Rigid Sphere Microphone Arrays for Spatial Recording and Holography

Peter Plessas

Nov 16th 2009

Advisor: Franz Zotter
Co-Advisors: Alois Sontacchi and Prof. David Wessel

Assessor: Prof. Robert Höldrich

Graz University of Technology, Austria

IEM Institute of Electronic Music and Acoustics
University of Music and Performing Arts Graz, Austria

In Collaboration with
CNMAT Center for New Music and Audio Technologies
University of California, Berkeley

Abstract

This work is a treatise on three-dimensional sound recording. It explores the suitability of spherical microphone arrays in music recording. A detailed discussion of spatial audio theory culminates in a unified representation of sound fields using the spherical harmonic transform. Diverse and alternative array architectures are simulated with regard to their performance. A mathematical model using a new error measure is given and employed in the evaluation of different array layouts and their possible imperfections. An implementation of the algorithms is shown and verified in test recordings using an actual array construction. The obtained results lead to an analysis and to possible improvements for the hardware and signal processing chain.

Kurzfassung

Diese Arbeit ist eine Abhandlung über die dreidimensionale Schallaufnahme. Sie basiert auf der Verwendung von kugelförmigen Mikrofonanordnungen um zu einer einheitlichen Darstellung von Schallfeldern zu gelangen. Die Eignung dieser Technik für die Aufnahme von Musik wird untersucht. Eine ausführliche Diskussion der Theorie räumlicher Schallausbreitung führt zu wichtigen Entwurfsrichtlinien für den Bau derartiger Apparaturen. Die Zerlegung eines Schallfeldes in sphärische harmonische Komponenten wird erklärt. Unterschiedliche Bauformen werden einer Simulation unterworfen um die Güte der Bearbeitung zu beurteilen. Ein mathematisches Modell unter Verwendung neuartiger Fehlergrössen wird vorgestellt und bewertet verschiedene Anordnungen und deren Toleranzen. Eine Implementation der Algorithmen zur Schallfeldzerlegung wird gezeigt und durch Aufnahmen mit einer neuartigen Mikrofonanordnung verifiziert. Die derart gewonnenen Ergebnisse werden analysiert und führen zu einer Aufstellung von Verbesserungen für die Apparatur wie auch für die Signalverarbeitungskette.

Acknowledgments

I want to thank Franz Zotter for being the best advisor this thesis can have. His inspiration and guidance have been invaluable to me as were his motivation and good humor.
I want to thank Brigitte Bergner, Gerhard Eckel, Robert Höldrich, Thomas Musil, Markus Noisternig, Winfried Ritsch, Alois Sontacchi and IOhannes Zmölnig at the IEM and my teachers and colleagues in Graz who make it the special place that has taught me so much.
I want to thank Rimas Avizienis, David Wessel and the staff at CNMAT in Berkeley for their great collaboration and ideas, John Meyer and Pete Soper at Meyersound for building the most beautiful microphone array, and Florian Hollerweger for tons of confidence.

I dedicate this work to my parents. My family's support and trust encouraged me to do what I love.

This thesis is submitted in partial fulfillment of the requirements for the degree "Diplomingenieur" in sound engineering (EE). This joint program between the University of Music and Performing Arts, and the Graz University of Technology is based on electrical engineering, acoustics and signal processing alongside music education, composition and recording.
Part of this work is the result of a collaboration between the IEM – Institute of Electronic Music and Acoustics, University of Music and Performing Arts Graz, and CNMAT – Center For New Music and Audio Technologies, University of California, Berkeley. It has been made possible with generous help from the Austrian Marshall Plan Foundation.
This document includes corrections as of November 24, 2009.

I-80 E, 37°46′36.23" N, 122°24′11.34" W

Contents

1 Introduction

Spatial recording of sound and music has a long and interesting history. Many promising attempts were received with varying success by the music and technology industry. A multichannel loudspeaker setup may still not suit the early 21st century living room and its inhabitants. But the availability of array processing knowledge and computing power leads to novel individual and institutional work. Spherical microphone arrays are an exciting progression in spatial recording techniques. They have been described by various authors giving design criteria and formulae for the processing of captured sound fields. Usable implementations, their exploration and audible results are still rare. There are many challenges in building microphone arrays. The robustness of the algorithms varies with the chosen application. Music recording imposes different requirements than room acoustics or speech processing. Measurements can be achieved with a single microphone mounted on a robotic arm as long as the system under test is time-invariant. The limited bandwidth suitable in speech processing does not satisfy the capture of an orchestra performance. It is the aim of this thesis to explain the theory and extend it to many different spherical arrays describing incident sound fields and to verify the suitability for music recording.

The applications for microphone arrays are manifold. Audio data can be treated in various ways after the recording has been done, permitting spatial selection of different sources and listening directions. An adjustable focus extends the stereo stage to 360 degrees, with microphone choice and placement adjustable in post-production. Virtual microphones can be modeled resembling different sensitivity patterns and steered into arbitrary directions. The inverse approach, canceling out unwanted sources or noise, is equally feasible. Since the complete spatial representation of a sound field is calculated, spherical arrays are perfect surround sound microphones independent of distribution standards and speaker layouts. It is possible to derive higher order Ambisonic signals as well as motion picture formats or stereo and binaural representations. In room acoustics, the measurement of a three dimensional impulse response does not only capture a room's parameters such as reverb length and frequency response but preserves a complete geometrical fingerprint identifying walls and objects through acoustic holography. These spatial impulse responses are crucial in authentic room simulation and reverberation. Adaptive filter techniques use the spatial information presented by a microphone array to locate sound sources in feedback suppression and noise cancellation algorithms.

2 Fundamentals: Exterior and Interior Problems

Most of the procedures and limitations presented here are applicable to loudspeaker and microphone arrays alike. The task common to all circumjacent microphone arrays is to determine a field according to one or several sources. Applied to acoustics this means that a sound field caused by one or several sources can be determined for a certain area. To be more precise, it is pressure changes or air particle velocity due to a sound source which is sampled by the array sensors. Once these values are known it is possible to deduct the field according to equations known from theory for any area vacant of additional sources [Wil99]. This task leads to the following two scenarios:

2.1 Exterior Problem

Sound sources cause a sound pressure and sound particle velocity distribution at any point in a volume. If this distribution is known along a surface enclosing these sources, it is possible to determine the entire sound field outside of the outermost source. This principle is shown in figure 1. Note that objects causing reflections are considered secondary sources. The challenge at hand is to determine the outer field, hence exterior problem. As an example, radiation analysis of sound sources requires the solution of this problem. Array geometries are not necessarily limited to spherical shells, but these allow for stable and elegant solutions as shown in this work.

2.2 Interior Problem

Similar conditions arise in the complementary task, the determination of a sound field caused by sources outside of the array. Provided that an interior volume is free of sources and objects, the entire field up to the innermost source can be deducted as sketched in figure 2. One application of interior problems is music recording preserving spatial information.

2.3 Mixed Problem

The two problem sets above can be combined into a mixed problem, where sources inside and outside of a volume to be described are known, or where two separated fields outside and inside of a source area are to be determined. In a later section, this mixed problem will be employed to compensate for the scattering effects of the microphone construction on the sound field.

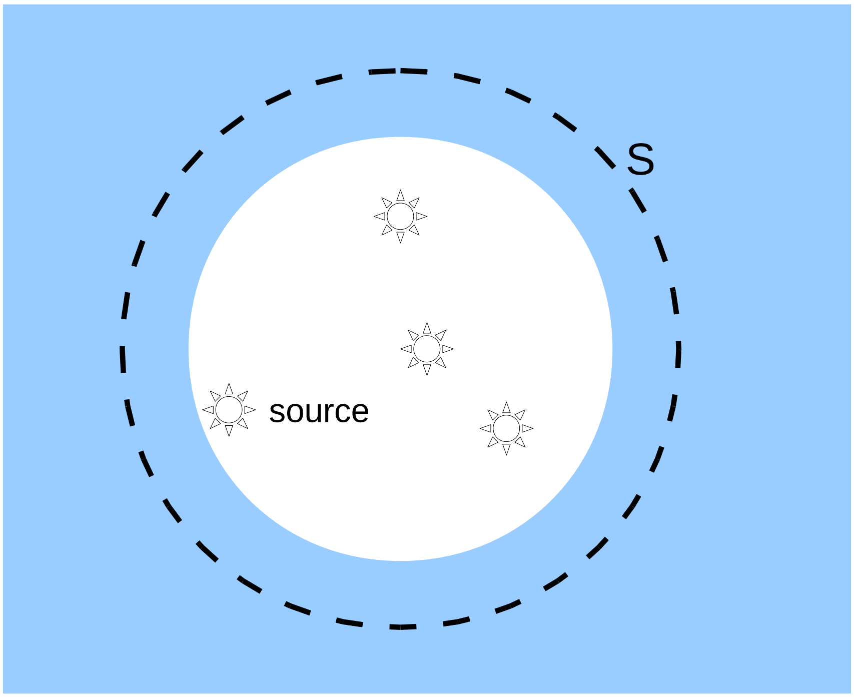

Figure 1: Exterior problem: The shaded source free volume exterior to all sources can be determined once its distribution is known on the entire surface S

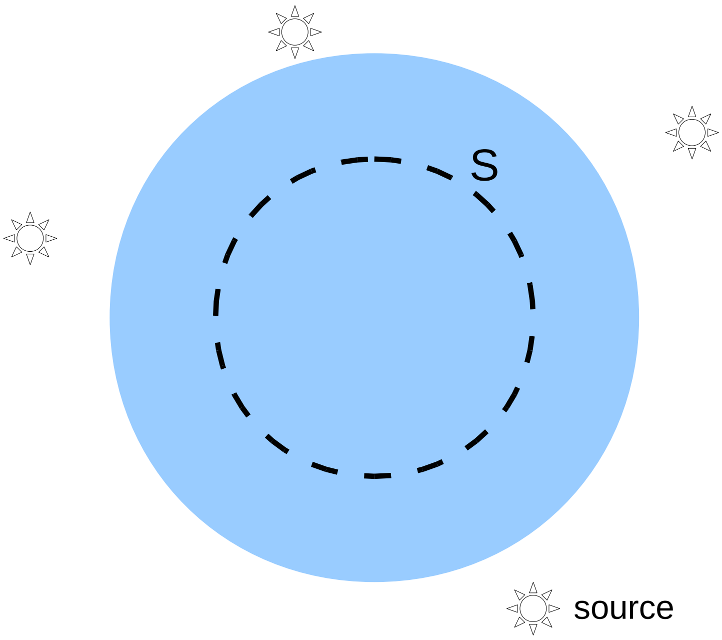

Figure 2: Interior problem: The shaded source free volume interior to all sources can be determined once its distribution is known along the surface S

2.4 Boundary Value Problems

The mathematical foundation for the tasks above is provided by the Dirichlet and Neumann boundary value problems for sound pressure and sound particle velocity, respectively [AW01, p.758]. In Dirichlet boundary value problems a given value (sound pressure) on a surface determines a valid solution from a set of partial differential equations. In the case of the Neumann boundary value problem, it is the value's radial gradient (radial sound particle velocity) which is used as a boundary condition [Wei09].

3 Spatial Fourier Transforms and the Spherical Wave Spectrum

3.1 Coordinate Systems

The spherical coordinate system used throughout this entire text is the ISO 33-11 standard [TT08], which is a right hand coordinate system with the thumb resembling the X-axis, index finger Y-axis and middle finger the Z-axis.
The corresponding angles in spherical coordinates are the inclination or zenith angle theta ϑ from the Z axis ranging from 0° to 180°, and the azimuthal angle phi φ counted positive in counterclockwise direction from the XZ plane along 0° to 360°.[2]
It is convenient to encode the two angles (ϑ, φ) into a unit vector $\boldsymbol{\theta}$ of radius one:

$$\boldsymbol{\theta} = \begin{pmatrix} \cos(\varphi)\sin(\vartheta) \\ \sin(\varphi)\sin(\vartheta) \\ \sin(\vartheta) \end{pmatrix} \tag{1}$$

3.2 Spherical Harmonics

A distribution on a spherical surface can be represented by a superposition of spherical harmonics. These harmonics are solutions to the spherical harmonic differential equation, the angular part of Laplace's equation in spherical coordinates [Wei09]. For most applications it is sufficient to use real-valued spherical harmonics [Zot09a]. Their argument is the angle $\boldsymbol{\theta}$. Spherical harmonics exist for different orders n due to their dependency on the degree of an associated Legendre polynomial. Every order n is represented by $(2n+1)$ modes, which are labeled m ranging from $-n$ to n, as shown in figure 3.

3.2.1 Real-valued spherical harmonics

The real-valued spherical harmonics are given as [Wil99, p.191]:

$$Y_n^m(\boldsymbol{\theta}) = N_n^m P_n^{|m|}(\cos(\vartheta))\sin(m\varphi) \quad \text{for } m < 0 \tag{2}$$

$$Y_n^m(\boldsymbol{\theta}) = N_n^m P_n^{m}(\cos(\vartheta))\cos(m\varphi) \quad \text{for } m \geq 0 \tag{3}$$

where $P_n^m(\cos(\vartheta))$ denotes an associated Legendre function.

[2] Note that some numerical computation programs such as GNU Octave [Oct] employ a different scheme here, with phi φ denoting an elevation angle reaching ±90° up and down from the XY plane, and the azimuthal angle designated theta ϑ.

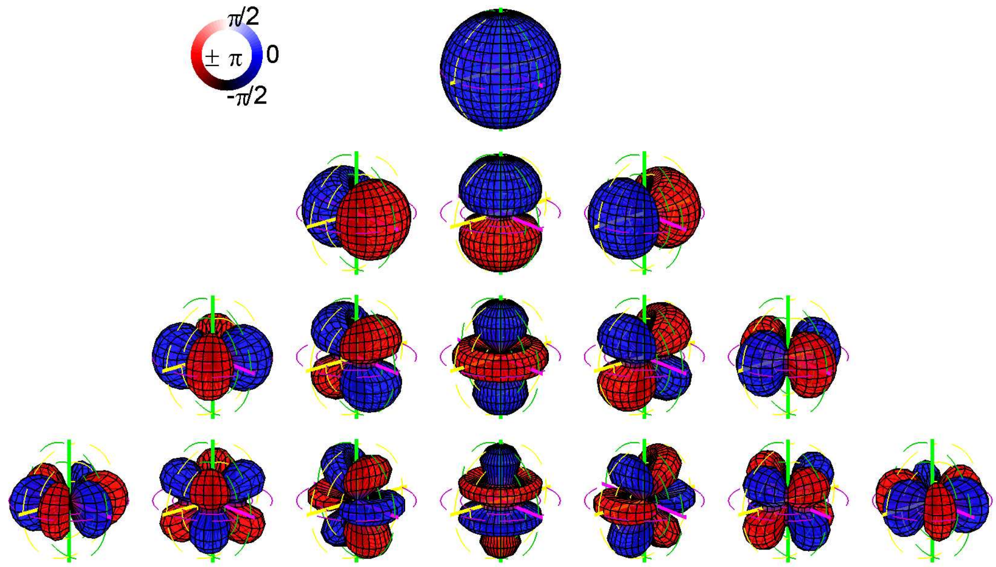

Figure 3: Magnitude of real-valued spherical harmonics plotted as radius, for different orders n and associated modes m [Pom08]

The normalization constant N_n^m is given as:

$$N_n^m = (-1)^{|m|} \sqrt{\frac{(2n+1)(2-\delta[m])}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} \tag{4}$$

These normalized real-valued spherical harmonics form a complete set of orthonormal base functions.

3.2.2 Orthonormality

The orthonormality of spherical harmonics is shown by the integral of two harmonics along a sphere, which equals zero for different indices and one for equal indices.

$$\int_{\varphi=0}^{2\pi} \int_{\vartheta=0}^{\pi} Y_n^m(\vartheta,\varphi) Y_{n'}^{m'}(\vartheta,\varphi) \sin(\vartheta)\, d\vartheta d\varphi = \delta(n-n')\delta(m-m')$$

And using the more concise unit vector $\boldsymbol{\theta}$ notation

$$\int_{\mathbb{S}^2} Y_n^m(\boldsymbol{\theta}) Y_{n'}^{m'}(\boldsymbol{\theta})\, d\boldsymbol{\theta} = \delta(n-n')\delta(m-m') \tag{5}$$

where δ denotes Kronecker's delta function, and the integral being

$$\int_{\mathbb{S}^2} d\boldsymbol{\theta} = \int_{\varphi=0}^{2\pi} \int_{\vartheta=0}^{\pi} \sin(\vartheta)\, d\vartheta d\varphi \tag{6}$$

3.3 Spherical Harmonic Transform

The transform of a distribution on a sphere into spherical harmonics is a transform of a periodic function into its orthonormal components, as familiar from Fourier transforms of time domain signals. For every order n and mode m the integral over all possible angular positions on a sphere gives the correlation of the function with the transform kernel, the respective spherical harmonic. It is therefore possible to refer to the spherical harmonic transform as a spatial Fourier transform. Instead of a frequency variable ω it is now the harmonic's indices n, m which allow to choose components of the resulting angular spectrum.

3.3.1 Spherical harmonic transform

The analysis or transform of a function $g(\vartheta, \varphi)$ into spherical harmonics coefficients γ_{nm} is defined as [Raf05]:

$$SHT_{nm}\left\{g(\vartheta,\varphi)\right\} = \gamma_{nm} = \int_{\varphi=0}^{2\pi}\int_{\vartheta=0}^{\pi} g(\vartheta,\varphi) Y_n^m(\vartheta,\varphi)\sin(\vartheta)\,d\vartheta d\varphi$$

And in unit vector notation:

$$SHT_{nm}\left\{g(\boldsymbol{\theta})\right\} = \gamma_{nm} = \int_{\mathbb{S}^2} g(\boldsymbol{\theta}) Y_n^m(\boldsymbol{\theta})\,d\boldsymbol{\theta} \tag{7}$$

Note that the unit vector notation will be used from now on. Refer to (1) and (6) for conversion.

In analogy to a Fourier transform resulting in a frequency spectrum, the arbitrary function $g(\boldsymbol{\theta})$ on a spherical surface is now given as γ_{nm}, being the result of a spherical harmonic transform (SHT). The initial function is now represented in the spherical harmonic spectrum.

It is important to note that the total amount of spherical harmonics is infinite here. It will be shown later that finite numbers of harmonics can be used in an implementation giving approximate results. The higher the order of spherical harmonics considered, the better the angular representation of the transformed function. This leads to the notion of angular bandwidth, which is infinite for an endless number of harmonics. The infinite transform results in a perfect representation for arbitrarily narrow functions in the angular sense. The function $g(\boldsymbol{\theta})$ is assumed to be defined and valid at any position on the sphere, which is not given using actual array sensors. The effects of a finite set of harmonics, and of functions sampled at discrete points are crucial to the performance of any microphone array and are discussed in this work.

3.3.2 Inverse spherical harmonic transform

The inverse spherical harmonic transform ($ISHT$) or expansion of a spherical harmonic spectrum into the function $g(\boldsymbol{\theta})$ is achieved by the infinite sum along all components m,n at the angular position $\boldsymbol{\theta}$:

$$ISHT_{nm}\{\gamma_{nm}\} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \gamma_{nm} \, Y_n^m(\boldsymbol{\theta}) = g(\boldsymbol{\theta}) \tag{8}$$

3.3.3 Completeness

The completeness of the spherical harmonics transform transform using infinite harmonics can be shown by a forward transform followed by an inverse transform resulting in the original function.

$$ISHT_{nm}\{SHT_{nm}\,\{g(\boldsymbol{\theta})\}\} = g(\boldsymbol{\theta}) \tag{9}$$

3.3.4 Parseval's theorem

The orthonormality property (5) and the completeness of the spherical harmonic transform [AW01] fulfill Parseval's theorem, which describes the unitarity of the spherical harmonic transform the same way as it does for other transforms [Zot09b].

$$\int_{\mathbb{S}^2} |g(\boldsymbol{\theta})|^2 d\boldsymbol{\theta} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} |\gamma_{nm}|^2 \tag{10}$$

3.3.5 Spherical harmonic spectra in audio engineering

One way of understanding the usefulness of spherical wave spectra in audio engineering is to think of them as an extension to the M/S microphone technique. In this microphone arrangement two capsules are mounted as closely together as possible, effectively recording the same sound field, but with different microphone pickup patterns as shown in figure 4. If for example an omnidirectional microphone is used along a figure-of-eight microphone, the listening direction in the stereo panorama can be determined at playback by combining the two signals at different levels and phases. This encoding of signals into mid and side components was invented by Alan Blumlein in his classic 1931 patent on stereophonic sound reproduction [Blu58]. An extension of this approach was described by Michael Gerzon in 1973 [Ger72] and became the system known as Ambisonics. It allows to capture the horizontal as well as the vertical dimension. First introduced as a procedure to reproduce a sound field using four loudspeakers, it is in fact a spherical harmonics representation of

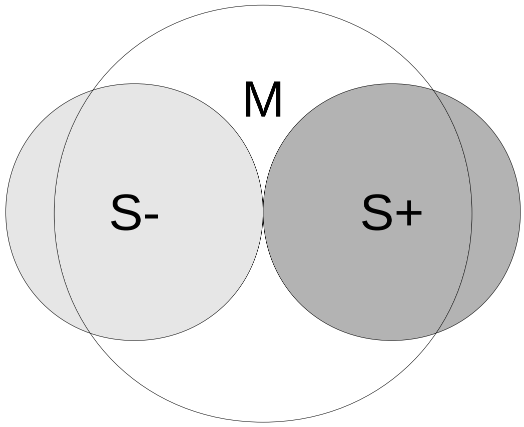

Figure 4: The M/S microphone technique: Combination of the omnidirectional microphone signal and the positive or negative figure-of-eight microphone signal allows to change the microphone's polar pattern and orientation after the recording has been done

order $N = 1$, requiring four channels of audio. Similarly to the M/S technique these four channels, labeled $WXYZ$, consist of an omnidirectional W channel and three figure-of-eight channels XYZ which are rotated according to the three modes m at order $n = 1$. This encoding scheme is known as the B-format. Due to the large physical extension of four microphone capsules it is not possible to place them in exactly the same spot. A microphone layout suited for order $N = 1$ was invented by Gerzon and Craven [Ger75] and built as a commercial product by Calrec and later marketed as the Soundfield microphone. It consists of four cardioid capsules arranged on the four surfaces of a tetrahedron. By matrixing the microphone signals and attempting to compensate for capsule distances with frequency filters, the first order Ambisonics B-format signals are derived. The origins of this approach can be seen as a first attempt at reconstructing sound fields using acoustic holography, which is discussed in a later section. Ambisonic microphones share the two challenges crucial to any spherical microphone array application: Decomposition of the sound field into spherical harmonics and filtering the signals according to the capsule placement.

4 Sources in the Spherical Harmonic Spectrum

To allow directivity by considering the radial properties of the recorded sound field, and to compensate for a microphone array's physical dimensions, the laws of sound propagation have to be taken into account. This is simplified by the representation of sound pressure and sound particle velocity in the spherical harmonic spectrum.

4.1 Sound Pressure and Particle Velocity

4.1.1 Spherical harmonics representation of sound pressure

The transformed sound pressure $\psi_n^m(kr)$ is represented by spherical harmonics n, m and is therefore independent on the angle $\boldsymbol{\theta}$. It incorporates the entire angular information and is merely dependent on (kr), the radius r at which the sound pressure is determined, and the wave number $k = \frac{\omega}{c}$ denoting frequency [Zot09a]:

$$\psi_n^m(kr) = SHT_{nm}\{p(k, r, \boldsymbol{\theta})\} = b_{nm} j_n(kr) + c_{nm} h_n^{(2)}(kr) \tag{11}$$

In this equation $j_n(kr)$ is the spherical Bessel function, and $h_n^{(2)}$ the spherical Hankel function of the second kind. b_{nm} are the coefficients of the incident wave $j_n(kr)$, and c_{nm} are the coefficients of the radiating wave $h_n^{(2)}(kr)$. Refer to appendix A for more details on the functions involved.

4.1.2 Spherical harmonics representation of sound particle velocity

The spherical harmonic transformed radial component of the sound particle velocity $\nu_n^m(kr)$ is given as [Zot09a]:

$$\nu_n^m(kr) = SHT_{nm}\{v(k, r, \boldsymbol{\theta})\} = \frac{\mathrm{i}}{\rho_0 c} \left[b_{nm} j_n'(kr) + c_{nm} h_n'^{(2)}(kr) \right] \tag{12}$$

The spherical Bessel and Hankel functions are given as derivatives with regard to (kr). They can be computed using the recurrence equation (69) derived in appendix A.

4.1.3 Example: Spherical harmonics expansion

The sound pressure $p(k, r, \boldsymbol{\theta})$ for frequency k at any point $(r, \boldsymbol{\theta})$ can be determined by expansion of a spherical spectrum into the pressure function. This expansion is the inverse spherical harmonic transform ($ISHT$):

$$p(k, r, \boldsymbol{\theta}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \psi_n^m(kr)\, Y_n^m(\boldsymbol{\theta})$$
$$= \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left[b_{nm} j_n(kr) + c_{nm} h_n^{(2)}(kr) \right] Y_n^m(\boldsymbol{\theta})$$

4.2 Point Sources and Plane Wave Sources

The coefficients b_{nm} and c_{nm} resulting from the spherical harmonic transform represent the components of the wave field. They can be formulated in the spherical harmonics domain:

4.2.1 Incident plane wave

An incident plane wave at a listening point $\boldsymbol{\theta}$ is caused by a source located at infinity. Coefficients b_{nm} for an incident plane wave arriving from source direction $\boldsymbol{\theta_s}$ are given as [Zot09a]:

$$b_{nm} = 4\pi \mathrm{i}^n Y_n^m(\boldsymbol{\theta_s}) \tag{13}$$

With this knowledge and the sound pressure given in (11) it is easy to state the sound pressure spectrum caused by a plane wave:

$$\psi_n^m = SHT_{nm}\{p(k, r, \boldsymbol{\theta}, \boldsymbol{\theta_s})\} = 4\pi \mathrm{i}^n j_n(kr) Y_n^m(\boldsymbol{\theta_s}) \tag{14}$$

The corresponding sound particle velocity of an incident plane wave is:

$$\nu_n^m = SHT_{nm}\{v(k, r, \boldsymbol{\theta}, \boldsymbol{\theta_s})\} = 4\pi \frac{\mathrm{i}^{n+1}}{\rho_0 c} j_n'(kr) Y_n^m(\boldsymbol{\theta_s}) \tag{15}$$

Using the inverse spherical harmonic transform, the actual sound pressure can now be computed for the listening point $(r, \boldsymbol{\theta})$

$$p(k, r, \boldsymbol{\theta}, \boldsymbol{\theta_s}) = ISHT_{nm}\{\psi_n^m\}$$

By definition, there is no such thing as a radiating plane wave because the listener would have to be positioned at infinity.

4.2.2 Spherical wave of a point source

The coefficients b_{nm} for an incident spherical wave of a point source located at source radius and angle $(r_s, \boldsymbol{\theta_s})$ and listening point $(r, \boldsymbol{\theta})$, where radius $r \leq r_s$, are [Zot09a]:

$$b_{nm} = -\mathrm{i}kh_n^{(2)}(kr_s)Y_n^m(\boldsymbol{\theta_s}) \tag{16}$$

When $r > r_s$, the wave is radiating and the coefficients result in:

$$c_{nm} = -\mathrm{i}kj_n(kr_s)Y_n^m(\boldsymbol{\theta_s}) \tag{17}$$

Hence for the spherical wave of a point source, the sound pressure in the spherical harmonic spectrum is:

$$\psi_n^m = -\mathrm{i}kh_n^{(2)}(kr_s)j_n(kr)Y_n^m(\boldsymbol{\theta_s}) \tag{18}$$

The sound particle velocity of an incident spherical wave can be computed in the same way, resulting in:

$$\nu_n^m = \frac{k}{\rho_0 c}h_n'^{(2)}(kr_s)j_n'(kr)Y_n^m(\boldsymbol{\theta_s}) \tag{19}$$

With these prerequisites taken, the analysis of a sampled sound field according to its wave nature is possible.

5 Capsules and Outer Space: Holography Filters and Radial Filters

Depending on the type of microphones used, the spherical harmonic transform of the microphone signals yields a pressure, or combination of pressure and velocity spectrum at the microphone radius r_d. The next step is to derive a holographic spectrum representing the entire sound field at any radius up to the innermost source. This holographic spectrum identifies the entire source free volume as given in the interior or exterior problem, and is the result of acoustic holography. The step from the sensor spectrum to the entire holographic representation is done by means of a holography filter. After this step, the holographic spectrum can be evaluated at a source radius r_s using a radial filter. The inverse spherical harmonics transform of the spectrum at this source radius gives the actual source amplitude. The holography filter extrapolates the description of the entire sound field from the sensor signals, while the radial filter selects a spectrum at one desired source radius.

5.1 Open versus Closed Spherical Arrays: A Reflective Subject

In cases where a spherical microphone array has dimensions causing scattering of a sound field, or if the array is based on a rigid sphere design, the presence of a physical object violates the requirement of a source free volume, as stated in the interior problem in section 2.2. A combination of the internal and external problem addresses this question as a mixed problem:

5.1.1 Reflections from a rigid sphere, mixed problem

The reflections from a rigid and sound-hard spherical surface of radius r_k is considered a secondary source within the measurement radius. Pressure ψ_n^m and velocity ν_n^m given in the spherical harmonic spectrum lead to the following formulation:
The sound particle velocity on a completely hard surface at radius r_k has to become zero [GW06]. In terms of incident and radiating wave it can be stated that

$$\nu_{n,incident}^m(kr_k) + \nu_{n,radiating}^m(kr_k) = 0 \tag{20}$$

With the radiating and incident velocities given in (12), this condition becomes:

$$\frac{\mathrm{i}}{\rho_0 c}\left[b_{nm} j_n'(kr_k) + c_{nm} h_n'^{(2)}(kr_k)\right] = 0 \tag{21}$$

Assuming the spectrum b_{nm} is known, the reflected (radiated) coefficients c_{nm} can be written as

$$c_{nm} = b_{nm} \frac{j_n'(kr_k)}{h_n'^{(2)}(kr_k)} \tag{22}$$

This is a simplification, since most physical material is not entirely sound-hard. For a more precise description the acoustic impedance of the object must be taken into account. It is favorable to achieve a representation of this impedance in the spherical harmonic spectrum.

A rigid sphere is assumed to be surrounded by microphone diaphragms at a radius $r_d > r_k$. The pressure $\psi_n^m(kr_d)$ and velocity $\nu_n^m(kr_d)$ at the microphone radius can be expressed as already shown in (11) and (12), now including the reflected radiating coefficients c_{nm} from above (22).

$$\psi_n^m(kr_d) = b_{nm} \left[j_n(kr_d) - \frac{j_n'(kr_k)}{h_n'^{(2)}(kr_k)} h_n^{(2)}(kr_d) \right] \tag{23}$$

$$\nu_n^m(kr_d) = \frac{\mathrm{i}}{\rho_0 c} b_{nm} \left[j_n'(kr_d) - \frac{j_n'(kr_k)}{h_n'^{(2)}(kr_k)} h_n'^{(2)}(kr_d) \right] \tag{24}$$

By using the values deducted here as a model for the propagation and scattering of incident waves, all effects of a spherical body can be compensated.

5.2 Holography filters for different array architectures

Depending on the design of the array construction the holography filter has very different properties. For arrays built around an open or rigid sphere, and for those using omnidirectional pressure microphones or cardioid ones, a different filter specification is required. It will be shown that the microphone radius r_d imposes a trade-off between a good signal-to-noise ratio for low frequencies and spatial resolution for high frequencies.

5.2.1 Open sphere with omnidirectional microphones

The spherical harmonic transformed measurement value at the microphone outputs is denoted $\chi_n^m(kr_d)$. For an open array consisting of omnidirectional pressure microphones at r_d the relation of this value to incident sound pressure waves is

$$\chi_n^m(kr_d) = \psi_n^m(kr_d) = b_{nm} j_n(kr_d) \tag{25}$$

To determine b_{nm}, the holographic spectrum for the entire source free volume, the measured value has to be divided by the term $j_n(kr_d)$, a spherical Bessel function dependent on the product of frequency k and microphone radius kr_d, which has several zeros. This would mean infinite gain at certain frequencies which can not be implemented.

5.2.2 Open sphere with cardioid microphones

The problem of inverting a function containing zeros can be avoided by using cardioid microphones facing outwards from an open sphere. Since cardioid microphones measure the sound pressure as well as the sound particle velocity their output $\chi_n^m(kr_d)$ is the following combination [BR07, Zot09a]:

$$\chi_n^m(kr_d) = \psi_n^m(kr_d) - \rho_0 c \nu_n^m(kr_d) \tag{26}$$

With substitution of pressure (11) and velocity (12) the above relation becomes

$$\chi_n^m(kr_d) = b_{nm}\left[j_n(kr_d) - \mathrm{i}j_n'(kr_d)\right] \tag{27}$$

The absolute difference of the Bessel function $j_n(kr_d)$ and its derivative $\mathrm{i}j_n'(kr_d)$ has no zeros. Division of the sensor spectrum $\chi_n^m(kr_d)$ by this difference is perfectly feasible. This filter's complex value (magnitude and phase) is dependent on frequency k as well as on the microphone radius r_d. The magnitude of the filter values for different orders N is shown in figure 5 for a capsule radius of 70 mm. Although the dimension of the array has influences on other performance parameters such as spatial resolution alike, the radius is inversely proportional to the magnitude of the filter at low frequencies. This filter gain in combination with the microphone's noise floor imposes a lower limit to the usable frequency range of an array. A comparison of different radii at a given order is shown figure 6.

5.2.3 Closed sphere with omnidirectional microphones

Inverting a function containing zeros can be also be avoided by placing microphones around a rigid sphere, and compensating for its effects on the sound field as shown in (23). For pressure microphones this results in the relation

$$\chi_n^m(kr_d) = \psi_n^m(kr_d) = b_{nm}\left[j_n(kr_d) - \frac{j_n'(kr_k)}{h_n'^{(2)}(kr_k)} h_n^{(2)}(kr_d)\right] \tag{28}$$

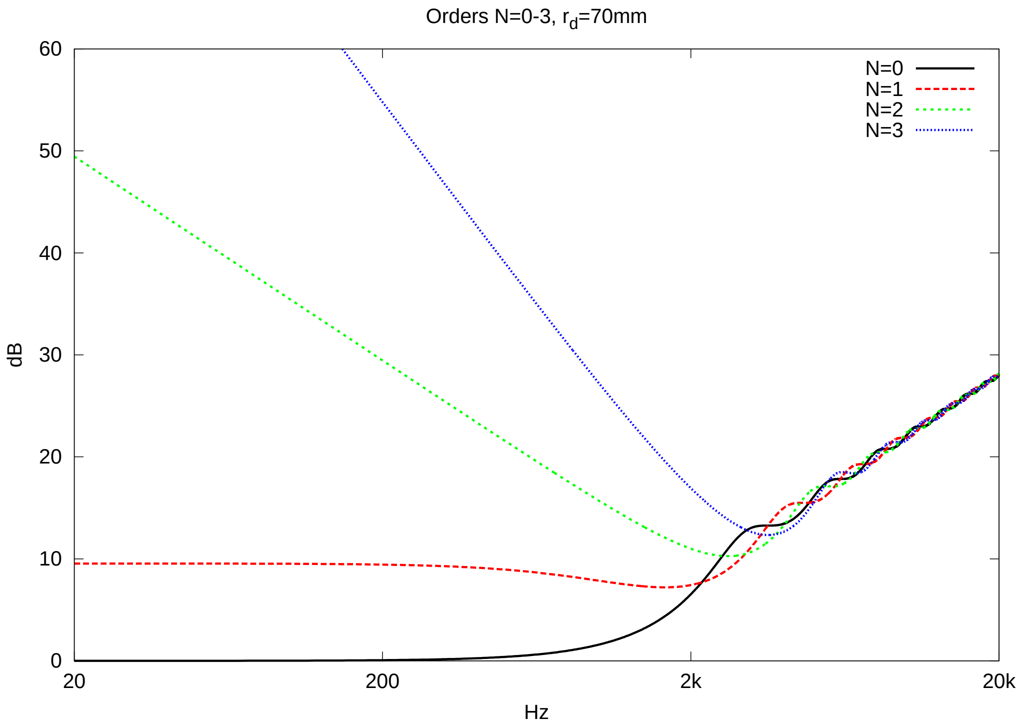

Figure 5: Open array with cardioid microphones, holography filter magnitude for orders N=0-3

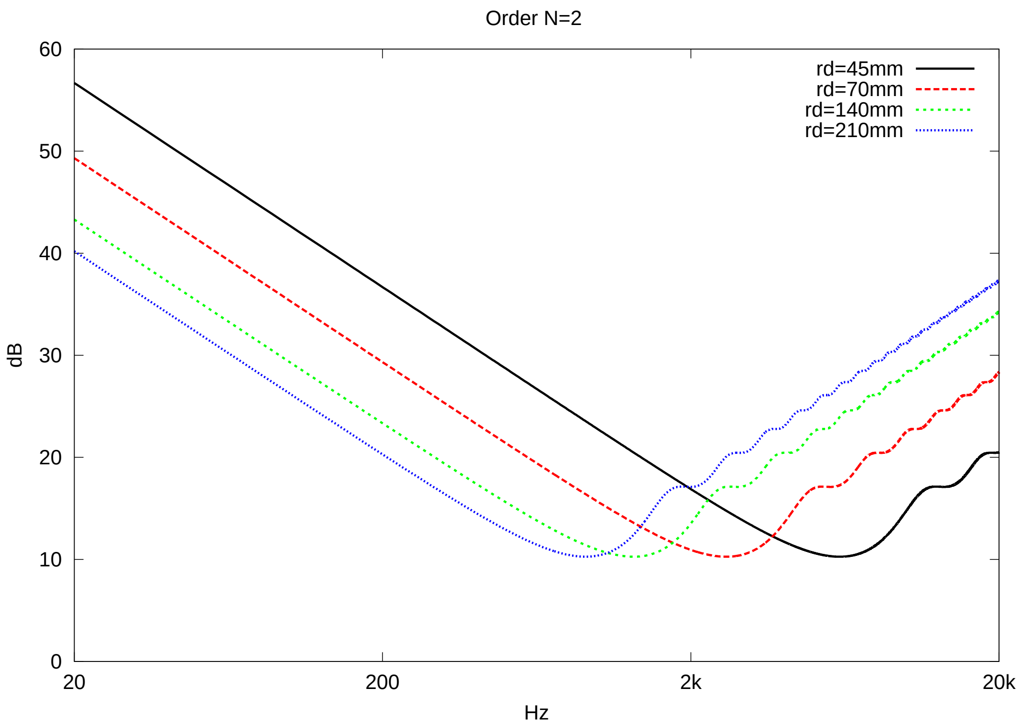

Figure 6: Open array with cardioid microphones, holography filter magnitude for different sensor radii and order N=2

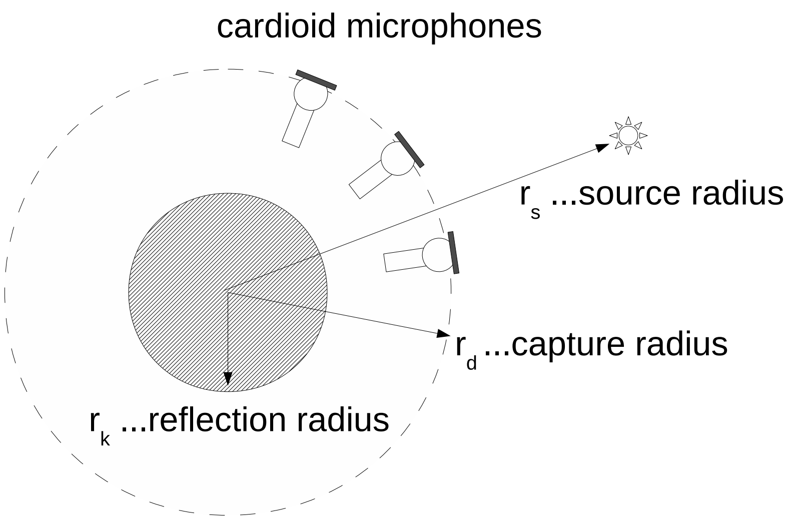

Figure 7: Cardioid microphones around a rigid sphere, with radii for reflective surface r_k, diaphragms r_d, and innermost source r_s given

which has no zeros and can be inverted. For a figure of this filter's magnitude refer to [BR07] and to figures 9 and 10 with $r_d = r_k$.

5.2.4 Closed sphere with cardioid microphones

In cases where cardioid microphones are used, and the microphone construction is large enough to be considered a reflective spherical object, it is important to keep a minimum distance between the rigid sphere and the back of the capsules. Cardioid microphones get their directional sensitivity from an opening in the casing at the back of the capsule. If such a microphone would be flush mounted into a hard sphere and no sound pressure would arrive at the back its response would be omnidirectional. Cardioid diaphragms at radius r_d combine pressure and velocity components of the incident and radiating field [BR07, Zot09a] into a measure value denoted $\chi_n^m(kr_d)$ as already shown in (26). The measured value can now be expressed in terms of the incident field and the reflected field off the sound-hard sphere at radius r_k, as derived in (23) and (24):

$$\chi_n^m(kr_d) = \left[j_n(kr_d) - \mathrm{i} j_n'(kr_d) + \left(\mathrm{i} h_n'^{(2)}(kr_d) - h_n^{(2)}(kr_d) \right) \frac{j_n'(kr_k)}{h_n'^{(2)}(kr_k)} \right] b_{nm} \qquad (29)$$

To determine the holographic spectrum b_{nm} the bracketed reflection and propagation term has to be inverted. This gives a filter function dependent on the wave number

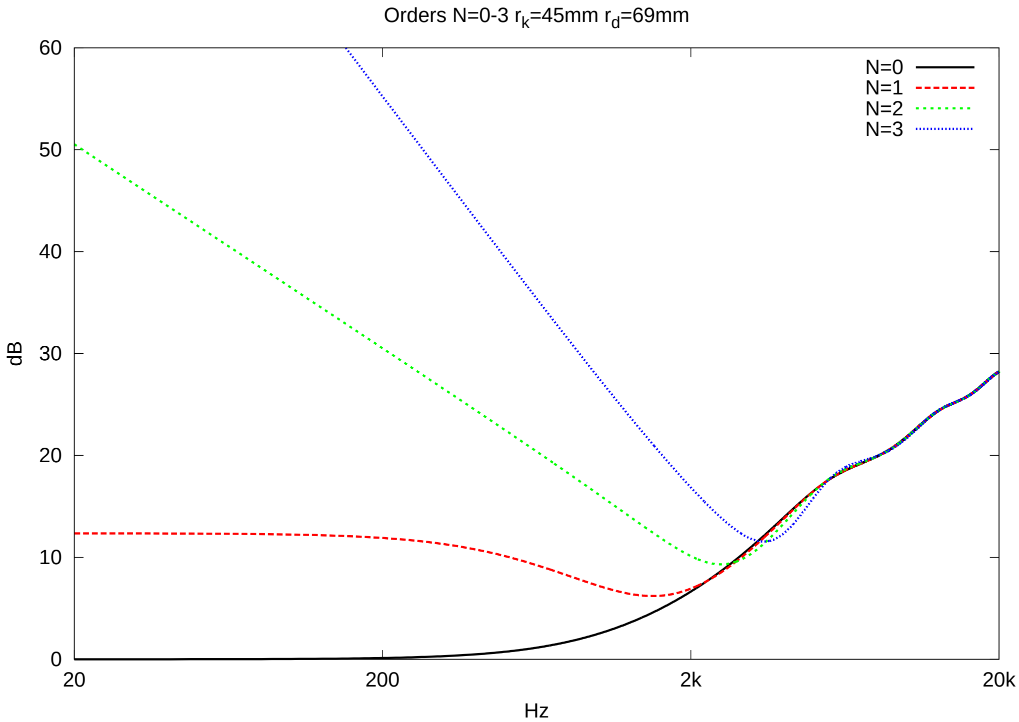

Figure 8: Cardioid microphones around a rigid sphere of radius r_k: Example of holography filter magnitude for different orders N

k with the two radii considered constant. An example of this filter's magnitude is shown in figure 8. The extreme gain for high orders at low frequencies is a challenge in an implementation and limits the usable frequency range. Even if the recorded source would not emit low frequencies, the microphone's own noise (thermal noise, quantization noise, etc.) is present in these low frequencies and would be amplified. It is inherently the microphone's signal-to-noise ratio which limits the use of high orders at low frequencies.

The effect of the reflective sphere is demonstrated best by giving the filter magnitude for different radii r_k as shown in figure 9 and 10 for orders N=1 and N=2 respectively.

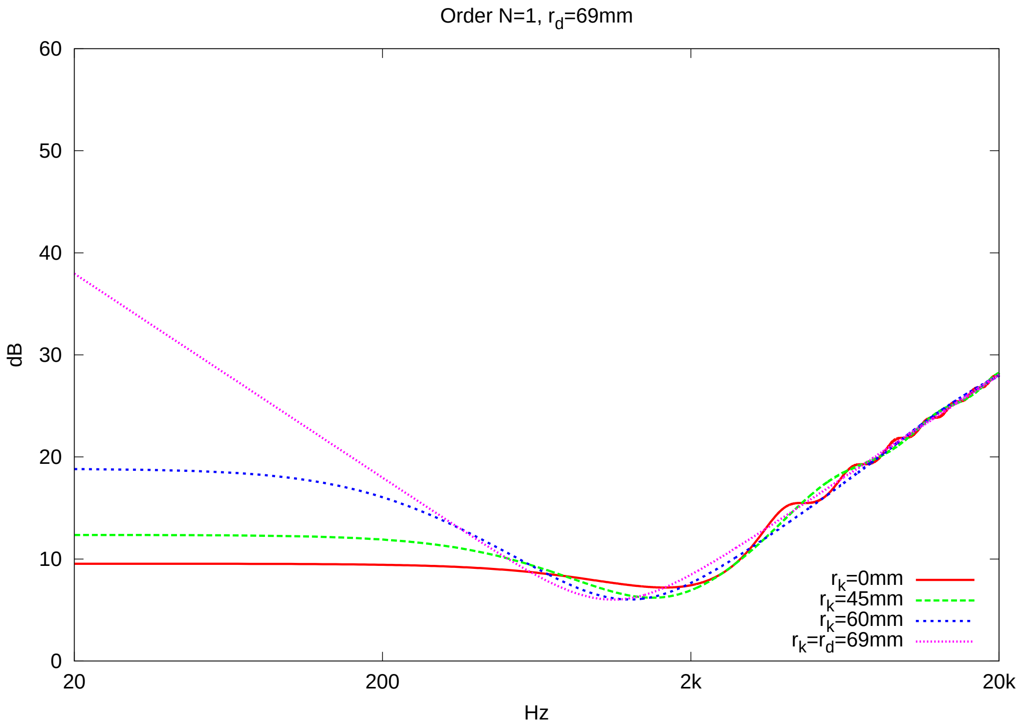

Figure 9: Cardioid microphones around rigid spheres of varying radius r_k: Holography filter magnitude for N=1

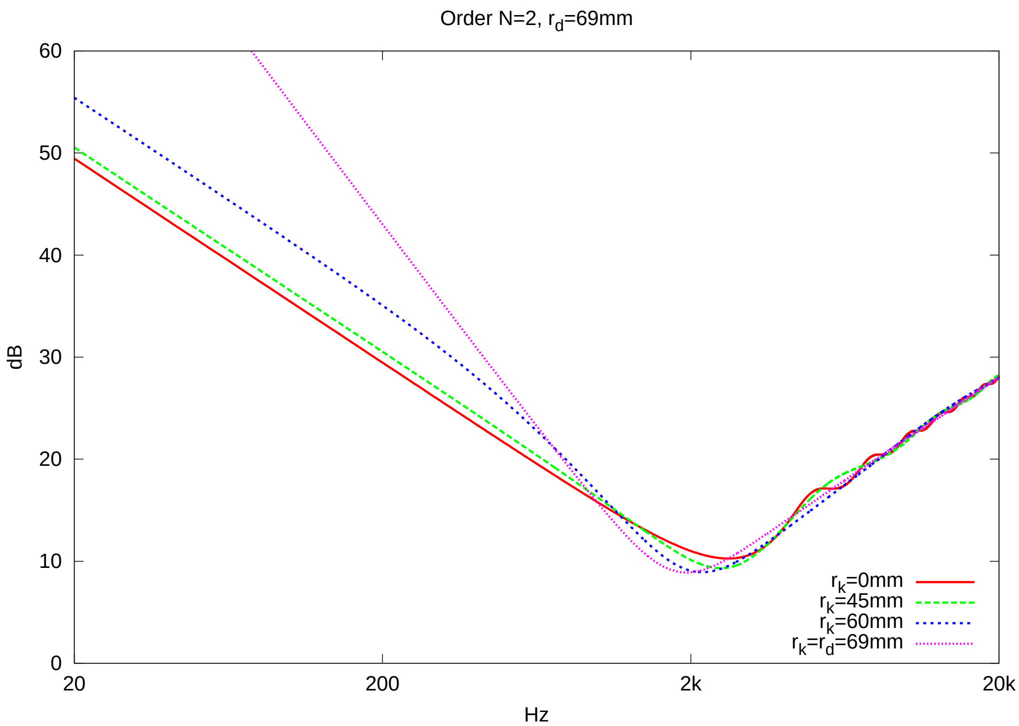

Figure 10: Cardioid microphones around rigid spheres of varying radius r_k: Holography filter magnitude for N=2

5.3 Radial Filters: Focusing on a Source

At this point the representation of incident waves in the spherical harmonics domain is complete. The holographic spectrum b_{nm} allows the evaluation of the source amplitudes at any radius up to the source radius itself. This evaluation is done with a radial filter. As shown in (16), the spectrum b_{nm} for spherical waves is dependent on the source radius r_s. To end up with a spherical harmonic spectrum ϕ_n^m representing the sources complex amplitude at the desired radius, the division

$$\frac{b_{nm}}{-\mathrm{i}kh_n^{(2)}(kr_s)} = \phi_n^m \tag{30}$$

defines this radial filter for spherical waves. It has a different value for different frequencies k and requires the definition of a source radius r_s. In accordance to theory this should be the radius of the actual source, or that of the innermost source if multiple sources are present. For plane waves the simpler term $b_{nm} = 4\pi\mathrm{i}^n Y_n^m(\boldsymbol{\theta_s})$ (15) is independent of a source radius. In the figures given so far as well as in the remainder of this thesis, this simplification was made under the assumption of plane waves for sources at radius $kr_s \geq 1$ and low orders N as discussed in appendix A.4. The results from this simplification are not universally valid but are assumed to be accurate enough for the discussion of the filters at hand.

5.4 Holography

It is important to relate to the term “acoustic holography” coined by Maynard, Williams and Lee [MWL85]. With b_{nm} deferred from the microphone data, it is possible to reproduce the sound field at any point in the source free volume according to its angular representation, the spherical harmonic spectrum and the selection of a radius with a radial filter. This ability and associated procedures are called spherical acoustic holography, in analogy to the well-known technique in optics. In accordance to reproduction systems in the line of the gramophone, the playback of the spatial sound field can be described as acoustic *holophony*. Ambisonic playback systems such as the IEM CUBE [ZSR03] form a subset of holophonic systems. Microphone arrays are the sensing element in acoustic holography. An application and visual representation of spherical holography follows in section 11.

6 Limited Order Harmonics: Discrete Transforms

The amount of points sampling a distribution on a surface is restricted by the size of the microphone capsules and physical dimensions of the array hardware. In order to obtain an infinite spherical harmonics representation as introduced in section 3, the number of microphones would need to be infinitely large, and their physical dimensions infinitely small. In every real microphone array the amount of sample points is therefore limited, and so is the maximum order of spherical harmonics N, which will provide $(N+1)^2$ harmonics in total, as can be derived from figure 3.

6.1 Discrete Spherical Harmonic Transform

A set of L measurement positions on the surface of a sphere samples the distribution at discrete positions. A given relation restricts this discrete set to be represented using a limited amount of spherical harmonics only. The result is an approximation of the sound field which lacks high angular resolution or bandwidth. Sources which are very narrow would need more and higher harmonics to be described completely. As with every sampling application, the highest possible frequency to be represented depends on the sampling rate. This relation is known as the Nyquist-Shannon sampling theorem. Its spatial and spherical version is given as the relation [Zot09b]

$$(N+1)^2 \leq L \tag{31}$$

Spatial aliasing is dependent on the amount and spacing of microphone capsules and increases with frequency. Wavelengths whose dimensions are small compared to the gap between microphones can not be sampled with sufficient angular resolution. This causes aliased copies to be mirrored into lower harmonics. It is this consequence limiting the usable frequency range towards high frequencies.

In order to study discrete spherical harmonic transforms, the inverse transform for limited order N is given first, using a matrix notation. Unlike the infinite transform (8) this finite sum is now limited by the relation $(N+1)^2 \leq L$. The result of expanding the spherical harmonic spectrum γ_{nm} using harmonics up to order N yields the value of the band-limited function $g(\boldsymbol{\theta})$ at the measurement point.

$$g(\boldsymbol{\theta}) = \sum_{n=0}^{N} \sum_{m=-n}^{n} Y_n^m(\boldsymbol{\theta}) \gamma_{nm} \tag{32}$$

To extend this expansion to multiple points a vector $\boldsymbol{g}_L$ holding all L measurement values is defined:

$$\boldsymbol{g}_L = \begin{pmatrix} g_0(\boldsymbol{\theta_0}) \\ g_1(\boldsymbol{\theta_1}) \\ . \\ . \\ g_L(\boldsymbol{\theta_L}) \end{pmatrix}$$

Accordingly, a vector $\boldsymbol{\gamma}_N$ with spherical harmonics coefficients for all orders N and modes $m = -n$ to n is given as

$$\boldsymbol{\gamma}_N = \begin{pmatrix} \gamma_{00} \\ \gamma_{1-1} \\ \gamma_{10} \\ \gamma_{11} \\ . \\ . \\ \gamma_{NM} \end{pmatrix} \quad \text{, holding } (N+1)^2 \text{ entries}$$

A matrix $\boldsymbol{Y}_N$ consisting of L rows holds values of spherical harmonics evaluated at positions $(\boldsymbol{\theta_L})$. All orders N and their modes $m = -n$ to n are represented within the matrix dimensions $L \, \mathrm{x} \, (N+1)^2$

$$\boldsymbol{Y}_N = \begin{pmatrix} Y_0^0(\boldsymbol{\theta_0}) & Y_1^{-1}(\boldsymbol{\theta_0}) Y_1^0(\boldsymbol{\theta_0}) Y_1^1(\boldsymbol{\theta_0}) & ... & Y_M^N(\boldsymbol{\theta_0}) \\ Y_0^0(\boldsymbol{\theta_1}) & Y_1^{-1}(\boldsymbol{\theta_1}) Y_1^0(\boldsymbol{\theta_1}) Y_1^1(\boldsymbol{\theta_1}) & ... & Y_M^N(\boldsymbol{\theta_1}) \\ & ... & & \\ & ... & & \\ Y_0^0(\boldsymbol{\theta_L}) & Y_1^{-1}(\boldsymbol{\theta_L}) Y_1^0(\boldsymbol{\theta_L}) Y_1^1(\boldsymbol{\theta_L}) & ... & Y_M^N(\boldsymbol{\theta_L}) \end{pmatrix} \tag{33}$$

6.1.1 Discrete inverse spherical harmonic transform

The finite sum as given in (32) can now be rewritten for L multiple sensor points as the inner product of matrix $\boldsymbol{Y}_N$ and vector $\boldsymbol{\gamma}_N$ [Zot09b]:

$$DISHT_N\{\boldsymbol{\gamma}_N\} = \boldsymbol{Y}_N \, \boldsymbol{\gamma}_N = \boldsymbol{g}_L \tag{34}$$

6.1.2 Discrete spherical harmonic transform

The discrete spherical harmonic transform ($DSHT$) requires the inversion of the matrix $\boldsymbol{Y}_N$:

$$DSHT_N\{\boldsymbol{g}_L\} = \boldsymbol{\gamma}_N = \mathbf{Y}_N^{-1}\,\boldsymbol{g}_L \tag{35}$$

The order N of the transform determines the number $(N+1)^2$ of spherical harmonics coefficients in the resulting vector $\boldsymbol{\gamma}_N$. Since the number of rows in the matrix $\boldsymbol{Y}_N$ is the amount of microphones L, only configurations with $L = (N+1)^2$ will allow for $\boldsymbol{Y}_N$ to be a square matrix and invertible. For non-square matrices, a pseudo-inverse has to be used, giving only an approximate result. A value indicating the accuracy of this approximation is the condition number of the matrix to be inverted.

6.1.3 Angular band limitation

Speaking of finite order spherical harmonics, the completeness property of the transform is no longer valid. The expansion of an infinite harmonic spectrum $Y_n^m(\boldsymbol{\theta_s})$ yields a spatial impulse for equal angles [Zot09a]:

$$\sum_{n=0}^{\infty}\sum_{m=-n}^{n} Y_n^m(\boldsymbol{\theta_s})\,Y_n^m(\boldsymbol{\theta}) = \delta\{\boldsymbol{\theta_s} - \boldsymbol{\theta}\} \tag{36}$$

For a finite number of harmonics this sum can be rewritten as inner product of two vectors, resulting in a band-limited angular impulse. This band limitation is represented by a function $\mathcal{B}_N$ without further definition.

$$\boldsymbol{y}_N(\boldsymbol{\theta}_s)\boldsymbol{y}_N(\boldsymbol{\theta}) = \mathcal{B}_N\{\delta(\boldsymbol{\theta}_s - \boldsymbol{\theta})\} \tag{37}$$

6.1.4 Finite order energy

Regarding the energy of a spherical harmonic spectrum, the surface integral over the band-limited function corresponds to a finite sum of its coefficients, which can be expressed by a vector norm:

$$\begin{aligned}
&\int_{\mathbb{S}^2} |\mathcal{B}_N\{g(\boldsymbol{\theta})\}|^2\, d\boldsymbol{\theta} \\
&= \sum_{n=0}^{N}\sum_{m=-n}^{n} |\gamma_{nm}|^2 \\
&= ||\boldsymbol{\gamma}_N||^2
\end{aligned}$$

6.2 Limited Order Holography Filters

The holography filter deriving the holographic spectrum $\boldsymbol{b}_N$ from the spectrum of the microphone signals as derived in section 5 can be given in finite resolution matrix notation as well:

$$\boldsymbol{b}_N = \boldsymbol{H}_N^{-1} \boldsymbol{\chi}_N \tag{38}$$

The filter matrix $\boldsymbol{H}_N$ consists of diagonal elements $h_n(k)$, which must not to be confused with the spherical Hankel functions $h_n^{(1)}$ and $h_n^{(2)}$. Since the same value $h_n(k)$ is repeated for the associated modes m in the diagonal, $\boldsymbol{H}_N$ has dimensions $(N+1)^2 \,\mathrm{x}\, (N+1)^2$ and is of the following structure:

$$\boldsymbol{H}_N = \begin{pmatrix} h_0(k) & 0 & \dots & 0 \\ 0 & h_1(k) & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & h_N(k) \end{pmatrix} \tag{39}$$

The values of the elements $h_n(k)$ depend on the holography filter for the respective array configuration as already shown in section 5.2. For an array of cardioid microphones around a rigid sphere these elements are (29):

$$h_n(k) = \left[j_n(kr_d) - \mathrm{i} j_n'(kr_d) + \left(\mathrm{i} h_n'^{(2)}(kr_d) - h_n^{(2)}(kr_d) \right) \frac{j_n'(kr_k)}{h_n'^{(2)}(kr_k)} \right] \tag{40}$$

6.3 Discrete Radial Filters

The source amplitude spectrum $\boldsymbol{\phi}_N$ itself can be derived from the coefficients $\boldsymbol{b}_N$ with further division by a term dependent on the source radius, the radial filter. This division, shown for spherical waves in (30), can be expressed as the inverse of the square matrix $\boldsymbol{P}_N$:

$$\boldsymbol{\phi}_N = \boldsymbol{P}_N^{-1} \boldsymbol{b}_N \tag{41}$$

For plane waves arriving from direction $\boldsymbol{\theta_s}$, the elements of $\boldsymbol{b}_N$ were given as (13)

$$b_{nm} = 4\pi \mathrm{i}^n Y_n^m(\boldsymbol{\theta_s})$$

and lead to a $\boldsymbol{P}_N$ having the structure:

$$\boldsymbol{P}_N = \begin{pmatrix} 4\pi & 0 & \dots & 0 \\ 0 & 4\pi\mathrm{i} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & 4\pi\mathrm{i}^N \end{pmatrix} \tag{42}$$

7 Sensor Layout on a Sphere

Sampling a distribution requires a dense and uniform arrangement of sensors on a spherical surface. The same is true for loudspeaker arrays consisting of many individual drivers. For numbers of up to 20 such elements, the platonic solids with their regular structure are ideal choices. With microphone capsules considered small points, they are best arranged on the corners of vertices. Compact spherical loudspeaker arrays generally take a somewhat different approach: Their emitted energy and frequency range is relative to the size of the driver employed, and the areas or faces of platonic solids are used to mount loudspeakers. A survey of isotropic radiation capabilities for the five platonic solids has been conducted by [Tar74]. Individual control of loudspeaker elements in radiation pattern synthesis has been explored by [WDC97] amongst others. A spherical loudspeaker array with 120 elements constructed in [AFKW06] is based on an icosahedron with each of its 20 triangular faces packed with six elements. Further theory and hardware regarding spherical loudspeaker arrays has been developed by [ZH07]. Different design strategies for loudspeaker layouts in periphonic sound spatialisation have been suggested by [Hol06]. As an alternative approach, a truncated icosahedron offering 12 pentagons and 20 hexagons as faces was used in [ME02] to mount a microphone capsule on the center of each face. An optimization demanding orthogonality is used in the T-designs introduced by [HS96], and explored for audio applications in [Li05] as well as in [Pet04], where an implementation with 64 microphones is presented. A sampling scheme on a Lebedev grid was used in [S+07] with a relocatable single microphone element. A layout optimized for square matrices and invertability has been discussed in [Zot09b] and employed in a 64 element enclosing array [Hoh09]. The spherical microphone array tested in the following section 11 is of the same layout as the 120 element icosahedral loudspeaker array mentioned above, which will be abbreviated “m120”. Its sensor positions are shown in figure 11.

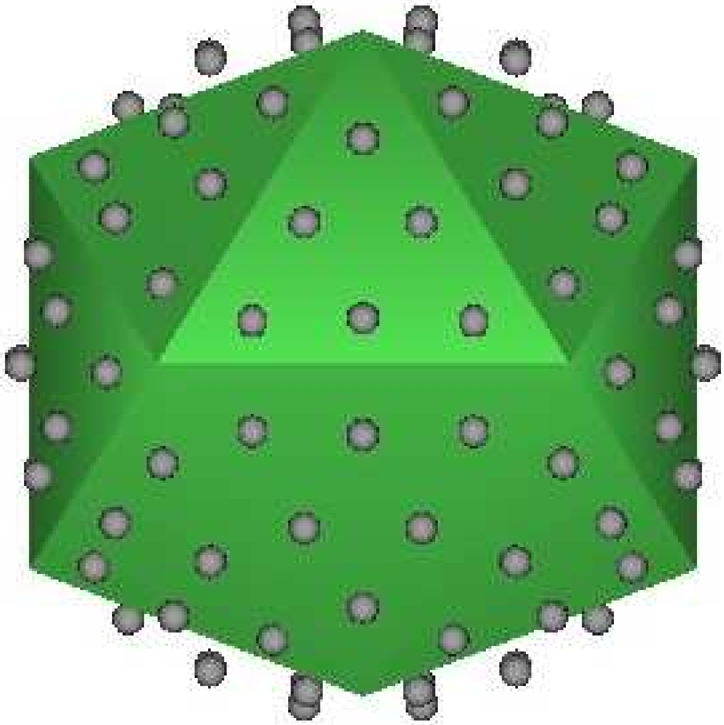

Figure 11: The “m120” microphone layout around an icosahedron

8 Finite Resolution Sampling and its Effects

Two important errors arise in the discrete spherical harmonic transform (DSHT) with its finite number of sampling points and therefore limited angular resolution: Narrow sources, or components thereof, are not included in the resulting spectrum which results in a truncation error as derived below, or shown for Ambisonic loudspeaker systems by [WA01].
Lower order spherical harmonic decomposition of high angular bandwidth distributions inevitably introduces spatial aliasing. Narrow components are mirrored into lower harmonics causing an aliasing error [Raf05].
A more universal measure as extension to the aliasing error is formulated in this work as a holographic error.

The introduction of these error measures is based on the analytic formulation of a known incident wave which is subjected to spatial sampling and discrete spherical harmonic transform. This reconstructed wave is then compared to the original. As part of this simulation positioning errors and gain deviations can be evaluated, as shown in the subsequent section 9. The analytic incident wave is synthesized using spherical harmonics expansion (8) of a plane wave in accordance with the far field condition $kr_s < 1$.

8.1 Truncation Error

The truncation error designates the discarded part of a spherical harmonic spectrum derived from a transformation of finite order N. No aliasing effects are taken into account. For sound pressure or velocity spectra the energy difference between the actual source and the decomposition result is determined. The normalized truncation error $\tau_N(kr_d)$ for plane waves is given in [WA01] as:

$$\tau_N(kr_d) = 1 - \sum_{n=0}^{N} (2n+1)|j_n(kr_d)|^2$$

It is clearly dependent on the decomposition order N, the frequency k and the microphone radius r_d. Figure 12 shows the truncation error in dB versus frequency for different orders N and microphone radius $r_d = 69\text{mm}$.
There is still a need for listening tests and evaluation data in order to find the biggest acceptable value of this error with regard to listener perception and source localization. Since truncated spectra merely lack angular detail but do not contain false directional information, the impact of truncation may be valued smaller than

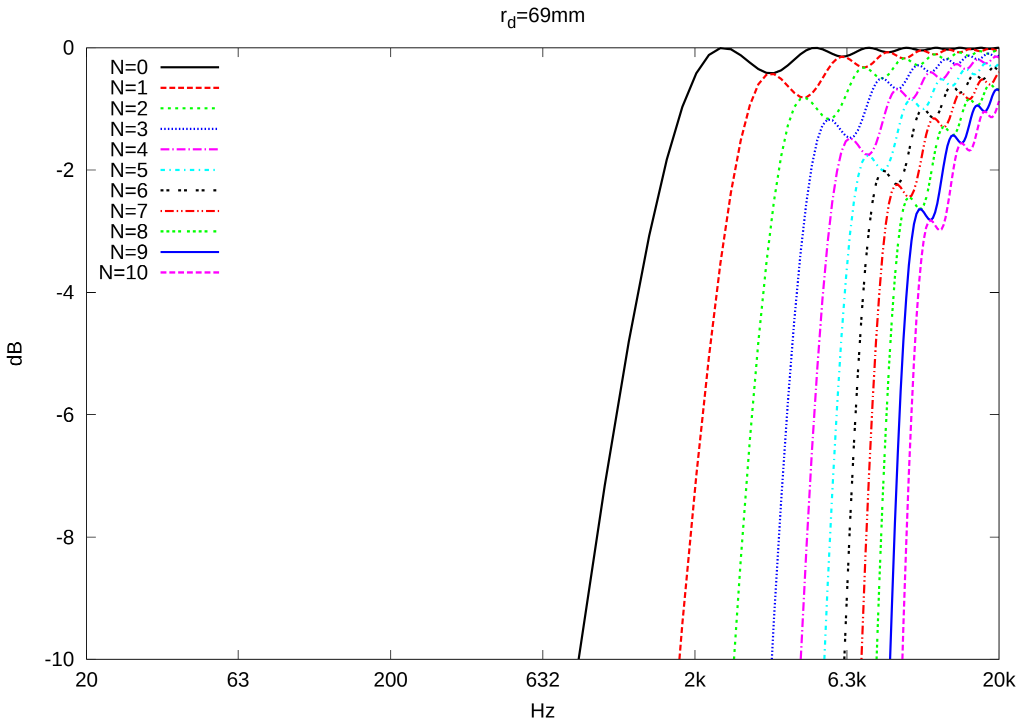

Figure 12: Truncation error $\tau_N(kr_d)$ for different orders N and capsule radius $r_d =$ 69mm

that of aliasing.

8.2 Aliasing as Matrix Product

In the survey of spatial aliasing the microphone signal vector $\boldsymbol{x}$ of length L is synthesized by spherical harmonics expansion. This requires the use of finite order harmonics and matrix notation. It is impossible to formulate a spherical harmonics matrix $\boldsymbol{Y}_\infty$ (33) of infinite dimensions. An approximation of infinite spherical harmonics expansion is achieved by choosing the dimensions exp of a matrix $\boldsymbol{Y}_{exp}$ high enough. The spherical harmonic spectrum $\boldsymbol{\chi}_{exp}$ of the microphone signals has the same high resolution and is derived from its continuous version, equation (26). This spectrum is expanded into the microphone signals $\boldsymbol{x}$.

$$\boldsymbol{x} = \boldsymbol{Y}_{exp}\boldsymbol{\chi}_{exp} \tag{43}$$

The order exp should be chosen as high as possible. A rule of thumb taking frequency and microphone radius into account is given in [WA01] as $exp > kr_d$, where exp is rounded to the next largest integer. This relation is shown in figure 13 for a microphone capsule radius of 70 mm.

The distribution $\boldsymbol{\chi}_{exp}$ is sampled at the L microphone positions, and transformed

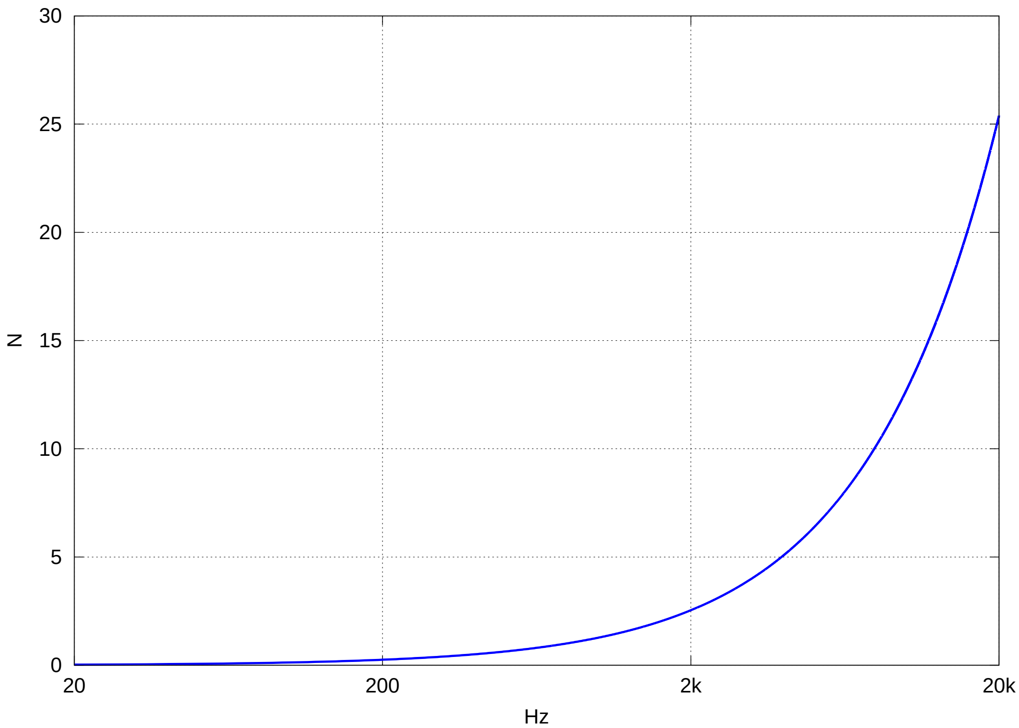

Figure 13: Recommended expansion order exp, sensor radius $r_d = 69$mm

into a spherical harmonic spectrum by $\boldsymbol{Y}_N^{-1}$, a matrix of smaller dimensions and more limited resolution.

$$\hat{\boldsymbol{\chi}}_N = \boldsymbol{Y}_N^{-1}\boldsymbol{Y}_{exp}\boldsymbol{\chi}_{exp} \tag{44}$$

The result is the lower resolution spectrum $\hat{\boldsymbol{\chi}}_N$, which contains coefficients distorted by high harmonics mirrored into the lower ones.

In analogy to aliasing in discrete time domain signal processing, this description of the sampling mechanism itself does not yet give information about the amount of aliasing appearing at the output. For time series sampling this amount is dependent on the input signal frequency. For spatial sampling this translates to high angular bandwidth at the input. The position and narrowness of the source at hand determine how many of the aliased coefficients are excited how much, and are therefore present in the output.

If the order of both transformation matrices was $N = exp$, the result would be an identity matrix due to the orthogonality of spherical harmonics. In this special case the sampled spectrum would be identical to the analytic spectrum. The accuracy of the retrieved coefficients is clearly dependent on the matrix product $\boldsymbol{Y}_N^{-1}\boldsymbol{Y}_{exp}$.

A sampling matrix of order $N < exp$ causes aliasing in the right-hand rows of the matrix product. A possible structure of spatial aliasing in the product $\boldsymbol{Y}_N^{-1}\boldsymbol{Y}_{exp}$ is

Figure 14: Spatial aliasing in columns representing orders $> N$

shown in figure 14. The columns right of the diagonal represent orders larger than N. Incident waves with higher angular bandwidth excite these columns, which are reflected into the coefficients $\hat{\boldsymbol{\chi}}_{\boldsymbol{N}}$ as aliased signals.

8.3 Condition Number in Matrix Inversion

In addition to the aliasing introduced by finite spherical harmonic transforms, the inverse $\boldsymbol{Y}_N^{-1}$ of any non-square matrix is inexact due to numerical approximation techniques. The matrix condition number c gives a measure for the precision of this result. It is defined as the ratio of the largest to smallest singular value of the matrix [Wei09]. The condition number for the inverse of a matrix combining $\boldsymbol{H}_N$ and $\boldsymbol{Y}_N$ has been shown to influence the orthogonality of the transform in [Raf08]. For the "m120" layout, the values for the condition number c of the matrix $\boldsymbol{Y}_N$ with regard to the order N are:

N	c
1	1
2	1.0001
3	1.0407
4	1.0496
5	1.0778
6	1.1882
7	1.1914
8	1.2349
9	3.2901

8.4 Aliasing Error

It is desirable to derive a more tangible measure for spatial aliasing. An error vector $\boldsymbol{\varepsilon}_N$ in the spherical harmonics domain can be defined as the difference between the sampled coefficients $\hat{\boldsymbol{\chi}}_N$ including aliasing, and the analytically derived clean and band-limited coefficients $\boldsymbol{\chi}_N$:

$$\boldsymbol{\varepsilon}_N = \hat{\boldsymbol{\chi}}_N - \boldsymbol{\chi}_N \tag{45}$$

The aliased coefficients $\hat{\boldsymbol{\chi}}_N$ are the result of a discrete spherical harmonic transform of the microphone signals as already introduced in equation (35).

$$\hat{\boldsymbol{\chi}}_N = \boldsymbol{Y}_N^{-1}\boldsymbol{x}$$

The microphone signals $\boldsymbol{x}$ on the right hand side of the equation are analytically generated using spherical harmonics expansion of the spectrum $\boldsymbol{\chi}_{exp}$ at high order exp as shown in (43).

$$\hat{\boldsymbol{\chi}}_N = \boldsymbol{Y}_N^{-1}\boldsymbol{Y}_{exp}\boldsymbol{\chi}_{exp}$$

In accordance with holography filters, as introduced in section 5.2, the analytic coefficients $\boldsymbol{\chi}_{exp}$ resembling the microphone signals can be rewritten as the product of a holography filter matrix $\boldsymbol{H}_{exp}$ and the holographic spectrum $\boldsymbol{b}_{exp}$. The elements of $\boldsymbol{b}_{exp}$ can be given for spherical (13) or plane waves (16).

$$\hat{\boldsymbol{\chi}}_N = \boldsymbol{Y}_N^{-1}\boldsymbol{Y}_{exp}\boldsymbol{H}_{exp}\boldsymbol{b}_{exp} \tag{46}$$

In the same fashion $\boldsymbol{\chi}_N$, the analytic version of the coefficients, is written as the product of filter matrix $\boldsymbol{H}_N$ and the vector $\boldsymbol{b}_N$:

$$\boldsymbol{\chi}_N = \boldsymbol{H}_N\boldsymbol{b}_N$$

At this point the aliasing vector for any incident sound field is expressed as the difference of analytic and aliased coefficients:

$$\boldsymbol{\varepsilon}_{N,exp} = \boldsymbol{Y}_N^{-1}\boldsymbol{Y}_{exp}\boldsymbol{H}_{exp}\boldsymbol{b}_{exp} - \boldsymbol{H}_N\boldsymbol{b}_N \tag{47}$$

This error spectrum is dependent on the sampling order N, on the maximum angular bandwidth exp allowed, on the frequency k of the filter matrices $\boldsymbol{H}$, on the microphone positions intrinsic to matrices $\boldsymbol{Y}_N$ and $\boldsymbol{Y}_{exp}$, and on the radial and angular position of the incident wave $\boldsymbol{b}_N$.
The holographic spectrum $\boldsymbol{b}_N$ can be separated into two parts, the radial filter matrix $\boldsymbol{P}_N$ and the spherical harmonics vector $\boldsymbol{y}_N(\boldsymbol{\theta_s})$ dependent on the source angle.

$$\boldsymbol{b}_N = \boldsymbol{P}_N\boldsymbol{y}_N(\boldsymbol{\theta_s}) \tag{48}$$

A simplified aliasing error spectrum for plane waves can thus be rewritten as:

$$\boldsymbol{\varepsilon}_{N,exp} = \boldsymbol{Y}_N^{-1}\boldsymbol{Y}_{exp}\boldsymbol{H}_{exp}\boldsymbol{P}_{exp}\boldsymbol{y}_{exp}(\boldsymbol{\theta_s}) - \boldsymbol{H}_N\boldsymbol{P}_N\boldsymbol{y}_N(\boldsymbol{\theta_s}) \tag{49}$$

By concatenating an $(N\mathrm{x}N)$ identity matrix with $(exp - N)$ rows of zeros, $\boldsymbol{y}_N(\boldsymbol{\theta_s})$ can be rewritten as a truncated version of $\boldsymbol{y}_{exp}(\boldsymbol{\theta_s})$

$$\boldsymbol{y}_N(\boldsymbol{\theta_s}) = [\boldsymbol{I}\,\boldsymbol{O}]\boldsymbol{y}_{exp}(\boldsymbol{\theta_s})$$

The aliasing error is now

$$\boldsymbol{\varepsilon}_{N,exp} = \left[\boldsymbol{Y}_N^{-1}\boldsymbol{Y}_{exp}\boldsymbol{H}_{exp}\boldsymbol{P}_{exp} - \boldsymbol{H}_N\boldsymbol{P}_N[\boldsymbol{I}\,\boldsymbol{O}]\right]\boldsymbol{y}_{exp}(\boldsymbol{\theta_s}) \tag{50}$$

This expression of the aliasing error describes the deviation between the real and sampled distribution on the surface of the array as defined in (45) above.

8.5 Holographic Error

The accuracy in depicting the original sound source demands an extension to the aliasing error: The error as presented in (50) is divided by the holography and radial filter matrices $\boldsymbol{H}_N^{-1}\boldsymbol{P}_N^{-1}$ to derive a holographic error $\boldsymbol{\sigma}_{N,exp}$. This error now denotes the difference between the actual sound source itself and its aliased replica from holography.

$$\boldsymbol{\sigma}_{N,exp} = \boldsymbol{P}_N^{-1}\boldsymbol{H}_N^{-1}\boldsymbol{\varepsilon}_{\boldsymbol{N,exp}} = \left[\boldsymbol{P}_N^{-1}\boldsymbol{H}_N^{-1}\boldsymbol{Y}_N^{-1}\boldsymbol{Y}_{exp}\boldsymbol{H}_{exp}\boldsymbol{P}_{exp} - [\boldsymbol{I}\,\boldsymbol{O}]\right]\boldsymbol{y}_{exp}(\boldsymbol{\theta_s}) \tag{51}$$

A scalar measure of the holographic error for all orders N can be expressed by a vector norm. Below, the trace

$$||\boldsymbol{\sigma}||^2 = Tr\{\boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathbf{T}}\} \tag{52}$$

will be used instead, to allow for an elegant simplification: As the result of the hermitian transposition, $\boldsymbol{y}_{exp}(\boldsymbol{\theta_s})\boldsymbol{y}_{exp}^{\mathrm{T}}(\boldsymbol{\theta_s})$ can be contracted. Preferably the scalar error measure includes all possible source positions $(\boldsymbol{\theta_s})$, which can be expressed by a surface integral. In this integral the orthonormality property (5) reduces this contracted term to an identity matrix.

$$\int_{\mathbb{S}^2} \boldsymbol{y}_{exp}(\boldsymbol{\theta_s})\boldsymbol{y}_{exp}^{\mathrm{T}}(\boldsymbol{\theta_s})\, d\boldsymbol{\theta_s} = \boldsymbol{I} \tag{53}$$

A normalization term $\frac{1}{4\pi}$ for plane waves is required to compensate for the squared absolute amplitude gained by the surface integral itself:

$$\int_{\mathbb{S}^2} d\boldsymbol{\theta_s} = 4\pi$$

The energy of the error signal is dependent on the spherical harmonics order N. Another normalization term with regard to energy is derived by calculating the squared norm of the spherical harmonics vector $||\boldsymbol{y}_N||^2$. Dividing the error by this norm introduces a normalization to the energy.
The trace introduced in (52) equals the squared Frobenius norm [Wei09]

$$Tr\{\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\} = ||\boldsymbol{A}||_F^2$$

and so the normalized scalar holographic error $||\boldsymbol{\sigma}_{N,exp}||$ is

$$||\boldsymbol{\sigma}_{N,exp}||^2 = \frac{1}{4\pi} \frac{1}{||\boldsymbol{y}_N||^2} ||\boldsymbol{P}_N^{-1} \boldsymbol{H}_N^{-1} \mathbf{Y}_N^{-1} \mathbf{Y}_{exp} \boldsymbol{H}_{exp} \boldsymbol{P}_{exp} - [\boldsymbol{I}\ \mathbf{0}]||_F^2 \tag{54}$$

This error vanishes for identical matrices $\boldsymbol{Y}_N^{-1}$ and $\boldsymbol{Y}_{exp}$, and is dependent on the sampling order N, on the frequency k, on the microphone positions itself, and on the allowed spatial bandwidth exp.

8.6 Interpretation of the Holographic Error

The formulation of a holographic error permits the comparison of different microphone arrays. The spatial sampling scheme determined by the layout of the microphones on the surface and its influence on the accuracy of the holography can be studied for every frequency k. Examination of various real-world effects can be simulated and the influence of gain and positional errors can be estimated prior to building the actual array, as shown below in section 9.

The very tempting comparison between arrays of different orders N is not valid though. The error identifies only the detected amount of aliased harmonics in the output signal and not the total amount possible. For example a spherical harmonic of order 24 could cause aliasing in another harmonic of order 12. If the microphone array was to decompose the distribution into harmonics of up to order 10, this aliased part will not get detected in the error. However, it is possible to compare layouts of identical orders, as shown in figure 15 for closed-sphere arrays using cardioid microphones arranged in either an icosahedron with six microphones on each tile, as

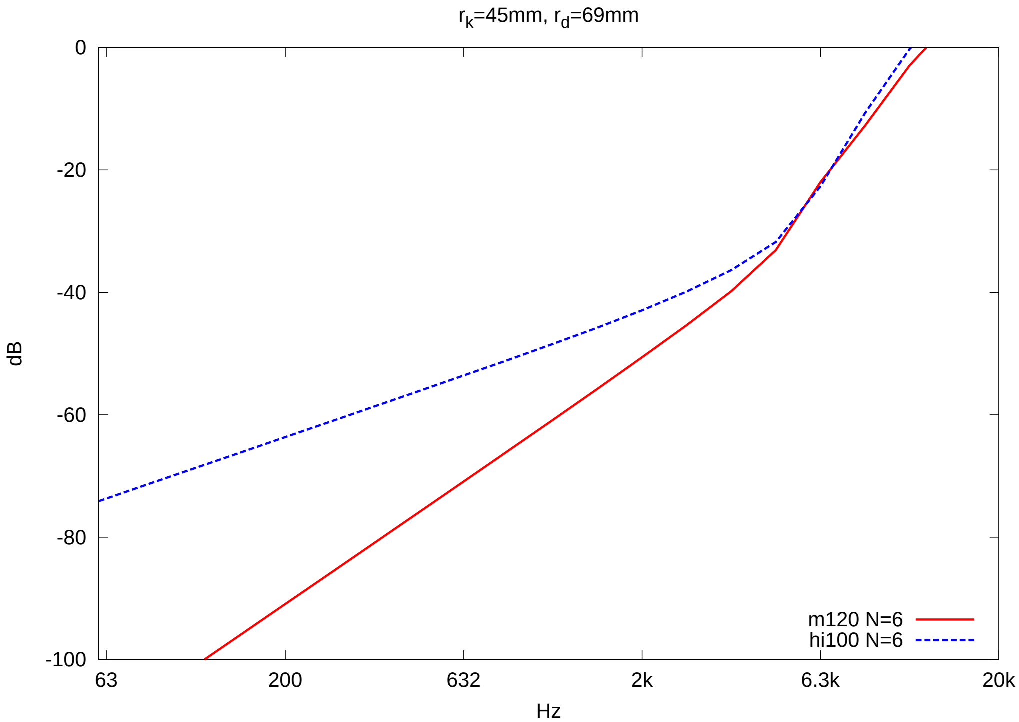

Figure 15: Holographic error $||\boldsymbol{\sigma}_{N,exp}||_{dB}$ for two different array geometries at order $N = 6$ and cardioids around a rigid sphere

shown in figure 11 and abbreviated "m120", and for a hyperinterpolation layout of 100 points, named "hi100". Both arrays are evaluated at a maximum decomposition order of $N = 6$.

The holographic error becomes zero for an exact reproduction of the incident wave, which translates to a value of $-\infty$ on a decibel scale. If the result bears no resemblance to the original wave, or if no signal is present at all, the error value of one equals 0 dB. A request for the holographic error to be smaller than a certain value will restrict the usable frequency range of both layouts. While this limit is the result of spatial aliasing for high frequencies, numerical problems and noise issues will drastically limit the array's lower frequency range as already shown in the discussion of holography filters in section 5.2. No psycho-acoustic evaluation of this error has been made so far, and a general limit for its value cannot be stated yet. It is also important to point out that the holographic error takes the entire processing chain into account, considering the influence of the holography and radial filter on the accuracy of the result.

The truncation error $\tau_N(kr_d)$ and holographic error $||\boldsymbol{\sigma}_{N,exp}||^2$ must not be combined into an overall error measure due to their different nature. The truncation error gives the deviation in energy caused by finite order sampling of a distribution on the array surface. The holographic error in turn describes the energy difference between the actual sound source and its replica due to aliasing. The effect of the holographic error may have a deeper impact on the listener, since aliasing can induce wrong spatial information, whereas truncated spectra merely lack angular resolution.

9 Array Imperfections

The error measures derived in this work are helpful in the evaluation of array imperfections inherent to an array and its hardware. The position of the microphones themselves can only be determined with a certain tolerance. As with any multichannel audio application, the similarity of the gain and transfer functions within channels is crucial. By modification of the virtual microphone signals in the error equations, different conditions and deviations are simulated.

9.1 Deviations in Actual Microphone Positions

Given the mechanical challenges when constructing spherical microphone arrays it is only possible to match the specified capsule positions with a finite degree of accuracy. In the formulation of the holographic error the capsule positions determine the values of the spherical harmonics matrix $\boldsymbol{Y}_{exp}$. Random variation of their angular arguments leads to a simulation of inaccurate capsule placement and results in a higher holographic error. Simulated deviations of up to $\pm 4°$ are shown in figure 16. The deviations have a higher impact on the low frequency performance. For big wavelengths, the phase difference between the microphones is generally very small. This leads to the high gains in the holography filters already discussed. Changes in microphone positions cause big perturbations in the phase differences for low frequencies. These effects are negligible if the holographic error is used to determine an upper frequency limit though.

9.2 Gain Mismatch

In every actual microphone array realization, imperfections in the signal path of individual channels are a major reason of concern. The effects of gain mismatch are examined by including a diagonal matrix $\boldsymbol{G}$ of random gain factors in the error equation (54).

$$||\boldsymbol{\sigma}_{N,exp}|| = \frac{1}{\sqrt{4\pi}} \frac{1}{||\boldsymbol{y}_N||} ||\boldsymbol{P}_N^{-1} \boldsymbol{H}_N^{-1} \mathbf{Y}_N^{-1} \boldsymbol{G}\, \boldsymbol{Y}_{exp} \boldsymbol{H}_{exp} \boldsymbol{P}_{exp} - [\boldsymbol{I}\, \mathbf{0}]||_F \qquad (55)$$

Figure 17 gives this holographic error for different random gain ranges. Here the influence is much more devastating. The misalignment is especially severe for low frequencies which is due to the small differences detected with closely spaced sensors at long wavelengths. The high gain of the holography filter causes the gain deviations to be amplified even more.

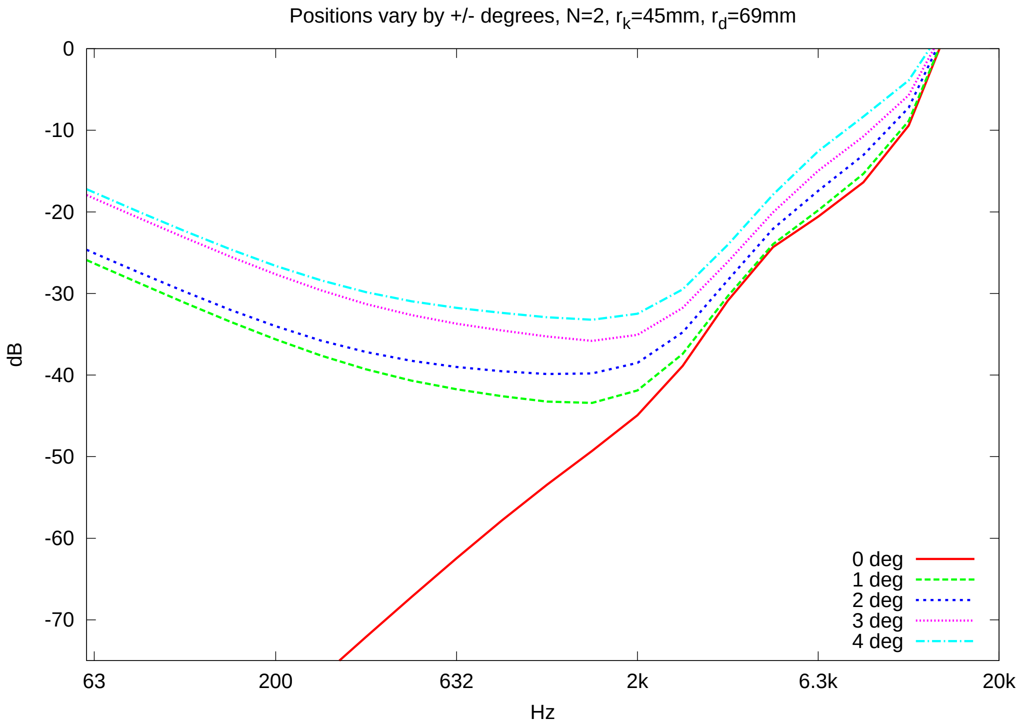

Figure 16: The effect of capsule position deviations on the holographic error for the "m120" layout at order $N = 2$ and cardioids around a rigid sphere

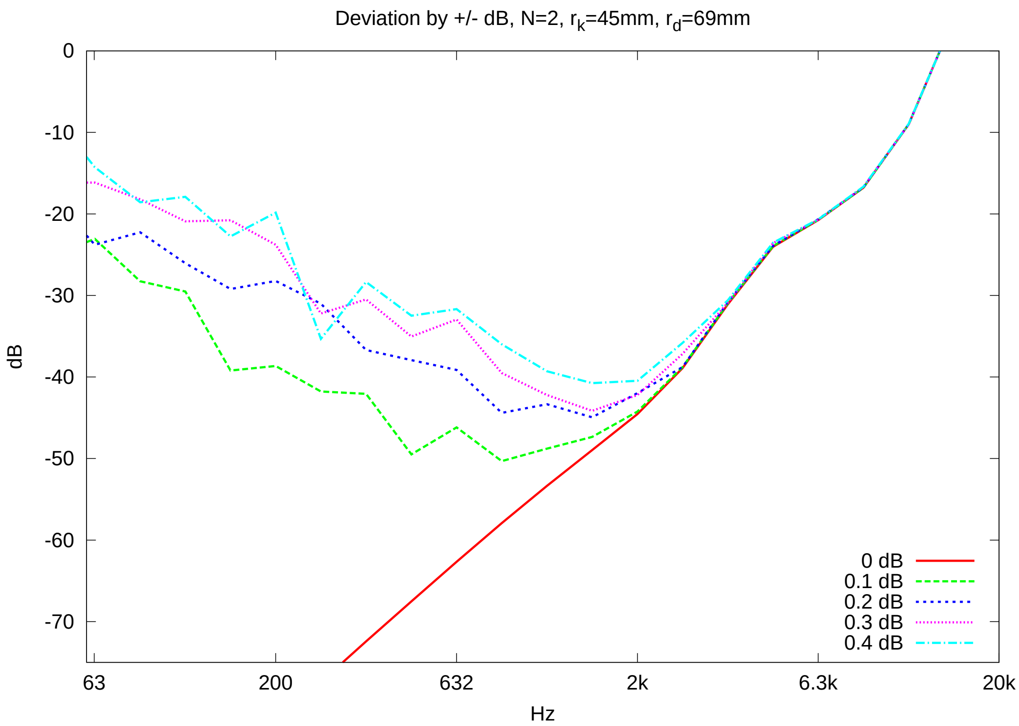

Figure 17: Holographic error considering gain mismatches for the "m120" layout at order $N = 2$ and cardioids around a rigid sphere

10 Implementation

The algorithms and transforms discussed in this thesis can be implemented using digital signal processing software on general purpose computers. An ideal numerical computation package suited for this task is GNU Octave [Oct]. Being an interpreted language it is not optimized for computation efficiency. Fast prototyping of algorithms, easy visualization of data and its open and cost-effective GPL license [Gpl07] and broad user-base make it a perfect solution. A software suite of several Octave functions and scripts has been compiled. The decomposition into spherical harmonics and radial filtering alongside scripted soundfile operations were implemented.[3] Figure 18 gives an overview of the steps required to listen to a holographic representation of a recorded sound field as presented in the last sections. The entire process can be separated into two parts: The decomposition and filtering on one side, and the holophonic reproduction for example via loudspeaker arrays or beamforming on the other side. The stage in the processing chain at which the audio data can be stored to memory is variable. The transform into the spherical harmonics domain is done according to matrix $\boldsymbol{Y}_N^{-1}$ whose elements are determined by the angular position of the microphones. The open or rigid structure of the array and the type of microphones influence the filter matrix $\boldsymbol{H}_N$. The source radius determines the values of the filter matrix $\boldsymbol{P}_N$, which can be simplified in a far field assumption (70).

10.1 Twofold Transform and Block Filters

The filter matrices $\boldsymbol{H}_N$ and $\boldsymbol{P}_N$ were already defined as diagonal matrices consisting of entries dependent on the spherical harmonic order and frequency. The filter equations give a spectrum in real and imaginary values which can be used as scaling factors for the Fourier transformed audio signals by complex multiplication. In the implementation presented here this complex multiplication of spectra was employed in a block filter approach. The block filter matrices $\boldsymbol{H}_N^\dagger$ and $\boldsymbol{P}_N^\dagger$ have a slightly different form, holding the corresponding values already inverted. The block filter technique has several drawbacks which are less prominent for long DFT sizes. More advanced filter design solutions such as the bilinear transform and the impulse invariance method are discussed in [Pom08].
The discrete time domain input samples $x_L(t)$ from the L microphone channels can be combined into a L x N_{samps} matrix $\boldsymbol{X}$.

[3] More information can be found on the author's website. See: http://plessas.mur.at

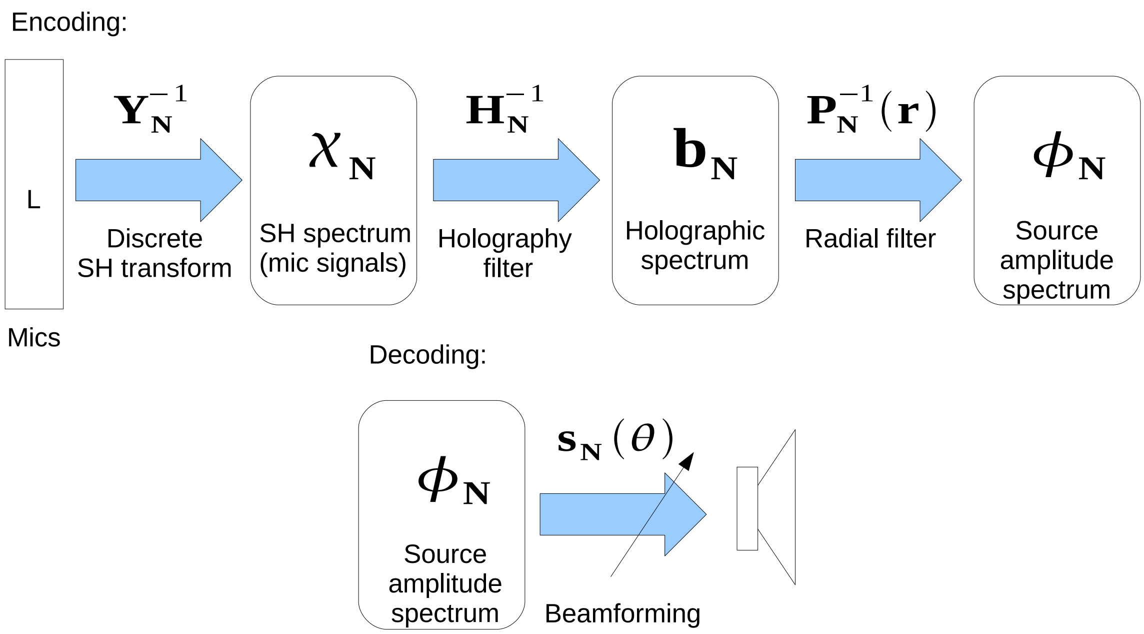

Figure 18: Processing structure for spherical harmonic transform, holography and radial filtering with optional beamforming. The spectrum $\boldsymbol{\chi}_N$ resembles the distribution at the array radius. The filter $\boldsymbol{H}_N$ returns the holographic spectrum $\boldsymbol{b}_N$. The radial filter matrix $\boldsymbol{P}_N$ then yields the source amplitude spectrum $\boldsymbol{\phi}_N$ at an outer radius. Beam forming with a steering vector $\boldsymbol{s}_N$ allows to selectively listen to the source amplitudes at an angular position on this outer radius.

$$\boldsymbol{X} = \begin{pmatrix} x_1(t_0) & \dots & x_1(N_{samps}) \\ \vdots & \ddots & \vdots \\ x_L(t_0) & \dots & x_L(N_{samps}) \end{pmatrix} \tag{56}$$

This matrix can be transformed into the frequency domain by a discrete Fourier transform (DFT), the result being a matrix with dimensions L x N_{DFT}:

$$DFT\{\boldsymbol{X}\} = \boldsymbol{X}_{DFT} = \begin{pmatrix} x_1(\omega_0) & \dots & x_1(N_{DFT}) \\ \vdots & \ddots & \vdots \\ x_L(\omega_0) & \dots & x_L(N_{DFT}) \end{pmatrix} \tag{57}$$

The discrete spherical harmonics matrix $\boldsymbol{Y}_N$ consists of elements determined by the microphone positions $\boldsymbol{\theta_L}$ and has a layout already described in equation (33), with dimensions L x $(N+1)^2$.

$$\boldsymbol{Y}_N = \begin{pmatrix} Y_0^0(\boldsymbol{\theta_0}) & Y_1^{-1}(\boldsymbol{\theta_0})Y_1^0(\boldsymbol{\theta_0})Y_1^1(\boldsymbol{\theta_0}) & ... & Y_M^N(\boldsymbol{\theta_0}) \\ Y_0^0(\boldsymbol{\theta_1}) & Y_1^{-1}(\boldsymbol{\theta_1})Y_1^0(\boldsymbol{\theta_1})Y_1^1(\boldsymbol{\theta_1}) & ... & Y_M^N(\boldsymbol{\theta_1}) \\ & ... & & \\ & ... & & \\ Y_0^0(\boldsymbol{\theta_L}) & Y_1^{-1}(\boldsymbol{\theta_L})Y_1^0(\boldsymbol{\theta_L})Y_1^1(\boldsymbol{\theta_L}) & ... & Y_M^N(\boldsymbol{\theta_L}) \end{pmatrix}$$

The pseudoinverse $\boldsymbol{Y}_N^{-1}$ of this matrix has the dimensions $(N+1)^2$ x L and is used in the discrete spherical harmonic transform ($DSHT$) to obtain $\boldsymbol{\chi}_{DFT}$, which has dimensions $(N+1)^2$ x N_{DFT} and holds the spherical wave spectrum of the Fourier transformed microphone signals.

$$\boldsymbol{\chi}_{DFT} = \boldsymbol{Y}_N^{-1}\boldsymbol{X}_{DFT}$$

This twofold transform provides a matrix layout which permits filtering using element wise multiplication. The required filter matrices hold a different coefficient for every frequency ω and every spherical harmonic. The holography filter matrix $\boldsymbol{H}_N$ got defined in section 6.2. Its block filter variant $\boldsymbol{H}_N^\dagger$ has dimensions $(N+1)^2$ x N_{DFT} and consists of column vectors with already inverted coefficients for every frequency ω.

$$\boldsymbol{H}_N^\dagger = \begin{pmatrix} h_0^{-1}(\omega_0) & h_0^{-1}(\omega_1) & \dots & h_0^{-1}(N_{DFT}) \\ h_1^{-1}(\omega_0) & \ddots & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ h_N^{-1}(\omega_0) & \dots & \dots & h_N^{-1}(N_{DFT}) \end{pmatrix} \tag{58}$$

The radial filter matrix $\boldsymbol{P}$ was already introduced in equation (42), has constant values for all frequencies and holds copies of an identical column vector, having dimensions $(N+1)^2$ x N_{DFT}. If plane waves are assumed, its block filter variant $\boldsymbol{P}_N^\dagger$ holding already inverted values has the shape

$$\boldsymbol{P}_N^\dagger = \begin{pmatrix} (4\pi)^{-1} & (4\pi)^{-1} & \dots & (4\pi)^{-1} \\ (4\pi\mathrm{i})^{-1} & (4\pi\mathrm{i})^{-1} & \dots & (4\pi\mathrm{i})^{-1} \\ \vdots & \vdots & \vdots & \vdots \\ (4\pi\mathrm{i}^N)^{-1} & (4\pi\mathrm{i}^N)^{-1} & \dots & (4\pi\mathrm{i}^N)^{-1} \end{pmatrix} \tag{59}$$

The entire processing chain using discrete Fourier transform and block filtering, with $\cdot$ denoting element wise multiplication, can be implemented as

$$\boldsymbol{\Phi}_{DFT} = \boldsymbol{P}_N^\dagger \cdot \boldsymbol{H}_N^\dagger \cdot \boldsymbol{\chi}_{DFT} \tag{60}$$

$$\boldsymbol{\Phi}_{DFT} = \boldsymbol{P}_N^\dagger \cdot \boldsymbol{H}_N^\dagger \cdot \boldsymbol{Y}_N^{-1} \boldsymbol{X}_{DFT} \tag{61}$$

The matrix $\boldsymbol{\Phi}_{DFT}$ can be transformed back into a time domain matrix $\boldsymbol{\Phi}$ of dimensions $(N+1)^2$ x N_{samps} holding the angular source amplitudes at the outer radius. This processing has been implemented using GNU Octave for short sample lengths captured in impulse response measurements, which are discussed in the following section 11.

10.2 Beamforming

In order to listen to the results of acoustic holography a holophonic loudspeaker layout such as Ambisonics can be used, sampling the spectrum $\boldsymbol{\Phi}_N$ at discrete points and reproducing it with loudspeakers. Another approach is to sample the spectrum at a single point only and change the angular position $\boldsymbol{\theta}_s$ of this point. This resembles a steerable beam, implemented by multiplication of $\boldsymbol{\Phi}_N$ with a static spherical harmonics vector $\boldsymbol{s}_N(\boldsymbol{\theta})$. This steering vector consists of all harmonics evaluated at the steering angle. The inner product resembles a weighted sum of spherical harmonics and results in a beam that can be freely positioned. It returns the playback signal, a time domain signal vector $\boldsymbol{l}(t)$:

$$\boldsymbol{l}(t) = \boldsymbol{\Phi}_N \, \boldsymbol{s}_N(\boldsymbol{\theta}) \tag{62}$$

For infinite spherical harmonics this beam would be a narrow impulse as shown in the orthogonality property (5). Limited order beams are an inner product of two spherical harmonics spectra and result in a band-limited angular impulse. These

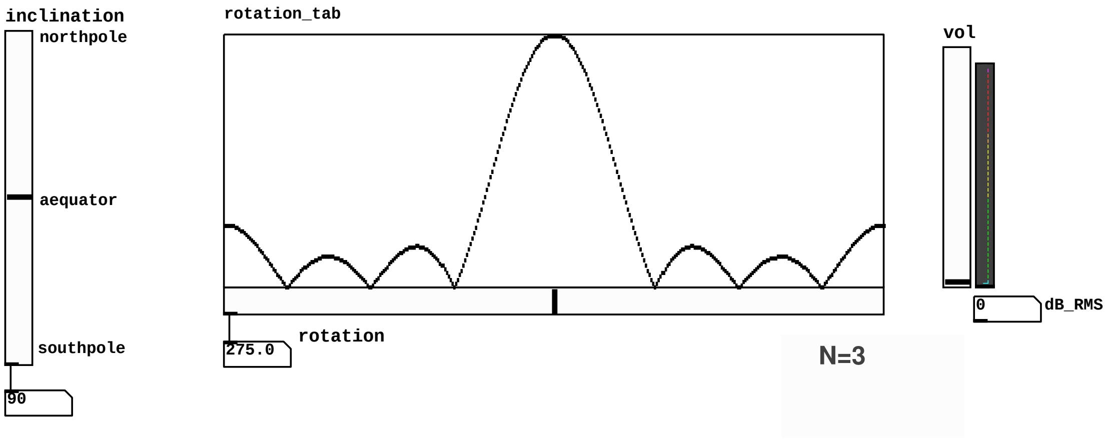

Figure 19: Implementation of a beam steering scenario of order $N = 3$ using Pure Data. Controls for the inclination and rotation angles of the beam $\boldsymbol{s}_N(\boldsymbol{\theta})$ are shown as sliders. A graphical representation of the beam pattern helps to identify sidelobes.

beams have a wider main lobe and sidelobes dependent on the maximum order N. The beam pattern for orders $N = 1 - 3$ is given in figure 20.
This beamforming approach has been implemented using the programming language Pure Data [Puc97]. A screenshot of its user interface is shown in figure 19.

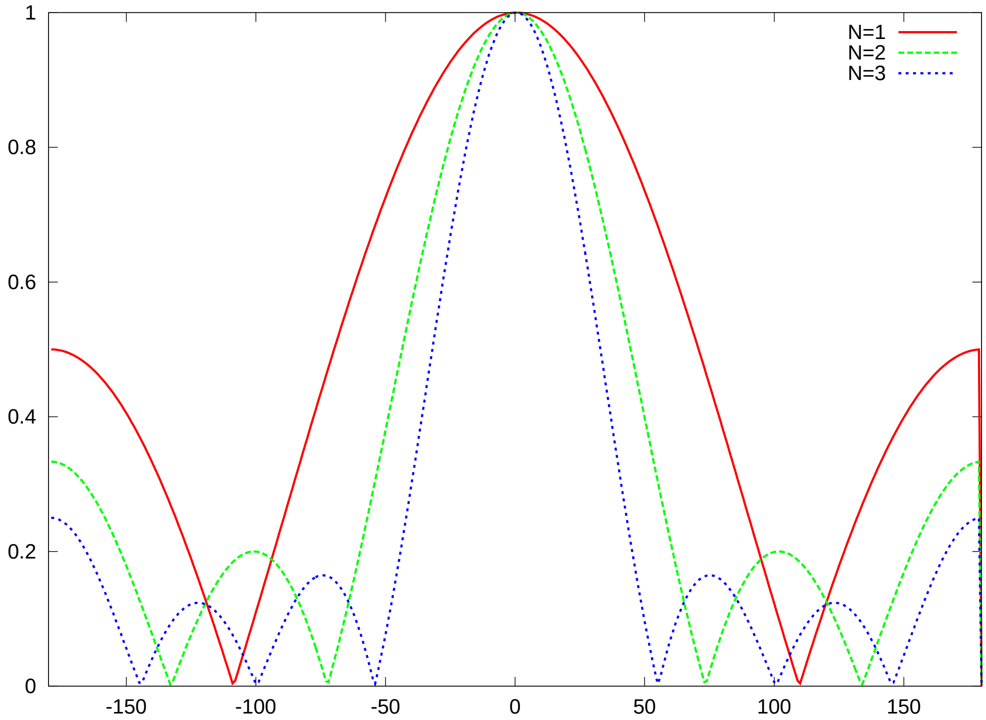

Figure 20: Beam pattern for orders up to $N = 3$ showing absolute values. Neighboring lobes have alternating signs and phases.

11 Array Hardware and Tests

One essential goal of this thesis is to explore the theory developed in the previous sections using actual hardware. Microphone arrays can be used in many different applications, for example in speech transmission, filtering and processing. The aim of this section is to verify the usability of the algorithms in music recording. This application imposes high demands on the frequency bandwidth and noise levels of the sensors.
As collaboration between the Institute of Electronic Music and Acoustics – IEM Graz Austria, the Center for New Music and Audio Technologies – CNMAT, and Meyer Sound Laboratories, both Berkeley, California, an actual array implementation has been tested. This array is of the "m120" layout already shown in figure 11 and consists of a rigid sphere with 120 cardioid microphones at a slightly larger radius. This core is complemented by four cantilevers holding 24 omnidirectional capsules at several bigger radii. The tests conducted in this thesis have focused on the cardioid "m120" core itself. The microphone signals are amplified and converted into the digital domain inside the array hardware itself. An ethernet protocol transfers multiplexed signals as UDP packets to a host computer for further processing and storage.

11.1 Impulse Response Measurements

As part of a test recording with this array, several 120-channel impulse responses have been captured. They identify the transfer functions of a system consisting of the loudspeaker, the room and the array itself. The anechoic chamber of the Hafter Auditory Perception Lab at the University of California in Berkeley provided a reflection free environment. Impulse responses were taken using the exponential swept sine technique as introduced by Farina in [Far00], played through a Meyer-sound HM-1 loudspeaker. Figure 21 gives an impression of the test setup. In the results given in the next section, the spacing between loudspeaker and microphone array is 1.29 meters.

11.2 Holographic Visualization

Evaluating the amplitude spectrum $\boldsymbol{\phi}_N$ by steering a beam along all angles and plotting the output signal magnitude, a visual and holographic representation of the entire sound field is obtained. This spectrum was derived with a plane wave model for the filter matrix $\boldsymbol{P}_N$. The horizontal axis corresponds to the rotational angle along the equator. The vertical axis denotes the elevation angle ranging up

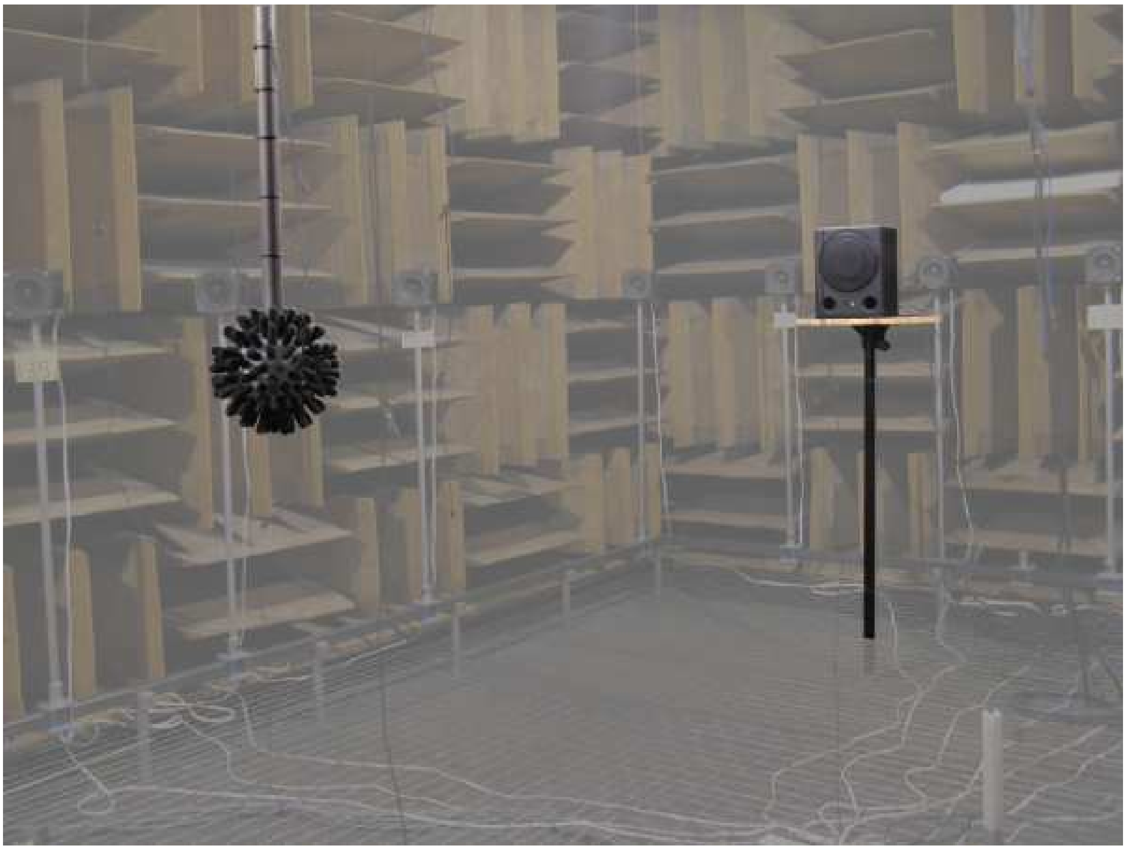

Figure 21: Test setup with array and coaxial loudspeaker in an anechoic chamber

to the north pole and down to the south pole. The color plots given in this section show the magnitude of the impulse response taken. The magnitude is evaluated for selected frequencies. The color chart given in figure 22 is used to identify the linear normalized magnitude.

Figure 22: Linear color chart ranging from 0 on the left to 1 on the right

The plot shown in figure 23 correctly identifies the loudspeaker position at rotation $270°$ and elevation $0°$. Processing with low spherical harmonic orders results in a wide main lobe and a prominent sidelobes. The mirror image observed in the left half of the plot is the result of a sidelobe. This is in accordance with the beam pattern already given in figure 20, identifying a major sidelobe at $N = 1$ and $\pm 180°$. The same frequency at increased order $N = 2$ is shown in figure 24, displaying a more narrow mainlobe and less prominent sidelobes.

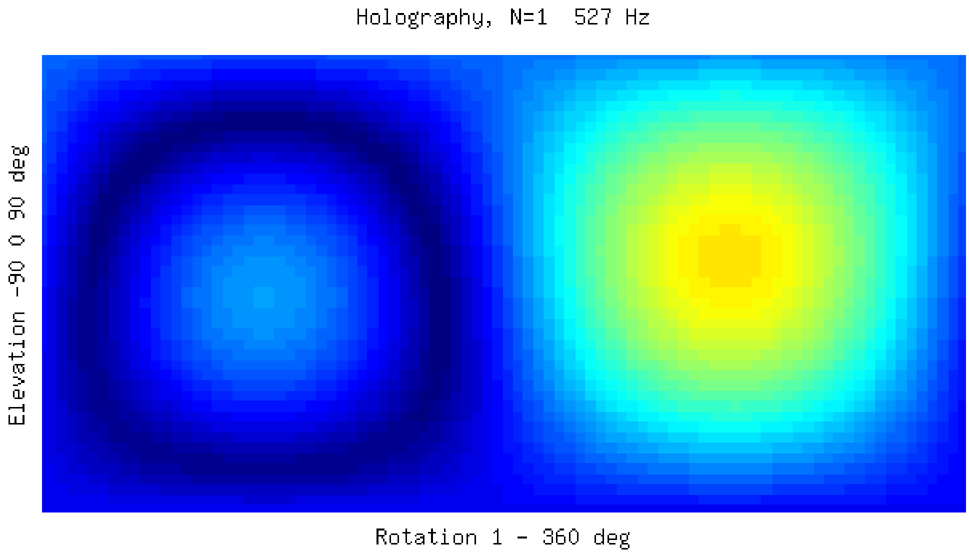

Figure 23: Low order decomposition at $N = 1$

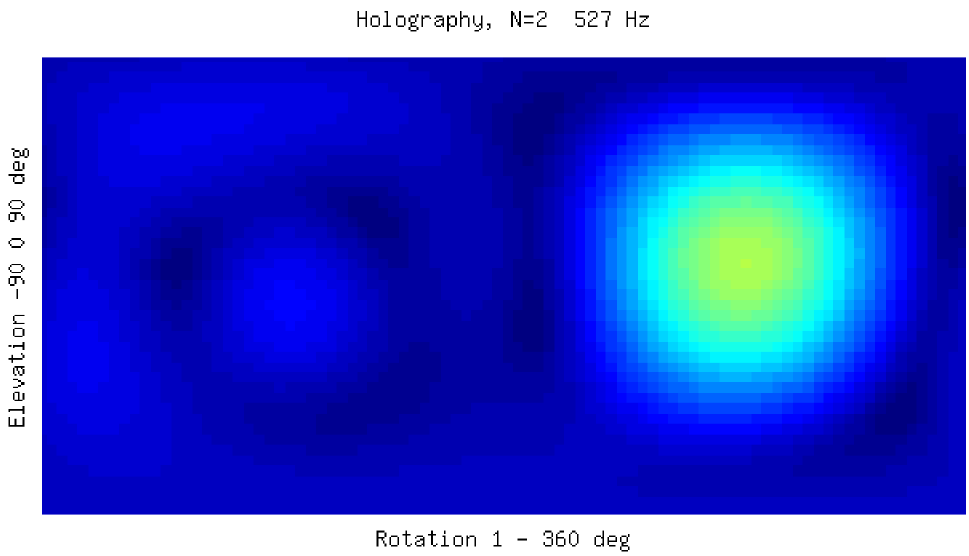

Figure 24: Decomposition at $N = 2$

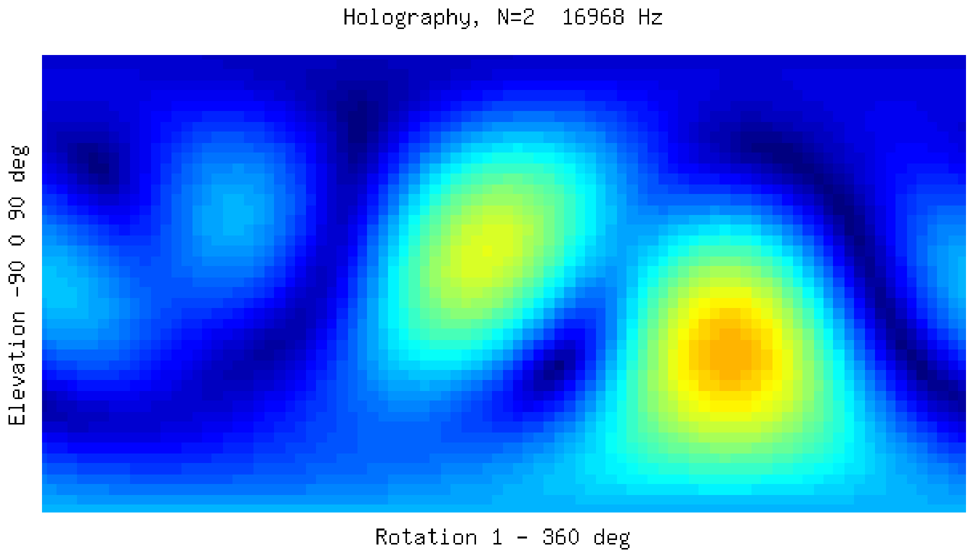

Figure 25: Spatial aliasing at a high frequency and order $N = 2$

Spatial aliasing is detected for a high frequency of 17 kHz and order $N = 2$, as can be seen in figure 25.
At order $N = 3$ the high gains of the holography filter matrix boost the noise floor at low frequencies as shown for 128 Herz in figure 26, rendering the result useless.

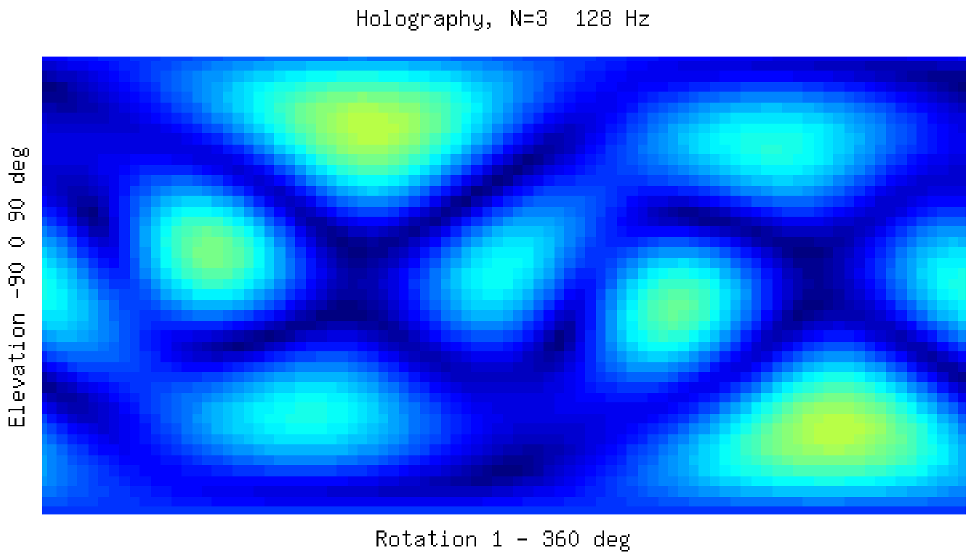

Figure 26: Noise floor at a low frequency and order $N = 3$

11.3 Results and Possible Improvements

The impulse response measurements are summarized in figure 27. Each column resembling a different spherical harmonic decomposition order and the plots are given for multiple frequencies. The trade-off between high spatial bandwidth and amplified background noise is inherently visible.

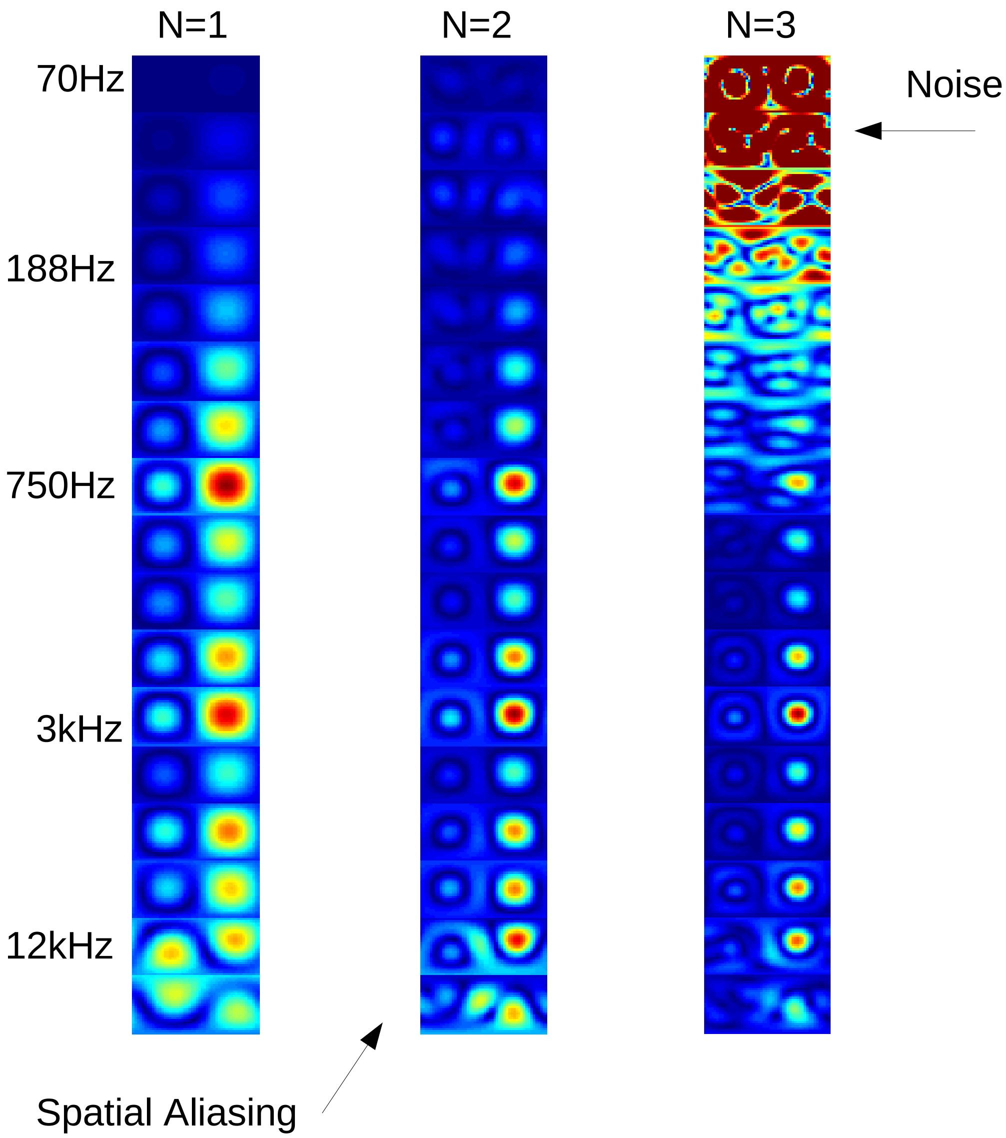

Figure 27: Comparison of angular magnitude at different orders

A solution to this problem is achieved by using the decomposition orders at their optimal frequency ranges. The result of this parallel processing is shown in 28. Low harmonic orders with moderate filter gains are used for low frequencies. Higher orders are applied in higher frequency ranges, giving more detailed spatial resolution. The noise floor of the microphone determines the lowest usable frequency in every band. The signal-to-noise ratio therefore constitutes the major limiting factor for the overall performance of the array.

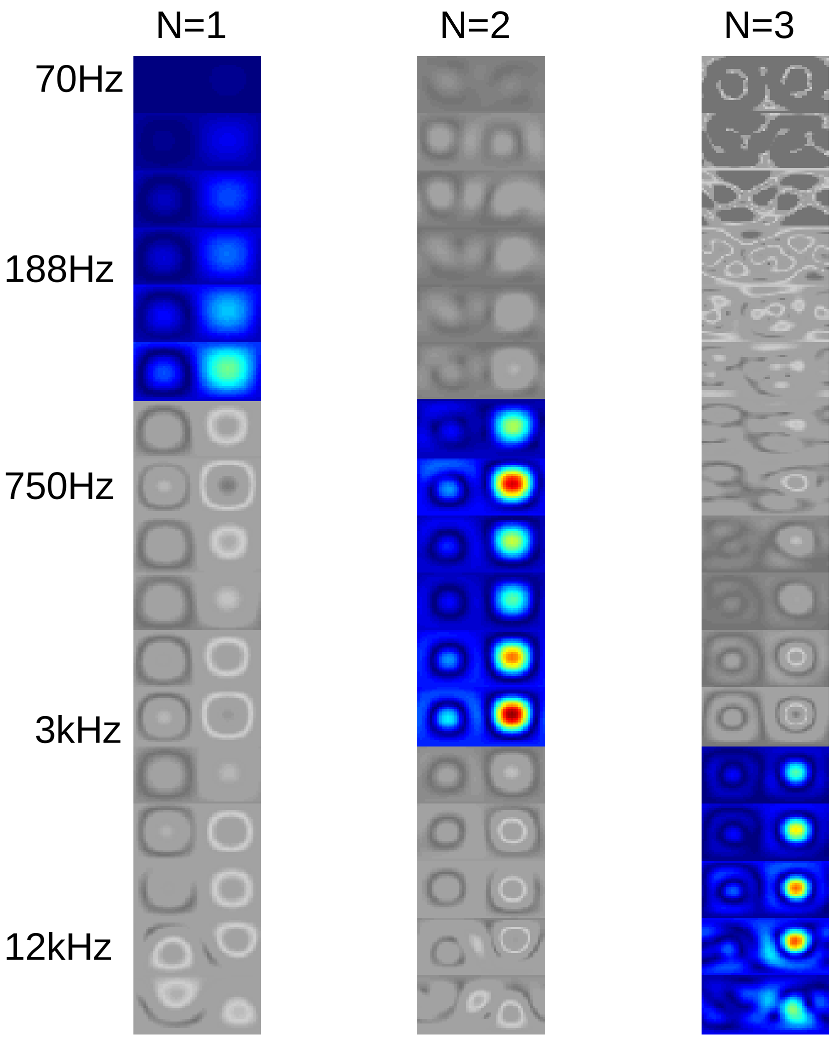

Figure 28: Parallel decomposition at different orders for different frequency bands. Gray areas are filtered out.

12 Summary

This thesis offered an in-depth review of spherical microphone arrays. The application of spherical harmonics in audio processing and their use in microphone arrays results in a complete spatial description of the recorded sound field via acoustic holography. Spherical arrays are universal sensors for measurements and sound recording alike. The independence between decomposition and playback schemes is a major strength, as well as the scalability of the harmonic order employed. The construction of a microphone array is a complex task and requires careful planning and simulation of the layout. This is simplified by the definition of holography filters for different architectures and the discussion of finite resolution sampling and limited spherical harmonics. The simulation of a variety of arrays helps to understand the relation between the design parameters. Error measures for spatial resolution as deducted in this work lead to a classification and invite further study and listening tests. The influence of noise inherent to any sensor is shown and a possible solution is suggested. The reproduction of a holographic recording using either multichannel loudspeaker setups or virtual microphones by means of modeling and beamforming opens an entirely new and exciting field of applications and future research. The holophonic reproduction of music in its performance space will get a more common sensation in the near future.

A Appendix: Functions and Figures

A.1 Spherical Bessel Function

The solution to the spherical Bessel equation exists as two types, the spherical Bessel function of the first kind with integer order n [Wei09]

$$j_n(x) = (-1)^n x^n \left(\frac{d}{x\,dx}\right)^n \frac{\sin(x)}{x} \tag{63}$$

and the spherical Bessel function of the second kind (also known as spherical Neumann function) [Zot09a]

$$y_n(x) = (-1)^{n+1} x^n \left(\frac{d}{x\,dx}\right)^n \frac{\cos(x)}{x} \tag{64}$$

A.2 Spherical Hankel Function

The spherical Hankel Function of the first kind with integer order n is defined as [Wil99, p.194]

$$h_n^{(1)}(x) = j_n(x) + \mathrm{i} y_n(x) \tag{65}$$

and of second kind, for real values x, and * denoting complex conjugation:

$$h_n^{(2)}(x) = h_n^{(1)}(x)^* \tag{66}$$

A.3 Derivatives of Spherical Bessel and Hankel Functions

Derivatives of the above functions exist as recurrence equation for $f_n = j_n$, y_n, $h_n^{(1)}$ and $h_n^{(2)}$

Since [Wil99, p.197]

$$\frac{2n+1}{x} f_n(x) = f_{n-1}(x) + f_{n+1}(x) \tag{67}$$

and

$$f_n'(x) = f_{n-1} - \frac{n+1}{x} f_n(x) \tag{68}$$

these two can be combined:

$$f_n'(x) = \frac{n}{x} f_n(x) - f_{n+1}(x) \tag{69}$$

A.4 Far Field Assumption

The assumption of plane waves is given by the following relation in [Zot09a]

$$kr \gg \frac{N(N+1)}{2} \tag{70}$$

describing the common asymptotic behavior of the spherical Bessel and Hankel functions as an indicator for far field conditions. For low orders n it simplifies to $kr \gg 1$, and is shown in figure 29.

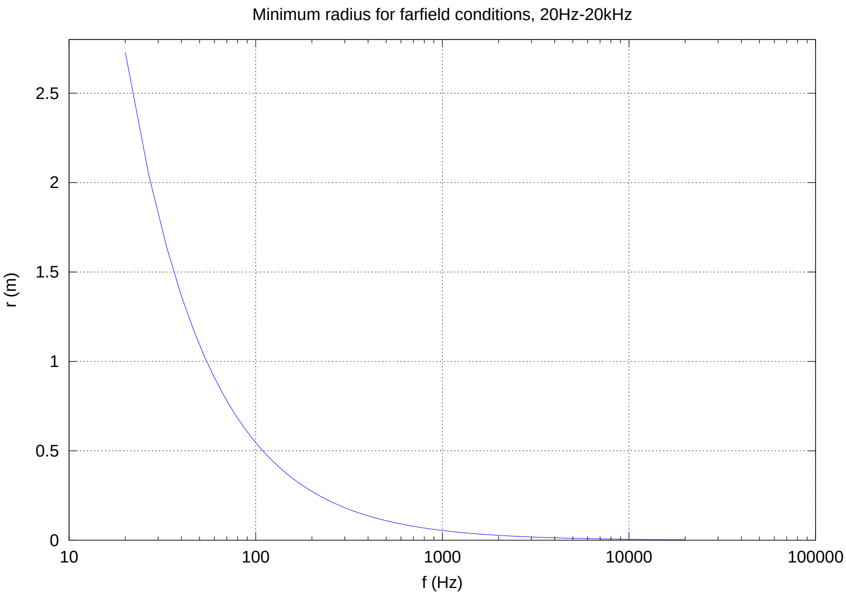

Figure 29: The condition $kr = 1$ is shown for frequency versus radius. All sources whose characteristics lie above the line can for low orders be simplified as emitting plane waves [Zot09a].

References

[AFKW06] Rimas Avizienis, Adrian Freed, Peter Kassakian, and David Wessel. Free Compact 120 Independent Element Spherical Loudspeaker Array with Programmable Radiation Patterns. *AES 120th Convention, Paris, France*, May 2006.

[AW01] George B. Arfken and Hans J. Weber. *Mathematical Methods for Physicists. 5th ed.* Orlando, FL: Academic Press. xiv, 2001.

[Blu58] Alan D. Blumlein. British Patent Specification 394325, submitted Dec 14 1931. *Journal of the Audio Engineering Society*, 6(2), April 1958. Reprint.

[BR07] Ilya Balmages and Boaz Rafaely. Open-sphere Designs for Spherical Microphone Arrays. *IEEE Transactions on Audio, Speech, and Language Processing*, 15(2):727–732, 2007.

[Far00] Angelo Farina. Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique. *Preprints-Audio Engineering Society*, 2000.

[Ger72] Michael Gerzon. Periphony: With-Height Sound Reproduction. *Journal of the Audio Engineering Society*, 21(1), January 1972.

[Ger75] Michael Gerzon. The Design of Precisely Coincident Microphone Arrays for Stereo and Surround Sound. *Proceedings of the 50th AES Convention*, March 1975.

[Gpl07] GPL - GNU General Public License, Version 3. Free Software Foundation, Inc. http://fsf.org, June 2007.

[GW06] Gerhard Graber and Werner Weselak. Raumakustik. Technical University Graz, Austria, 2006.

[Hoh09] Fabian Hohl. Spherical Microphone Array for Radiaton Analysis of Musical Instruments. Master's thesis, Technical University Graz, Austria, December 2009.

[Hol06] Florian Hollerweger. Periphonic Sound Spatialization in Multi-User Virtual Environments. Master's thesis, Technical University Graz, Austria, March 2006.

[HS96] Ronald Hardin and Neil Sloane. McLaren's Improved Snub Cube and other New Spherical Designs in Three Dimensions. *Discrete Computational Geometry*, 15, 1996.

[Li05] Zhiyun Li. *The Capture and Radiation of 3D Auditory Scenes*. PhD thesis, University of Maryland, 2005.

[ME02] Jens Meyer and Gary Elko. A Highly Scalable Spherical Microphone Array Based on an Orthonormal Decomposition of the Soundfield. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2, 2002.

[MWL85] Julian D. Maynard, Earl G. Williams, and Y. Lee. Nearfield Acoustic Holography: I. Theory of Generalized Holography and the Development of NAH. *J. Acoust. Soc. Am.*, 78(4), October 1985.

[Oct] GNU Octave, a High Level Language for Numerical Computation by John Eaton and others. http://www.gnu.org/software/octave/.

[Pet04] Svend Oscar Petersen. Localisation of Sound Sources using 3D Microphone Array. Master's thesis, University of Southern Denmark, 2004.

[Pom08] Hannes Pomberger. Angular and Radial Directivity Control for Spherical Loudspeaker Arrays. Master's thesis, IEM - Institute of Electronic Music and Acoustics, University of Music and Performing Arts Graz, Austria, April 2008.

[Puc97] Miller S. Puckette. Pure Data: Another Integrated Computer Music Environment. *Reprinted from Proceedings, Second Intercollege Computer Music Concerts, Tachikawa*, pages pp. 37–41., May 1997.

[Raf05] Boaz Rafaely. Analysis and Design of Spherical Microphone Arrays. *IEEE Transactions on Speech and Audio Processing*, 13(1):135–143, 2005.

[Raf08] Boaz Rafaely. The Spherical-Shell Microphone Array. *IEEE Transactions on Audio, Speech, and Language Processing*, 16(4), May 2008.

[S+07] Anton Schlesinger et al. Holographic Sound Field Analysis with a Scalable Spherical Microphone Array. *AES 122th Convention, Vienna, Austria*, May 2007.

[Tar74] Viggo Tarnow. Sound Radiation from Loudspeaker System with the Symmetry of the Platonic Solids. *Bruel&Kjaer Technical Review*, (4), 1974.

[TT08] Ambler Thompson and Barry N. Taylor. NIST - Guide for the Use of the International System of Units (SI). NIST - National Institute of Standards and Technology Gaithersburg, MD, March 2008.

[WA01] Darren B. Ward and Thushara D. Abhayapala. Reproduction of a Plane-Wave Sound Field Using an Array of Loudspeakers. *IEEE Transactions on Speech and Audio Processing*, 9(6):697–707, 2001.

[WDC97] Olivier Warusfel, Philippe Derogis, and Rene Causse. Radiation Synthesis with Digitally Controlled Loudspeaker. *Proceedings of the 103th AES Convention*, September 1997.

[Wei09] Eric W. Weisstein. Wolfram Mathworld, 2009.

[Wil99] Earl G. Williams. *Fourier acoustics: Sound Radiation and Nearfield Acoustical Holography*. Academic Press, 1999.

[ZH07] Franz Zotter and Robert Höldrich. Modelling Radiation Synthesis with Spherical Loudspeaker Arrays. *19th International Congress on Acoustics, Madrid, Spain*, September 2007.

[Zot09a] Franz Zotter. *Analysis and Synthesis of Sound-Radiation with Spherical Arrays*. PhD thesis, IEM - Institute of Electronic Music and Acoustics, University of Music and Performing Arts Graz, Austria, November 2009.

[Zot09b] Franz Zotter. Sampling Strategies for Acoustic Holography/Holophony on the Sphere, March 2009. Presentation, NAG-DAGA, 23-26 March, Rotterdam, The Netherlands.

[ZSR03] Johannes Zmölnig, Alois Sontacchi, and Winfried Ritsch. The IEM-Cube, A Periphonic Re-/Production System. *AES 24th International Conference on Multichannel Audio, The New Reality, Banff, Canada*, June 2003.