}_{pw}(\omega) e^{-i\frac{\omega}{c}\boldsymbol{n}_{pw}^T \boldsymbol{x}_S}$$

$$= \hat{S}_{pw}(\omega) \times 4\pi \sum_{n=0}^{\infty} (-i)^n j_n\left(\frac{\omega}{c}r\right) \sum_{m=-n}^{n} Y_n^m(\theta, \varphi) Y_n^m(\theta_{pw}, \varphi_{pw})^* \qquad (3.31)$$

The spherical harmonics expansion of the wave field due to a point source with position $\boldsymbol{x}_{Sw} = [\theta_{Sw}, \varphi_{Sw}, r_{Sw}]^T$ and with $r < r_S$ is given as

$$S_{sw}(\boldsymbol{x}, \omega) = \hat{S}_{sw}(\omega) \frac{e^{-i\frac{\omega}{c}|\boldsymbol{x}-\boldsymbol{x}_S|}}{4\pi|\boldsymbol{x}-\boldsymbol{x}_S|} =$$

$$= \hat{S}_{sw}(\omega) \times \left(-i\frac{\omega}{c}\right) \sum_{n=0}^{\infty} j_n\left(\frac{\omega}{c}r\right) h_n^{(2)}\left(\frac{\omega}{c}r\right) \sum_{m=-n}^{n} Y_n^m(\theta, \varphi) Y_n^m(\theta_{sw}, \varphi_{sw})^* \qquad (3.32)$$

3.5.3 Mode-Matching

Polleti [Pol05] has summarized the mode-matching based Ambisonics approach, which introduces the reproduction of the three-dimensional wave field (expressed by its finite spherical harmonics expansion) using L secondary sources with amplitudes w_l and distributed on a sphere with radius R.

For secondary sources modelled as plane wave sources where the L loudspeakers in a surround sound system are at a large distance from the listener the loudspeakers weights can be calculated as solution of linear equation system. In this case the weights w_l can be interpreted as panning functions (pure weighting).

We aim to synthesize an outgoing plane wave from arbitrary direction $(\theta_{pw}, \varphi_{pw})$ using L plane wave sources at (θ_l, φ_l, R) with amplitudes w_l. Using Equation 3.31 as the synthesized plane wave field of the loudspeaker (assuming $\hat{S}_{pw}(\omega) = 1$) [Ahr08] allows the approximated reproduced wave field to be formulated as

$$P_{pw}(\theta, \varphi, r, \omega) = 4\pi \sum_{n=0}^{\infty} (-i)^n j_n\left(\frac{\omega}{c}r\right) \sum_{m=-n}^{n} Y_n^m(\theta, \varphi) \times \sum_{l=1}^{L} w_l(\theta_l, \varphi_l) Y_n^m(\theta_l, \varphi_l)^* \qquad (3.33)$$

This must equal the plane wave expansion of the desired plane wave arriving from direction $(\theta_{pw}, \varphi_{pw})$ (Eq. 3.31). The resulting matching equation for each n and m is

$$\sum_{l=1}^{L} w_l(\theta_l, \varphi_l) Y_n^m(\theta_l, \varphi_l)^* = Y_n^m(\theta_{pw}, \varphi_{pw})^* \qquad (3.34)$$

Equation 3.34 introduces linear equation system be solved for the L weights $w_l(\theta_{pw}, \varphi_{pw})$ for all Spherical Harmonics up to some order N. Since the total number of Spherical Harmonics up to order N is $(N+1)^2$, this requires $(N+1)^2 \geq L$ [Pol05 and references there in]. This case represents the so-called Amplitude Panning Ambisonics (APA)[3].

Assuming that the L loudspeakers generate spherical waves at positions (θ_l, φ_l, R), the synthesis of a plane wave as the sum of L spherical waves (Eq. 3.32) to an outgoing plane wave from direction $(\theta_{pw}, \varphi_{pw})$, yields

$$\sum_{l=1}^{L} w_l(\theta_l, \varphi_l) Y_n^m(\theta_l, \varphi_l)^* = \frac{(-i)^n}{-i\frac{\omega}{c} h_n^{(2)}(\frac{\omega}{c}R)} Y_n^m(\theta_{pw}, \varphi_{pw})^* \ , \tag{3.35}$$

where R is the radius of the spherical loudspeaker array. If the source to be synthesized using spherical source loudspeakers at (θ_l, φ_l, R) is a single spherical source $\boldsymbol{x_S} = [\theta_S, \varphi_S, r_S]$, then the matching equations become

$$\sum_{l=1}^{L} w_l(\theta_l, \varphi_l) Y_n^m(\theta_l, \varphi_l)^* = \frac{h_n^{(2)}\left(\frac{\omega}{c} r_S\right)}{h_n^{(2)}\left(\frac{\omega}{c}R\right)} Y_n^m(\theta_S, \varphi_S)^* \tag{3.36}$$

The loudspeaker weights obtained from this set of equations (Eq. 3.36) are now panning-functions, which vary with the source radius. When the spherical source has the same radius as the loudspeakers ($r_S = R$), the panning functions reduce to the plane wave functions in Eq. (3.34).

For the general free-field case, the field due to spherical source loudspeakers at (θ_l, φ_l, R) is equated to the general solution of the wave equation in yielding

$$\sum_{l=1}^{L} w_l(\theta_l, \varphi_l) Y_n^m(\theta_l, \varphi_l)^* = \frac{A_n^m(\omega)}{-i\frac{\omega}{c} h_n^{(2)}\left(\frac{\omega}{c}R\right)} \tag{3.37}$$

The weights w_l can no longer be interpreted as panning functions since they provide reproduction of multiple sources [Pol05].

---

[3] Amplitude Panning Ambisonics (APA) is an abbreviation used in this thesis to donate the typical panning based Ambisonics (see Section 3.5.4)

3.5.4 Amplitude Panning Ambisonics (APA)

The Mode-Matching based Ambisonics approach is usually divided into an encoding and a decoding stage to allow for storing and transmission of content independently from the loudspeaker setup. In this Section we will consider the loudspeaker as wave plane sources. The virtual source could be modelled as plane wave or spherical wave.

Encoding stage involves encoding the directional information of a virtual source (θ_S, φ_S) to a set of signals B_n^m by multiplying the source signal S_s with the value of the respective spherical harmonic function of the virtual source direction. The signals B_n^m are refered to as the Ambisonics channels. Encoding APA-System of N-th order can be expressed in an equation system as follows

$$B = C_s . S_s \quad , \tag{3.38}$$

where B is a $((N+1)^2 \times 1)$ matrix of Ambisonics channels, C_s is the $((N+1)^2 \times 1)$ Encoding Matrix and S_s the virtual source signal.

Decoding stage aims to find the L loudspeaker weights (driving function) S, which can be superposed to reproduce the encoded sound field. In order to express this in an equation system, we use a representation of the loudspeaker signals equivalent to the encoded sound field B_n^m, i.e. a spherical harmonic decomposition of the virtual source, depending on the position of the respective loudspeaker as follows [Wei08]

$$B = C . S$$

$$\rightarrow \quad S = C^{-1} . B = D . B \quad , \tag{3.39}$$

where S is a (L×1) matrix of loudspeakers weights, D is the $((N+1)^2 \times L)$ Decoding Matrix. If $(N+1)^2 = L$, then the Decoding Matrix can be found from the inverse of the Matrix C (re-encoding matrix) whose elements are the encoding gains associated to the loudspeaker directions.

$$D = C^{-1} \tag{3.40}$$

If the number of modes to be matched is smaller than the number of loudspeakers $((N+1)^2 < L)$, then the system of mode-matching equations is under-determined and the Decoding Matrix can be found from the pseudo- inverse of the Encoding Matrix C

$$D = pinv(C) = C^T (C . C^T)^{-1} \tag{3.41}$$

Finally, if the number of modes exceeds the number of loudspeakers $((N+1)^2 < L)$, then there is no exact solution. Typically, the solution with the minimum least-squared error is derived or the pseudo- inverse Matrix

$$D = (C^T C .)^{-1} C^T \tag{3.42}$$

3.5.5 Amplitude Panning Ambisonics: Plane Wave as Virtual Source

Daniel [Dan00] has introduced the basics formulation of panning based Ambisonics. This will be used for derivation of APA-encoding and -decoding equations.

1. Encoding

The Ambisonics channels of a virtual plane wave source S_v arriving from incident angles $(\theta_{pw}, \varphi_{pw})$ is given as follows

$$B_{nm}^{\sigma}(\theta, \varphi) = Y_{nm}^{\sigma}(\theta_{pw}, \varphi_{pw})S \tag{3.43a}$$

Or matrix-wise

$$\begin{pmatrix} B_{00}^1 \\ B_{10}^1 \\ B_{11}^1 \\ B_{11}^{-1} \\ \vdots \\ B_{NN}^{-1} \end{pmatrix} = \begin{pmatrix} Y_{00}^1(\theta_{pw}, \varphi_{pw}) \\ Y_{10}^1(\theta_{pw}, \varphi_{pw}) \\ Y_{11}^1(\theta_{pw}, \varphi_{pw}) \\ Y_{11}^{-1}(\theta_{pw}, \varphi_{pw}) \\ \vdots \\ Y_{NN}^{-1}(\theta_{pw}, \varphi_{pw}) \end{pmatrix} S_v \quad , \tag{3.43}$$

where Y_{nm}^{σ} denotes the real valued spherical Harmonics of the n-th order and positive m-th degree and defined as follows

$$Y_{nm}^{\sigma}(\theta, \varphi) = \sqrt{2n + 1} \, \hat{P}_n^m(\sin \varphi) \times \begin{cases} \cos m\theta & if \ \sigma = +1 \\ \sin m\theta & if \ \sigma = -1 \end{cases} \quad , \tag{3.44}$$

where $\hat{P}_n^m(\sin \varphi)$ the "Schmidt semi-normalized" Legendre's function P_n^m and related to the associated Legendre's function as

$$\hat{P}_n^m(\sin \varphi) = \sqrt{\left(2 - \delta_{0,m}\right) \frac{(n - m)!}{(n + m)!}} \, P_n^m(\cos \theta) \tag{3.45}$$

$$where \ \delta_{ij} = 1 \ if \ i = j \ \ and \ 0 \ otherwise$$

Daniel distinguishes the following types of normalization [Dan00]:

- SN3D: semi-normalization 3D.
- N3D: full normalization 3D.
- MaxN: max normalization.
- FuMa: Furse-Malham set.

The SN3D encoding functions are the Spherical Harmonic functions (see Figure 3.6) with semi-normalization as applied in equation (3.42). Daniel suggests using them as the reference encoding convention. The N3D convention is of interest regarding the evaluation of loudspeaker layouts in Ambisonics sound field reproduction.

The MaxN convention features weighing factors of the spherical harmonics, which ensure that the maximum absolute value of each function is 1, which is of relevance regarding digital signal representations [Mal03].

The FuMa set, which is equivalent to the MaxN convention except for an additional $1/\sqrt{2}$ weighing factor for the zeroth order channel, takes practical considerations into account.

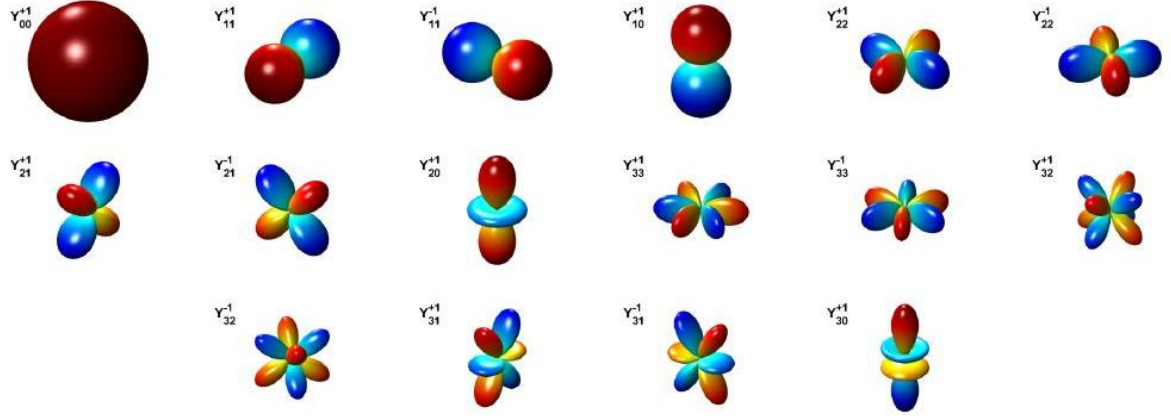In table 3.2, the SN3D channels and their equivalent FuMa weights are presented up to third order.

Figure 3.6: Spherical harmonic functions up to third order [Dan03].

Order	n,m,σ	Channel	SN3D	FuMu weight
0	0,0,1	W	1	$1/\sqrt{2}$
1	1,1,1	X	$\cos\theta\cos\varphi$	1
	1,1,-1	Y	$\sin\theta\cos\varphi$	1
	1,0,1	Z	$\sin\varphi$	1
2	2,0,1	R	$(3\sin^2\varphi - 1)/2$	1
	2,1,1	S	$(\sqrt{3}/2)(\cos\theta\sin 2\varphi)$	$\sqrt{2}/3$
	2,1,-1	T	$(\sqrt{3}/2)(\sin\theta\sin 2\varphi)$	$\sqrt{2}/3$
	2,2,1	U	$(\sqrt{3}/2)(\cos 2\theta\cos^2\varphi)$	$\sqrt{2}/3$
	2,2,-1	V	$(\sqrt{3}/2)(\sin 2\theta\cos^2\varphi)$	$\sqrt{2}/3$
3	3,0,1	K	$\sin\varphi\ (5\sin^2\varphi - 3)/2$	1
	3,1,1	L	$(\sqrt{3}/8)\cos\theta\cos\varphi\ (5\sin^2\varphi - 3)$	$\sqrt{45}/32$
	3,1,-1	M	$(\sqrt{3}/8)\sin\theta\cos\varphi\ (5\sin^2\varphi - 3)$	$\sqrt{45}/32$
	3,2,1	N	$(\sqrt{15}/2)\cos 2\theta\sin\varphi\cos^2\varphi$	$\sqrt{3}/5$
	3,2,-1	O	$(\sqrt{15}/2)\sin 2\theta\sin\varphi\cos^2\varphi$	$\sqrt{3}/5$
	3,3,1	P	$(\sqrt{5}/8)\cos 3\theta\cos^3\varphi$	$\sqrt{8}/5$
	3,2,-1	Q	$(\sqrt{5}/8)\sin 3\theta\cos^3\varphi$	$\sqrt{8}/5$

Table 3.2: SN3D normalized Ambisonics channels and their equivalent FuMa weights up to third order.

2. Decoding

In Section 3.5.4 we have introduced the general formulation of decoding matrix $\boldsymbol{D}$. Let L be the number of loudspeakers at the positions $(\theta_i, \varphi_i, i = 1 \dots L)$. In theory, the signals from loudspeakers in directions other than the one of the source cancel each other out in the sweet spot. However, diffraction effects due to the presence of the listener's head disturb this destructive interference, which causes the signals of loudspeakers in the opposite direction of the sound source to be perceived as independent sources. This has led to a modification of the basic decoder, originally suggested for horizontal systems by David Malham [Mal99] and extended to periphonic systems by Daniel [Dan00]. It is referred to as the in-phase decoder, which avoids contributions from opposite loudspeakers for the price of an increased spread of the sound source (see Figure 3.8).

Other modifications have been suggested, which can generally be described as a multiplication of the decoding matrix $\boldsymbol{D}$ with a Modifying diagonal matrix. The elements in the diagonal matrix represent additional weighting factors for the spherical harmonics components of the loudspeaker signals, depending on their respective order n.

3.6 Single Layer Potential based Higher Order Ambisonics (NFC-HOA)

The Ambisonics approach is usually divided into an encoding and a decoding stage to allow for storing and transmission of content independently from the loudspeaker setup. However in this Section the loudspeaker driving signals will be directly derived from the initial, virtual wave field description.

The Single Layer Potential based higher order Ambisonics approach implicitly includes the near-field correction introduced by Daniel [Dan03]. Thus, Ambisonics will refer implicitly to the corrected higher order Ambisonics (NFC-HOA). Ahrens [Ahr08] has presented basics and real world implementation strategies for sound field reproduction using Simple-Source Approach based higher order Ambisonics approach for circular and spherical loudspeaker arrays. His work will be the base for this Section.

The formulation of the basic Ambisonics equation for a continuous, secondary source distribution on a sphere whose center is placed in the coordinate origin (Eq. 3.7) reads

$$P(\boldsymbol{x}, \omega) = \oint_{\partial V} D(\boldsymbol{x_0}, \omega) G(\boldsymbol{x}|\boldsymbol{x_0}, \omega) \, dS_0 \,^4 , \qquad (3.48)$$

where ∂V denotes the surface of a sphere with radius r_0 on which the secondary sources are located at a point on the sphere $\boldsymbol{x_0}$ and $G(\boldsymbol{x}|\boldsymbol{x_0}, \omega)$ can be interpreted as three-dimensional free-field Green's function (Eq. 2.45).

---

[4] The minus in the integral in Equation (2.7) can be included in the driving function term

A propagating wave field can be described by its spherical harmonics expansion as

$$P_S(x_S, \omega) = P(\theta, \varphi, r, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \underbrace{A_n^m(\omega) j_n\left(\frac{\omega}{c}r\right) Y_n^m(\theta, \varphi)}_{A_n^m(r, \omega)} \qquad (3.49)$$

Equation (3.48) can be interpreted as a convolution [Ahr08 and references therein] along the surface of a sphere. Thus, for spherical secondary sources the distribution case yields

$$A_n^m(r, \omega) = 2\pi r_0 \sqrt{\frac{4\pi}{2n+1}} D_n^m(\omega) G_n^0(r, \omega) \, , \qquad (3.50)$$

and therefore

$$D_n^m(\omega) = \frac{1}{2\pi r_0} \sqrt{\frac{2n+1}{4\pi}} \frac{A_n^m(r, \omega)}{G_n^0(r, \omega)} \qquad (3.50a)$$

$$D_n^m(\omega) = \frac{1}{2\pi r_0} \sqrt{\frac{2n+1}{4\pi}} \frac{A_n^m(\omega)}{G_n^0(\omega)} \frac{j_n\left(\frac{\omega}{c}r\right)}{j_n\left(\frac{\omega}{c}r\right)} \qquad (3.50b)$$

In Equation (3.50b) the parameter r appears both in the numerator as well as in the denominator in the Bessel's function $j_n\left(\frac{\omega}{c}r\right)$. For the case where $j_n\left(\frac{\omega}{c}r\right) \neq 0$ in both numerator and denominator of r will be canceled out. For the case where $\frac{\omega}{c}r = 0$ it can be proved (de l'Hôpital's rule [Ahr08 and references therein]) that $j_n(0)$ also cancelled out. Thus, the driving function is independent from the receiver position in these cases.

However, in particular situations, i.e. when $j_n\left(\frac{\omega}{c}r\right) = 0$ and $r \neq 0$, (3.50b) can be undefined. In this case forbidden frequencies arise. These forbidden frequencies are discrete and represent the resonances of the spherical cavity. In simple words, resonances exhibit a zero on the boundary of the cavity. Since the secondary sources are positioned in these zeros, they cannot excite the respective mode. To get rid of forbidden frequencies it is suggested to reference the reproduced wave field to the center ([0 0 0]) of the secondary source distribution. Thus, in the following we can assume the cancelation of Bessel function effect [Wil99]. Then the secondary source driving function $D_{3D}(\alpha, \varphi, r_0, \omega)$ for three-dimensional reproduction of a desired wave field with expansion coefficients $A_n^m(\omega)$ is

$$D_{3D}(\theta, \varphi, r_0, \omega) = \frac{1}{2\pi r_0} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \sqrt{\frac{2n+1}{4\pi}} \frac{A_n^m(\omega)}{G_n^0(\omega)} Y_n^m(\theta, \varphi) \qquad (3.51)$$

For a circular distribution of secondary point sources, equation (3.48) degenerates to

$$P(\boldsymbol{x}, \omega) = P(\theta, r, \omega) = \int_0^{2\pi} D(\boldsymbol{x_0}, \omega) G(\boldsymbol{x} - \boldsymbol{x_0}, \omega) r_0 d\theta_0 \qquad (3.52)$$

A propagating wave field can be described by its spherical harmonics expansion as

$$P_S(\boldsymbol{x_S}, \omega) = P\left(\theta, \varphi = \frac{\pi}{2}, r, \omega\right) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \underbrace{A_n^m(\omega) j_n\left(\frac{\omega}{c} r\right) Y_n^m\left(\theta, \varphi = \frac{\pi}{2}\right)}_{A_n^m(r,\omega)} \qquad (3.53)$$

Equation (3.52) can be interpreted as a circular convolution and thus the convolution theorem [Wil99]

$$A_m(r, \omega) = 2\pi r_0 D_m(\omega) G_m(r, \omega) \ , \qquad (3.54)$$

and therefore

$$D_m(\omega) = \frac{1}{2\pi r_0} \frac{A_m(r, \omega)}{G_m(r, \omega)} \qquad (3.55)$$

Unlike the case of spherical secondary source distributions, the radius r does not cancel out. The driving function is therefore dependent on the receiver position. Therefore, we choose a radius as a reference (r>0). The investigation of forbidden frequencies is not obvious. However, choosing the radius (r=0) as a reference will help to avoid formation of these frequencies [Ahr08]. For circular geometries the Fourier series expansion (Eq. 3.56) constitutes a useful mathematical tool. The Fourier series expansion of a two-dimensional propagating wave field is

$$P(\theta, r, \omega) = \sum_{m=-\infty}^{\infty} A_m(r, \omega) e^{im\theta}, \qquad (3.56)$$

where the expansion coefficients are given as

$$A_m(r, \omega) = \sum_{n=|m|}^{\infty} \underbrace{A_n^m(\omega) j_n\left(\frac{\omega}{c} r\right) Y_n^m\left(\theta, \varphi = \frac{\pi}{2}\right)}_{A_n^m(r,\omega)} \qquad (3.57)$$

Introducing the center of the array (*r=0*) as reference position and applying the of de l'Hôpital's rule [Ahr08 and references there in] yields the driving function $D_{2.5D}(\theta, \omega)$ for (2.5)[5]-dimensional reproduction as

$$D_{2.5D}(\theta, \omega) = \frac{1}{2\pi r_0} \sum_{m=-\infty}^{\infty} \frac{A_{|m|}^m(\omega)}{G_{|m|}^m(\omega)} e^{im\theta} \qquad (3.58)$$

---

[5] The approach of employing secondary sources which are intended for three-dimensional reproduction in such an imperfect two-dimensional scenario is typically referred to as 2.5-dimensional reproduction

3.6.1 Derivation of Driving Function

Real-world implementations of audio reproduction systems will always employ a limited number of discrete, secondary sources. The spatial discretisation constitutes spatial sampling and thus may produce spatial aliasing[6]. In this Section it will be illustrated how wave field reproduction according to single layer potential based Ambisonics can be accomplished. The procedure yields the secondary source driving function for rendering a virtual plane wave and point sources.

3.6.2 Derivation of the driving function for a virtual plane wave

The wave field of a plane wave with propagating direction $(\theta_{pw}, \varphi_{pw})$ can be expanded into its radial and angular dependencies around the origin of the coordinate system as

$$
S_{pw}(\boldsymbol{x}_S, \omega) = \hat{S}_{pw}(\omega) e^{-i\frac{\omega}{c}\boldsymbol{n}_{pw}^T \boldsymbol{x}_S} =
$$

$$
= \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \underbrace{\hat{S}_{pw}(\omega) . 4\pi(-i)^n j_n\left(\frac{\omega}{c}r\right) Y_n^m(\theta_{pw}, \varphi_{pw})^*}_{A_n^m(r,\omega)} Y_n^m(\theta, \varphi) ,
\tag{3.59}
$$

where $\hat{S}_{pw}(\omega)$ denotes the temporal spectrum of the plane wave and $j_n(.)$ the n-th order spherical Bessel's function. In case of spherical source distribution modeling the spatial transfer function of a loudspeaker at position $\boldsymbol{x_0} = [\theta_0, \varphi_0, r_0]$ on a sphere ∂V as the free-field Green's function and expanding it into its radial and angular dependencies leads to

$$
G(\boldsymbol{x}|\boldsymbol{x_0}, \omega) = \frac{1}{4\pi} \frac{e^{-i\frac{\omega}{c}|\boldsymbol{x}-\boldsymbol{x_0}|}}{|\boldsymbol{x}-\boldsymbol{x_0}|} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \underbrace{\left(-i\frac{\omega}{c}\right) j_n\left(\frac{\omega}{c}r\right) h_n^{(2)}\left(\frac{\omega}{c}r_0\right) Y_n^m(\theta_0, \varphi_0)^*}_{G_n^m(r,\omega)} Y_n^m(\theta, \varphi)
\tag{3.60}
$$

Combining Equations (3.49), (3.50a), (3.59) and (3.60) yields

$$
D(\boldsymbol{x}_S, \omega) = D_{pw}^{3D}(\theta_0, \varphi_0, r_0, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} D_n^m(\omega) j_n\left(\frac{\omega}{c}r_0\right) Y_n^m(\theta_0, \varphi_{0,})
$$

$$
D_n^m(\omega) = \frac{1}{2\pi r_0} \sqrt{\frac{2n+1}{4\pi}} \frac{A_n^m(r,\omega)}{G_n^0(r,\omega)} = \frac{1}{2\pi r_0} \sqrt{\frac{2n+1}{4\pi}} \frac{\hat{S}_{pw}(\omega) . 4\pi(-i)^n j_n\left(\frac{\omega}{c}r\right) Y_n^m(\theta_{pw}, \varphi_{pw})^*}{\left(-i\frac{\omega}{c}\right) j_n\left(\frac{\omega}{c}r\right) h_n^{(2)}\left(\frac{\omega}{c}r_0\right) Y_n^0(0,0)^*}
$$

$$
= \hat{S}_{pw}(\omega) \frac{i}{2\pi r_0 \frac{\omega}{c}} \sqrt{\frac{2n+1}{4\pi}} \frac{4\pi(-i)^n Y_n^m(\theta_{pw}, \varphi_{pw})^*}{h_n^{(2)}\left(\frac{\omega}{c}r_0\right) \sqrt{\frac{(2n+1)}{4\pi}\frac{(n-0)!}{(n+0)!}} \underbrace{P_n^0(\cos 0)}_{1} \underbrace{e^{-i0}}_{1}} =
$$

$$
D_{pw}^{3D}(\theta_0, \varphi_0, r_0, \omega)
$$

$$
= \hat{S}_{pw}(\omega) \times 2 \frac{i}{\frac{\omega}{c}r_0} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{(-i)^n}{h_n^{(2)}\left(\frac{\omega}{c}r_0\right)} Y_n^m(\theta_{pw}, \varphi_{pw})^* Y_n^m(\theta_0, \varphi_0)
\tag{3.61}
$$

---

[6] For more information about investigation of spatial aliasing in two-dimensional higher order Ambisonics we refer to [Ahr08 and references there in].

For a planar circular distribution of secondary point sources to render horizontal virtual plane wave we set the elevation angle φ in all position vectors to $\frac{\pi}{2}$. As advisable for circular geometries we adopt the expansion of the loudspeaker wave fields in (3.60) by exchanging the order of summations to arrive at such a Fourier series expansion:

$$G(\boldsymbol{x}|\boldsymbol{x_0},\omega) = \frac{1}{4\pi}\frac{e^{-\mathrm{i}\frac{\omega}{c}|\boldsymbol{x}-\boldsymbol{x_0}|}}{|\boldsymbol{x}-\boldsymbol{x_0}|} = \sum_{m=-\infty}^{\infty}\sum_{n=|m|}^{\infty}\underbrace{\left(-\mathrm{i}\frac{\omega}{c}\right)j_n\left(\frac{\omega}{c}r\right)h_n^{(2)}\left(\frac{\omega}{c}r_0\right)Y_n^m\left(\theta_0,\frac{\pi}{2}\right)^{*}Y_n^m\left(\theta,\frac{\pi}{2}\right)}_{G_n^m(\mathrm{r},\omega)}$$

$$= \sum_{m=-\infty}^{\infty}e^{\mathrm{i}m(\theta-\theta_0)}\times$$

$$\times\sum_{n=|m|}^{\infty}\left(-\mathrm{i}\frac{\omega}{c}\right)j_n\left(\frac{\omega}{c}r\right)h_n^{(2)}\left(\frac{\omega}{c}r_0\right)\frac{(2n+1)}{4\pi}\frac{(n-m)!}{(n+m)!}P_n^m(0)^2 \tag{3.62}$$

The according Fourier expansion of the plane wave field yields

$$S_{pw}(\boldsymbol{x_S},\omega) = \hat{S}_{pw}(\omega)e^{-\mathrm{i}\frac{\omega}{c}\boldsymbol{n}_{pw}^T\boldsymbol{x_S}} = \sum_{m=-\infty}^{\infty}\sum_{n=|m|}^{\infty}\underbrace{\hat{S}_{p\omega}(\omega).4\pi(-\mathrm{i})^n j_n\left(\frac{\omega}{c}r\right)Y_n^m\left(\theta_{pw},\frac{\pi}{2}\right)^{*}Y_n^m\left(\theta,\frac{\pi}{2}\right)}_{A_n^m(\mathrm{r},\omega)}$$

$$= \sum_{m=-\infty}^{\infty}e^{\mathrm{i}m(\theta-\theta_{pw})}$$

$$\times\hat{S}_{pw}(\omega).\sum_{n=|m|}^{\infty}4\pi(-\mathrm{i})^n j_n\left(\frac{\omega}{c}r\right)\frac{(2n+1)}{4\pi}\frac{(n-m)!}{(n+m)!}P_n^m(0)^2 \tag{3.63}$$

Combining Equations, (3.58), (3.62) and (3.63) yields

$$D_{pw}^{2.5D}\left(\theta_0,r_0,\omega\right) =$$

$$= \frac{1}{2\pi r_0}\sum_{m=-\infty}^{\infty}\frac{A_{|m|}^m(\omega)}{G_{|m|}^m(\omega)}e^{\mathrm{i}m\theta_0}$$

$$= \frac{1}{2\pi r_0}\sum_{m=-\infty}^{\infty}\frac{4\pi(-\mathrm{i})^{|m|}Y_{|m|}^m\left(\theta_{pw},\frac{\pi}{2}\right)^{*}}{\left(-\mathrm{i}\frac{\omega}{c}\right)h_{|m|}^{(2)}\left(\frac{\omega}{c}r_0\right)Y_{|m|}^m\left(0,\frac{\pi}{2}\right)^{*}}e^{\mathrm{i}m\theta_0}$$

$$= \frac{1}{2\pi r_0}\sum_{m=-\infty}^{\infty}\frac{4\pi(-\mathrm{i})^{|m|}\sqrt{\frac{(2|m|+1)}{4\pi}\frac{(|m|-m)!}{(|m|+m)!}}P_{|m|}^m\left(\cos\frac{\pi}{2}\right)e^{-\mathrm{i}m\theta_{pw}}}{\left(-\mathrm{i}\frac{\omega}{c}\right)h_{|m|}^{(2)}\left(\frac{\omega}{c}r_0\right)\sqrt{\frac{(2|m|+1)}{4\pi}\frac{(|m|-m)!}{(|m|+m)!}}P_{|m|}^m\left(\cos\frac{\pi}{2}\right)\underbrace{e^{-\mathrm{i}0}}_{1}}e^{\mathrm{i}m\theta_0}$$

$$D_{pw}^{2.5D}\left(\theta_0,r_0,\omega\right) = \hat{S}_{pw}(\omega)\times\sum_{m=-\infty}^{\infty}\frac{2\mathrm{i}}{\frac{\omega}{c}r_0}\frac{(-\mathrm{i})^{|m|}}{h_{|m|}^{(2)}\left(\frac{\omega}{c}r_0\right)}e^{\mathrm{i}m(\theta_0-\theta_{pw})} \tag{3.64}$$

For a finite number of discrete loudspeakers L the most suitable choice is to use as many order yields

$$D_{pw}^{2.5D}\left(\theta_0,r_0,\omega\right) = \hat{S}_{pw}(\omega)\times\sum_{m=-\frac{L-1}{2}}^{\frac{L-1}{2}}\frac{2\mathrm{i}}{\frac{\omega}{c}r_0}\frac{(-\mathrm{i})^{|m|}}{h_{|m|}^{(2)}\left(\frac{\omega}{c}r_0\right)}e^{\mathrm{i}m(\theta_0-\theta_{pw})} \tag{3.65}$$

49

3.6.3 Derivation of the driving function for a virtual point source

The wave field of a spherical wave with center position $x_{Sw} = [\theta_{sw}, \varphi_{sw}, r_{sw}]^T$ can be expanded into its radial and angular dependencies around the origin of the coordinate system as

$$S_{sw}(x, \omega) = \hat{S}_{sw}(\omega) \frac{1}{4\pi} \frac{e^{-i\frac{\omega}{c}|x-x_{Sw}|}}{|x - x_{Sw}|} =$$

$$= \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \underbrace{\hat{S}_{sw}(\omega) \left(-i\frac{\omega}{c}\right) j_n\left(\frac{\omega}{c}r\right) h_n^{(2)}\left(\frac{\omega}{c}r_{sw}\right) Y_n^m(\theta_{sw}, \varphi_{sw})^*}_{A_n^m(r,\omega)} Y_n^m(\theta, \varphi) \quad , \qquad (3.66)$$

where $\hat{S}_{pw}(\omega)$ denotes the temporal spectrum of the plane wave and $j_n(.)$ the n-th order spherical *Bessel's* function. In case of spherical source distribution modeling the spatial transfer function of a loudspeaker at position $x_0 = [\theta_0, \varphi_0, r_0]$ on a sphere ∂V as the free-field Green's function and expanding it into its radial and angular dependencies leads to

$$G(x|x_0, \omega) = \frac{1}{4\pi} \frac{e^{-i\frac{\omega}{c}|x-x_0|}}{|x - x_0|} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \underbrace{\left(-i\frac{\omega}{c}\right) j_n\left(\frac{\omega}{c}r\right) h_n^{(2)}\left(\frac{\omega}{c}r_0\right) Y_n^m(\theta_0, \varphi_0)^*}_{G_n^m(r_0,r,\omega)} Y_n^m(\theta, \varphi) \quad (3.67)$$

Combining Equations (3.49), (3.50a), (3.66) and (3.67) yields

$$D(x_S, \omega) = D_{sw}^{3D}(\theta_0, \varphi_0, r_0, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} D_n^m(\omega) j_n\left(\frac{\omega}{c}r_0\right) Y_n^m(\theta_0, \varphi_{0,})$$

$$D_n^m(\omega) = \frac{1}{2\pi r_0} \sqrt{\frac{2n+1}{4\pi}} \frac{A_n^m(r, \omega)}{G_n^0(r, \omega)}$$

$$= \frac{1}{2\pi r_0} \sqrt{\frac{2n+1}{4\pi}} \frac{\hat{S}_{sw}(\omega) \left(-i\frac{\omega}{c}\right) j_n\left(\frac{\omega}{c}r\right) h_n^{(2)}\left(\frac{\omega}{c}r_{sw}\right) Y_n^m(\theta_{sw}, \varphi_{sw})^*}{\left(-i\frac{\omega}{c}\right) j_n\left(\frac{\omega}{c}r\right) h_n^{(2)}\left(\frac{\omega}{c}r_0\right) Y_n^0(0,0)^*}$$

$$= \frac{1}{2\pi r_0} \sqrt{\frac{2n+1}{4\pi}} \frac{\hat{S}_{sw}(\omega) \left(-i\frac{\omega}{c}\right) h_n^{(2)}\left(\frac{\omega}{c}r_{sw}\right) Y_n^m(\theta_{sw}, \varphi_{sw})^*}{\left(-i\frac{\omega}{c}\right) h_n^{(2)}\left(\frac{\omega}{c}r_0\right) \sqrt{\frac{(2n+1)}{4\pi}\frac{(n-0)!}{(n+0)!}} \underbrace{P_n^0(\cos 0)}_{1} \underbrace{e^{-i0}}_{1}}$$

$$D_{sw}^{3D}(\theta_0, \varphi_0, r_0, \omega)$$

$$= \hat{S}_{sw}(\omega) \times \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{1}{2\pi r_0} \frac{h_n^{(2)}\left(\frac{\omega}{c}r_{sw}\right) Y_n^m(\theta_{sw}, \varphi_{sw})^*}{h_n^{(2)}\left(\frac{\omega}{c}r_0\right)} Y_n^m(\theta_0, \varphi_0) \qquad (3.68)$$

For a planar circular distribution of secondary point sources to render horizontal virtual plane wave we set the elevation angle φ in all position vectors to $\frac{\pi}{2}$. As advisable for circular geometries we adopt the expansion of the loudspeaker wave fields in (3.60) by exchanging the order of summations to arrive at such a Fourier series expansion:

$$G(\boldsymbol{x}|\boldsymbol{x_0},\omega) = \frac{1}{4\pi} \frac{e^{-i\frac{\omega}{c}|\boldsymbol{x}-\boldsymbol{x_0}|}}{|\boldsymbol{x}-\boldsymbol{x_0}|} = \sum_{m=-\infty}^{\infty} \sum_{n=|m|}^{\infty} \underbrace{\left(-i\frac{\omega}{c}\right) j_n\left(\frac{\omega}{c}r\right) h_n^{(2)}\left(\frac{\omega}{c}r_0\right) Y_n^m\left(\theta_0,\frac{\pi}{2}\right)^* Y_n^m\left(\theta,\frac{\pi}{2}\right)}_{G_n^m(r,\omega)}$$

$$= \sum_{m=-\infty}^{\infty} e^{jm(\theta-\theta_0)}$$

$$\times \sum_{n=|m|}^{\infty} \left(-i\frac{\omega}{c}\right) j_n\left(\frac{\omega}{c}r\right) h_n^{(2)}\left(\frac{\omega}{c}r_0\right) \frac{(2n+1)}{4\pi}\frac{(n-m)!}{(n+m)!} P_n^m(0)^2 \qquad (3.69)$$

The according Fourier expansion of the spherical wave field yields

$$S_{sw}(\boldsymbol{x},\omega) = \hat{S}_{sw}(w)\frac{e^{-i\frac{\omega}{c}|\boldsymbol{x}-\boldsymbol{x_{sw}}|}}{|\boldsymbol{x}-\boldsymbol{x_{sw}}|}$$

$$= \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \underbrace{\hat{S}_{sw}(\omega).\left(-i\frac{\omega}{c}\right) j_n\left(\frac{\omega}{c}r\right) h_n^{(2)}\left(\frac{\omega}{c}r_{sw}\right) Y_n^m(\theta_{sw},\varphi_{sw})^* Y_n^m(\theta,\varphi)}_{A_n^m(r,\omega)}$$

$$= \hat{S}_{sw}(\omega).\sum_{m=-\infty}^{\infty} e^{jm(\theta-\theta_{sw})} \times$$

$$\times \sum_{n=|m|}^{\infty} \left(-i\frac{\omega}{c}\right) j_n\left(\frac{\omega}{c}r\right) h_n^{(2)}\left(\frac{\omega}{c}r_{sw}\right) \frac{(2n+1)}{4\pi}\frac{(n-m)!}{(n+m)!} P_n^m(0)^2 \qquad (3.70)$$

Combining Equations (3.58), (3.69) and (3.70) yields

$$D_{sw}^{2.5D}(\theta_0,r_0,\omega) =$$

$$= \frac{1}{2\pi r_0} \sum_{m=-\infty}^{\infty} \frac{A_{|m|}^m(\omega)}{G_{|m|}^m(\omega)} e^{im\theta_0}$$

$$= \frac{1}{2\pi r_0} \sum_{m=-\infty}^{\infty} \frac{\left(-i\frac{\omega}{c}\right) j_n\left(\frac{\omega}{c}r\right) h_n^{(2)}\left(\frac{\omega}{c}r_{sw}\right) Y_n^m\left(\theta_{sw},\frac{\pi}{2}\right)^*}{\left(-i\frac{\omega}{c}\right) j_n\left(\frac{\omega}{c}r\right) h_{|m|}^{(2)}\left(\frac{\omega}{c}r_0\right) Y_{|m|}^m\left(0,\frac{\pi}{2}\right)^*} e^{im\theta_0}$$

$$= \frac{1}{2\pi r_0} \sum_{m=-\infty}^{\infty} \frac{h_n^{(2)}\left(\frac{\omega}{c}r_{sw}\right) \sqrt{\frac{(2|m|+1)}{4\pi}\frac{(|m|-m)!}{(|m|+m)!}} P_{|m|}^m\left(\cos\frac{\pi}{2}\right) e^{-im\theta_{sw}}}{h_{|m|}^{(2)}\left(\frac{\omega}{c}r_0\right) \sqrt{\frac{(2|m|+1)}{4\pi}\frac{(|m|-m)!}{(|m|+m)!}} P_{|m|}^m\left(\cos\frac{\pi}{2}\right) \underbrace{e^{-i0}}_{1}} e^{im\theta_0}$$

$$D_{sw}^{2.5D}(\theta_0,r_0,\omega) = \hat{S}_{sw}(\omega) \times \sum_{m=-\infty}^{\infty} \frac{1}{2\pi r_0} \frac{h_{|m|}^{(2)}\left(\frac{\omega}{c}r_{sw}\right)}{h_{|m|}^{(2)}\left(\frac{\omega}{c}r_0\right)} e^{im(\theta_0-\theta_{sw})} \qquad (3.71)$$

For a finite even number L of discrete loudspeakers the most suitable choice is to use as many order yields (N=L-1)

$$D_{sw}^{2.5D}(\theta_0,r_0,\omega) = \hat{S}_{sw}(\omega) \times \sum_{m=-\frac{L-1}{2}}^{\frac{L-1}{2}} \frac{1}{2\pi r_0} \frac{h_{|m|}^{(2)}\left(\frac{\omega}{c}r_{sw}\right)}{h_{|m|}^{(2)}\left(\frac{\omega}{c}r_0\right)} e^{im(\theta_0-\theta_{sw})} \qquad (3.72)$$

3.6.4 Sampling Requirements for Reproduction

A rule of thumb is given in [Pol05 and references therein] for the number of loudspeakers L required to produce an accurate reconstruction of a plane wave for a given head radius r and frequency $\omega = k/c$,

$$L \geq (\lceil kr \rceil + 1)^2,$$

where $\lceil . \rceil$ denotes rounding up to the nearest integer. The sampling requirements for recording and reproducing a sound field up to a given order N are the same. It can be shown that an accurate construction of 3D fields is not practical across the entire audio range [Pol05]. However, relatively large arrays of around 100 loudspeakers would allow reconstruction for a single listener up to 4 kHz, which would allow the majority of the sound field energy to be reproduced accurately and many of the directional cues to be recreated.

3.6.5 Error Performance

The performance of Higher Order Ambisonics can be quantified analytically assuming ideal spherical loudspeaker layout (Sphere) and an anechoic environment by determining the normalized radial error [Pol05]

$$\varepsilon(r,\omega) = \frac{\int_0^{2\pi} \int_0^{\pi} \left| P(\theta,\varphi,r,\omega) - \tilde{P}(\theta,\varphi,r,\omega) \right|^2 \sin(\theta)\, d\theta d\varphi}{\int_0^{2\pi} \int_0^{\pi} |P(\theta,\varphi,r,\omega)|^2 \sin(\theta)\, d\theta d\varphi} \quad , \qquad (3.73)$$

where $\tilde{P}(\theta,\varphi,r,\omega)$ is the field approximated by the reproduction system. There are two normalized radial errors of interest. The first is the theoretical truncation error caused by the truncation of the spherical source expansion in Eq. (3.2) to a maximum order $n = N$. This ignores the aliasing caused by having a finite number of loudspeakers. The truncation error therefore represents a lower bound on the reproduced error [Pol05].

The second error of interest is the actual, reproduced field error for a given source position and set of loudspeaker weights (or driving functions). This error includes both the truncation error and the aliasing error caused by the finite number of loudspeakers used.

3.7 Vector Base Amplitude Panning (VBAP)

Vector Base Amplitude Panning (VBAP) has been introduced by Ville Pulkki [Pul97] as an extension of stereophonic techniques. Panning, i.e. suitably weighted playback of a signal on a loudspeaker setup can produce the impression of one single sound source. VBAP is a sound-source oriented approach in the sense that it aims at the creation of phantom sound sources rather than at global sound field reconstruction as by Wave Field synthesis (WFS) or Higher order Ambisonics (HOA). However, VBAP enables the creation of two- or three-dimensional sound scenes where any number of loudspeakers can be placed arbitrarily.

In simple amplitude two/dimensional panning method (see Figure 3.9) two loudspeakers are used, which formulate a two-dimensional vector base. The base is defined by the unit-vectors $l_1 = [l_{11}\ l_{12}]^T$ and $l_2 = [l_{21}\ l_{22}]^T$ pointing toward loudspeakers. The unit-vector $\boldsymbol{p} = [p_1\ p_2]^T$ pointing toward the virtual source can be treated as linear combination of the loudspeakers vectors as

$$\boldsymbol{p} = g_1 l_1 + g_2 l_2 \ , \tag{3.74}$$

where g_1 and g_2 denote nonnegative gain factors. The Equation 1 can be formulated in matrix form as

$$\boldsymbol{p} = \boldsymbol{g}\boldsymbol{L_{12}} \ , \tag{3.75}$$

where $\boldsymbol{g} = [g_1\ g_2]$ and $\boldsymbol{L_{12}} = [l_1\ l_2]^T$. The Equation 2 can be solved if the invers matrix $\boldsymbol{L_{12}^{-1}}$ exists, yielding

$$\boldsymbol{g} = \boldsymbol{p}^T \boldsymbol{L_{12}^{-1}} = [p_1\ p_2]\begin{bmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{bmatrix}^{-1} \tag{3.76}$$

The gain factors that control the channel levels have to be normalized. The sound power can be set a constant value C, whereby the following approximation can be stated

$$g_1^2 + g_2^2 = C \tag{3.76}$$

Thus, the normalized gain factors are given as

$$\boldsymbol{g_{norm}} = \frac{\sqrt{C}\boldsymbol{g}}{\sqrt{g_1^2 + g_2^2}} \tag{3.78}$$

The virtual source can be produced by the loudspeaker base on the active arc of which the virtual source is located. Thus the sound field that can be produced with VBAP is a union of the active arcs of the available loudspeaker bases.
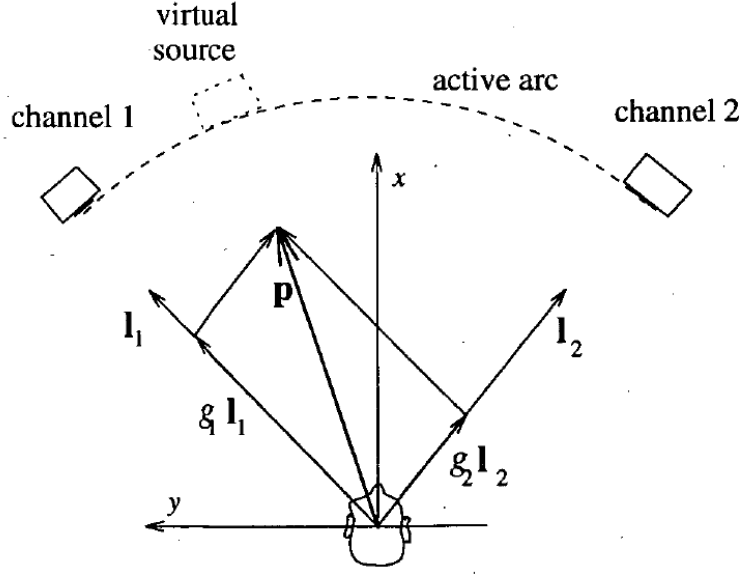
Figure 3.9: Illustration of the basic two-dimensional VBAP [Pul97]

The two-dimensional VBAP method may now be generalized to the three-dimensional VBAP method utilizing loudspeaker triplet. Let the loudspeakers be positioned on the surface of a sphere, equidistant from the listener. The three-dimensional unit-vectors l_1, l_2 and l_3 define the directions of loudspeakers 1, 2 and 3 respectively. The unit-vector pointing at the virtual source $= [p_1 \ p_2 \ p_3]^T$ defines the virtual source direction. In analogy to the two-dimensional VBAP we derive the normalized gain factors as follows

$$\boldsymbol{p} = g_1 l_1 + g_2 l_2 + g_3 l_3 \ , \tag{3.79}$$

and therefore

$$\boldsymbol{p} = \boldsymbol{g} \boldsymbol{L_{123}} \ , \tag{3.80}$$

where $\boldsymbol{g} = [g_1 \ g_2 \ g_3]$ the gain factors and $\boldsymbol{L_{12}} = [l_1 \ l_2 \ l_3]^T$. Vector $\boldsymbol{g}$ can be solved,

$$\boldsymbol{g} = \boldsymbol{p}^T \boldsymbol{L_{12}^{-1}} = [p_1 \ p_2 \ p_3] \begin{bmatrix} l_{11} & l_{12} & l_{13} \\ l_{21} & l_{22} & l_{23} \\ l_{31} & l_{32} & l_{33} \end{bmatrix}^{-1} \tag{3.81}$$

and finally

$$\boldsymbol{g}_{norm} = \frac{\sqrt{C}\boldsymbol{g}}{\sqrt{g_1^2 + g_2^2 + g_3^2}} \ , \tag{3.82}$$

where C is given as $C = g_1^2 + g_2^2 + g_3^2$ .The virtual source can be produced by the loudspeaker base on the active triangle, which is region on the surface of the sphere onto which the virtual source can be positioned (see Figure 3.10).
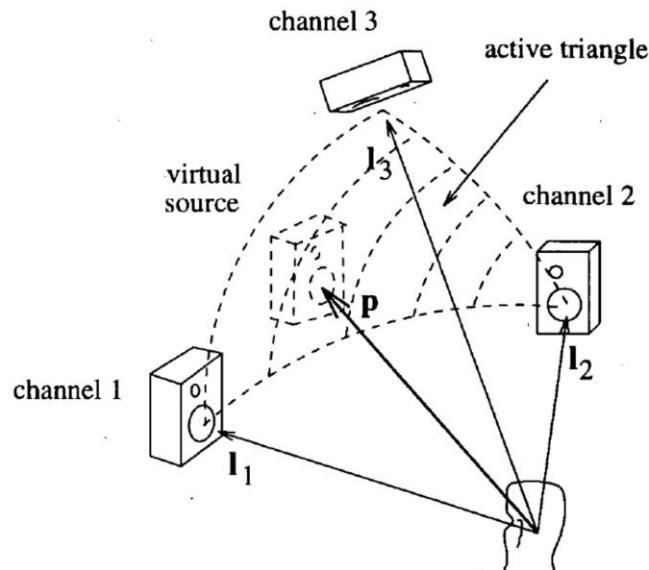
Figure 3.10: Illustration of the basic three-dimensional VBAP [Pul97].

The three-dimensional VBAP can be applied to three-dimensional array- based systems to represent virtual sources. The number of loudspeakers in a base is obviously three, and each loudspeaker can belong to several bases. The active triangles of bases should not be intersecting, and they should be selected so that maximum localization accuracy in each direction is provided. When good localization accuracies on a large listening area are desired, the dimensions of the active regions must be decreased (i.e. more loudspeakers) [Pul97].

Some Features of VBAP:

Vector Base Panning is a very simple and efficient way of achieving periphonic sound spatialisation. It is very flexible regarding the loudspeaker layout: a homogeneous loudspeaker distribution will provide more homogeneous localization quality, but irregularities in the layout will only locally affect the reproduction, due to the angular limitation of active loudspeakers. Daniel has pointed out that this limitation also gives a relative good stability of localization at off-center listening positions [Dan00], since a sound source can generally not be 'dragged' towards the loudspeakers that are closer to the listener like in holophonic techniques (Wave Field Synthesis, Ambisonics, etc.), where all loudspeakers are active at each moment.

The distance of a sound source cannot be reproduced by VBAP. Generally, the virtual sources appear to be moving on the surface of the loudspeaker layout. Outside sources (with a radius bigger than the radius of the loudspeaker array) can be simulated by adding artificial reverberation. However, it is not possible to reproduce sound sources within the loudspeaker layout, like in the case of holophony or Higher Order Ambisonics. The distance effect can be implemented by controlling the amplitude and delay of sound signal. The amplitude is calculated with $1/r$ law to simulate the distance attenuation. The sound signal is delayed to simulate the air propagation delay.

Chapter 4

Loudspeaker Array Design

Introduction

The process of designing a suitable loudspeaker layout for a reproduction sound system is not a trivial task. This thesis will support multiple reproduction methods, which have different requirements regarding the loudspeaker layout. In this Section, we will define some common layouts and present different strategies for finding suitable layouts considered for reproduction methods presented in Chapter 3.

4.1 Wave Field Synthesis

In this thesis only two-dimensional WFS-systems are considered corresponding to so-called WFS 2.5 (Section 3.3.2). These systems allow flexible layouts, which basically consist of a defined convex contour (two-dimensional) surrounding the listening area of a plane (typically horizontal). Circular and rectangular layouts will be basically of interest. However, for practical applications the contour of a layout can be adapted to a given room geometry in predefined plane.

4.2 Higher Order Ambisonics

In Higher Order Ambisonic systems, the regularity of the loudspeaker layout plays a more important role than in the case of Vector Base Panning. However, the homogeneity of localization in a VBAP system will also benefit from a regular layout. Since in Ambisonics systems generally all loudspeakers contribute to the reproduction of a sound field, irregularities in the layout globally affect the sound field reconstruction. Daniel has given definitions regarding regularity of a loudspeaker layout in the Ambisonics sense in [Dan00]. A loudspeaker layout is referred to as *regular*, if it preserves orthonormality. This can be examined using the N3D encoding to check [Dan00] as

$$\frac{1}{M} C . C^T = I_M \,, \tag{4.1}$$

where M is the number of Ambisonic channels, $. C^T$ is the transposed Ambisonic re-encoding matrix from (Section 3.5.4), and I_M is the M-by-M unity matrix. Equation (4.1) depends on the Ambisonics order N, but not on the order of rows (loudspeakers) and columns (Ambisonics channels) in C, nor on the orientation of the loudspeaker layout in the

coordinate system. Also note that layouts that are regular for an order N are always regular for all lower orders as well, and that regularity of a layout implies that the $L \geq N$ criterion is fulfilled. Daniel has noted that only five layouts fulfill the regularity criterion defined in Equation (3.1).These layout are known as the *Platonic solids*, discussed in Section 4.4.1. However, these provide 20 vertices at the most and regularity in the Ambisonics sense for a maximum order of M = 2. Daniel has thus suggested looking for quasi-regular layouts for higher orders using the method of *geodesic spheres* (Section 4.2.). This technique is indeed suitable for the design of well-conditioned Higher Order Ambisonics loudspeaker layouts.

4.3 Vector Base Panning

For implementation of a Vector Base Panning system, a triangulation of the surface defined by the loudspeaker layout has to be performed (Section 3.7). To achieve a homogeneous localization quality, the shape and size of the different triangles should not differ too much from each other. Rather than first defining a loudspeaker layout, which is then triangulated, we might find suitable layouts defined as structures of vertices and facets from the theory of *geodesic spheres* (Section 4.5), keeping the differences between the triangles at a minimum. Best results are obtained if the loudspeakers are placed at equal distances from the sweet spot.

4.4 Polyhedra

Generally, it is possible to approximate a spherical surface by means of various polyhedra, which can be classified according to the properties of their vertices, faces, and edges [Wik4]

• A polyhedron is convex if the line segment joining any two points of the polyhedron is contained in the polyhedron or its interior.

• A polyhedron is vertex-uniform if all vertices are the same, in the sense that for any two vertices symmetry of the polyhedron mapping the first isometrically onto the second exists.

• A polyhedron is edge-uniform if this symmetry is also given for its edges, and face-uniform if given for its faces.

• A polyhedron is regular if it is vertex-uniform, edge-uniform and face-uniform. This implies that every face is a regular polygon and all faces have the same shape.

• A polyhedron is quasi-regular if it is vertex-uniform and edge-uniform, and every face is a regular polygon.

• A polyhedron is semi-regular if it is vertex-uniform and every face is a regular polygon. The convex ones consist of the prisms and antiprisms and the Archimedean solids.

• A polyhedron is uniform if it is vertex-uniform and every face is a regular polygon, i.e. it is regular, quasi-regular, or semi-regular, but not necessarily convex.

Note that the definitions regarding regularity and semi-regularity do not match the according definitions provided by Daniel to evaluate the regularity of an Ambisonics loudspeaker system (Section 4.2). For example, the dodecahedron (Section 4.4.1) is regular according to the above definition, but it is not regular in the Ambisonics sense for an Ambisonics order of $N = 3$, although it does provide enough loudspeakers for a third order system according to the $L \geq N$ criterion.

An interesting property of polyhedra is the Euler characteristic, which relates the number of edges E, vertices V, and faces F of a simply connected polyhedron:

$$V - E + F = 2$$

In this Section, we will present different polyhedra with interesting properties regarding loudspeaker distribution in different reproduction systems. Since we are primarily interested in layouts approximating a spherical surface, we will only consider convex polyhedra, including the five regular convex polyhedra, which are also called the Platonic solids

4.4.1 Platonic Solids

There are only nine polyhedra which are regular by the definition given above. The five convex ones are also known as the Platonic solids. These are the tetrahedron, the hexahedron (i.e. the cube), the octahedron, the dodecahedron, and the icosahedron. They are shown in Figure 4.1, and their properties are summarized in table 4.1
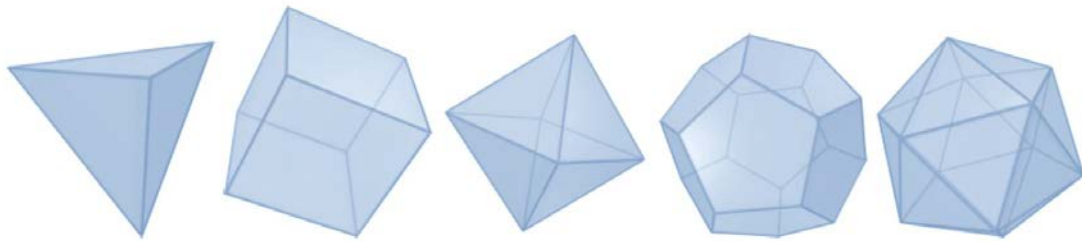

Figure 4.1: The Platonic solids: tetrahedron, hexahedron, octahedron, dodecahedron, icosahedron [Wik4].

Polyhedron	Vertices	Facets	Facet Shape
Tetrahedron	4	4	triangular
Hexahedron (cube)	8	6	square
Octahedron	6	8	triangular
Dodecahedron	20	12	pentagonal
Icosahedron	12	20	triangular

Table 4.1: Platonic solids.

Obviously, the Platonic solids offer only a very limited amount of vertices with a maximum of 20 vertices in the case of the dodecahedron. For a periphonic sound spatialisation system in a large-scale virtual environment, this does not satisfy our requirements. Also, only the octahedron and the icosahedron provide a ring of loudspeakers in the horizontal plane. We will discuss possible extensions of these arrangements by the method of *geodesic spheres* in the next Section.

4.5 Geodesic Spheres

In previous Section we have seen that mathematical regularity of a polyhedron unfortunately translates to a rather low number of vertices. Arrangements with more vertices can be found among the semi-regular and other classes of polyhedra. However, the number of loudspeakers in these arrangements is still pre-determined by the solid itself. A way to introduce a higher number of vertices (i.e. loudspeakers) while retaining the advantages regarding the regularity of the Platonic solids is given by the method of *geodesic spheres*. This method is used in architecture to distribute stress on 'geodesics', i.e. large circles around a spherical structure. Usually, it is described as the tessellation of the faces of a Platonic solid. This process creates new vertices which are then 'pushed out' to the radius of the polyhedron's circumscribed sphere. The faces are rearranged as well in order to include the new vertices. This process can be repeated in an iterative way. Figure 4.2 shows the process of building a geodesic sphere from an icosahedron. The first picture shows the original icosahedron, being triangulated in the second picture. The third picture shows the triangulated icosahedron with the new vertices pushed out to its circumsphere. Repeated application of the process of pictures 1 and 2 on the polyhedron from picture 3 gives the resulting geodesic sphere in picture 4.
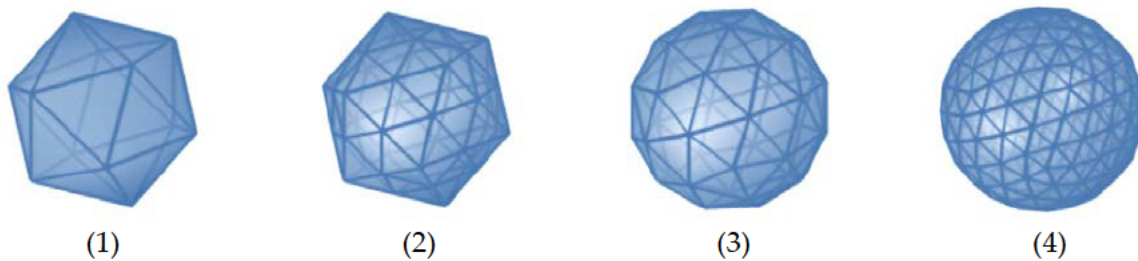

(1)   (2)   (3)   (4)

Figure 4.2: Building a geodesic sphere from an icosahedron

Since we aim at a choice of the total number of vertices (i.e. loudspeakers) as arbitrarily as possible, we will generalize the concept of geodesic spheres, so that it can be applied to other structures than the Platonic solids as well. For this purpose, we can define the following tessellation rules for differently shaped faces of a polyhedron:

- Triangles can either be midpoint-triangulated or triangulated at an arbitrary 'frequency'.
- Rectangles can be midpoint-triangulated or rectangulated at an arbitrary 'frequency'
- Polygons with more than four vertices (pentagons, hexagons, etc.) can only be midpoint-triangulated.

The term midpoint-triangulation is used to describe the process of adding a new vertex in the center of an arbitrary polygon and connecting it to each of the existing vertices in the polygon, resulting in a number of new facets which is equal to the number of vertices in the original polygon. Midpoint-triangulations of a triangle, a rectangle, and a pentagon are shown in Figure 4.3. Another tessellation strategy bases on the subdivision of a polygon's

edges, resulting in a triangulation of triangles or a rectangulation of rectangles. The tessellation frequencies $f - 1 > 2$ refers to the number of subSections in one edge. Additionally, we use f = 2 to denote its midpoint-triangulation.
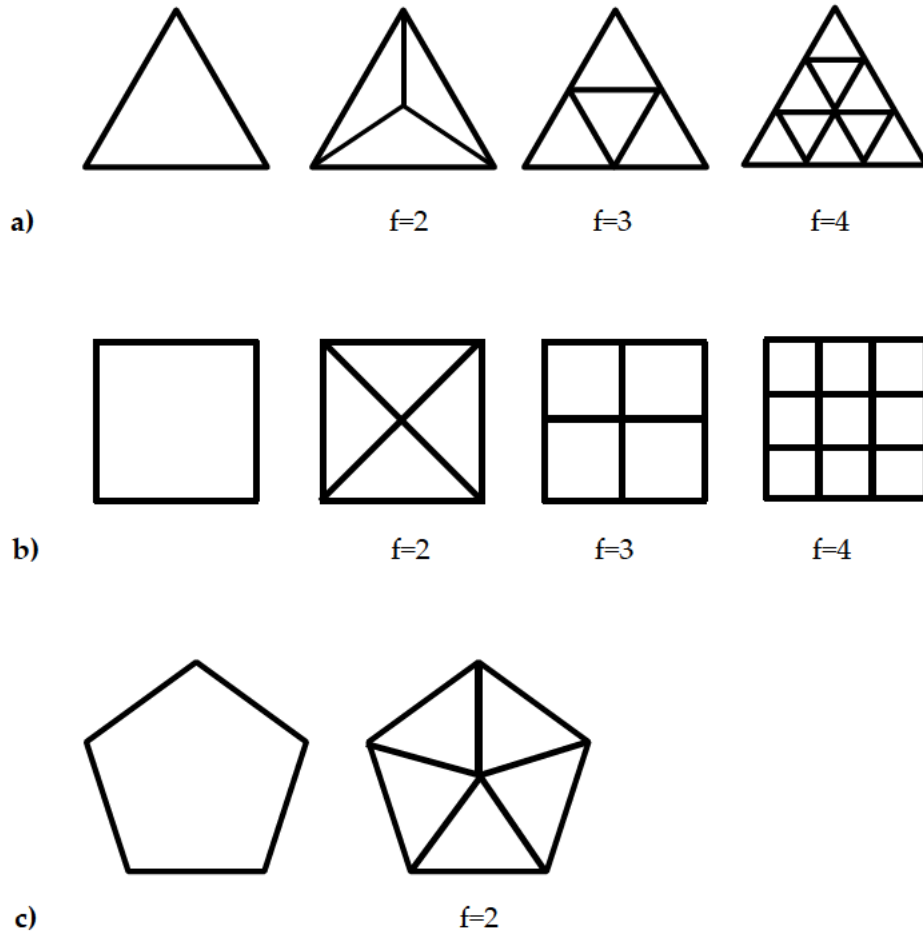

a)   f=2   f=3   f=4

b)   f=2   f=3   f=4

c)   f=2

Figure 4.3: **a)** A triangle, its midpoint-triangulation and two triangulations at different frequencies. **b)** A rectangle, its midpoint-triangulation and two rectangulations at different frequencies. **c)** A pentagon and its midpoint-triangulation.

4.6 Lebedev Quadrature Nodes

In numerical analysis, Lebedev quadrature [Wik5], named after Vyacheslav Ivanovich Lebedev, is an approximation to the surface integral of a function over a three-dimensional sphere. The grid is constructed so to have octahedral rotation and inversion symmetry. The number and location of the grid points together with a corresponding set of integration weights are determined by enforcing the exact integration of polynomials (or equivalently, spherical harmonics) up to a given order.
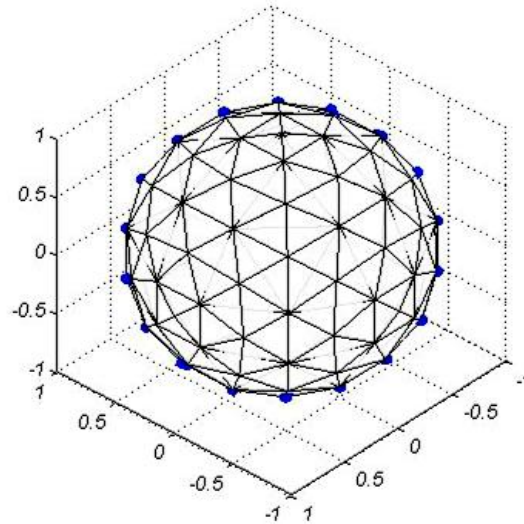
Figure 4.4: 17-th order Lebedev grid (110 nodes) on unit sphere.

The grid points can also represent an interesting layout where the distance between the neighboring points (vertexes) are approximately equal. This property will be advisable by VBAP method. The Nodes positions are computed up to the 131-th order [Leb99].The table 4.2 shows equivalent number of nodes for given orders.

Table 4.2: Equivalent number of nodes for orders

Order	Number of Nodes	Order	Number of Nodes	Order	Number of Nodes
3	6	25	230	77	2030
5	14	27	266	83	2354
7	26	29	302	89	2702
9	38	31	350	95	3074
11	50	35	434	101	3470
13	74	41	590	107	3890
15	86	47	770	113	4334
17	110	53	974	119	4802
19	146	59	1202	125	5294
21	170	65	1454	131	5810
23	194	71	1730		

4.7 Sphere Discretization

Other spherical layouts can be derived from the angular discretization of a sphere. Two cases will be considered

- Equiangular discretization: Equal discretization with N point for Elevation and Azimuth angles yields N^2 points (Figure 4.5).
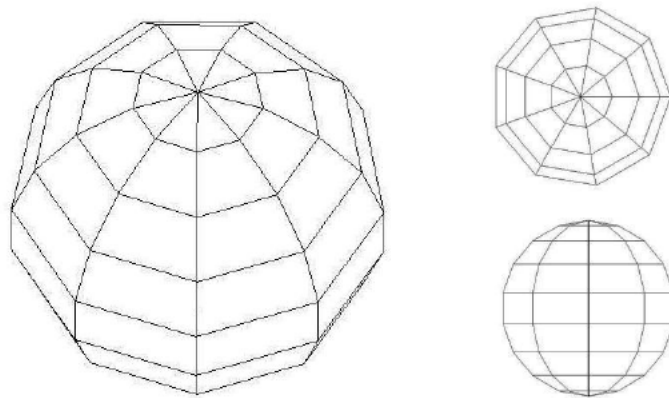

Figure 4.5: Illustration of Equiangular discretization ($N = 10$)

- Gaussian grid: Discretization with N points for Elevation and $2N$ for Azimuth angle yields $2N^2$ points (poles are not considered) (Figure 4.6).
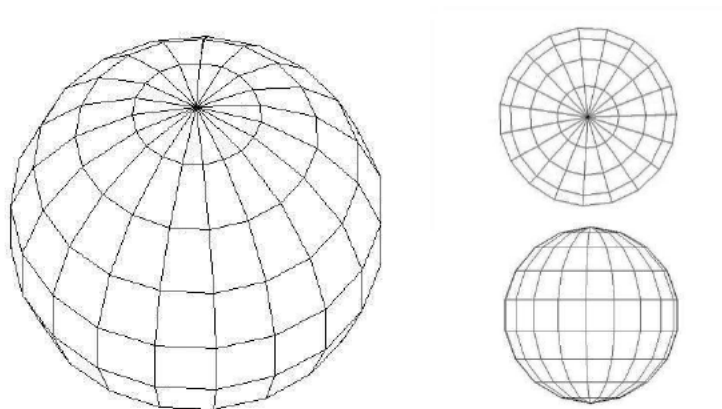

Figure 4.6: Illustration of regular Gaussian Grid ($N = 10$).

Chapter 5

Auralization

Introduction

The aim of auralization in room acoustics is to make the acoustics of a virtual (simulated) room audible. That is to say the result of an auralization is an aural impression of a source playing inside an existing or virtual room. A general definition of auralization is given by [Vor08]:"Auralization is the technique of creating audible sound files from numerical (simulated, measured, or synthesized) data."

Today, different room acoustics simulation software, for example EASE, presents a powerful tool for acoustic engineers. Three components have to be defined and modeled for auralization:

- source
- medium (room)
- receiver

During the simulation, the source radiates rays according to its directivity. The room acoustics, defined by its geometry and surface properties, transmits the rays. Finally, the receiver, modeled by applying so-called head-related transfer functions (HRTFs), keeps track of the hits. Hence, a binaural room impulse response (BRIR) can be calculated. The last step is to convolve the BRIR with an anechoic signal. The result can be listened to via headphones or a cross-talk canceled stereo loudspeaker[7] set.

In this Chapter the auralization process will be introduced, this involves the representation of sound sources, Simulation methods of room acoustics and the binaural reproduction. The related details to the work will be accentuated.

---

[7] CTC (Cross Talk Cancelation) is a technique for reproduce binaural signals by means of a stereo loudspeaker setup. For more Information we refer to [Vor08].

5.1 Representation of Sound Sources

Characterization and modelling of sound sources is the first component in the auralization process. There are different options to represent the radiation patterns of different types of sources (e.g. musical instruments, the human voice). However in this work the sources can be simplified and modeled as a point source (Section 2.5.1) or plane wave (far field).

5.2 Simulation Methods of Room Acoustics

The simulation of the room can be modeled by means of an LTI-system. It is fully characterized by its impulse response, which can be obtained by different methods. To calculate the room impulse response the sound propagation in room must be modeled. The quality of the auralization is closely related to the quality of the room acoustic model used for calculation of the impulse response .The computational modeling of sound propagation can be divided into wave based algorithms, geometrical acoustics and artificial methods [Vor08] (Figure 5.1).
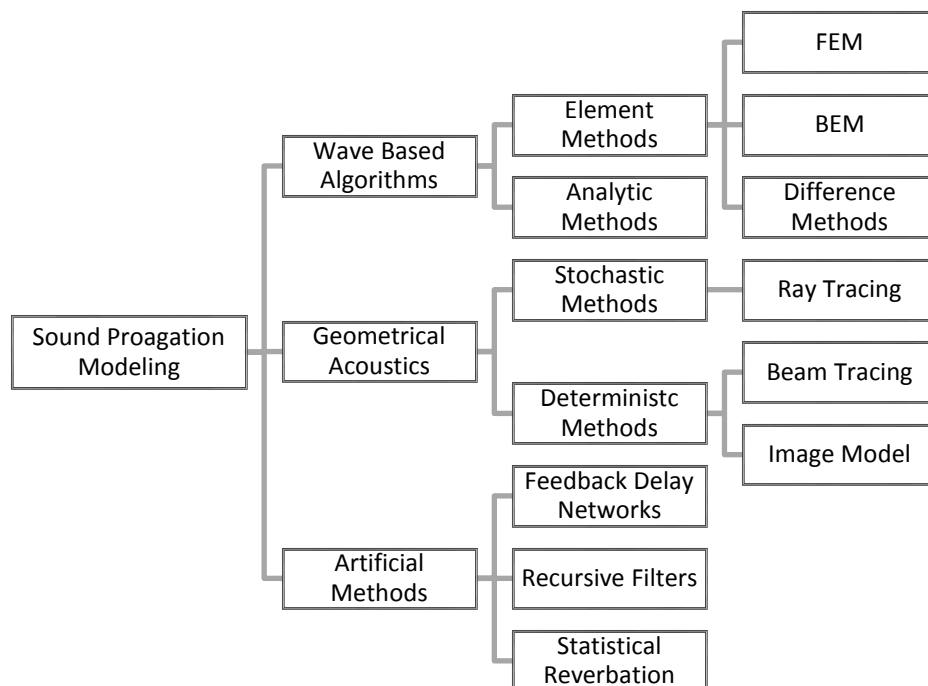
Figure 5.1: Simulation models for sound propagation [Vor08]

Due to computational constraints, not all of these techniques can be used for auralization purposes, especially not in real-time applications.

5.3 Geometrical Acoustics

The concept of geometrical acoustics is based on the assumption of sound propagation along straight lines comparable to ray optics. A sound ray is an energy bundle perpendicular to the wave front. It represents a spherical wave with an infinitely small opening angle $d\Omega$ (see Figure 5.2). The intensity of the ray decreases with law $1/r^2$, where r is the distance from the origin. Generally, two different algorithms of implementing geometrical acoustics, ray tracing and the image source method, can be distinguished. Ray tracing is a stochastic method (energy spreading by ray density, energy detection by volumes), whereas the image source algorithm belongs to the deterministic methods (energy spreading according to distance law, energy detection by points).

Figure 5.2: sound ray [Vor08]

5.3.1 Stochastic Ray Tracing

In the basic algorithm, the source radiates sound rays at $t = 0$ in various directions. Every time a surface is hit, their initial amount of energy is reduced and their direction is changed. If a detector (mostly spheres) is hit by a ray, the energy, the direction of incidence and the time difference since radiation are detected. The result can be presented as a histogram. The absorption and scattering effects at the boundary could be modeled [Vor08].
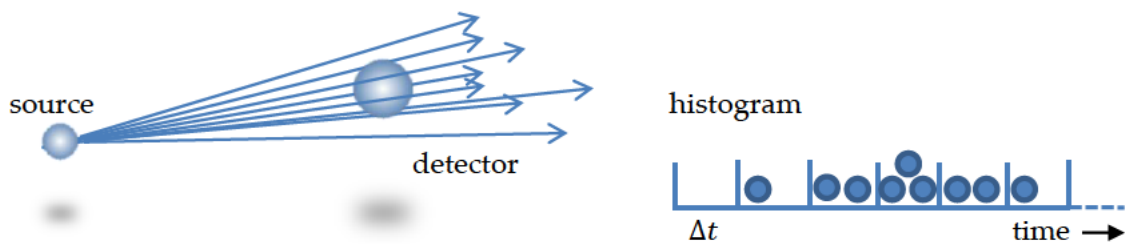

Figure 5.3: Left: In stochastic ray tracing energy spreading is implemented by counting. Right: The impulse response is created by counting events. An array called a histogram stores the time, angle and energy of each detected ray [Vor08].

5.3.2 Image Source Method

In the basic image source principle the sound source is mirrored at each plane or the boundary. The obtained image sources are again mirrored, which leads to image sources of second, third, etc. order. As a result the original room can be represented by an infinite pattern. Now, the idea is to replace the tracing of individual sound rays by adding the contributions of all image sources (see Figure 5.4). In the process the energy spreading by distance is included by the $\frac{1}{r^2}$ law. Since this model is strictly deterministic, receivers are points.

Not all of image sources are audible at the receiver position. Therefore, a so-called audibility check has to be performed for each image source.
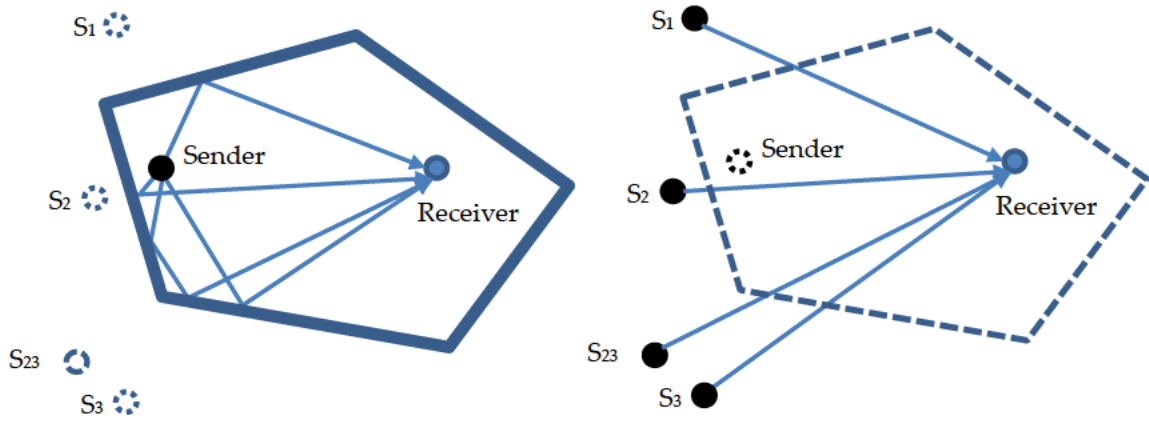
Figure 5.4: Left: Illustration of the reflections and the image sources of the first order (S_1, S_2, S_3) and of second order (S_{23}). Right: Modeling the reflections through the audible image sources.

5.4 Binaural Room Impulse Response (BRIR)

The next step towards auralization in room acoustics is the calculation of the binaural room impulse response (BRIR). Applying the room acoustic simulation delivers the binaural room impulse response, which will be presented through calculated reflections, described for each channel (ear) by time of arrival, sound pressure level for each octave band, and angle of incidence.

Let H_j be the spectrum of the *j-th* reflection, respectively the *j-th* image source, is calculated by

$$H|_{left,right} = \frac{e^{-i\omega t_j}}{ct_j} H_{source}(\theta,\varphi).H_{air}(\theta,\varphi).HRTF(\theta,\varphi)|_{left,right}\cdot\prod_{i=1}^{n_j} R_i \qquad (5.1)$$

Where t_j denotes the delay, ωt_j the phase, $1/ct_j$ decay law of the spherical wave, $H_{source}(\theta,\varphi)$ the directivity dependent transfer function of the source, $H_{air}(\theta,\varphi)$ the air attenuation (low pass behavior), R_i the complex geometrical reflection factors of the involved

wall i, and $HRTF(\theta, \varphi)|_{left,right}$ the head-related transfer function (see Section 5.5) of the left/ right ear at a certain orientation of the listener. Subsequently, the component $H|_{left,right}$ is transformed into the time domain by an inverse Fourier transformation (Eq. 2.6b) and then added to the BRIR [Vor08].

5.5 Reverberant Sound Wave Field in Auralization

The room impulse responses are usually considered to comprise three successive parts: Direct sound, early reflections and a tail of stochastic reverberation (see Figure 5.5). Historically, reverberation in a room has always been characterized only by the reverberation time (T_{60}). It is defined as the time in which the sound pressure level drops 60 dB after switching off a stationary sound source. The research on reverberation has traditionally focused on this reverberation time; for instance on how to calculate the reverberation time from the dimensions and material properties of the room. The relationship between reverberation time and absorption area is expressed through Sabine equation $A = 0.163\, V/T$ (absorption area A in m^2, volume V in m^3, T in s) [Vor08].
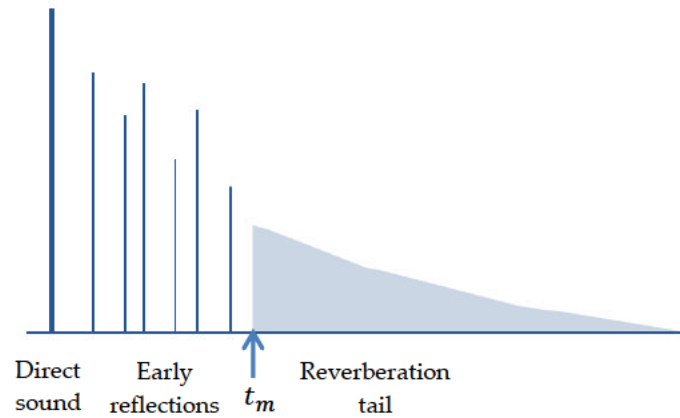
Figure 5.5: Schematic representation of room impulse response (time energy curve)

In literature reverberation is generally assumed to be *diffuse* [Kut91]. Diffuse means that the wave field is built up from an unlimited number of uncorrelated and equally strong plane waves, arriving from all directions. This property can be split into two parts: *homogeneity* and *isotropy*. However, many researchers considered the isotropy of a wave field only, and sometimes even considered a wave field to be diffuse when only the isotropy quality was determined.

The transition point between early reflections and the stochastic reverberation tail is called the physical mixing time (t_m) [Lin10a]. Due to increasing reflection density and diffuseness of the decaying sound field individual reflections become perceptively less distinguishable. This provides an opportunity for reducing the length of binaural impulse responses that are dynamically exchanged in virtual acoustic environments [Lin10a].

5.5.1 Modeled Reverberation Tail

In the context of Virtual Acoustic Environments it is useful reduce the length of binaural impulse responses that are dynamically exchanged in virtual acoustic environments. An obvious method to achieve this reduction would be to replace the individual reverberation tail of the BRIRs – after an instant in time when perceptual discrimination is no longer possible – with an arbitrary and constant reverberation tail, which can be stochastically generated [Lin10 and references therein]. The reverberation tail must be mixed properly (Alignment in time and level, in- out fading) into early part of BRIRs.

5.5.2 Reverberant Sound Wave Field Generation by Plane Wave Synthesis

Alternatively to the modeled reverberation tail approach, the reverberation can be generated using plane wave synthesis of the desired reverberant sound wave field. Maat [Maa97] has determined the minimum number of uncorrelated plane waves needed to synthesize a perceptually reverberant sound field in the horizontal plane. Maat considered that —as an isotropic wave field is rotation invariant— a reverberant wave field should be considered *perceptually isotropic* if an observer cannot distinguish the original sound field from a rotated version of the original sound field. Maat concluded that at least 10 plane waves are needed to construct a reverberant, perceptually isotropic sound field in the horizontal plane. This number was even found to be independent of the presence or absence of direct sound and early reflections. The minimum number of waves needed varied between 6 and 10 for the different subjects. Therefore Maat determined 10 plane components to be a very safe upper limit, well above the observed error margins. Reverberant sound wave field generation by plane wave synthesis is widely used by WFS systems. In concept of Auralization the needed uncorrelated plane waves will be additional, virtual sources to be rendered. Figure 5.6 illustrates the presented approach.
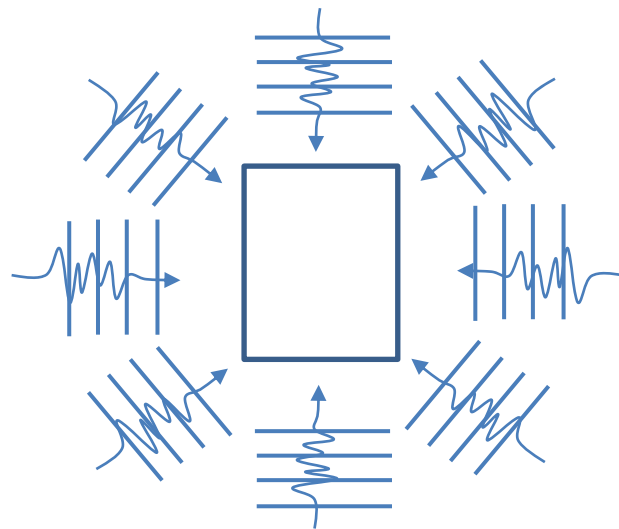

Figure 5.6: Schematic representation of 8 uncorrelated reverberant waves for rectangular layout.

5.6 Dynamic Auralization and Head Tracking

The head-related coordinate system is introduced in Figure 5.7, where the azimuthal angle θ between 0° (frontal direction) and 360° degrees (counterclockwise) describes the horizontal plane. The median plane is described by the polar angle φ between 0° (frontal direction) and +90°/ -90° (upper hemisphere/ lower hemisphere).

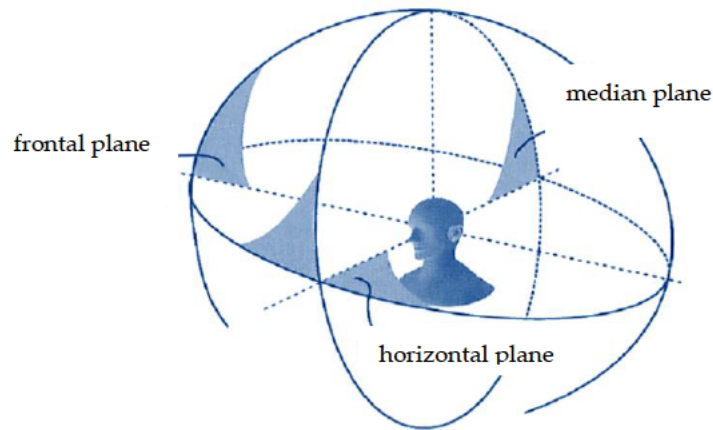Figure 5.7: A head-related coordinate system [Vor08].

Sound can be localized due to spatial hearing. For example a sound wave incident from the right-hand side of a listener has a shorter travelling time to the right ear than to the left one. These direction-dependent differences are called interaural time differences (ITD). Furthermore, the (frequency- and direction dependent) sound wave will be reflected, absorbed and diffracted by the head and torso. These differences in amplitude between the two ears are called interaural level differences (ILD). In the horizontal plane the localization is much better than in the median plane [Vor08]. The total linear distortions caused by head, pinna and torso are described formally by the HRTFs which are, by the way, individual since the size and shape of heads differ. HRTFs are defined by the proportion of the sound pressure at the eardrum or the entrance of ear canal to the sound pressure in the center position of the head(in the absence of the head). This relation is of course dependent on the direction of sound incidence. Figure 5.8 shows a HRTF and the corresponding HRIR of sound incident from the left-hand side.
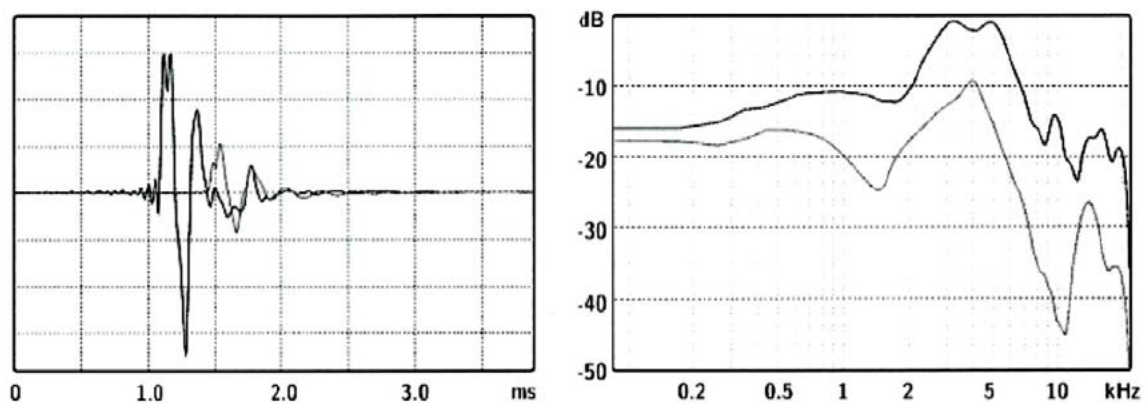

Figure 5.8: Left: A set of HRIRs. Right: The corresponding HRTFs [Vor08].

Binaural reproduction can be achieved directly by headphones or by two loudspeakers with crosstalk cancellation (transaural). In both methods, the aim is to reproduce the same signals at the listener's ears as if the listener were in the real room. This is obtained by listener modeling with the aid of head related transfer functions (HRTF) (corresponding to the according head-related impulse responses (HRIR)). The basic idea is outlined in Figure 5.8 and it is applied to calculate the BRIR (see Section 5.4). A mono signal can be shifted to any direction by filtering (convolution) it with an equivalent pair of HRIRs [Vor08]. This process can be adapted to Head movement denoting the dynamic auralization technique. An important feature of dynamic auralization is the so-called head tracking. A head tracker usually consists of a receiver somehow mounted at the listener and a transmitter located in front or above the head. This enables the determination of the current observer position and orientation relative to the environment. Hence, the BRIR can be synthesized according to Eq. (4.1); always containing the correct set of HRTFs. Devices for tracking head movements may be based on different technologies: ultrasound, electromagnetic, optical or mechanical.

A new approach used by the dynamic auralization is based on modified BRIRs, where the ITDs in used HRTFs dataset will be individualized and re-established between left and right ear signals [Lin10b]. This method will be used by the dynamic auralization experimeted in this thesis. Figure 5.9 illustrates the dynamic binaural synthesis with ITD individualization [Lin10b].
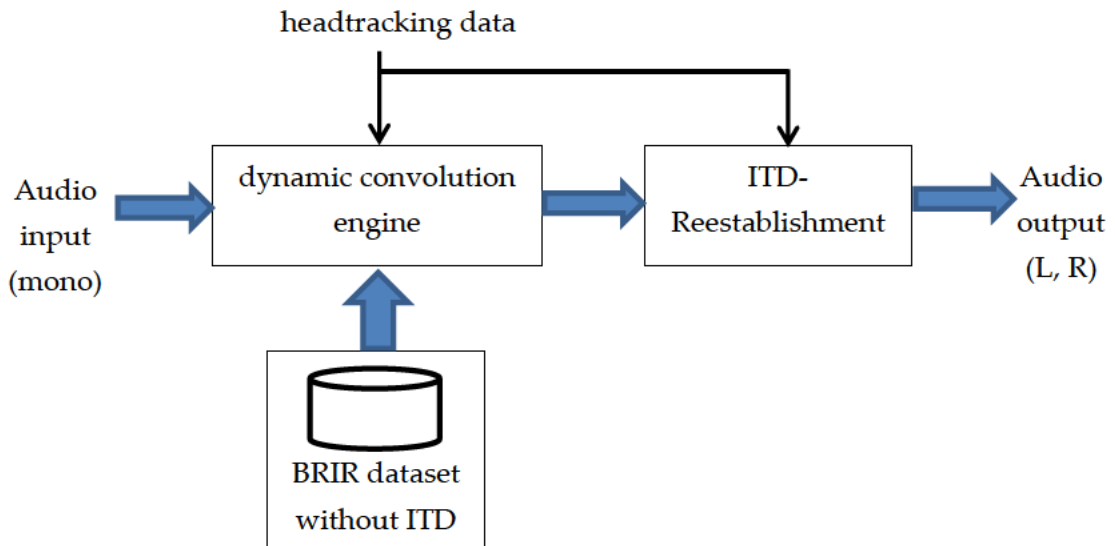

Figure 5.9: Schematic depiction of dynamic binaural synthesis with ITD individualization.

Chapter 6

Results & Discussion

This thesis project is dedicated to the introduced sound field synthesis methods. We have presented the theoretical foundations of sound field propagation (Chapter 2) as well as aspects of sound field reproduction (Chapter 3). The basics for loudspeaker layouts design were introduced (Chapter 4). Furthermore the aspect of Auralization was introduced (Chapter 5).

In this Chapter we will present the practical implementation steps done to build a software tool implemented in MATLAB. This tool enables designing different loudspeaker arrays and derivation of the driving functions for the different reproduction techniques (Chapter 2). The result will be used to dynamically auralize a modeled room acoustics simulated using the simulation program (EASE) via the virtual loudspeaker array.

6.1 Loudspeaker Arrays

A Matlab function was implemented (get_loudspeaker_Positions.m) to calculate the loudspeakers coordinates in Spherical Coordinate system for a given configuration. For two-dimensional configuration circular, rectangular and adapted distribution based on Klangakademie geometry (Figure 6.1) were implemented. Array dimensions, spatial positions and number of loudspeakers are configurable. Figure 6.2 illustrates the basic two-dimensional loudspeaker layouts in the modeled Klangakademie geometry. These Graphics are also done in Matlab with additional self-programmed scripts.

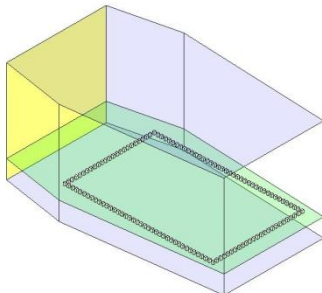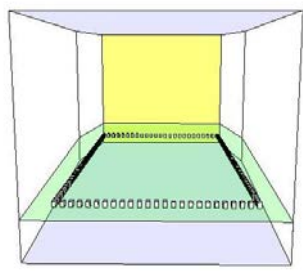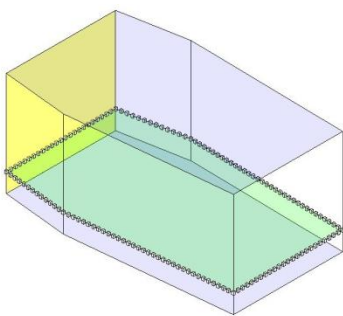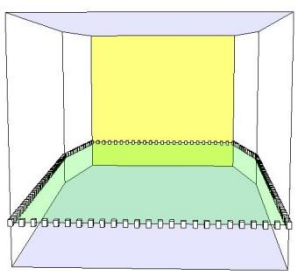| Klangakademie Berlin			
Dimensions	Max. width	14 m	
	Max. length	23.5 m	
	Mean height	11 m	
Area	312m²	Stage	91 m²
		Audience	221 m²
Volume	ca. 3432 m³		
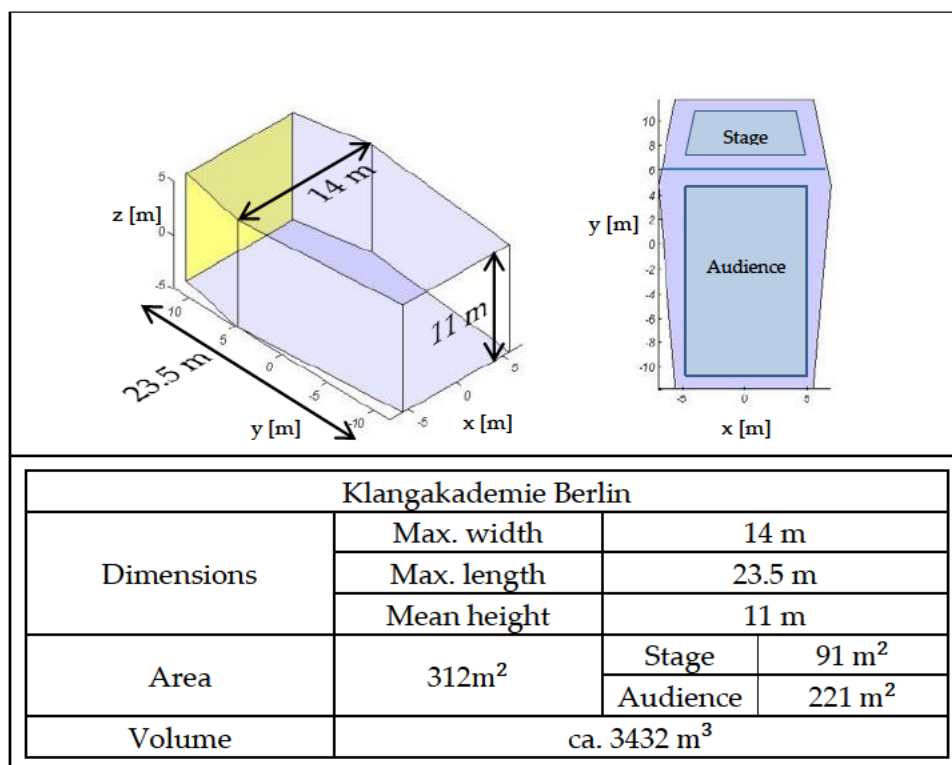
Figure 6.1: Schematic representation of Klangakademie geometry
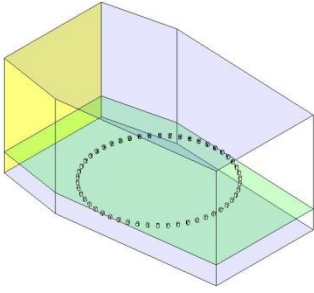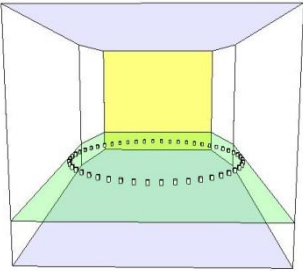
ID	Configuration	3D-view	Perspective 3D-view
1	Circular		
2	Rectangular		
3	Adapted		

Figure 6.2: Schematic representation of the two-dimensional loudspeaker layouts.

Three-dimensional spherical loudspeaker layouts introduced in Chapter 3 were implemented. The radius of the sphere (or hemisphere) and the desired number of loudspeakers can be adjusted. However, depending on the layout there are some limitations of choosing the number of the loudspeakers (e.g. Lebedev grid). Figures 6.3, 6.4 and 6.5 illustrate some three-dimensional arrays implemented.
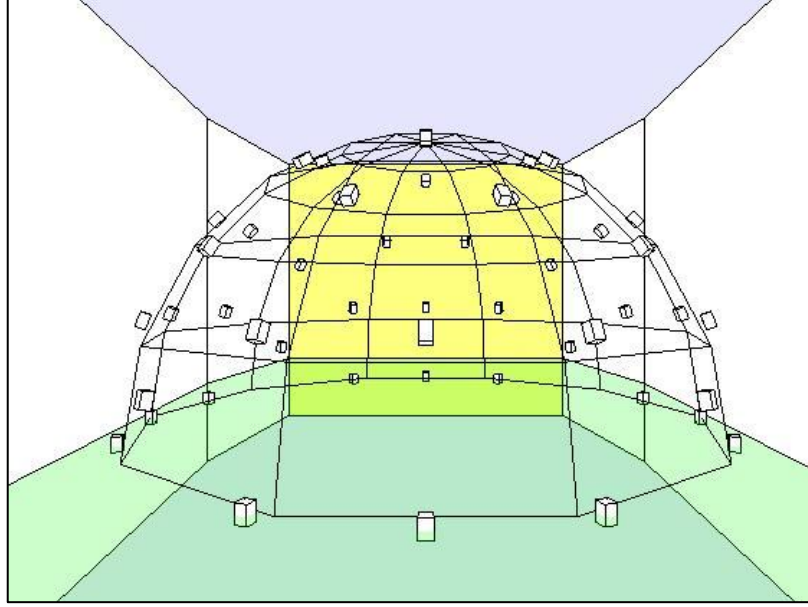

Figure 6.3: Lebedev grid of 49 nodes. The radius of hemisphere is 6.5 m and the array center is at the position $x_C = [0\ 0\ 2]^T$. For ease of Illustration a hemisphere grid is also shown.
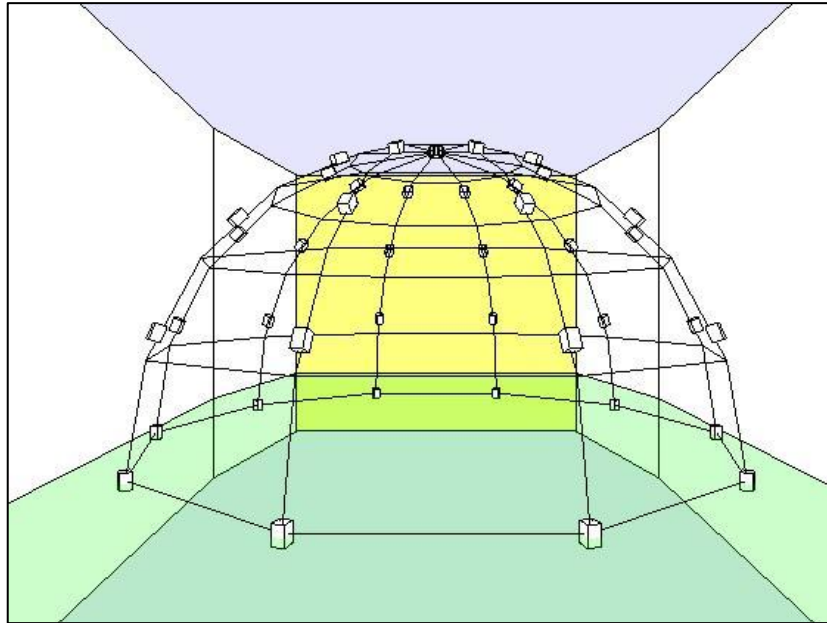

Figure 6.4: Gaussian grid of 50 nodes (N=5). The radius of hemisphere is 6.5 m and the array center is at the position $x_C = [0\ 0\ 2]^T$.
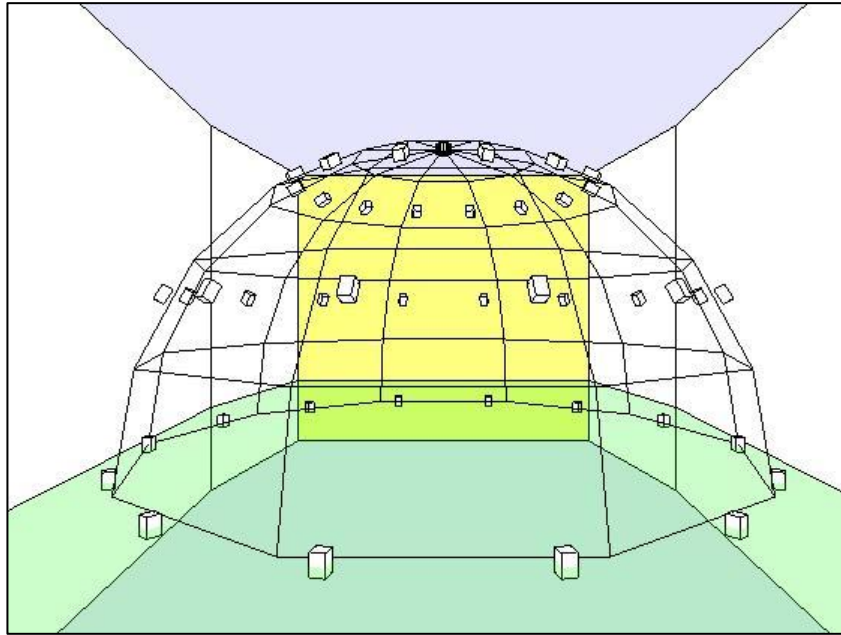
Figure 6.5: Equiangular grid of 49 nodes (N=7). The radius of hemisphere is 6.5 m and the array center is at the position $x_c = [0\ 0\ 2]^T$.

Platonic solids layouts and tessellation function are found for free by the Matlab central[8].

Matlab Implementation:

In the following we introduce the Matlab function implemented (get_Loudspeaker_position.m) for layout design.

```
[LS_Positions]= get_loudspeakers_Positions(Dimensions,…
                loudspeakers_layout_mode, halve_the_space)
```

Input:
```
Dimensions: Number of Dimensions:
       Values:
       -  Two dimensions  :   2
       -  Three dimensions:   3

loudspeakers_layout_mode: this vector defines the loudspeakers layout
description and takes the from

      For two dimensional case(Dimensions = 2) the loudspeakers_layout_mode
      vector is defined as:

      -  Circle       : [1, radius, Number_of_loudspeakers]
      -  Rectangle    : [2, Width, length, distance_between_2_loudspeakers]
      -  Klangakademie: [3, desired_Number_of_loudspeakers]
```

---

[8]Matlab central is a platform for exchanging knowledge and experience:
URL: http://www.mathworks.com/matlabcentral/fileexchange/

For two dimensional case(**Dimensions** = 3) the **loudspeakers_layout_mode**
vector is defined as:

- Lebedev quadrature: [1, radius, desired_Number_of_loudspeakers]
- Gauss grid   : [2, radius, desired_Number_of_loudspeakers]
- Equiangular
  Discretization   : [3, radius, desired_Number_of_loudspeakers]
- Platonic solid   : [4, radius, platonic_solid, Tesselation_factor]

radius: denotes the radius of sphere or hemisphere.
platonic_solid: defines the platonic solid used as follows

        - Tetrahedron : 1
        - Cube   : 2
        - Dodecahedron : 3
        - Octahedron   : 4
        - Icosahedron : 5

Tesselation_factor: defines the integer tessellation frequency, which
takes the minimum value = 2 middle point (Section 4.5).

halve_the_space: defines space domain as follows

        - Sphere   : 0
        - Hemisphere   : 1

Output:

LS_Positions: loudspeakers Spherical Coordinates in m✕3, where m the number
of loudspeakers. The position obey Matlab convention is defined as follows

```
               90                        90
               |                         |
      (L)180 -   - 0(R)     (Rear)180 -   - 0(Front)
        -180   |                         |
              -90                       -90


          Theta(Azimuth)          Phi(Elevation)
```

In the implementation Loudspeaker positions should be adapted to different
conventions used by derivation of driving function or HRTFS data set.

Dependencies:
The function "get_Loudspeaker_position.m" invokes help functions (tools) or
data structures (Data) for special layout geometries

Data:
- Lebendev_Nodes.mat: mat-structure including pre-calculated lebedev
quadrature nodes (table 3.2)

Tools:
        - Calculation of different platonic solids nodes:
                - tetrahedron.m
                - cube.m
                - dodecahedron.m
                - octahedron.m
                - icosahedron.m

        - Tessellation option
                - tesselatePoly.m
                - tesselateTri.m
                - addVertex.m

6.2 Calculation of Driving Functions

After designing the loudspeaker layouts the next step will be the implementation of the driving functions. Driving functions of loudspeakers should be calculated and implemented as FIR-filter (Finite Impulse Response filter) for every reflection (image source) of the room acoustic model. We will start with VBAP, which is relatively simple to implement. VBAP implementation will not consider the secondary source (loudspeaker) model. In contrast, wave field synthesis and higher order Ambisonics take the virtual source model and the used secondary source into account. The virtual source can be modeled as a plane wave or as a spherical wave and the secondary source will be modeled as point source (WFS, NFC-HOA) or as plane wave (APA). However, for ease of demonstration we will adapt to the simulated sound field description of used EASE program and represent the image sources as plane wave sources as a case of study. Additionally for ease of the final demonstrator some graphical tools were implemented in Matlab to illustrate the loudspeaker arrays and some examples of simulated reproduced sound field.

6.2.1 Vector Base Amplitude Panning (VBAP)

Calculating VBAP-gains for a given virtual source on a sphere (or hemisphere) and given spherical loudspeaker array involves finding the nearest 2 or 3 representing loudspeakers and then calculation of the gain factors (Section 3.7). Figure 6.6 illustrates the process where a virtual source (blue colored) is at the position $x_S = [\frac{\pi}{4} \ \frac{\pi}{4} \ 6.5]^T$. The active loudspeakers are colored red.
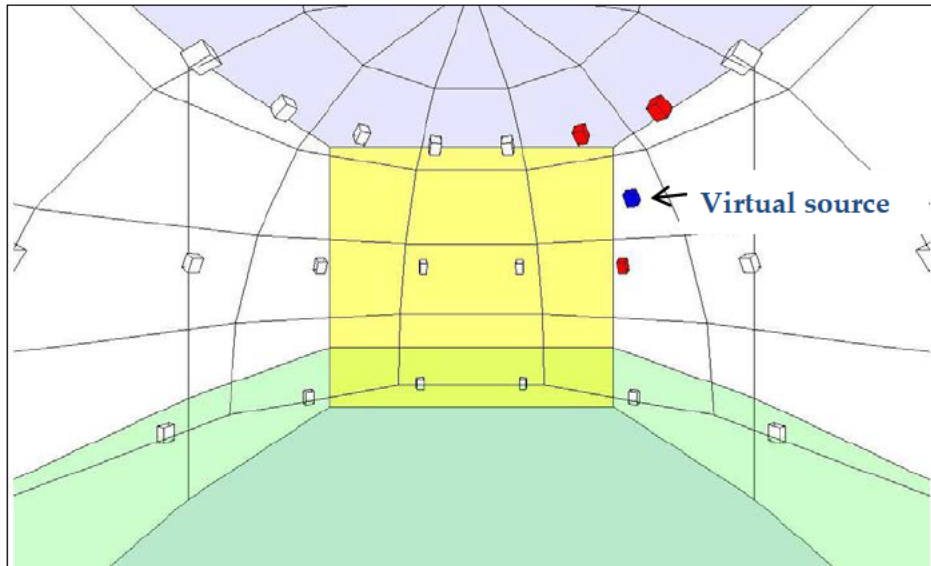

Figure 6.6: illustration of 3D-VBAP Implementation.

The driving function by VBAP will take the form of weighted delayed Dirac impulses, where the delay is defined by the arrival time of the plane wave of the corresponding virtual source or image source (reflection) and the weighting are adapted to the image source signal (reflection impulse response) .

6.2.2 Amplitude Panning Ambisonics (APA)

In Section 3.5.4 the APA approach was introduced. This assumes that the loudspeakers are far enough from the listener position. The derivation of the driving function can be split to encoding and decoding stages, which are performed as matrix operations. Similarly to VBAP APA-driving function will be interpreted as weighted delayed Dirac impulses.

a) Plane Wave as Virtual Source

The implementation of APA was tested for reproduction of a monochromatic virtual plane wave in the x-y plane. Figure 6.7 illustrates the effect of the system order N (expansion order) by reproduction of a virtual plan wave (f=500 Hz,). Here, the circular array used consists of 100 loudspeakers and has a radius of 2m. Obviously the higher the order the better will be the reconstruction over a bigger area (sweet spot).
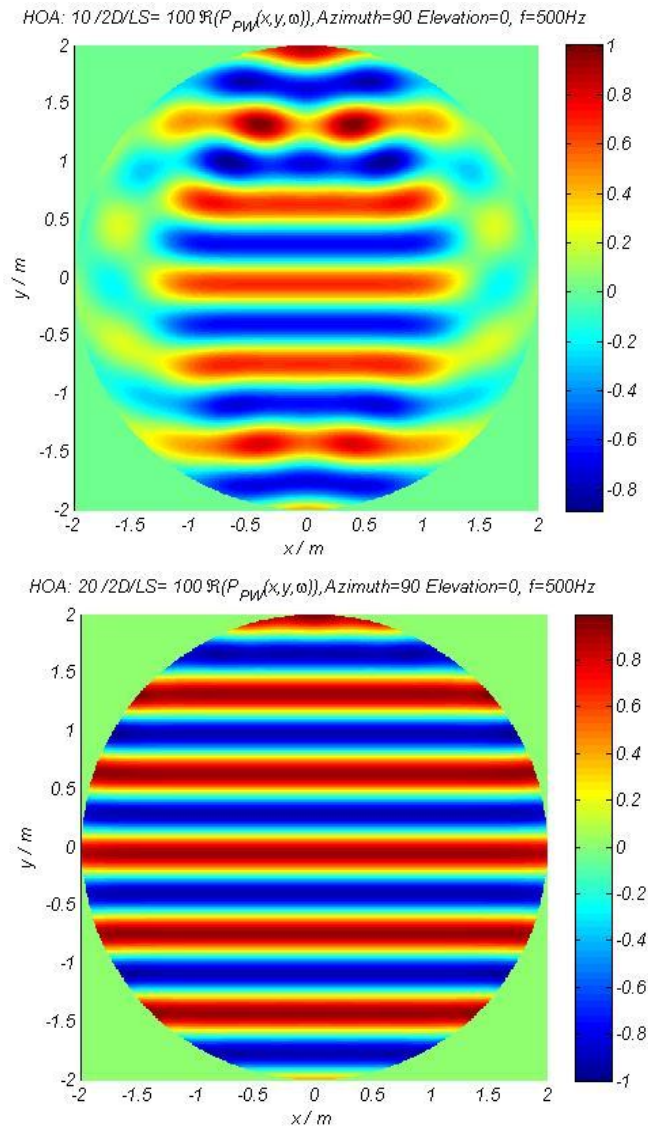

Figure 6.7: Virtual plane wave reproduction for N=10 and 20, respectively.

b) Near Field Effect

Using the Encoding equation (2.43a) without considering the near field effect leads to reproduce spherical wave (Figure 6.8(a)) instead of a plane wave (Figure 6.9(b)). The differences between spherical and plane wave fronts propagation yield that the finite distance of the loudspeakers results in a bass boost effect [Dan03]. It has already been suggested by Gerzon to compensate for this *bass boost* effect at decoding time [Ger92] by adjusting the decoding matrix with a set of filters [Dan03]. Introducing the secondary sources as point sources instead of plane wave sources in APA approach will cause artifact by reproducing wave field of plane wave
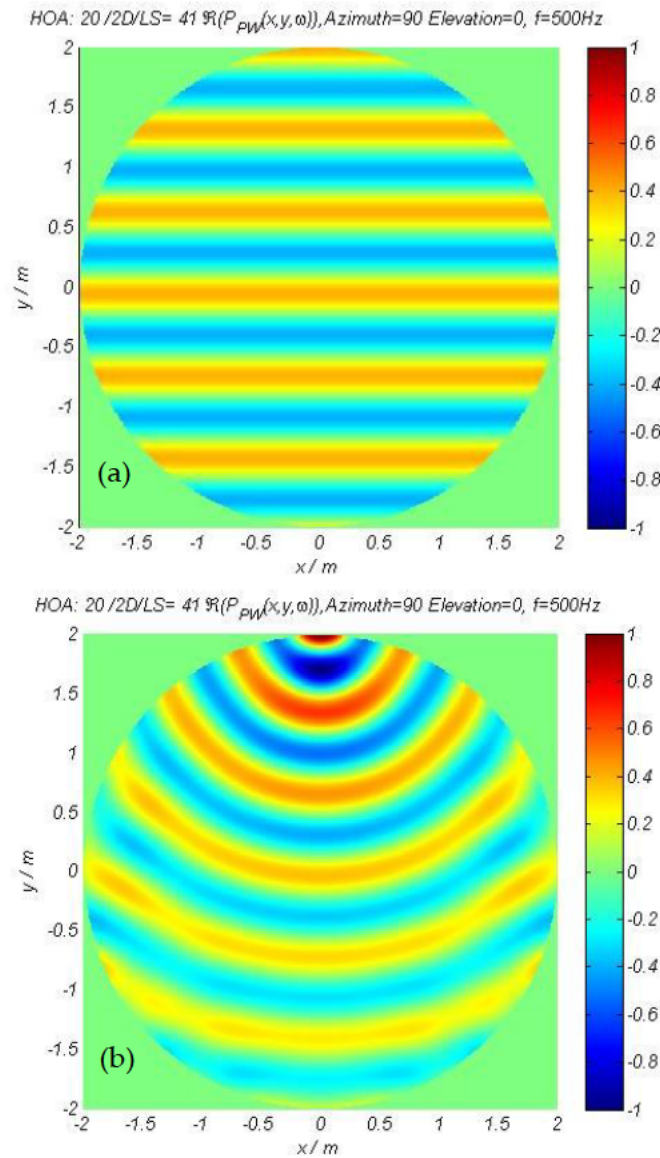

Figure 6.8: (a) reproduction of a plane wave with plane wave sources (N=20, L=41), (b) reproduction of a plane wave with point sources (near field effect).

6.2.3 Higher Order Ambisonics with Near Field Compensation (NFC-HOA 2.5)

a) Plane Wave as Virtual Source

Figure 6.9 illustrates the reproduced wave field for a monochromatic virtual plane wave (f= 500 Hz) with circular NFC-HOA reproduction system of the order N=20 with a radius of R = 2 m consisting of L = 50 secondary monopole sources. The incidence angle of the plane wave is $\theta_{pw} = 90°$.
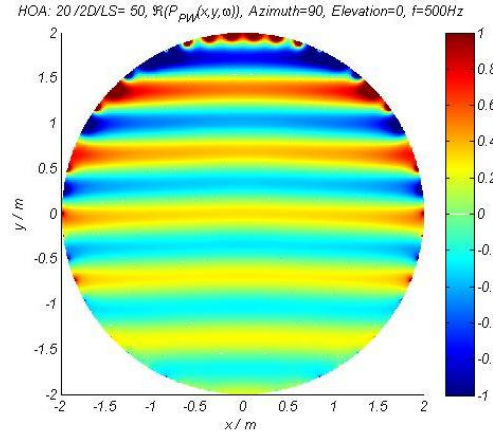

Figure 6.9: Illustration of implemented NFC-HOA for reproduction monochromatic virtual plane wave.

b) Spherical wave as Virtual source

Figure 6.10 illustrates the reproduced wave field for a monochromatic virtual spherical wave (f= 500 Hz) with circular NFC-HOA reproduction system of the order (N=20) with a radius of R = 2 m consisting of L = 50 secondary monopole sources. The center position of the virtual source is $x_C = [0\ 3\ 0]^T$.
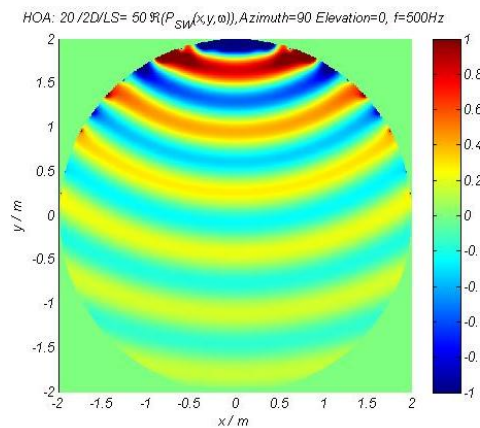

Figure 6.10: Illustration of implemented NFC-HOA for reproduction of virtual monochromatic spherical wave.

6.2.4 Wave Field Synthesis (WFS 2.5)

a) Plane Wave as Virtual Source

In Section 3.3.2 the driving function for a virtual plane wave was formulated in Equation (2.16).

$$D_{pw,2.5D}(x_0, \omega) = -2a_{pw}(x_0)\sqrt{2\pi|x_{ref} - x_0|}\, n_{pw}^T\, n(x_0)\, \sqrt{j\frac{\omega}{c}}\, \hat{S}_{pw}(\omega)\, e^{-j\frac{\omega}{c}n_{pw}^T x_0}\,, \quad (6.1)$$

where $\hat{S}_{pw}(\omega)$ denotes the temporal spectrum of the plane wave with incidence direction n_{pw}^T . For circular planer array the normal vectors $n(x_0)$ will point at the center of the array, which will be chosen as the reference position x_{ref}. Note that the reproduced wave field will exhibit amplitude decay within the listening area resulting amplitude decay of approximately 3 dB per doubling of distance. This is a result of using three-dimensional Green's function as an approximation for the two-dimensional one. Figure 6.11 illustrates the reproduced wave field for a monochromatic virtual plane wave (f= 500 Hz) with circular WFS reproduction system of the order with a radius of R = 2 m consisting of L = 50 secondary monopole sources. The incidence angle of the plane wave is $\theta_{pw} = 90°$.
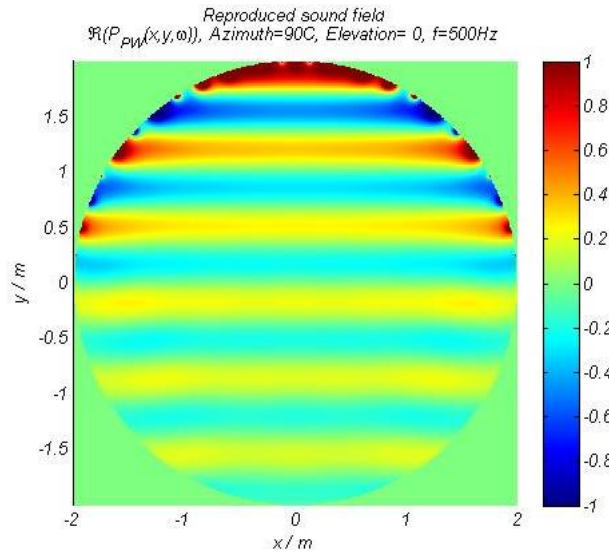

Figure 6.11: Illustration of implemented WFS 2.5 for reproduction monochromatic virtual plane wave.

b) Spherical wave as Virtual source

In Section 3.3.2 the driving function for a virtual plane wave was formulated in Equation (2.20).

$$D_{sw,2.5D}(x_0, \omega) = -2a_{sw}(x_0) \frac{(x_0 - x_S)^T n(x_0)}{|x_0 - x_S|} \sqrt{2\pi|x_{ref} - x_0|} \times$$

$$\times \left(\frac{1}{\sqrt{j\frac{\omega}{c}}|x_0 - x_S|} + \sqrt{j\frac{\omega}{c}} \right) \hat{S}_{sw}(\omega) \frac{e^{-j\frac{\omega}{c}|x_0 - x_S|}}{|x_0 - x_S|} , \tag{6.2}$$

where $\hat{S}_{sw}(\omega)$ the temporal spectrum of a spherical wave with center position x_S .For circular planar array the normal vectors $n(x_0)$ will point at the center of the array, which will be chosen as the reference position x_{ref}. Figure 6.12 illustrates the reproduced wave field for a monochromatic, virtual spherical wave (f= 500 Hz) with a circular WFS reproduction system with a radius of R = 2 m consisting of L = 50 secondary monopole sources. The center position of the virtual source is $x_C = [0\ 4\ 0]^T$.
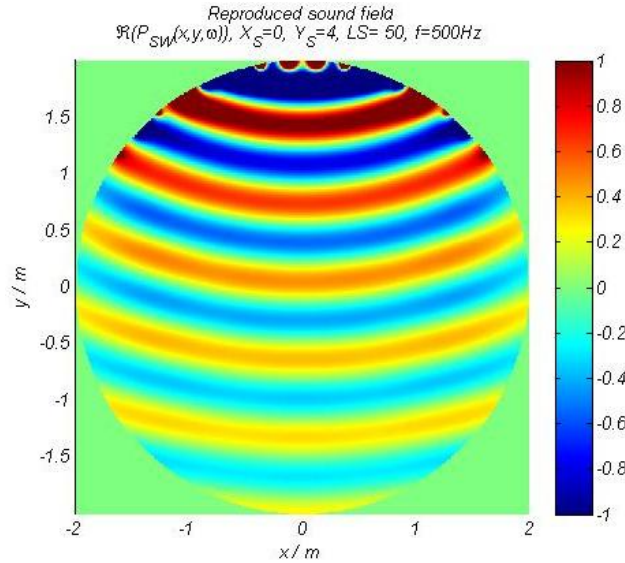

Figure 6.12: Illustration of implemented WFS 2.5 for reproduction monochromatic virtual spherical wave.

Matlab Implementation:

The calculation of driving functions will be implemented in the main function (`get_driving_function`), which is defined as follows

```
driving_function = get_driving_function (Method, settings, virtual_source,
loudspeakers_positions)
```

Input:
`Method`: name of the reproduction technique and takes the string values

```
    - Vector base Amplitude Panning : 'VBAP'
    - Ambisonics(APA, NFC-HOA)       : 'HOA'
    - Wave Field synthesis           : 'WFS'
```

`Settings`: a vector defines the settings parameters for the given `Method`

```
    - VBAP: no additional settings (future development)
    - HOA : [HOA_Order, type, encoding_config, vitual_source_type]
        - HOA_Order: the order of the HOA-system
        - type: HOA-Type : defines Ambisonics approach
                -APA      : 1
                -NFC-HOA : 2
        - encoding_config: defines the normalization used by Spherical
        Harmonics calculation accept
                - Semi- normalization (SN3D, SN2D) : 1 (implemented)
                - Fully normalization (N3D, N2D)    : 2 (future work)
                - Furse Malham normalization (FuMH): 3 (future work))
                - Maximum normalization (MaxN)      : 4 (future work)

        - virtual_source_type: defines
                - Plane wave        : 1
                - Spherical wave    : 2
    -WFS: [vitual_source_type, array_design]
        - virtual_source_type: defines the virtual source
                -Plane wave        : 1
                -Spherical wave    : 2
        - array_design: indicates array layout
                - Circle            : 1
                - rectangular       : 2
                - Klangakdemie Berlin: 3
```

`virtual_source`:defines the Spherical Coordinates of the virtual source

`loudspeakers_positions`: Array of the Spherical Coordinates of the loudspeakers and calculated by the function (get_loudspeakers_Positions)

Output:

`driving_function`: Matrix (L×FFT_SIZE) includes the driving function of L loudspeakers as impulse response of the length FFT_SIZE=2^14, which defines the frequency resolution of Fs/FFT_SIZE, where Fs is the sampling rate equal to 41100 Hz.

Dependencies

Tools:

```
-   Calculation of spherical Harmonics
        Ynm_APA.m:  calculates real Spherical Harmonics [Dan03] with semi-
        normalization
        get_Ymn.m: calculates complex Spherical Harmonics [Ahr08]
-   Mathematical functions
        sph_hankel_2nd.m: calculate the spherical n-th order hankel
        function of second kind (Eq. 2.37)
```

Important note:

Considering the virtual source (image source) as plane (Demonstration) the driving functions will include delay simulating the arrival time of the plane wave front of the source at loudspeaker array boundary as illustrated below.
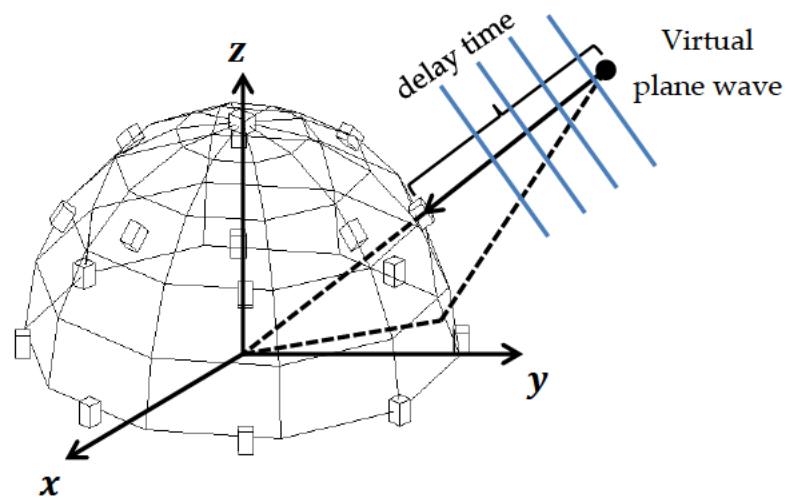

Figure 6.13: Schematic representation plane wave arrival time at boundary hemisphere shaped loudspeaker layout.

6.3 Auralization

Auralization of virtual environment was introduced in Chapter 4. The modeled room acoustics used has been performed for Audimax auditorium at the Technical University Berlin. The binaural room impulse responses (BRIRs) for 5 sources (primary sources) located on the stage were calculated with the Program EASE. These BRIRs are given as reflections described for each channel (ear) by time of arrival, sound pressure level for each octave (or third octave) band and angle of incidence (see Section 5.4). These reflections will be treated as image sources[9] representing the respective virtual source. For every reflection (image source) the equivalent driving function for a given loudspeaker array (secondary sources) must be calculated (Figure 6.14).
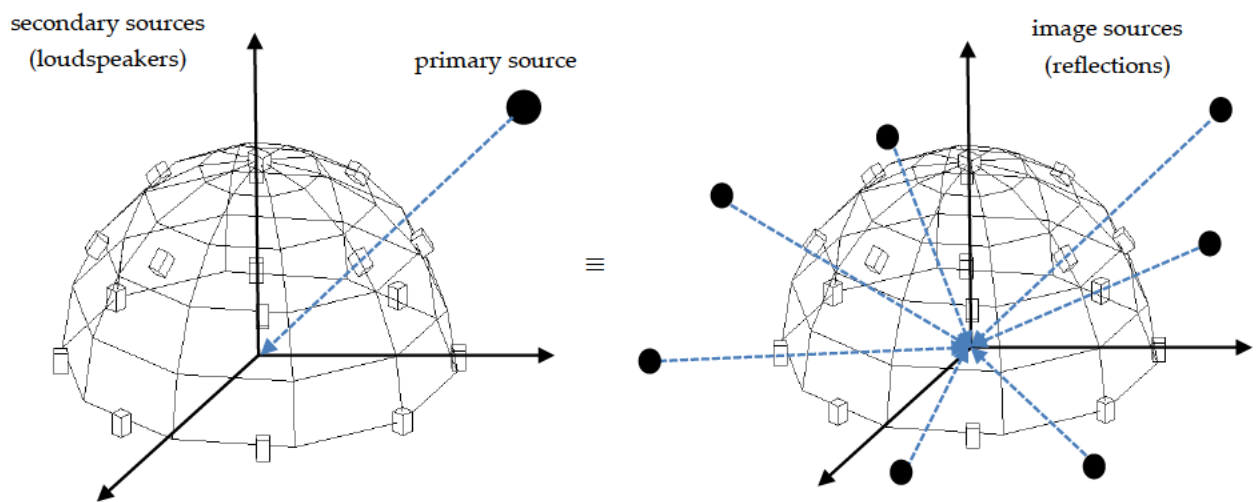

Figure 6.14: Representation of a primary source by its simulated reflections

The dynamic BRIR sub-dataset for a reflection in the horizontal plane will be achieved involving the respective HRTFs/HRIRs dataset for given head orientation and listening position. The final dynamic BRIR dataset for a primary source results by accumulation of all its reflections BRIR sub-datasets. Figure 6.15 illustrates the generated BRIR sub-dataset for of image source for different listening positions. To achieve the auralization experience at different positions within the array the delays and decays caused by the different paths between loudspeakers and listener must be considered (6.16). This is applied as delay and weighting (damping/amplification) on the loudspeaker signal. The decay will follow the $1/r$ law of the monopole, where r denotes the distance between the loudspeaker and listener where the listener is supposed to be far enough (far field).The secondary sources will be simplified as monopoles (omni-directional) for all reproduction methods.

---

[9] The reflection (image source) is interpreted as virtual source for the reproduction technique used
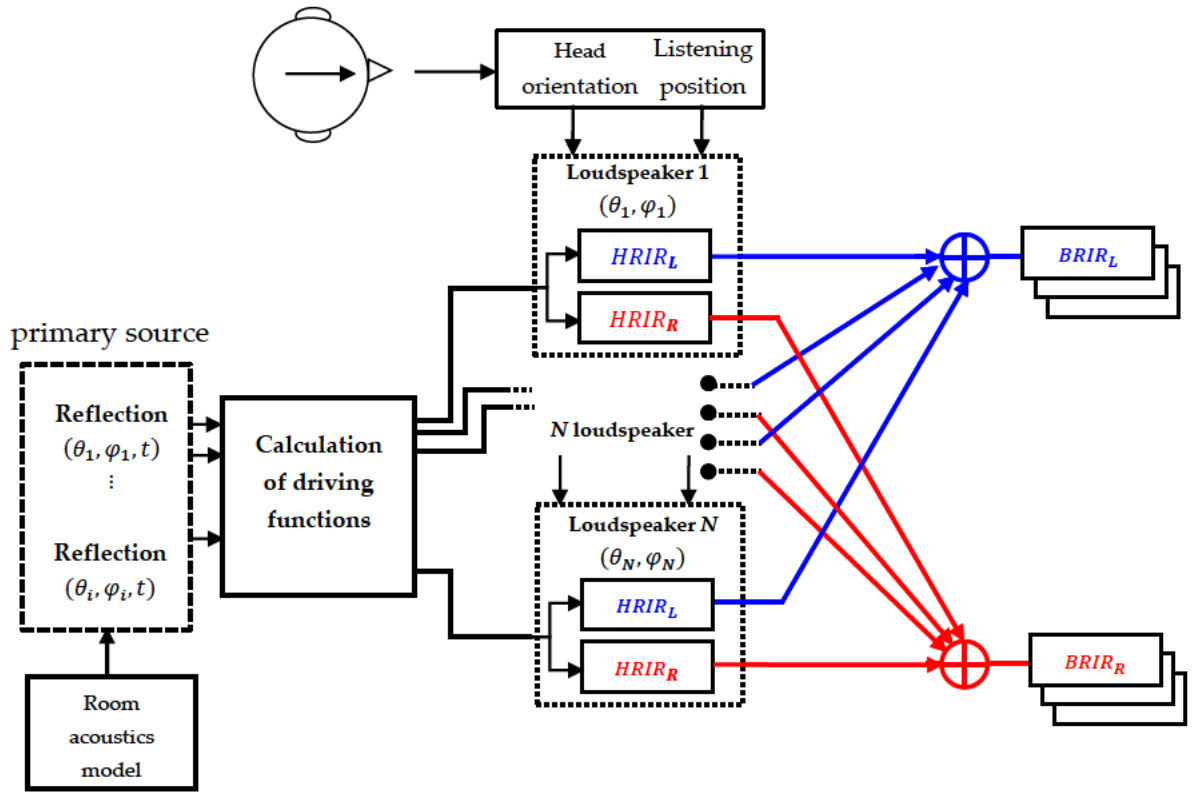
Figure 6.15: Schematic representation of generation dynamic BRIR dataset for a primary source in virtual acoustic environment.
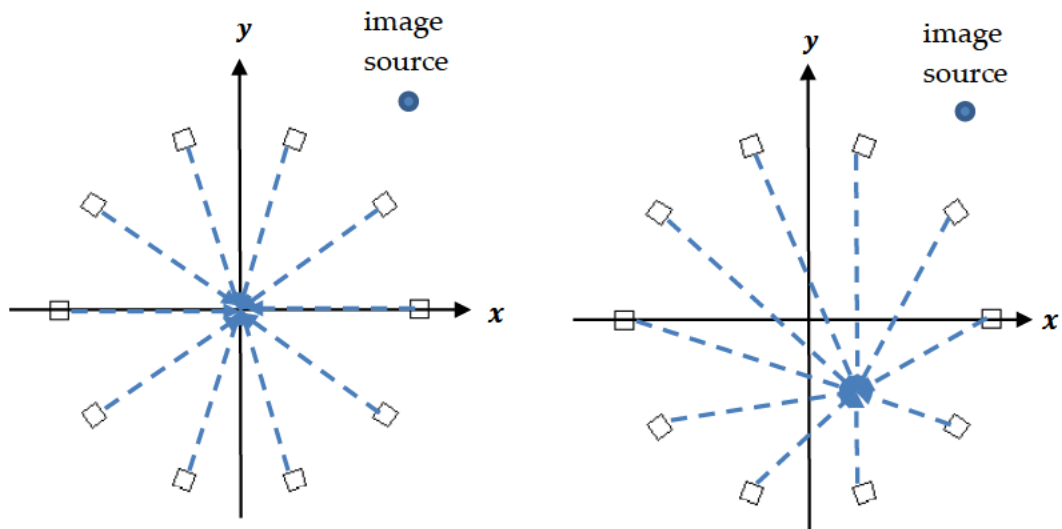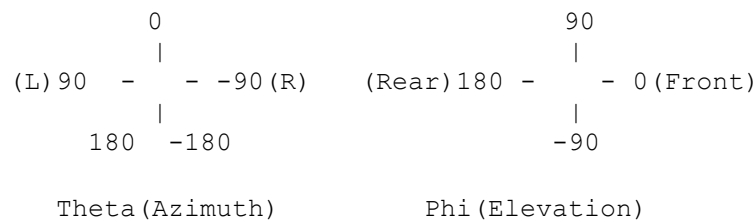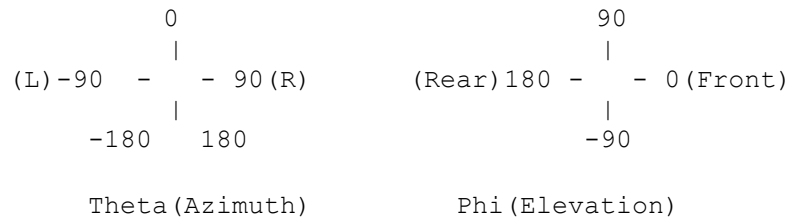

Figure 6.16: Schematic illustration of the paths between loudspeakers and listening position.

The reflections signals (spectra) delivered by EASE program are given in 21 third octave bands (Table 6.1) and they will be interpolated over the audio frequency range (20 Hz - 20 kHz) with frequency resolution of $(44100/2^{12})$ and normalized (0dB) at the frequency (1kHz) as minimum-phase FIR-Filter. The reflections signals in EASE are characterized by:

- Arrival time at the listener position in EASE model. This position isset to the center of the simulated loudspeaker array.
- Weight factor, representing the reflection pressure level normalized to the direct sound, which takes the weight factor of 1.
- Angles of incidence (Azimuth and Elevation): giving the incidence direction of the reflection (plane wave) in EASE Spherical coordinate system, which is defined as follows

```
            0                                 90
            |                                  |
    (L)90  -   - -90(R)      (Rear)180 -     - 0(Front)
            |                                  |
          180  -180                          -90

        Theta(Azimuth)              Phi(Elevation)
```

The BRIR dataset for a primary source consist of separate .wav files. The number of files is equal to horizontal resolution of head rotation between. The maximum resolution in horizontal plane is equal to 1°, which results 161 files for a 160° range in the horizontal plane. The Coordinate system convention (so-called Fabian or fWonder) is referred to and defined as

```
            0                                 90
            |                                  |
    (L)-90  -   - 90(R)      (Rear)180 -     - 0(Front)
            |                                  |
          -180   180                          -90

        Theta(Azimuth)              Phi(Elevation)
```

The wav.file name takes the format (fWonder format): sign(**)sign(***).wav, where Sign defines the direction (where to look):
- (**): 2 digits for Elevation (-40°, 90°) , we set to 00 (just horizontal)
- (***): 3 digites for Azimuth (-80°,80°)
- For azimuth (***) : P(Positive); turn to the right, N(Negative);turn to the left.
- For elevation (**) : P(look to /), N(look down\)

Example: P00P000.wav (to the front), P00N800.wav (80° to the left), P00P800.wav (80° to the right)

Important Note:

By generation of dynamic binaural data set of aprimary the convertion between different Coordinates systems (Matlab, EASE and HRTFs conventions) must be applyed.

The dynamic auralization implementation can be split into two steps. In the first step the dynamic binaural dataset for every primary source must be separately generated (Figure 6.15). The process can be described as a pseude-code as follows

```
% Create dynamic BRIR set (in fwonder format) from modular 'AK sound field'
% description over virtual array
%
%     - Enter source & destination paths (settings)
%     - Select ranges and step sizes for head movements (settings)
%     - Start script (processing) ...
%----------------------------------------------------------------------
% Settings:
%
% - define folders for data (sound field data, metadata, destination for
results)
% - define sources
% - define head movements
& - define listening Cartesian position (in Matlab convention)
% - Define simulated sound field reproduction (method, array)
% - Define folders for: metadata, decomposed sound field, and results
% - determine arrival time of last reflection (for BRIR length)
% - find minimum direct sound arrival time (for reflections validation)
% - save settings in .mat file
%
% Processing:
% - calculate loudspeaker positions in (in Matlab convention)
% - calculate distance between loudspeakers and listener
& - calculate the loudspeakers weighting factors to simulate listening
%   position variation
%
% for each head orientation & each source calculate complete brir:
%   find valid reflections (audible, outside the array)
%   for all reflections of a source...
%       1. read reflection
%       2. apply amplitude weight
%       3. calculate driving function for a given loudspeaker array
%          accumulate driving functions for all reflections
%
%   for all loudspeakers
%       5. calculate loudspeaker signals for all reflection summation
%       6. determine & get HRIR for direction of incidence corrected for
%          current head orientation and listening position
%          ATTENTION: (two incompatible coordinate systems:
%           -> convention adaption
%       7. copy hrir into a matrix with ITDmargin/2 at front and end
%       8. re-introduce ITD (10x oversampled ...)
%       9. convolve delay arrival time from loudspeaker to Listener
%          position
%      10. Weighting the loudspeaker signal according to listener position
%      11. convolve HRIR with loudspeakersignal /loudspeaker/
%      12. save binaural reflection 'refl' into in BRIR-accumulator
%             save BRIR for each head orientation & each source calculate
% ... and repeat till end
%
% log some things to text file
```

The second step towards dynamic Auralization of a stimulus (anechoic signal) is to render the dynamic binaural dataset with a real- time, head tracked convolution machine. The real-time rendering application, running on Linux OS, is implemented as a JACK[10] Audio server client. It is able to render multiple sources at a time depending on computational power and RAM size. It uses fast non-uniform partitioned block convolution and a double static/dynamic caching algorithm to account for latency and memory limitations [Lin07]. All primary sources (5) are rendered continuously. For instantaneous switching between multiple sets of BRIRs OSC (Open Sound Control) commands are sent from the graphical user interface (Matlab, pure data[11] GUIs) of the listening test software switching audio outputs on and off (Figure 6.17). This can be done one system or distributed system (distributed processing). Since the computing power demand is high, it is not advisable to render many primary sources (BRIR datasets), which leads to limitation by demonstration part done in this thesis.
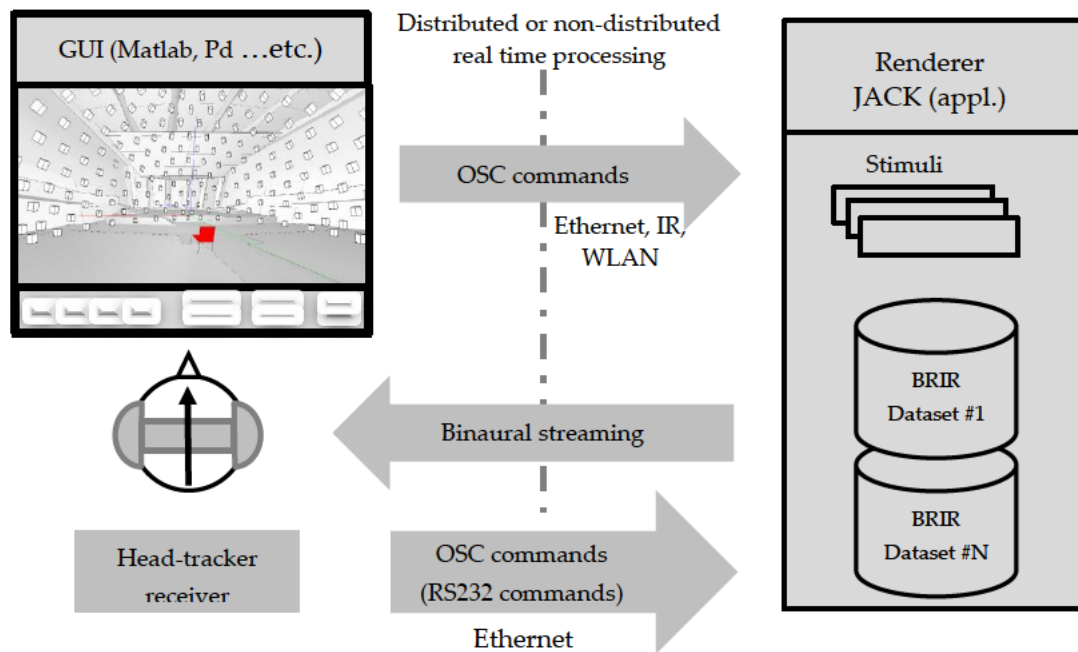

Figure 6.17: flow diagram of control and audio data in the rendering process

For introducing reverberation to BRIR datasets during rendering process the modeled reverberation tail approach will be used for all methods. The modeled tail will generated based on the reverberation tail in; However, For WFS the reverberation with plane wave synthesis (Section 5.5.2) will be available. This will involve 8 uncorrelated plane wave primary sources. Each of the plane waves will represented as mixing of the original 5 sources $(S_1 \ldots S_5)$.

---

[10] JACK: a professional sound server daemon that provides real-time, low latency connections for both audio and MIDI data between applications. http://jackaudio.org

[11] Pure Data (or Pd) is a visual programming language for the creation of interactive computer music and multimedia works. http://puredata.info/

6.3.1 Room Acoustics Model

The room modeled was the auditorium maximum (Audimax) at the technical university of Berlin. The impulse response for five primary sources at one listening position was simulated. In these impulse responses we mainly consider the early reflections within the first 0.3 sec starting from the direct sound arrival time (direct sound arrives at t=0). The reverberation part is meant to be stochastically generated and adapted to the simulated reverberation frequency response. In order to improve the quality of the simulation results the acoustical properties of the room model were adapted to have similar behavior of real measurement done in auditorium. The table 6.1 and Figure 6.18 show the parameter and settings used for the simulation.

Table 6.1: Simulation settings and parameter

| Parameter	Value/type						
Volume	$69614m^3$						
Surface	$46853 \ m^2$						
Air Temperature	20°C						
Humidity	60 %						
Simulation method	Ray tracing Eyring						
	100	125	160	200	250	315	400
Frequency bands [Hz]	500	630	800	1000	1250	1600	2000
	2500	3150	4000	5000	6300	8000	10000
Source	Source position [m]			Number of reflections within first 0.3 sec			
	X	Y	Z				
S_1	1.45	-3.22	1.1	3287			
S_2	3.45	-3.72	1.1	3306			
S_3	2.2	-4.22	1.1	3318			
S_4	0.7	-4.22	1.1	3384			
S_5	-0.55	-3.72	1.1	3266			
Listener position	1.45	4.06	0.48				
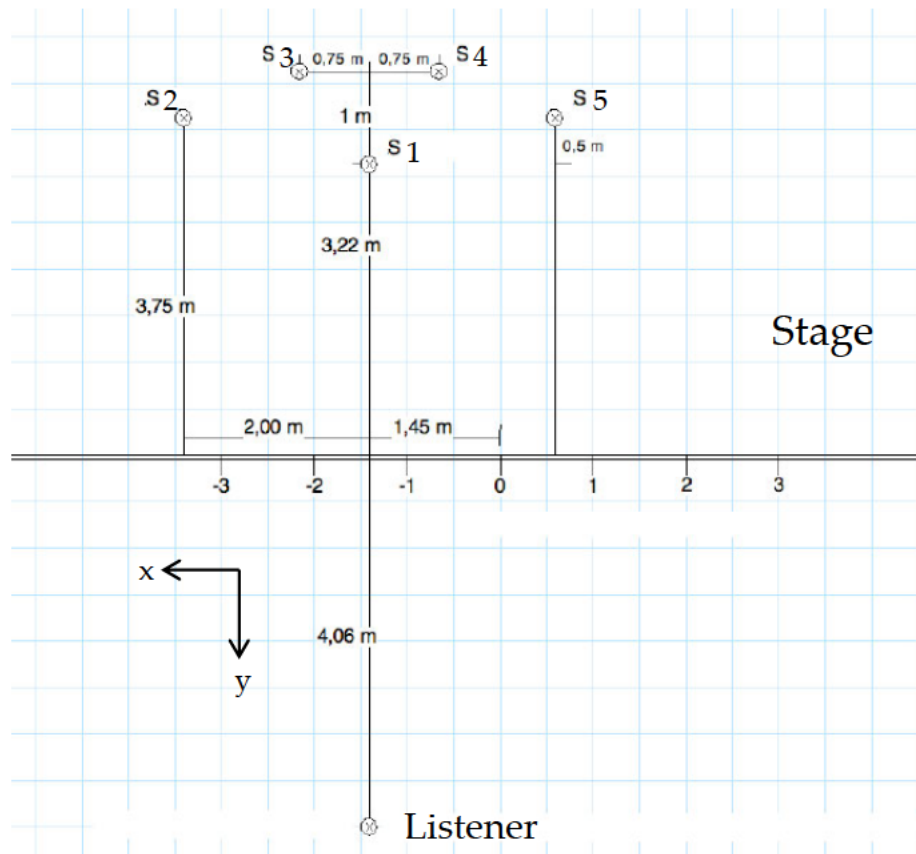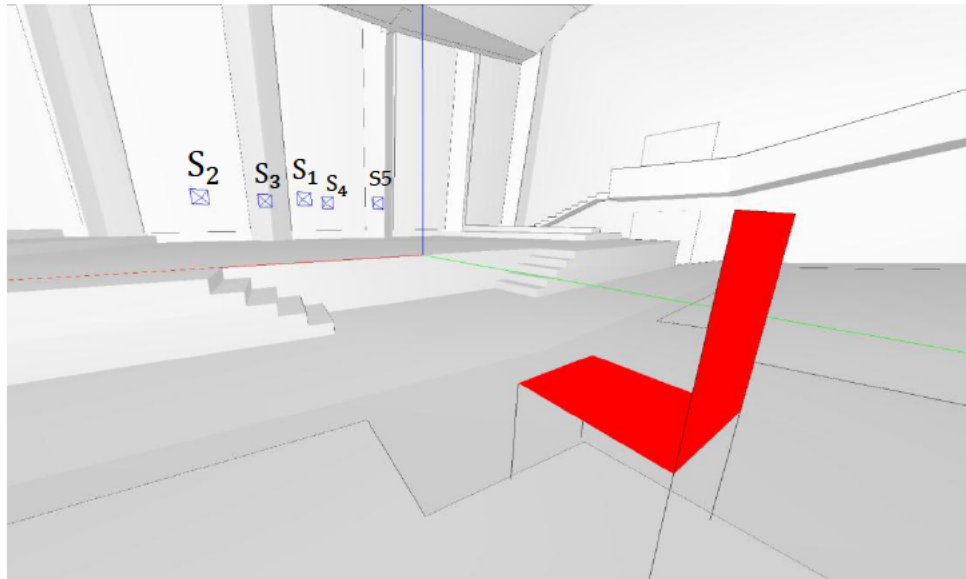
Figure 6.18: Graphical illustration of EASE program; Top: CAD-Model. Bottom: primary sources and listener positions in x-y plane.

6.3.2 Auditory Model

One part of the project was to develop an auditory model to check the Audibility of the pulse-shaped Reflections (clicks). This part is done by a colleague Max Röhrbein [Röh10]. This Auditory model takes into account the psychoacoustic properties of human hearing like masking and human hearing sensitivity, which are meant to be simulated. Applying the Auditory model will reduce the number of reflection to be rendered (Figure 6.20). This could extremely reduce computational power needed to render. The Model was based on the study [oli89] and delivers a similar behavior, which defines the audible reflection threshold relative to the direct sound level as function of reflection delay time (Figure 6.19). The parameter (*at*) defines the step width and range of auditory model sensitivity. The bigger the parameter (*at*) is, the less reflections will be considered.
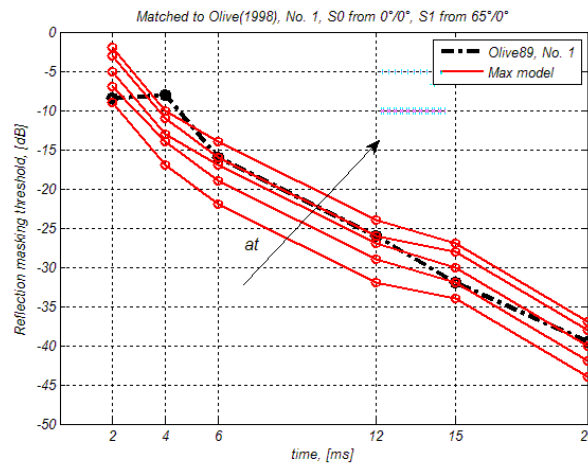

Figure 6.19: The auditory model of Max Röhrbein matched to Olive's. The model of Olive was experimented using a source representing the direct sound at (0°, 0°) and a source representing the reflection with the incidence direction (Azimuth=+65° (to the right), Elevation=0°) in an anechoic chamber (reverberation time = 0 sec).
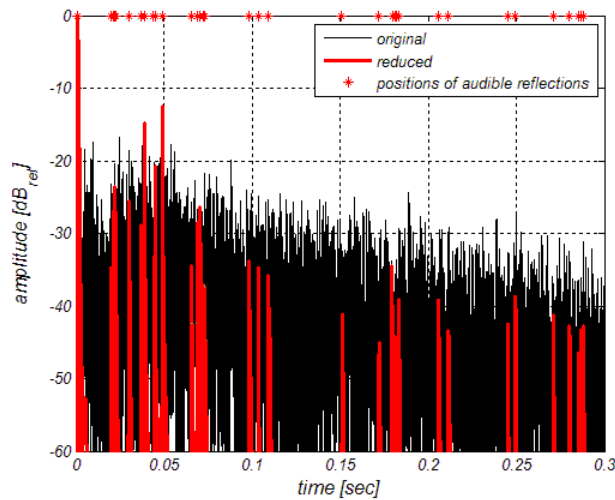

Figure 6.20: Illustration of the audible reflections (red) detected by the auditory model for the source (S1).

6.4 HRIRs Pre- and Post-Processing

For this project a special dataset was used, the so-called Moldrzyk HRIRs [Mol04]. These were measured in the anechoic chamber of the Institute of Technical Acoustics at the RWTH Aachen. The horizontal range (azimuth) of the data varies from 0° to 359.5° (anticlockwise) with a 0.5° step resolution. Measurements were also conducted for elevations between -60° to +90° with a 5° step resolution.

For Auralization purpose the HRIRs have been processed. This was the work of the colleague Antonio Cardenas. In his work the HRIRs were first preprocessed including the compensation of the possible measurement errors like compensation of the loudspeaker directivity, unwanted reflections and measurement-microphones. Afterwards the HRIRs were meant to be presented as a minimum-phase filter and pure delay [Nic10]. This enables the separation of ITD and ILD cues, which is useful to compensate for the dummy head's wrong morphology or head size and can be also used to get rid of long pre-delay in the used HRIR-dataset. This will reduce dataset size and simplify later interpolation. Figure 6.21 illustrates an example of the post processed HRIRs.
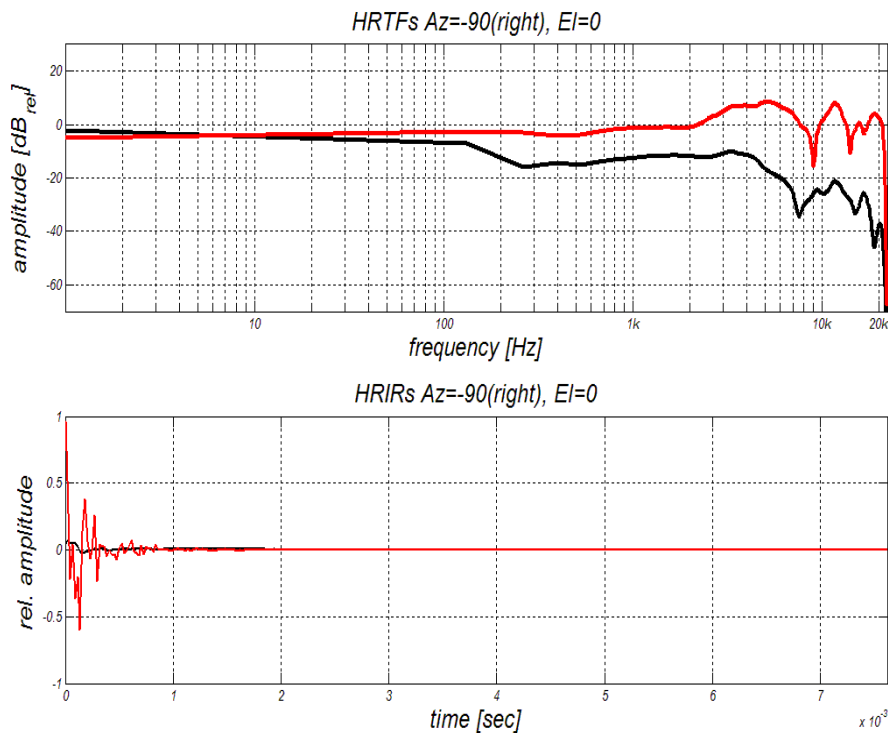

Figure 6.21: post processed HRIRs and equivalent HRTFs (azimuth=-90, elevation=0). The ITD is equal to 625.85 μ sec.

6.5 Demonstrator

The implementation carried out in this thesis was the first step of developing a simulation environment for array-based sound field reproduction systems. Thus the results are presented by means of a working demonstrator. The simulated room acoustics (Audimax model) was dynamically auralized based on a set of 4 reproduction techniques that were implemented. This considered two listening positions and two methods of modeling of the reverberation (Section 5.5). The subject can interactively experience the dynamic auralization of different stimulus in virtual environment. This will be done using a graphical user interface (GUI). Perspective views (snapshots) of the virtual environment/loudspeaker-array graphics can be displayed to give the subject a visual feeling about the environment meant to be simulated (Figure 6.22).

Demonstration setup:

Table 6.2: Loudspeaker layouts:

Reproduction (rendering) method	Loudspeaker layout		
	Type	Mode	Number of loudspeakers
VBAP	Lebedev grid	hemisphere (R=6.5m)	309
APA	Lebedev grid	hemisphere (R=6.5m)	309
NFC-HOA	Lebedev grid	hemisphere (R=6.5m)	105
WFS 2.5	Klangakademie	Rectangular (5m×5m)	200

Table 6.3: Reproduction method settings:

Vector Base Amplitude Panning (VBAP)	
Parameter	**Value/comment**
Number of primary sources	5
Listening positions (within the array)	$[0\ 0\ 0\ ]^T$ and $[3\ -3\ 0\ ]^T$
Auditory model parameter at	[0.3]
Reverberation synthesis method	Modeled reverberation tail
Amplitude Panning Ambisonics (APA)	
Parameter	**Value/comment**
Expansion order n	16
Virtual source type (primary source)	Plane wave
Secondary source type (loudspeaker)	Plane wave
Encoding normalization type	SN3D
Decoding type	Pseudo invers (pinv)
Number of primary sources	5
Listening positions (within the array)	$[0\ 0\ 0\ ]^T$ and $[3\ -3\ 0\ ]^T$
Auditory model parameter at	[0.3]
Reverberation synthesis method	Modeled reverberation tail

Near Field Compensation Higher Oder Ambisonics (NFC-HOA)	
Parameter	**Value/comment**
Expansion order n	9
Virtual source type (primary source)	Plane wave
Secondary source type (loudspeaker)	Point source (3D Green function)
Number of primary sources	5
Listening positions (within the array)	$[0\ 0\ 0\ ]^T$ and $[3\ -3\ 0\ ]^T$
Auditory model parameter at	[0.3]
Reverberation synthesis method	Modeled reverberation tail
Wave Field Synthesis (WFS 2.5)	
Parameter	**Value/comment**
Virtual source type (primary source)	Plane wave
Secondary source type (loudspeaker)	Point source (3D Green function)
Reference position x_{ref} (Cartesian)	$[0\ 0\ 0\ ]^T$
Number of primary sources	5+8(reverberation)
Listening positions (within the array)	$[0\ 0\ 0\ ]^T$ and $[3\ -3\ 0\ ]^T$
Auditory model parameter at	[0.3]
Reverberation synthesis method	Modeled reverberation tail and plane wave synthesis

Table 6.4 shows the reduced number of image sources (early reflections) for possible used parameter at values.

Table 6.4: reduced number of audible image sources (reflections) for given values of the parameter at.

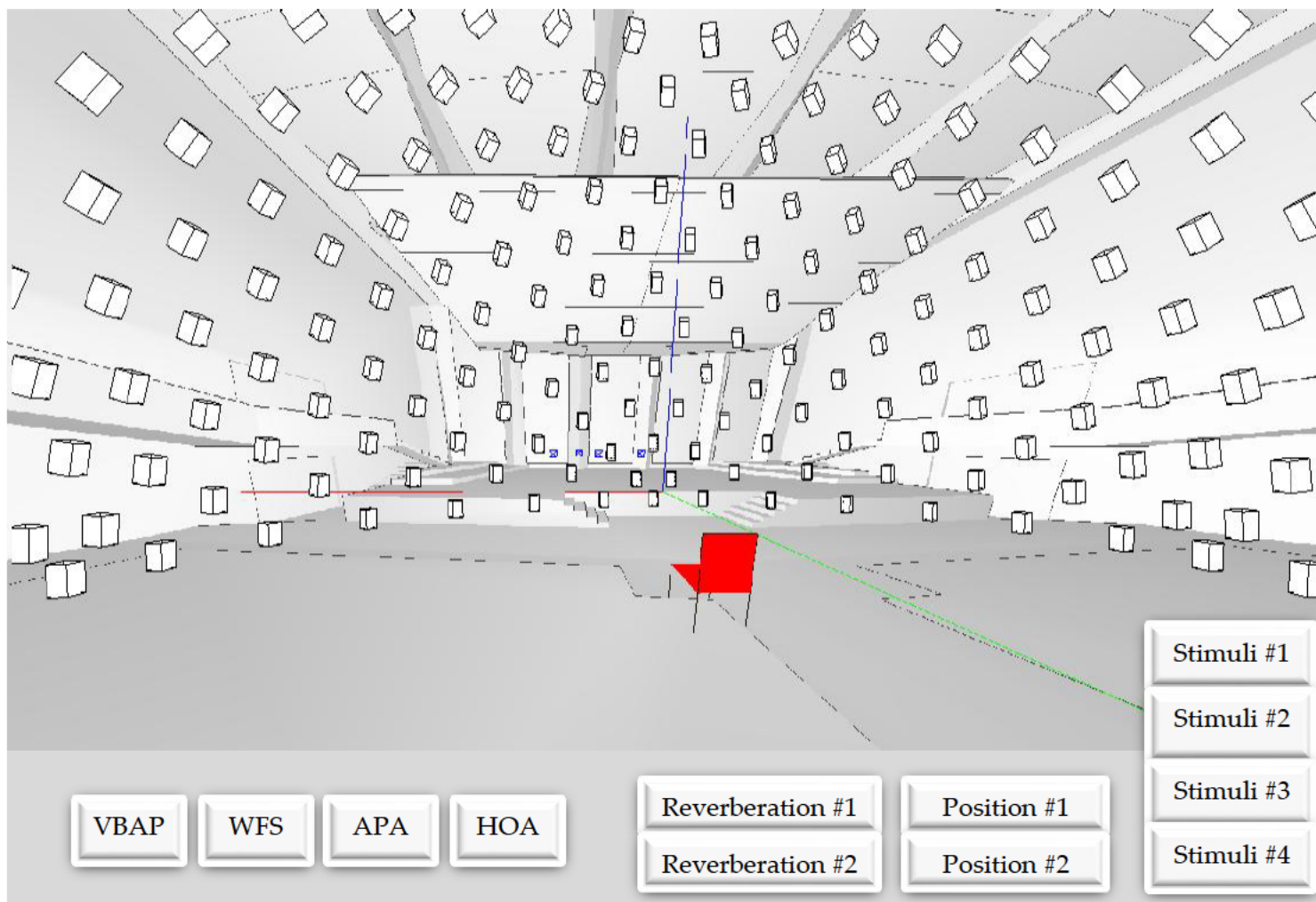| Source	Auditory model parameter	
	$at = 0.3$	$at = 0.6$
	Number of audible image sources	**Number of audible image sources**
S_1	36	2
S_2	28	2
S_3	52	5
S_4	32	4
S_5	46	1

Figure 6.22: Graphical User Interface of the demonstrator (3D array).The room displayed represents the simulated room acoustics (simulated Audimax in EASE). The loudspeaker array should reproduce the simulated room acoustics.

6.6 Evaluation

This Section deals with the final Evaluation of the end results acquired.

In this Section, a BRIR dataset as generated by different reproduction techniques will be evaluated compared with the simulated, reference binaural dataset generated by the EASE program. This will be achieved according to certain criteria. More precisely, the evaluation will be achieved by comparing the energy time curve (energy decay) and the frequency responses of the resulted BRIR datasets with respect to the simulated BRIR (in EASE), where the energy time curve of a BRIR gives an overview of the level and alignment of the reflections in time.

Figure 6.23 and 6.24 illustrate, respectively, the binaural frequency responses and energy time curves (ETCs) of resulted BRIRs for the primary source S_1 evaluated at the reference position [0 0 0]. Settings for the reproduction methods were defined in Tables 6.2 and 6.3. The evaluation will consider 2 listening positions and reference head orientation of (0°, 0°).

a) Evaluation of binaural frequency responses:

Since the source S_1 is located at the front of the listener (Figure 6.18), we expected a high similarity between the both right and left channels. This is noticed in the simulated BRIR in EASE (Figure 6.23).The results illustrated in Figure 6.23 can be summarized as follows:

- **VBAP**: Amplification at low frequencies (100-1000 Hz) is noticed.
- **WFS**: because of the high-pass term in the formulation of the driving function in Equation (3.16) we could notice amplification at high frequencies (3-20 kHz), which is more noticeable by the left channel.
- **APA**: the main characteristic of the frequency response was amplification at low frequency and a damping gap at middle frequencies. This can be explained as artifacts caused by possibly unsuitable loudspeaker layout.
- **NFC-HOA**: exhibited a strange characteristic where the left channel response was remarkably amplified. A similar simulation was carried out for equiangular array (49 loudspeakers, N=6) and delivered better frequency responses and energy decay. This highlights the issue of the geometrical loudspeaker array design.

b) Evaluation of energy time curve:

The Energy Time Curve (ETC) describes the energy decay of the BRIR signals (Figure 6.24). In an ETC the reflections will be demonstrated as peaks at their arrival times. Figure 6.24 illustrates the ETC of the energy decay curve of source S_1 evaluated at the center position of the loudspeaker array$[0\ 0\ 0]^T$. A similarity of the reflections events can be observed for the resulted BRIR datasets. A low-pass filter (Butterworth 8th order, $f_{cutoff} = 1700 Hz$) was applied (anti-aliasing filter) in **WFS** to be able to see peaks of the energy time curve. This can also be done in **NFC-HOA** case.

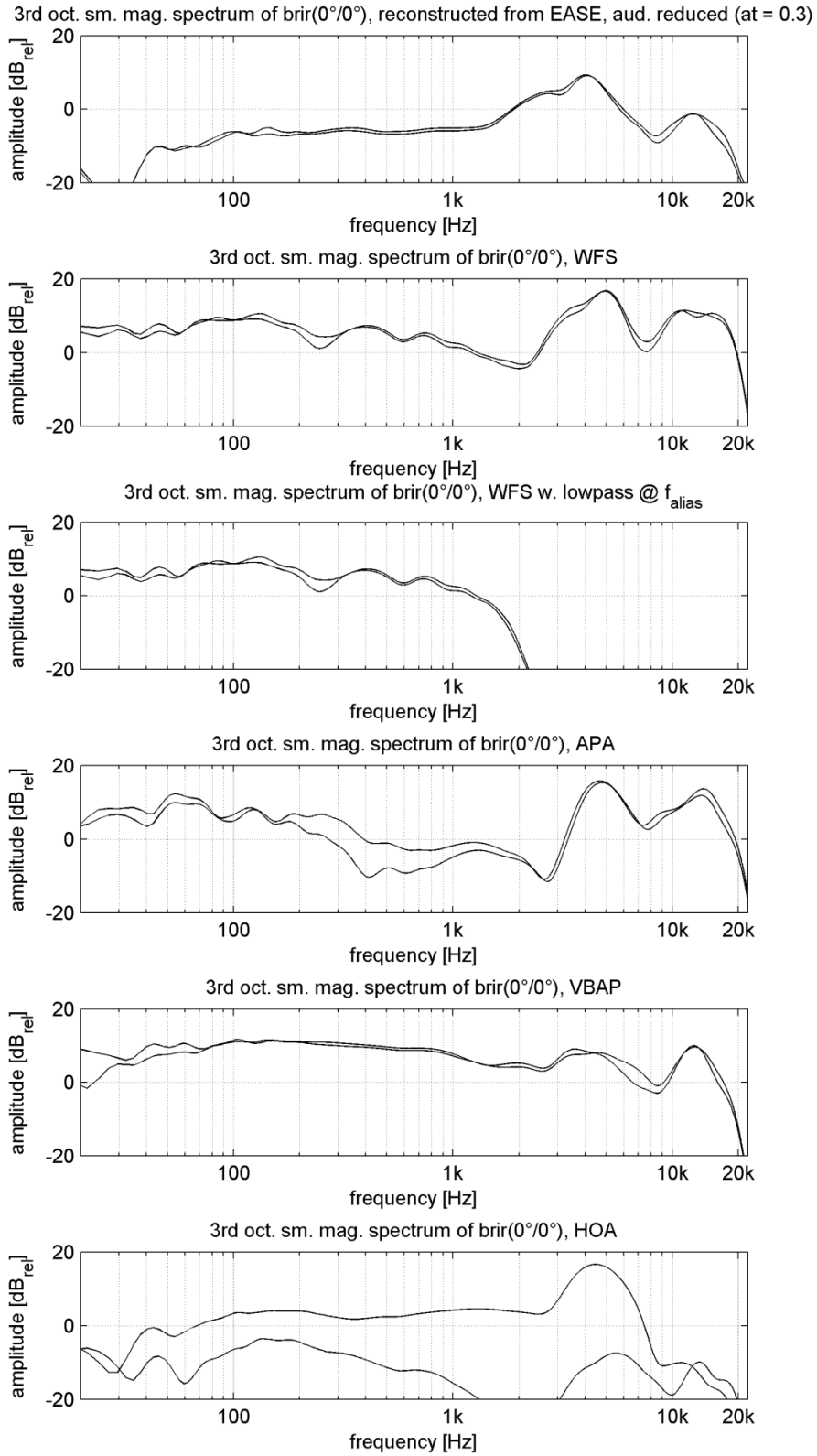**Figure 6.23**: Frequency responses of resulted BRIRs of source S_1 evaluated at the center position of the loudspeaker array $[0\ 0\ 0]^T$.
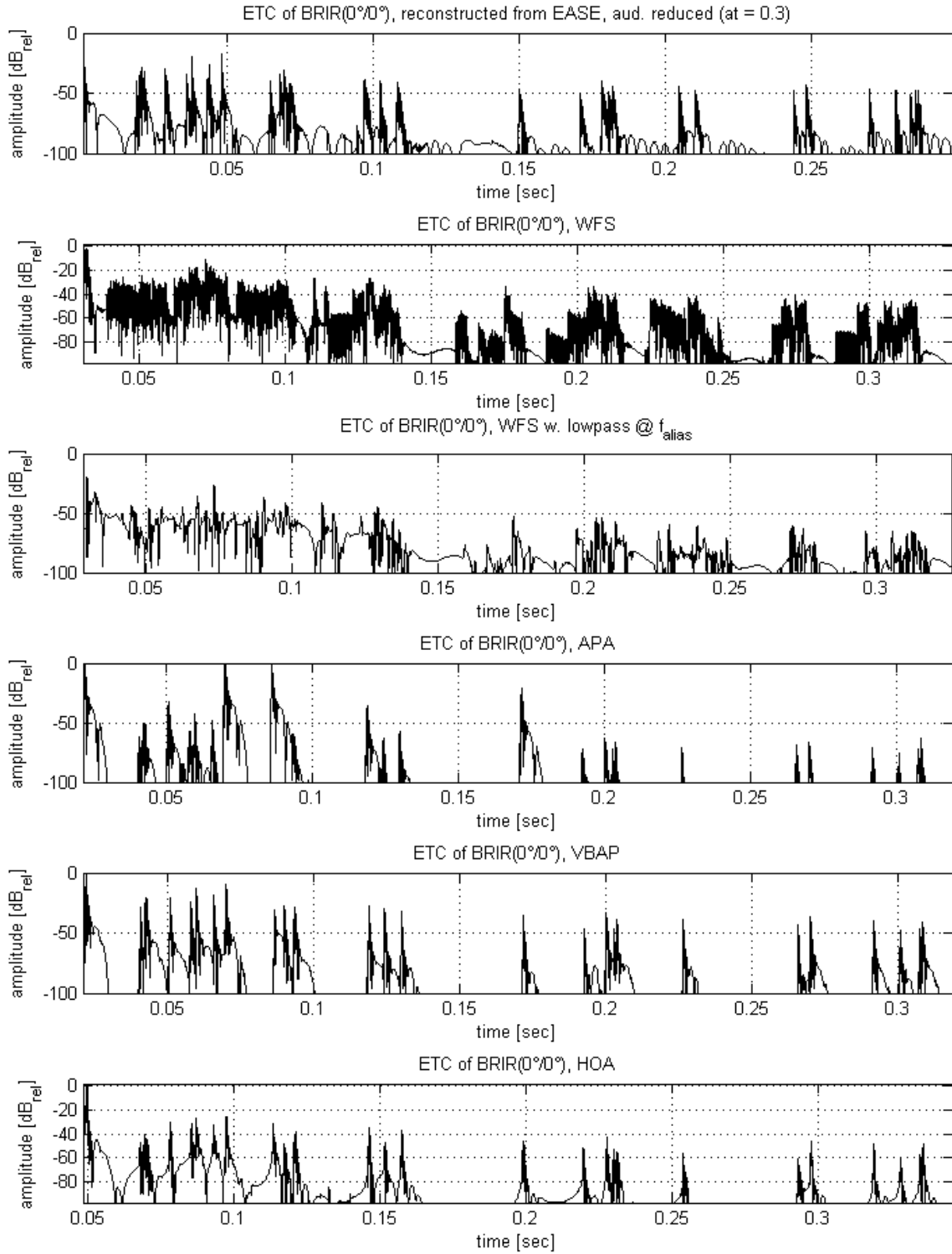
Figure 6.24: Energy decay curve of source S_1 evaluated at the center position of the loudspeaker array$[0\ 0\ 0]^T$.

Note: the different time scaling in Figure 6.24 is caused by mismatch of time alignment of the resulted BRIR datasets. This error has to be debugged.

The evaluation of resulting BRIRs frequency responses and ETCs was made at another listening position $[3 \ -3 \ 0]^T$. The frequency responses and ETC are illustrated in Figure 6.25 and 6.26, respectively.

In the frequency responses of resulting BRIRs it was noticed that the amplitude of the left channel was higher. In the ETCS the density of the peaks was spread in time that reflections become perceptively less distinguishable.
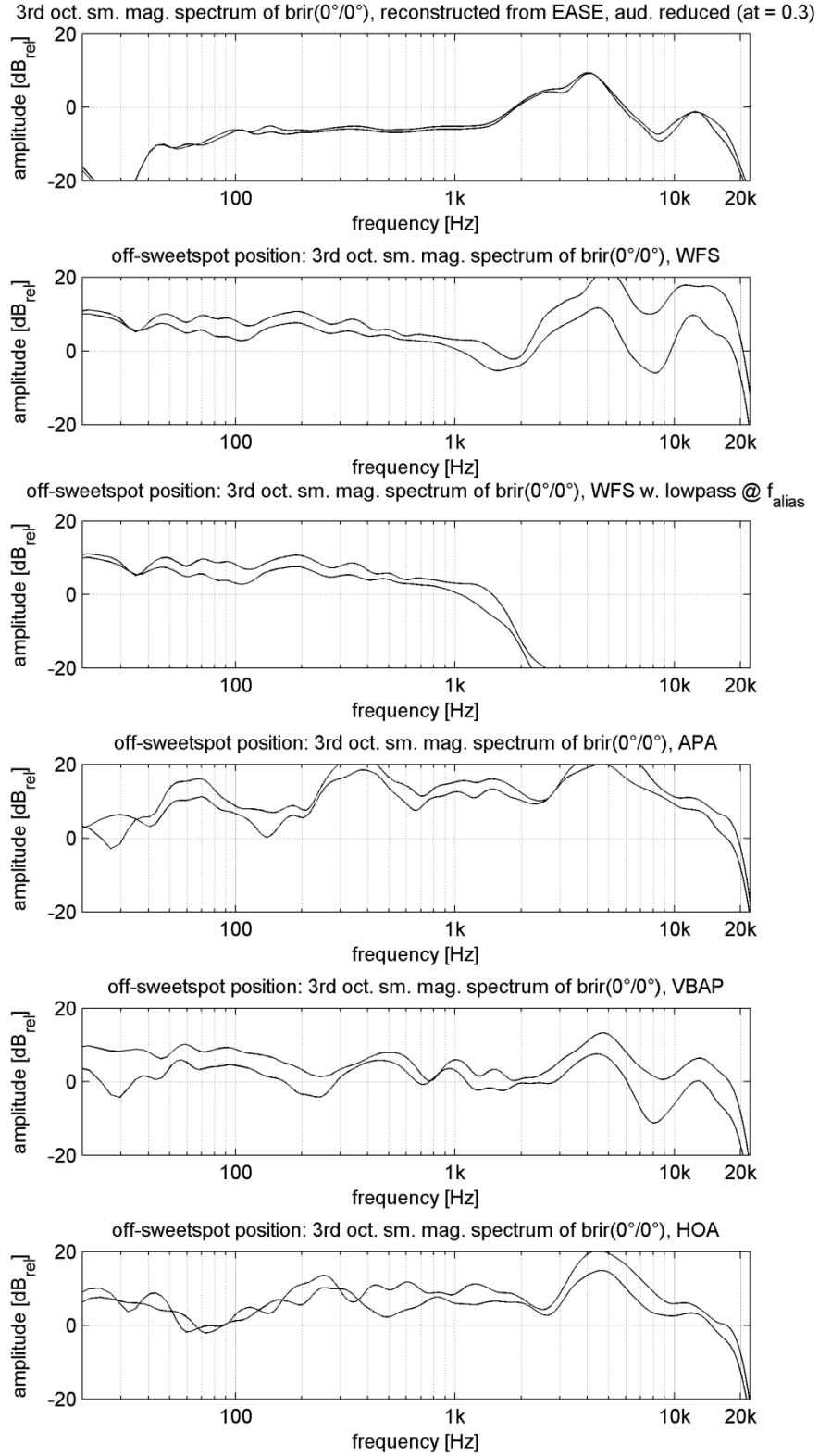
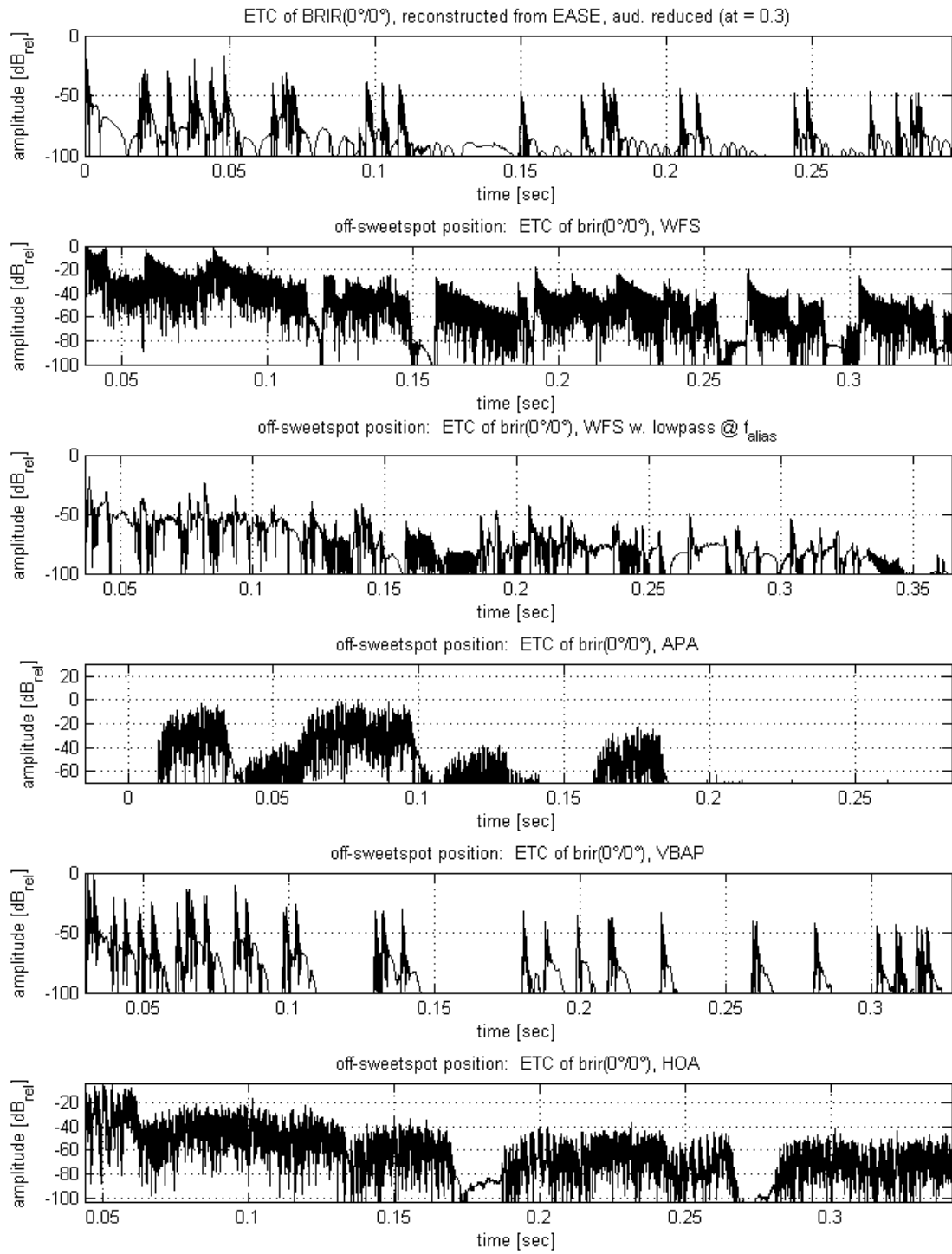**Figure 6.25**: Frequency responses of resulting BRIRs of source S_1 evaluated at position $[3 \ {-}3 \ 0]^T$.

Figure 6.26: Energy decay curves of source S_1 evaluated at the position$[3 − 3\ 0]^T$.

6.7 Discussion

In order to achieve the overall goal of this thesis, i.e. reproduction of room acoustics via array based field synthesis methods; we have implemented and evaluated the 3 reproduction methods WFS, HOA and VBAP. We have also considered APA, a special case of HOA.

The implementation was done using Matlab programming language and making use of the room acoustics simulation program EASE for modelling a given room acoustics. The input parameter and settings or the simulation by EASE program, e.g. (primary) sources, listener position and room acoustic/geometrical properties were listed in Table 6.1. The simulated sound field description (reflections) was reduced using the Auditory model (Section 6.3.2)

To design the sound system the loudspeaker array must be defined precisely. This was done using the implemented Matlab function "get_loudspeaker_Positions.m". The next step was to calculate the driving function for a given system. This was achieved by the function "get_driving_function.m". Consequently, the generation of a dynamic BRIR dataset for all the primary sources could be carried out as described and illustrated in Figure 6.15. The resulting BRIR datasets as illustrated in Figure 6.23 und 6.24 were compared to the simulated BRIR generated in EASE and the matching quality was evaluated according to a number of criteria, which were elaborated in Section 6.6.

The overall results, as expected, revealed a better performance of the 4 implemented examples at the center position (sweet point) of the used arrays. The reflections in the resulting BRIRs were reconstructed properly and yielded curves of a quality that was comparable to the reference simulated BRIRs.

WFS is broadband approach, i.e. no band-width limitation exist in the formulation of the diving function. However, this technique suffers mostly from the spatial aliasing caused by the limited number of loudspeakers used, which again limits the operating band width. This will affect the dimensions the sweet spot, which will be not to desribe.

The calculation of **NFC-HOA** driving function has requested the most calculation time in comparison to other methods. However it was expected to give a better result.

APA should offer a centralized sweet spot independent of the frequency band limitation caused by a given order of the system. Some modification in basic decoder could improve the performance of the system over a sufficient listening area. The assumption of plane wave secondary source can be easily made for larger musical venues.

VBAP was the simplest method to implement yet has delivered good results, especially for homogeneous arrays with a sufficient number of loudspeakers. However, the sweet spot by VBAP is not expected to be large enough to meet the demands of contemporary musical venues.

Thus, it seems that the APA method is the most appropriate reproduction technique for the system under investigation, since it unifies complexity and performance. Still, for perspective evaluation, listening tests are necessary (future work).

Chapter 7

Conclusions

In this thesis we have demonstrated the principle of room acoustics reproduction by array-based sound field synthesis methods. This case was simulated using an implemented software tool and experimented in term of the dynamic auralization demonstrator. We introduced some examples of different reproduction methods.

The results can be useful for simulation of array-based sound field synthesis methods to investigate its performance in virtual and- real acoustic applications (e.g. virtual acoustic environment, computer games and cinemas ….etc.). The software tool implemented in Matlab can be optimized to meet the needs of real world applications.

Since the sound field synthesis methods demand high computational power, optimization computing solutions considering systems properties and requirements can be developed for different applications. In this thesis the loudspeaker is assumed to be monopole. Considering the directivity pattern of the real world loudspeaker can improve the results. The formulation of spatial aliasing consideration would help by finding the best array design and therefore enhance the performance of the system. The results acquired were dependent on resources, which have to be evaluated, for example in auditory models.

Bibliography

[AW01]   Arfken , G. B., and H. J. Weber. *Mathematical Methods for Physicists.* Academic Press, 2001.

[Ahr08]   Ahrens, J., and S. Spors. "*An Analytical Approach to Sound Field Reproduction Using Circular and Spherical Loudspeaker Distributions*" (Acta Acustica United With Acustica) 94 (2008): 988 – 999.

[Ahr10]   Ahrens, J. *The Single-layer Potential Approach Applied to Sound Field Synthesis Including Cases of Non-enclosing Distributions of Secondary Sources.* Berlin: PhD Thesis, Institut für Softwaretechnik und Theoretische Informatik, Fachgebiet Quality and Usability Lab der Technischen Universität Berlin, 2010.

[Dan00]   Daniel, J. "Représentation de champs acoustiques, application à la transmission et à la reproduction de scénes sonores complexes dans un concontexte multimédia." PhD Thesis, Université Paris 6, 2000.

[Dan03]   Daniel, J.. *Spatial Sound Encoding Including Near Field Effect Introducing Distance Coding Filters and a Viable, New Ambisonic Format.* Copenhagen: In Proceedings of the 23rd International Conference of the Audio Engineering Society, 2003.

[Faz09]   Fazi, F. M., P. A. Nelson, und R. Potthast. „Analogies and Differences Betweeen Three Methods for Sound Field Reproduction." *AMBISONICS SYMPOSIUM.* Graz, June 25-27 2009.

[Ger85]   Gerzon, M. A. *Ambisonics in Multichannel Broadcasting and Video.* Vol. 11. 33 vols. Journal of the Audio Engineering Society, 1985.

[Ger92]   Gerzon, M. *General metatheory of auditory localisation.* Vienna: In Proceedings of the 92th convention of the Audio Engineering Society, 1992.

[Hul04]   Hulsebos, E. *Auralization using Wave Field Synthesis.* PhD thesis, Delft University of Technology, 2004.

[Kut91]   Kuttruff, H. *Room acoustics.* London: Elsevier Science Publishers, 1991.

[Leb99]   Lebedev, V. I., and D. N. Laikov. "*A quadrature formula for the sphere of the 131st algebraic order of accuracy*." Doklady Mathematics 59, no. 3 (1999): 477-481.

[Lin07]   Lindau, A., T. Hohn, and S. Weinzierl. "Binaural resynthesis for comparative studies of acoustical environments." *Journal of the Audio Engineering Society* (Presented at the 122nd Convention), 5-8 May 2007.

[Lin10a]   Lindau, A., L. Kosanke, and S. Weinzierl. *Perceptual evaluation of physical predictors of the mixing time in binaural room impulse responses.* London: presented at the 128th Audio Engineering Society Convention, 22-25 May 2010.

[Lin10b]   Lindau, A., J. Estrella, and S. Weinzierl. *Individualization of dynamic binaural synthesis by real time manipulation of the ITD.* London: presented at the 128th Audio Engineering Society Convention, 22-25 May 2010.

[Maa97]   Maat, I. *Reduction of perceptual redundancy in reverberation synthesis.* Master's thesis, Delft University of Technology, 1997.

[Mal99]   Malham, D. *Second Order Ambisonics The Furse-Malham set.* 1999. http://www.york.ac.uk/inst/mustech/3d_audio/secondor.html (accessed 2010).

[Mal03]   Malham, D. "Space in Music Music in Space." Mphil Sythesis, University of York, 2003.

[Mol04]   Moldrzyk, C., W. Ahnert, S. Feistel, T. Lentz, and S. Weinzierl. *Headtracked auralization of acoustical simulation.* San Francisco: Presented at the 117th AES Convention, 2004.

[Nic10]   Nicol, R. *Binaural technology.* 1. AES Monograph, 2010.

[OL10]   Ohm, J. R., und H Lüke. *Signalübertragung.* 11. Springer, 2010.

[Oli89]   Olive, S. E., and F. E. Toole. "*The Detection of Reflections in Typical Rooms*" (Journal of the Audio Engineering Society) 37, no. 7/8 (July/August 1989).

[Pie91]   Pierce, A. D. *Acoustics. An Introduction to its Physical Principles and Applications.* Acoustical Society of America, 1991.

[Pol05]   Poletti, M. A. *Three-Dimensional Surround Sound Systems Based on Spherical Harmonics.* Journal of the Audio Engineering Society, 27 September 2005.

[Pul97]   Pulkki, V. . *Virtual sound source positioning using vector base amplitude panning.* Journal of the Audio Engineering Society, 1997.

[Pul99]   Pulkki, V. „Uniform Spreading of Amplitude Panned Virtual Sources." *IEEE* (In Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics), 1999: W99.1- W99.4.

[Spo05]   Spors, S. M. *Active Listening Room Compensation for Spatial Sound Reproduction Systems.* PhD Thesis, Friedrich-Alexander-University Erlangen-Nürnberg, 2005.

[Spo06]   Spors, S. M., and R. Rabenstein. *Spatial Aliasing Artifacts Produced by Linear and Circular Loudspeaker Arrays used for Wave Field Synthesis.* Journal of the Audio Engineering Society, 2006.

[Spo07]   Spors, S. M. *Extension of an Analytic Secondary Source Selection Criterion for Wave Field Synthesis.* 123th AES Convention, New York: Audio Engineering Society, 2007.

[Spo08]   Spors, S. M., R. Rabenstein, and J. Ahrens. *The Theory of Wave Field Synthesis.* Amsterdam: presented at the 124th Audio Engineering Society Convention, 17–20 May 2008.

[Spo081]   Spors, S. M., and J. Ahrens. *A Comparison of Wave Field Synthesis and Higher-Order Ambisonics with Respect to Physical Properties and Spatial Sampling.* Journal of the Audio Engineering Society, 2008.

[Vor08]   Vorländer, M. *Auralization.* Vol. 1st edition. Springer-Verlag Berlin Heidelberg, 2008.

[Wei08]   Weinzierl, Stefan. *Handbuch der Audiotechnik.* Springer Verlag, 2008.

[Wik1]   Wikipedia. *Bessel function.* 2008. http://en.wikipedia.org/wiki/Bessel_function (accessed 2010).

[Wik2]   Wikipedia. *Huygen's principle.* 2010. http://en.wikipedia.org/wiki/Huygens-Fresnel_principle (accessed 2010).

[Wik3]   Wikipedia. *Zugeordnete Legendrepolynome.* 2009. http://de.wikipedia.org/wiki/Zugeordnete_Legendrepolynome (accessed 2010).

[Wik4]   Wikipedia. *Polyhedron.* 2010. http://en.wikipedia.org/wiki/Polyhedron (accessed 2010).

[Wik5]   Wikipedia. *Lebedev Quadrature.* 2010. http://en.wikipedia.org/wiki/Lebedev_quadrature (accessed 2010).

[Wil99]   Williams, E. G. *Fourier Acoustics: Sound Radiation and Nearfield Acoustical.* Academic Press, 1999.

[Zot09]   Zotter, F. *Analysis and Synthesis of Sound-Radiation with Spherical Arrays.* Austria: PhD Thesis, Institute of Electronic Music and Acoustics, University of Music and Performing Arts, 2009.

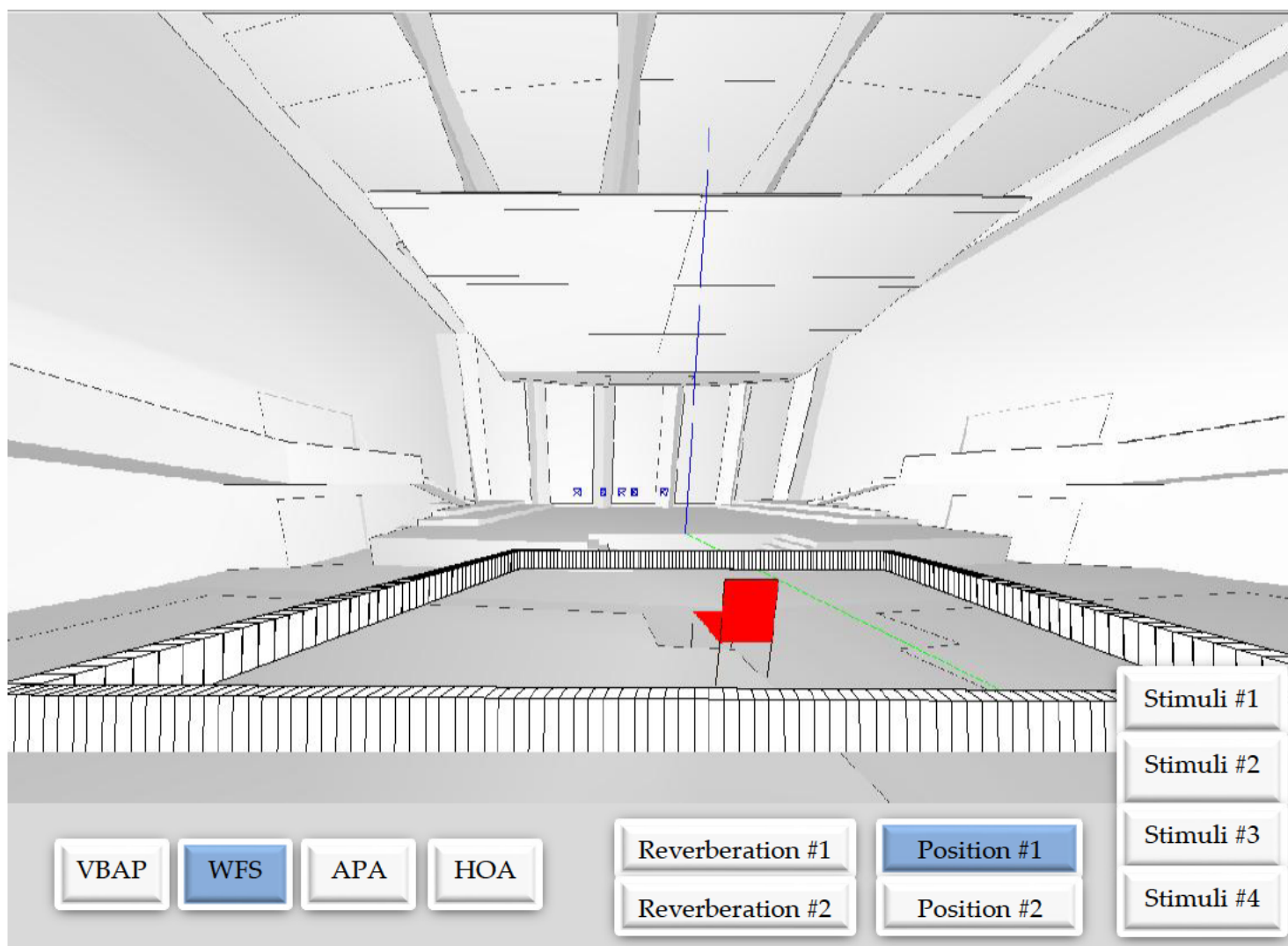# Appendix A:

Demonstrator GUI-Design

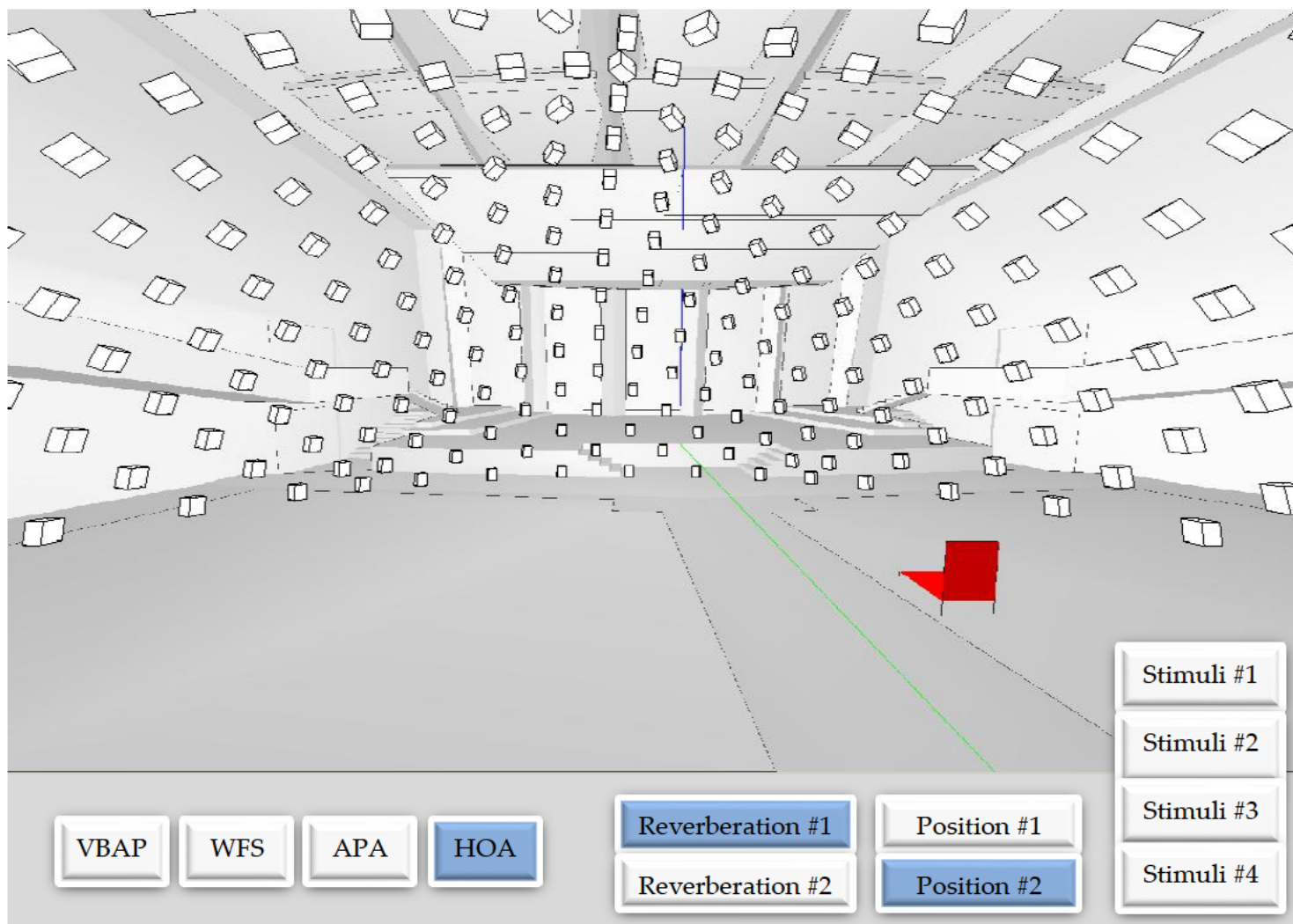**Figure A.1**: Graphical User Interface of the demonstrator (WFS, position #1: $[0\ 0\ 0]^T$)

Figure A.2: Graphical User Interface of the demonstrator (HOA, position #2:$[3 - 3\ 0]^T$ and Reverberation #1 based on modeled reverberation tail).
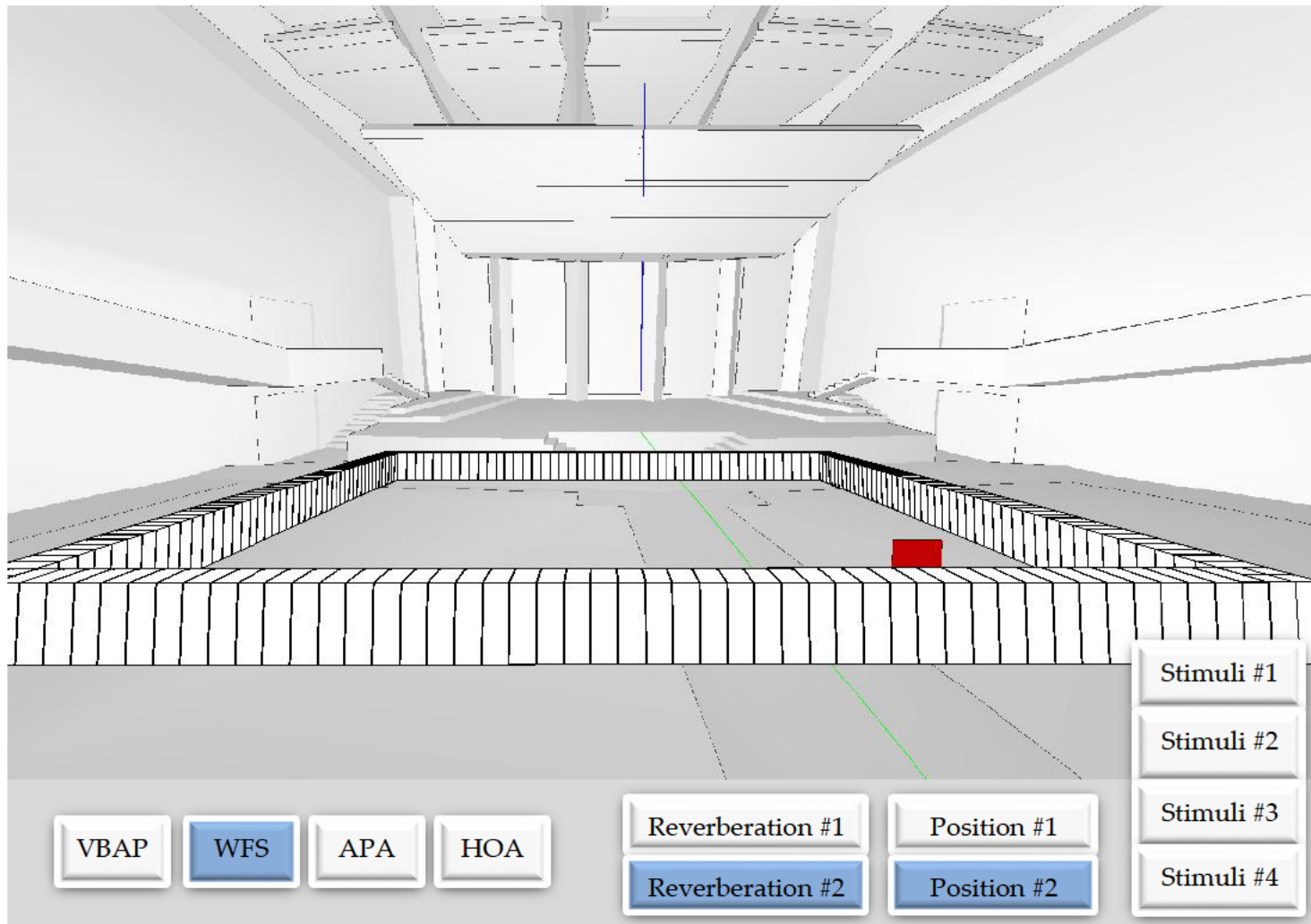
Figure A.3: Graphical User Interface of the demonstrator (WFS, position #2:$[3 - 3\ 0]^T$ and Reverberation #2 based on plane wave synthesis).