Acoustic Centering and Rotational Tracking in Surrounding Spherical Microphone Arrays

Daniel Deboy

Graz, March 2010

Advisor: Dr. Franz Zotter
Assessor: Prof. Robert Höldrich

Institute of Electronic Music and Acoustics

University of Music and Performing Arts Graz, Austria
Graz University of Technology, Austria

UNIVERSITÄT
FÜR MUSIK UND
DARSTELLENDE KUNST
GRAZ - AUSTRIA

TU Graz
Graz University of Technology

iem

Declaration

All parts of this work, including this documentation and the implementation of the presented algorithms were autonomously written by the author, except where explicitly otherwise stated.

....

Abstract

This diploma thesis considers the development and testing of methods applicable for the analysis of radiation patterns from real-world instruments. The goal is the estimation of geometric parameters of the radiation-patterns from instrumental sounds. These parameters shall be retrieved from sound-radiation data that have been captured by a surrounding microphone array.

In a first step, solution strategies for different problems are developed that arise in the recording situation of surrounding microphone arrays. For example, the correct orientation of the analysis data using a spherical microphone array of which the orientation is known imprecisely. Moreover, a major goal is to find simple measures for detecting the acoustical center of a sound source inside the microphone array. For a decomposition of the measured data into spherical harmonics, we assume a limited spatial bandwidth due to the finite number of discrete angular microphone positions. This assumption yields problems whenever a sound source is not perfectly located at the center of the array. A cost function shall be minimized in order to find a center of decomposition that is superior for decomposition in the analysis.

Furthermore, this work investigates the tracking of a rotating sound source using correlation measures. For this purpose successive sound-radiation data is compared by rotational matching.

Kurzfassung

Im Rahmen dieser Diplomarbeit werden Verfahren zur Untersuchung der Abstrahlcharakteristik von realen Instrumenten entwickelt und getestet. Ziel ist die Extraktion von geometrischen Parametern der Schallabstrahlung von Instrumentalklängen. Diese Parameter sollen aus Abstrahlungsdaten gewonnen werden, die rund um das Instrument aufgezeichnet wurden.

Zunächst werden Lösungsstrategien für verschiedene Probleme entwickelt, die bei der Aufnahme mit umgebenden Kugelmikrofonanordnungen auftreten. Dazu gehören die korrekte (Rotations-)Ausrichtung der Analyse bei verdrehter Kugelmikrofonanordnung und vor allem der Versuch, ein einfaches Maß der Zentriertheit des akustischen Zentrums einer Schallquelle innerhalb der Mikrophonanordnung zu finden. Für eine Zerlegung der Messwerte in Kugelharmonische wird aufgrund der endlich vielen, im Winkel diskreten Mikrofonpositionen die Annahme einer räumlichen Bandbegrenzung getroffen. Das kann aber genau dann zu Problemen führen, wenn an sich räumlich bandbegrenzte akustische Strahler nicht perfekt im Zentrum sitzen. Durch die Minimierung einer Kostenfunktion wird das Zerlegungszentrum vor der Analyse verschoben, um diese Einschränkung weitgehend aufzulösen.

Weiters wird versucht, eine Rotationsverfolgung der Quellsignale mit Hilfe von Korrelationsmaßen zu realisieren. Hierbei werden zeitlich aufeinanderfolgende Abstrahlungsdaten relativ zueinander betrachtet.

Acknowledgements

My sincere thanks to Franz Zotter for being a superb advisor throughout the whole period this thesis has been elaborated.

I want to thank Fabian Hohl for providing the recorded samples and additional documentation material.

I want to thank Nino Skiljic for his implementation of the spherical base-solution transforms.

Thanks to my fellow students and friends for the past terrific years of joy and hard work. I hope we will stay in touch.

Dedicated to my parents.

Contents

Chapter 1

Introduction

Motivation

A surrounding microphone array consists of several microphones that capture the sound pressure at certain positions around a musician playing a musical instrument. This kind of array provides a comfortable and easy way to gather information about the behavior of musical instruments. A musical instrument consists of one or more oscillating parts radiating sound to the surrounding air. The radiated wave may be omni-directional or focused to a certain direction, depending on its wavelength, i.e. the frequency. The radiated waves form a radiation pattern on a closed surface surrounding the musical instrument in the far-field.

Since there seem to exist no feasible solutions for recording the sound-pressure at a continuos spherical surface around sound sources, the sound-pressure on this surface can only be recorded at discrete angular locations. From discrete observations, we can describe the sound-pressure continuously at the surface by employing a mathematical interpolation method. The spherical harmonics decomposition (SHD) provides a suitable method of interpolation in the special case of the hyperinterpolation method.

One of the main problem arises because of the discrete sampling of the surface. As it is convenient to assume a limit of the bandwidth before discretization, spherical harmonics spectra components will be aliased. Fortunately, for most applications the aliasing error is acceptable as long as the sound source lies in the center of the sphere. However, shifting the source generates higher order components in the spherical harmonics spectrum and spatial aliasing cannot be neglected anymore.

Centering the sound source minimizes higher order components in the spherical harmonics spectrum. Different cost functions are proposed to estimate localization maps from the recorded sound pressure. Minimizing these cost functions intends to yield a precise translation vector which can be used to center the source signal by simply changing the coordinate origin of the SHD.

Chapter 2

Spherical Exterior Problem

This chapter provides a brief overview on continuous wave field theory in spherical co-
ordinates. An intensive study of this theory is necessary for the understanding of the
discrete case we are dealing with when using a spherical microphone array. Details on
the theoretical aspects are found in [Zot09] and [Wil99].

2.1 Wave field description in spherical coordinates

The Helmholtz equation (eq. 2.1) can be used to describe sound pressure fields $p(\boldsymbol{r},\omega)$
in the frequency domain. It is derived from the wave equation in time domain and can
be written as

$$(\triangle + k^2)p(\boldsymbol{r},\omega) = 0. \tag{2.1}$$

In order to solve this equation in spherical coordinates, the cartesian coordinates in $\mathbb{R}_3$
have to be transformed according to

$$\boldsymbol{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = r\boldsymbol{\theta}, \quad \mathring{\boldsymbol{r}} = \begin{pmatrix} r \\ \varphi \\ \vartheta \end{pmatrix} = \begin{pmatrix} \sqrt{x^2+y^2+z^2} \\ \arctan(\frac{y}{x}) \\ \arctan\frac{\sqrt{x^2+y^2}}{z} \end{pmatrix}, \quad \boldsymbol{\theta} = \begin{pmatrix} cos(\varphi)sin(\vartheta) \\ sin(\varphi)sin(\vartheta) \\ cos(\vartheta) \end{pmatrix}, \tag{2.2}$$

where $k = \omega/c$ is the wave number and $r, \varphi\,\vartheta$ are called the radius, azimuth, respectively
zenith angle. The vector $\boldsymbol{\theta}$ describes a unit sphere and is proposed for a more concise
notation [Zot09].
Throughout this thesis, the Fourier coefficient of sound-pressure at a certain frequency
ω is denoted as

$$p(kr, \boldsymbol{\theta}) = p(\boldsymbol{r},\omega) = p(\boldsymbol{r},t)e^{i\omega t}. \tag{2.3}$$

2.2 Spherical harmonics

The spherical harmonic functions can be seen as the modes of vibration on a spherical surface. Fig. 2.1 shows the real-valued spherical harmonic functions up to an order of $N = 2$. We write $Y_n^m(\boldsymbol{\theta})$ for the real-valued spherical harmonic function with n being the order and m the degree of the spherical harmonic

$$Y_n^m(\boldsymbol{\theta}) = \sqrt{\frac{(2n+1)(2-\delta_m)}{4\pi}\frac{(n-m)!}{(n+m)!}}P_n^m(\cos(\vartheta)) \cdot \left\{ \begin{array}{ll} \sin(m\varphi), & for\ m\ < 0, \\ \cos(m\varphi), & for\ m\ \geq 0. \end{array} \right.$$

(2.4)
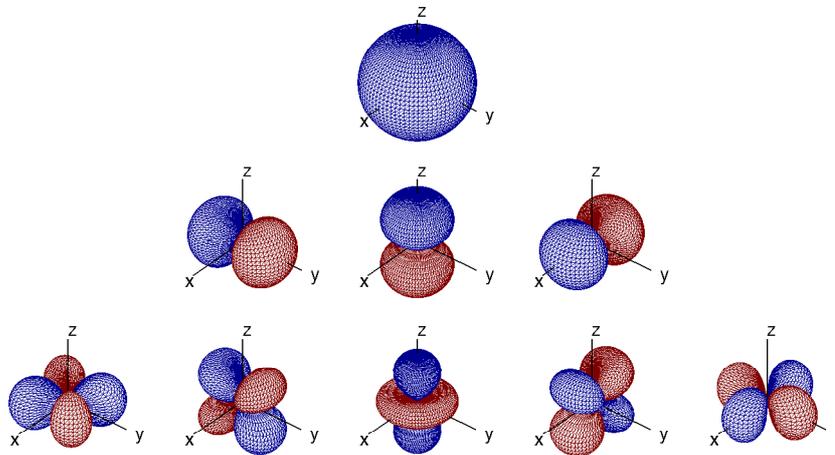
Figure 2.1: Real valued spherical harmonic functions for n=0...2

The spherical harmonics are orthonormal and are the basis of the spherical harmonics domain. The orthonormality is shown by the integral

$$\int_{\mathbb{S}^2} Y_n^m(\boldsymbol{\theta})Y_{n'}^{m'}(\boldsymbol{\theta})d\boldsymbol{\theta} = \delta_{nn'}\delta_{mm'}.$$

(2.5)

P_n^m are the associated Legendre functions

$$P_n^m(x) = (-1)^m(1-x^2)^{m/2}\frac{d^m}{dx^m}P_n(x).$$

(2.6)

With the Legendre polynomials $P_n(x)$ which can be determined by recurrence relations.

2.3 Spherical base solutions

Several solutions of the Helmholtz equation in spherical coordinates are known. However, not all of them are physical. According to [Wil99] the solution for standing waves is

written as

$$p(kr, \boldsymbol{\theta}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} (a_{nm} j_n(kr) + b_{nm} y_n(kr)) Y_n^m(\boldsymbol{\theta}), \tag{2.7}$$

and the solution for traveling waves is written as:

$$p(kr, \boldsymbol{\theta}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} (c_{nm} h_n^{(2)}(kr) + d_{nm} h_n^{(1)}(kr)) Y_n^m(\boldsymbol{\theta}), \tag{2.8}$$

where $a_{nm}, b_{nm}, c_{nm}, d_{nm}$ are coefficients, j_n, y_n are the spherical Bessel and Neumann functions and $h_n^{(1)}, h_n^{(2)}$ are the spherical Hankel functions, cf. fig. 2.2,

$$j_n(x) = (-x)^n \left(\frac{1}{x}\frac{d}{dx}\right)^n \left(\frac{sin(x)}{x}\right), \tag{2.9}$$

$$y_n(x) = -(-x)^n \left(\frac{1}{x}\frac{d}{dx}\right)^n \left(\frac{cos(x)}{x}\right), \tag{2.10}$$

$$h_n^{(1)}(x) = (-x)^n \left(\frac{1}{x}\frac{d}{dx}\right)^n \left(\frac{e^{ix}}{ix}\right), \tag{2.11}$$

$$h_n^{(2)}(x) = h_n^{(1)}(x)^*. \tag{2.12}$$

It is easy to see that $h_n^{(1)}$ is a complex valued composite of the spherical Bessel functions

$$h_n^{(1)}(x) = j_n(x) + i y_n(x). \tag{2.13}$$

Regarding singularities and the Sommerfeld radiation condition we may state the physical solutions as [Zot09]

$$p(kr, \boldsymbol{\theta}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} (a_{nm} j_n(kr) + c_{nm} h_n^{(2)}(kr)) Y_n^m(\boldsymbol{\theta}), \tag{2.14}$$

and I will refer to the coefficients a_{nm} and c_{nm} as the *wave spectra* in this work. The spherical Bessel functions describe standing waves whereas traveling waves that are singular at $r = 0$ are characterized by the spherical Hankel functions of the second kind. From the spherical Hankel functions, a condition for the far-field depending on the order of the spherical harmonic can be derived,

$$kr' \gg \frac{n(n+1)}{2}. \tag{2.15}$$
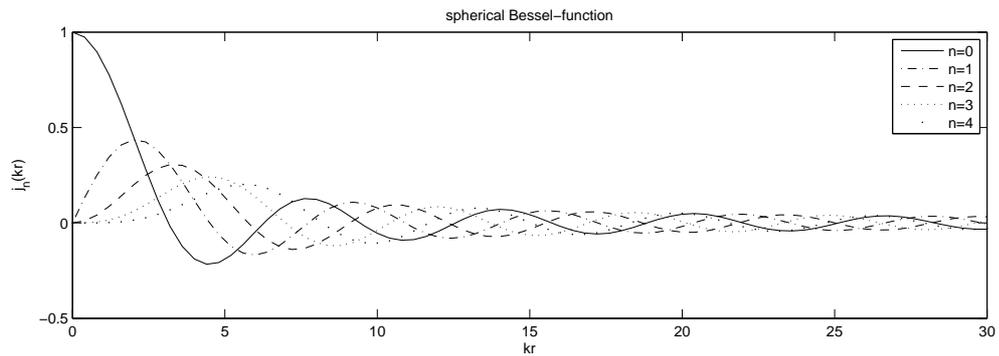
(a) Spherical Bessel function
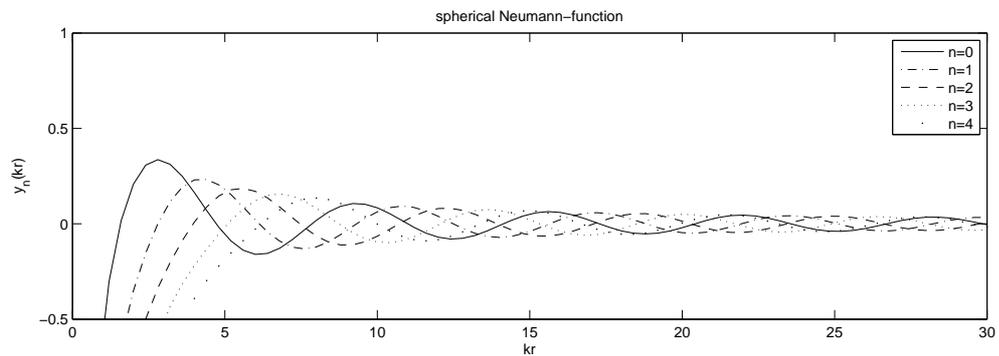
(b) Spherical Neumann function
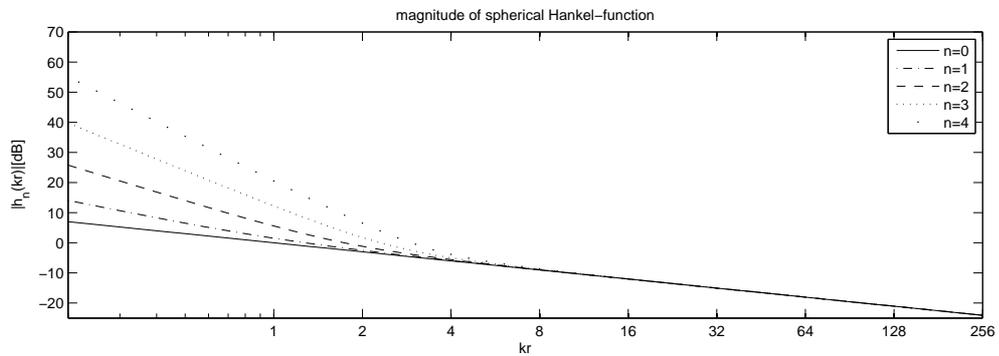
(c) Magnitude of spherical Hankel function

Figure 2.2: Spherical Bessel, Neumann and Hankel function for n=0...4

2.4 Spherical wave spectrum

The spherical wave spectrum as introduced in [Wil99] describes the spherical harmonics decomposition of a sound particle velocity on a spherical surface. It allows us to describe spherical boundary values.

Spherical harmonics transform

The spherical harmonics transform is a decomposition of a continuous distribution $g(\boldsymbol{\theta})$ given at a unit sphere $\mathbb{S}^2$ into its components of spherical harmonics

$$SHT_{nm}\{g(\boldsymbol{\theta})\} = \gamma_{nm} = \int_{\mathbb{S}^2} g(\boldsymbol{\theta}) Y_n^m(\boldsymbol{\theta}) d\boldsymbol{\theta}. \tag{2.16}$$

The expansion of γ_{nm} is complete, we can write the inverse transform as

$$ISHT = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \gamma_{nm} Y_n^m(\boldsymbol{\theta}) = g(\boldsymbol{\theta}). \tag{2.17}$$

The expansion fulfills Parseval's theorem

$$\int_{\mathbb{S}^2} |g(\boldsymbol{\theta})|^2 d\boldsymbol{\theta} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} |\gamma_{nm}|^2. \tag{2.18}$$

Spherical wave spectrum of the sound-pressure

Decomposing a continuous sound-pressure distribution of a sphere into spherical harmonics yields the *spherical wave spectrum* [Wil99]

$$\psi_n^m(kr) = SHT\{p(kr, \boldsymbol{\theta})\}. \tag{2.19}$$

With the expansion we can regain the sound-pressure distribution at the sphere

$$p(kr, \boldsymbol{\theta}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \psi_n^m(kr) Y_n^m(\boldsymbol{\theta}). \tag{2.20}$$

Note that the coefficients of spherical wave spectra still feature a radial weighting by the spherical Bessel and Hankel functions and can be further decomposed into the wave spectra a_{nm}, c_{nm}.

$$\psi_n^m(kr) = a_{nm} j_n(kr) + c_{nm} h_n^{(2)}(kr). \tag{2.21}$$

2.5   Boundary value problem
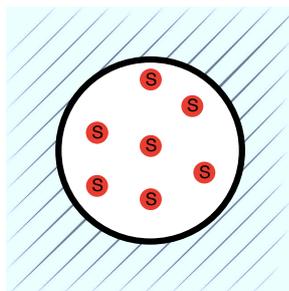
Figure 2.3: Exterior problem

A given sound pressure distribution on a closed surface S of a Volume V is called Dirichlet boundary condition [Zot09] for the Helmholtz equation. Since every sound source of a surrounding microphone array should be located inside of V, the solutions of this boundary value problem are valid outside the array. Therefore it is referred to as "exterior" problem in [Zot09]. The divergence theorem proves that the sound field outside of the Volume V is fully determined by the distribution on the closed surface . Note that even if a source free field outside of the array is ensured, reflections of surrounding surfaces can form mirrored sources and the assumption of an exterior problem becomes invalid. Hence the exterior problem needs an anechoic environment to be evaluated properly, or must separate between exterior and interior solutions using either two layers of microphones or the velocity and pressure distribution [WA80].

If we know the sound-pressure distribution on a sphere $\mathbb{S}^2$ with radius r we can apply the SHT. The outmost sound source at radius r_0 therefore defines the smallest possible sphere surrounding all sound sources by $r > r_0$.

In terms of a pure exterior problem, the spherical wave spectrum reduces to $(a_{nm} \equiv 0)$

$$\psi_n^m(kr) = c_{nm} h_n^{(2)}(kr) = SHT\{p(kr, \boldsymbol{\theta})\}. \tag{2.22}$$

Physically, $a_{nm} \equiv 0$ means that there are no incident waves superimposed to the free-field outside r.

Spherical acoustic holography

If the purpose of a measured spherical sound-pressure distribution is the evaluation of the wave field at other radii, we refer to this as *acoustic holography* [Zot09]. Acoustic holography for the spherical exterior problem extrapolates from the spherical wave spectrum at radius r to the spherical wave spectrum at any other radius r' that still encloses all sound sources:

$$\psi_n^m(kr') = \psi_n^m(kr) \frac{h_n^{(2)}(kr')}{h_n^{(2)}(kr)}. \tag{2.23}$$

The sound-pressure distribution at r' extrapolated by spherical acoustic holography becomes:

$$p(kr', \boldsymbol{\theta}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \psi_n^m(kr') Y_n^m(\boldsymbol{\theta}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \psi_n^m(kr) \frac{h_n^{(2)}(kr')}{h_n^{(2)}(kr)} Y_n^m(\boldsymbol{\theta}). \tag{2.24}$$

Deriving radiation patterns

We can derive a far-field radiation pattern of the sources located inside r by extrapolating the sound-pressure distribution $p(kr, \boldsymbol{\theta})$ to a new radius r'>r that lies in the far-field, cf. eq. (2.15).
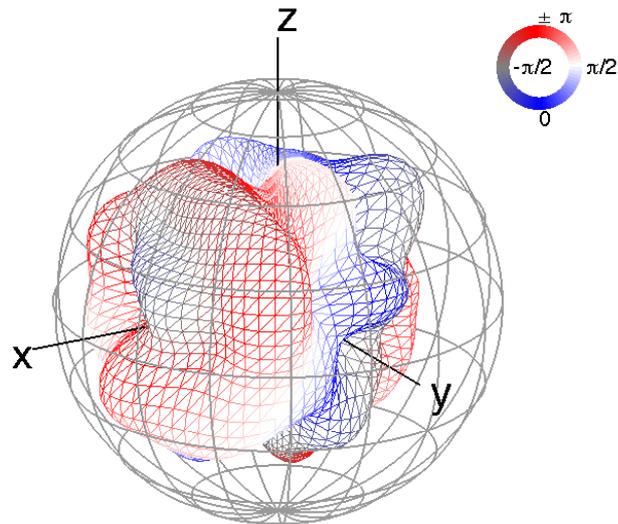
Figure 2.4: Example of a far-field radiation pattern: Trumpet playing A5, observed partial at 888Hz, shown is the extrapolated sound-pressure distribution within a dynamic range of the top 20dB.

Chapter 3

Surrounding Microphone Arrays

A surrounding microphone array captures the sound pressure $p(t)$ or particle velocity $v(t)$ at spatial sampling points on a sphere surrounding the sound source.

For this thesis, recordings have been taken using the 64 channel surrounding microphone array at IEM Graz, Austria. The array frame has a radius of about $R_a = 1.35m$ at which it mounts 64 omni-directional microphones. Omni-directional microphones produce a voltage proportional to the sound pressure $p(\boldsymbol{r}_{ch}, t)$ at the position $\boldsymbol{r}_{ch}$ of their capsule depending on the time instant t. In the discrete-time domain, the Fourier coefficients $p(\boldsymbol{r}, \omega)$ can be derived according to eq. (2.3).

Simulations for estimating the performance of the algorithms are based on the structure of this array. A detailed description of the array can be found in [Hoh09].

Figure 3.1: Picture of the surrounding spherical microphone array at IEM Graz, Austria

3.1 Estimation of the spherical wave spectrum from discrete observations

In order to preserve as much information as possible when discretizing a spherical distribution, it is convenient to assume it having a finite-resolution, i.e. angular band-limit. The following sections describe how to obtain wave spectra from discrete observation utilizing this assumption.

3.1.1 Band-limited wave spectrum

In vector notation the spherical wave spectrum

$$p(kr, \boldsymbol{\theta}) = \sum_{n=0}^{N} \sum_{m=-n}^{n} (a_{nm} j_n(kr) + c_{nm} h_n^{(2)}(kr)) Y_n^m(\boldsymbol{\theta})$$

can be expressed briefly as:

$$\psi_N(kr) = \boldsymbol{J}_N \boldsymbol{b}_N + \boldsymbol{H}_N \boldsymbol{c}_N. \tag{3.1}$$

Truncating the spherical wave spectrum at the order N generates a truncation error which can be expressed by:

$$\epsilon_{trunc} = \boldsymbol{\psi}_\infty - \boldsymbol{\psi}_N = \boldsymbol{H}_\infty \boldsymbol{c}_\infty - \boldsymbol{H}_N \boldsymbol{c}_N. \tag{3.2}$$

This kind of error is important when regarding sources which produce high-order components $N' > N$, or when dealing with translation, cf. section 4.2 and eq. 5.4.

Vector and matrix notation

To fill the coefficients b_{nm} into vectors and matrices, the following conventions are useful:

$$idx = n^2 + n + m + 1, \tag{3.3}$$
$$b_{idx} = b_{nm}, \tag{3.4}$$

$$\boldsymbol{b}_N = vec\{b_{idx}\} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_{(N+1)^2} \end{pmatrix}, \tag{3.5}$$

$$diag_N\{b_{nm}\} = diag_N\{b_{idx}\} = \begin{pmatrix} b_{0,0} & 0 & 0 & 0 & \cdots & 0 \\ 0 & b_{1,-1} & 0 & 0 & \cdots & 0 \\ 0 & 0 & b_{1,0} & 0 & \cdots & 0 \\ 0 & 0 & 0 & b_{1,1} & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & b_{N,N} \end{pmatrix}, \tag{3.6}$$

$$mtx\{\boldsymbol{b}_N^{(L)}\} = \left(\boldsymbol{b}_N^{(1)}, \cdots, \boldsymbol{b}_N^{(L)}\right). \tag{3.7}$$

If b_{idx} is independent of m, we can use the above description and replace m with zero,

$$\boldsymbol{H}_N := diag_N\{h_n^{(2)}(kr)\}, \tag{3.8}$$

$$\boldsymbol{J}_N := diag_N\{j_n(kr)\}. \tag{3.9}$$

3.1.2 DSHT by hyperinterpolation on the sphere

According to [Zot09], angular sampling nodes for hyperinterpolation on the sphere $\mathbb{S}^2$ use quadratic matrices $\boldsymbol{Y}_N$ and require good condition of $\boldsymbol{Y}_N$. [WS01] gives a detailed description of how to derive $L = (N+1)^2$ angular sampling nodes

$$\boldsymbol{Y}_N := mtx\{\boldsymbol{y}_N^{(l)T}\} = \begin{pmatrix} \boldsymbol{y}_N^{(1)T} \\ \boldsymbol{y}_N^{(2)T} \\ \vdots \\ \boldsymbol{y}_N^{(L)T} \end{pmatrix}, \tag{3.10}$$

where $\boldsymbol{y}_N^{(l)}$ is a vector of the spherical harmonics evaluated at sampling node l:

$$\boldsymbol{y}_N^{(l)} = \begin{pmatrix} Y_{0,0}^{(l)} \\ Y_{1,-1}^{(l)} \\ Y_{1,0}^{(l)} \\ Y_{1,1}^{(l)} \\ \vdots \\ Y_{n,n}^{(l)} \end{pmatrix}. \tag{3.11}$$

The computation of the discrete spherical harmonics transform (DSHT) requires an inversion of the spherical harmonics matrix $\boldsymbol{Y}_N$. Thus the matrix has to be of full rank

to be invertible [Bög03]. The order N determines the number of rows of the spherical harmonics matrix with $(N+1)^2$, and the number of sampling nodes L directly yields the number of columns.

$$DSHT_N\{\boldsymbol{g}_L\} = \boldsymbol{\gamma}_N = \boldsymbol{Y}_N^{-1}\boldsymbol{g}_L, \tag{3.12}$$

$$DISHT_N\{\boldsymbol{\gamma_N}\} = \mathbf{Y_N}\boldsymbol{\gamma_N} = \mathbf{g_L}. \tag{3.13}$$

Only a quadratic matrix of $((N+1)^2 \times L)$ with $L = (N+1)^2$ can have full rank. The determinant $det(\boldsymbol{Y}_N)$ must not become zero ($det(\boldsymbol{Y}_N) \neq 0$) and for numerical computations should not even get close to zero. The determinant of a matrix $\boldsymbol{Y}$ is defined as the product of its eigenvalues λ_i. Another possible measure of numerical invertibility is the condition number defined as the relation between the largest and the smallest eigenvalue of $\boldsymbol{Y}_N$.

3.1.3 Approximation using SVD

It is due to the invariance of the resolution of spherical harmonics decompositions with respect to rotation that the condition number won't be affected by rotations. However, translating the coordinate system by a vector $\boldsymbol{d} = \begin{bmatrix} dx & dy & dz \end{bmatrix}^T$ distorts the angular positions and may hereby yield a bad condition number. In such cases, hyperinterpolation becomes numerically infeasible. Nevertheless, an approximation of the inverse by regularization can be achieved. Hereby, linear combinations of spherical harmonics are omitted that become linearly dependent because they are weakly sampled.

The singular value decomposition transforms the spherical harmonics matrix into a diagonal matrix $\boldsymbol{S}_N$ and two unitary matrices $\boldsymbol{U}$ and $\boldsymbol{V}^*$. The entries of $\boldsymbol{S}_N$ are called singular values of $\boldsymbol{Y}$

$$\boldsymbol{Y}_N = \boldsymbol{U}\boldsymbol{S}_N\boldsymbol{V}^*, \tag{3.14}$$

$$\boldsymbol{S}_N = \begin{pmatrix} s_0 & 0 & 0 & \cdots & 0 \\ 0 & s_1 & 0 & \cdots & 0 \\ 0 & 0 & s_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & s_N \end{pmatrix}. \tag{3.15}$$

Commonly SVD algorithms sort the singular values in descending order. Setting all singular values that fall below a certain threshold relative to the largest s_0 to zero forms a new diagonal matrix $\tilde{\boldsymbol{S}}_N$ of rank $(N - \#s_i < level)$. The hereby obtained matrix $\tilde{\boldsymbol{Y}}$

approximates $\boldsymbol{Y}_N$ as

$$\tilde{\boldsymbol{Y}}_N = \tilde{\boldsymbol{U}}\tilde{\boldsymbol{S}}_N\tilde{\boldsymbol{V}}^*, \tag{3.16}$$

and allows us to calculate the so called *pseudo-inverse* of $\boldsymbol{Y}_N$

$$\boldsymbol{Y}_N^+ = \tilde{\boldsymbol{V}}\tilde{\boldsymbol{S}}_N^{-1}\tilde{\boldsymbol{U}}^*. \tag{3.17}$$

3.1.4   Estimation of spherical wave spectrum

Using eq. (3.12) to obtain the spherical wave spectrum from $\boldsymbol{p}_L$ for the exterior problem $(\boldsymbol{a}_{nm} = 0)$, we find $\hat{\boldsymbol{\psi}}_N$ as

$$\hat{\boldsymbol{\psi}}_N(kr) = \boldsymbol{Y}_N^{-1}\boldsymbol{p}_L. \tag{3.18}$$

The coefficients $\boldsymbol{c}_N$ are obtained by inversion of eq. (3.18),

$$\boldsymbol{c}_N = \boldsymbol{H}_N^{-1}\hat{\boldsymbol{\psi}}_N(kr) = \boldsymbol{H}_N^{-1}\boldsymbol{Y}_N^{-1}\boldsymbol{p}_L. \tag{3.19}$$

3.2   Spatial aliasing

In general, an unlimited wave spectrum $\boldsymbol{c}_\infty$ of a sound source can be expanded to the measured discrete sound-pressure distribution $\boldsymbol{p}_L$. Hence, estimating the spherical wave spectrum up to order N using discrete observations $\boldsymbol{p}_L$ can be expressed by

$$\hat{\boldsymbol{\psi}}_N = \boldsymbol{Y}_N^{-1}\boldsymbol{p}_L = \boldsymbol{Y}_N^{-1}\boldsymbol{Y}_\infty\boldsymbol{H}_\infty\boldsymbol{c}_\infty, \tag{3.20}$$

whereas a band-limited truncated spherical wave spectrum could be derived from the unlimited wave spectrum by truncation

$$\boldsymbol{\psi}_N = \boldsymbol{H}_N\boldsymbol{c}_N = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} \boldsymbol{H}_\infty\boldsymbol{c}_\infty. \tag{3.21}$$

In eq. (3.20), higher order components of c_∞ may be misinterpreted (aliased) as components of orders equal or below N due to sparse observation by the discrete sampling nodes. A spatial aliasing error is generated and can be expressed by

$$\begin{aligned} \epsilon_{sa} &= \hat{\boldsymbol{\psi}}_N - \boldsymbol{\psi}_N \\ &= (\boldsymbol{Y}_N^{-1}\boldsymbol{Y}_\infty - \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \end{bmatrix})\boldsymbol{H}_\infty\boldsymbol{c}_\infty. \end{aligned} \tag{3.22}$$

The aliasing component coefficients are depicted in the matrix multiplication $\boldsymbol{Y}_N^{-1}\boldsymbol{Y}_\infty$. Fig. 3.2 shows an example of the matrix product. The first $(N+1)^2$ components are correctly mapped while components of higher orders are aliased.

A similar aliasing effect is well known from discrete time sampling of signals having frequency components above the Nyquist frequency.
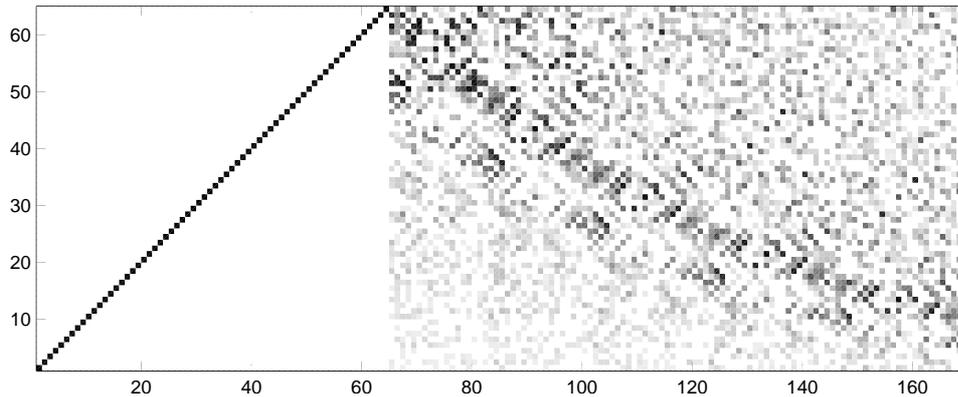
Figure 3.2: Example for matrix product $\boldsymbol{Y}_N^{-1}\boldsymbol{Y}_\infty$ ($\boldsymbol{Y}_\infty$ is of finite size here due to numerical calculation, a small value was chosen for illustration purpose) [Ple09]

3.3   Array imperfections

Several errors may influence the analysis of wave spectra in a surrounding spherical microphone array including diffraction, reflection, noise, gain mismatch and microphone displacement. The effects of these errors on the tracking algorithms are not investigated in this thesis.

However, the gain mismatch of the signal paths including microphone, preamplifier and analog-digital converter have been adjusted using a level calibrating tool.

As examined in [Hoh09] the structure of the microphone array leads to comb filter effects. The comb filter effect is minimized by installing concentric foam absorbers around the microphone locking.

3.4   Calibrating the orientation of the coordinate system

The orientation of the coordinate system sometimes may not match the orientation of the real-world coordinate system or sometimes one may chose an arbitrary coordinate system. A robust method how to achieve a rotation matrix that matches the coordinate systems has been adopted from optical tracking systems and can be found in appendix A.

Chapter 4

Discretization Of The Kinematic Space

This chapter describes ways of discretization for displacements (translation, rotation) in a near-uniform way. In particular, this facilitates rotational and translational tracking tasks by providing a smaller number of discrete nodes with good coverage of the regarded kinematic space.

4.1 Rotation

In cartesian coordinates a rotation in $\mathbb{R}_3$ has three degrees of freedom. This section intends to provide a better understanding of rotations and their application and presents different formats thereby.

4.1.1 Angle axis format

We can fully describe a rotation by an rotation axis $\boldsymbol{n}$ and a rotation angle α.

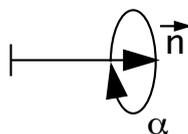Figure 4.1: Angle axis representation of rotations

A clockwise rotation $g = \alpha \, \boldsymbol{n}$ and its counter-clockwise foil $g = -\alpha \, - \boldsymbol{n}$ describe the same movement. Since we are only interested in the resulting rotation, we can limit α to $\{0...\pi\} \widehat{=} \{0...180°\}$ without any restrictions. This reduces ambiguities only to $\alpha = \pi \widehat{=} 180°$.

4.1.2   The special group of rotation matrices SO(3)

In $\mathbb{R}_3$ it is obvious that any rotation linearly transforms an orthonormal basis to another orthonormal basis. Thus we can represent rotations with an orthogonal 3×3 matrix $\boldsymbol{Q}$. Matrices of proper rotations have the determinant $det = +1$ (those of the improper rotations have the determinant $det = -1$ due to a succeeding inversion operation). All proper rotations form the special group of rotation matrices SO(3), and fulfill

$$\boldsymbol{Q}^T \boldsymbol{Q} = \boldsymbol{I}. \tag{4.1}$$

The cartesian coordinates are rotated using

$$\boldsymbol{r}' = \boldsymbol{Q}\boldsymbol{r}. \tag{4.2}$$

A distance measure of two rotation matrices $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$ is defined as [Mit07]

$$\alpha_{dist} = acos\big(\frac{1}{2}(\boldsymbol{Tr}(\boldsymbol{Q}_1\boldsymbol{Q}_2^{-1}) - 1)\big). \tag{4.3}$$

4.1.3   Euler angles

An arbitrary rotation $g = \alpha\,\boldsymbol{n}$ can be decomposed in three rotations, one rotation $g_z = \alpha_1\,\boldsymbol{e}_z$ around the z-axis, followed by a rotation $g_y = \alpha_2\,\boldsymbol{e}_y$ around the y-axis and a third rotation $g_z = \alpha_3\,\boldsymbol{e}_z$ around the new z-axis. The angles $\alpha_1, \alpha_2, \alpha_3$ are named Euler angles [Wei10].
In cartesian coordinates, the rotation is represented by three 3×3 rotation matrices $\boldsymbol{Q}_z(\alpha_1), \boldsymbol{Q}_y(\alpha_2)$ and $\boldsymbol{Q}_z(\alpha_3)$.

$$\boldsymbol{Q}_z = \begin{pmatrix} \cos(\alpha_1) & -\sin(\alpha_1) & 0 \\ \sin(\alpha_1) & \cos(\alpha_1) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \boldsymbol{Q}_y = \begin{pmatrix} \cos(\alpha_2) & 0 & -\sin(\alpha_2) \\ 0 & 1 & 0 \\ \sin(\alpha_2) & 0 & \cos(\alpha_2) \end{pmatrix}. \tag{4.4}$$

4.1.4   Mapping SO(3) to $\mathbb{S}^3$

Rotations can be represented as points on a hypersphere $\mathbb{S}^3$. A coordinate system analogue to the spherical coordinate system for the $\mathbb{R}_3$ can be defined for the $\mathbb{R}_4$ by adding another angular component. Hence the cartesian coordinates can be expressed using the angular components [Wik10]:

$$\begin{aligned} w &= \cos(\phi_1), \\ z_h &= \sin(\phi_1)\cos(\phi_2)\cos(\phi_3), \\ y_h &= \sin(\phi_1)\sin(\phi_2)\cos(\phi_3), \\ x_h &= \sin(\phi_1)\sin(\phi_2)\sin(\phi_3), \end{aligned} \tag{4.5}$$

where $\phi_1, \phi_2 = \{0...\pi\}$ and $\phi_3 = \{0...2\pi\}$.

Sampling the hypersphere with sampling nodes distributed uniformly may yield a set of rotations that have equally spaced distance relations. A sampling method is shown in section 4.1.5.

Representation on w- and x-axis

For an easier understanding of this representation, we consider a unit circle on the w- and x_h-axis in cartesian coordinates. The projection of v onto w describes the rotation angle $\alpha = 2\arccos(v_w)$ and the projection onto x_h embodies the rotation axis by

$$n = \frac{v_x}{|v_x|} e_x.$$

No rotation is described at $w = \{-1/1\}$. The above representation only formulates rotations around the positive or negative x-axis in $\mathbb{R}_3$. According to section 4.1.1 it is still achievable to formulate all possible rotations around e_x if we reduce the unit circle to the upper half. As in eq. (4.5) we can scale the normalized rotation axis by $\cos(\phi_1)$ with $\phi_1 = \{0...\pi\}$.
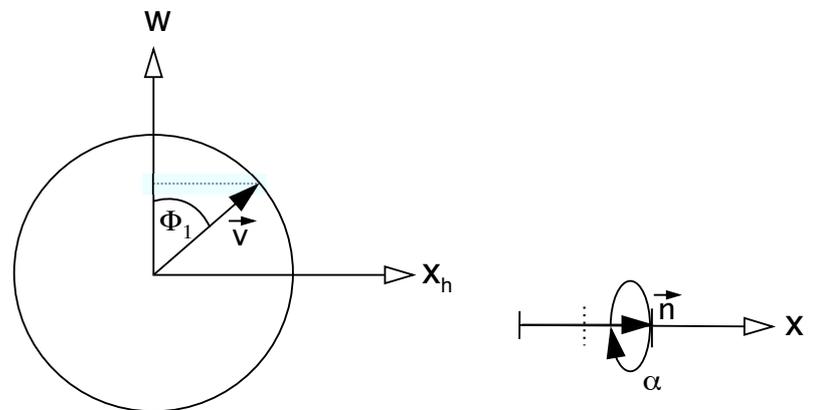
(a) 2 dimensions of the hypersphere representation

(b) Angle axis format in $\mathbb{R}_1$

Figure 4.2: Rotation representation on w- and x_h-axis

Representation on w-axis and x_h-y_h-plane

If we extend the representation to another axis y_h we consider a unit sphere. Again the projection of v onto w describes the rotation angle and the projection onto the x_h-y_h-plane formulates the normalized rotation axis to

$$n = \frac{v_x e_x + v_y e_y}{\sqrt{v_x^2 + v_y^2}}.$$

The north- and the south-pole of the unit sphere describe no rotation at all. Therefore we do not need to differentiate between different orientations of the rotation axis. At all other positions the rotation axis lies in the x-y-plane in $\mathbb{R}_3$. Reducing the unit sphere to a half sphere still allows us to describe every possible rotation around an axis on the x-y plane. The normalized rotation axis can be scaled with $\cos(\phi_1)$ with $\phi_1 = \{0...\pi\}$.

(a) 3 axes of the hypersphere representation   (b) Angle axis format in $\mathbb{R}_2$
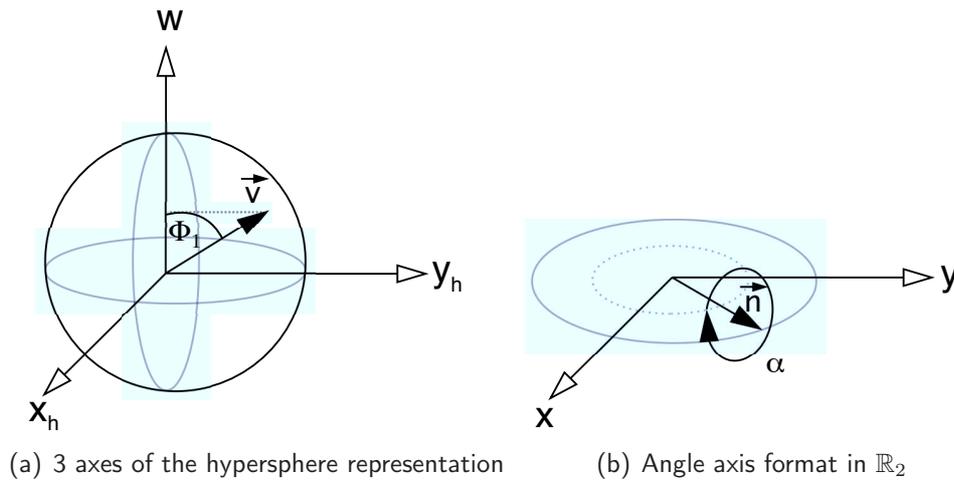
Figure 4.3: Representation on w-axis and x_h-y_h-circle

Representation on w-axis and x_h-y_h-z_h-sphere

Finally, we add another axis z_h and consider a unit hypersphere. Now, all rotation axes in $\mathbb{R}_3$ are allowed

$$n = \frac{v_x e_x + v_y e_y + v_z e_z}{\sqrt{v_x^2 + v_y^2 + v_z^2}}.$$

Again, we can reduce the unit hypersphere to positive w only and still gather all possible rotations in $\mathbb{R}_3$. The normalized rotation axis can be scaled with $\cos(\phi_1)$ with $\phi_1 = \{0...\pi\}$.

(a) 4 axes of the hypersphere representation (w not shown)
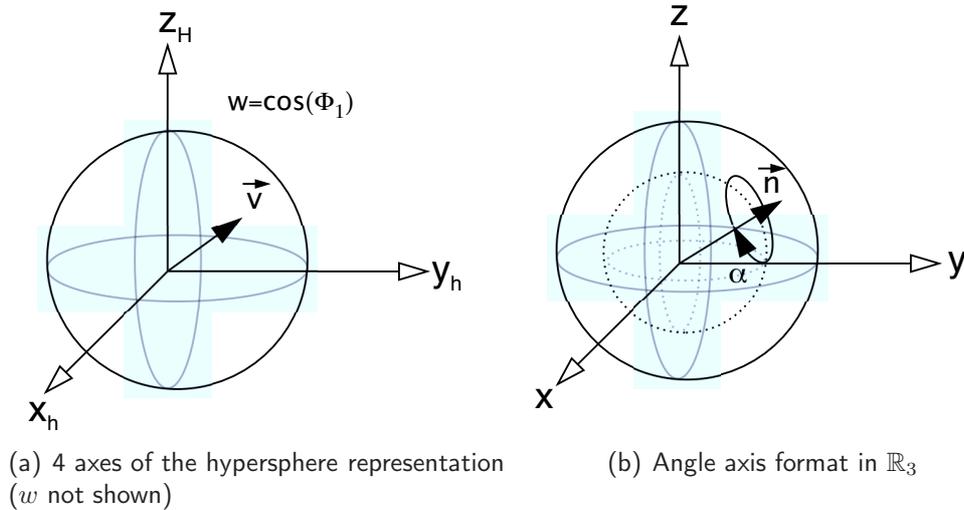
(b) Angle axis format in $\mathbb{R}_3$

Figure 4.4: Representation on w-axis and x_h-y_h-z_h-sphere

4.1.5 Sampling the hypersphere $\mathbb{S}^3$

The representation of rotations on a hypersphere becomes useful if we want to obtain a set of rotations which are uniformly distributed in the sense of the distance between rotation matrices. The longitudes represent different rotation axes. They converge at the north- and south-pole and diverge towards the equator. By uniformly sampling the hypersphere representation, rotation-axes directions are discretized using more finely-resolved directions for large rotations than for small ones.

Equal area partitioning

Equal area partitioning provides a near-uniform sampling of a sphere $\mathbb{S}^n$. It follows the objective of partitioning the sphere in equal area rectangular compartments. The $\mathbb{S}^n$ sphere can be partitioned into a freely chosen number of sampling points. [Leo06]. Fig. 4.1.5 shows an example of equal area partitioning of a $\mathbb{S}^2$ sphere.
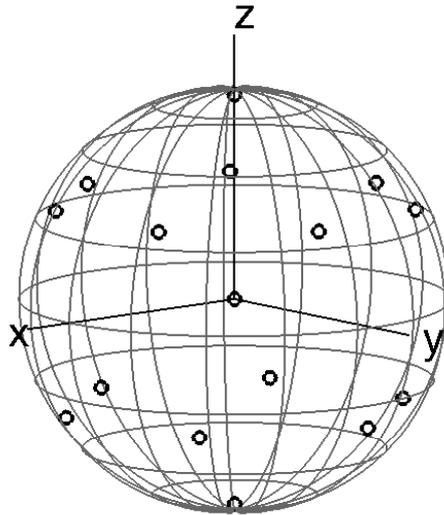
Figure 4.5: Example of equal area partitioning using the recursive zonal equal area partitioning toolbox for Matlab offered under the MIT license [Leo10]

Rotational testset

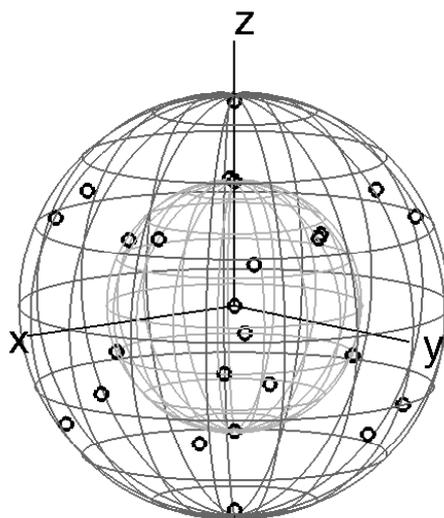With the above representation we can illustrate a set of rotations in $\mathbb{R}_3$:

Figure 4.6: Example rotational test set (three layers representing $\alpha = \{0, \frac{\pi}{2}, \pi\}$, the layers have an increasing number of sampling nodes representing rotation axes)

4.2   Translation

A translation is a simple shift of the coordinate system by a vector $\boldsymbol{d}$ in cartesian coordinates. A position vector in the shifted coordinate system is denoted by:

$$\boldsymbol{r}' = \boldsymbol{r} + \boldsymbol{d} \tag{4.6}$$

Sampling spherical volumes

In order to sample the inner volume of the microphone array we need to choose a suitable sampling method. The preferable method may depend whether we need linear distances in cartesian coordinates or in spherical coordinates. Therefore a few methods have been chosen to fulfill the demands of different applications.

Linear sampling in cartesian coordinates

The most obvious method seems to be the cubic grid obtained by sampling a volume with discrete incremental distances d_x, d_y, d_z from the origin (0,0,0).

Linear discretization of z-axis and rotation with equally spaced distance

Every position $\boldsymbol{r}$ on the sphere $\mathbb{S}^2$ can be rotated to the z-axis of the coordinate system by a rotation axis on the x-y-plane being orthogonal to the z-axis and the position vector $\boldsymbol{r}$. It is therefore sufficient to linearly sample the radius r of the sphere by a given d_z on the z-axis.

$$\boldsymbol{d}_z = \begin{pmatrix} d_z & 2d_z & 3d_z & \dots & sd_z \end{pmatrix} \tag{4.7}$$

A set of rotation axes can be obtained with the equal area partition method for $\mathbb{S}^2$ described in section 4.1.5. To achieve uniformly distribution of sampling nodes of the volume V, the number of rotations has to grow with $\boldsymbol{d}_z$.

Cubic close sphere packing

There are different possibilities for packing spheres [Mat10b]. The tightest sphere packings are the cubic and the hexagonal close packing illustrated in fig. 4.7. Both yield a density of 74.048 percent. Fig. 4.8 shows an example of the sampling nodes obtained by cubic tightest sphere packing of a spherical volume.

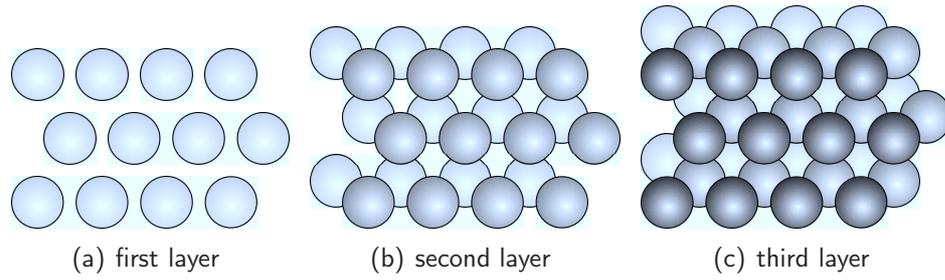(a) first layer (b) second layer (c) third layer

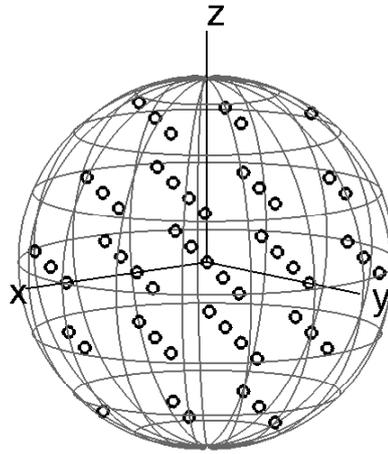Figure 4.7: Layers of cubic close sphere packing (2 layers) and hexagonal close sphere packing (3 layers)

Figure 4.8: Example of cubic tightest sphere packing of a spherical volume.

4.3 Re-expansion of spherical base solutions

A detailed description of the numerical computation of transformed spherical base solutions can be found in [Zot09]. The following recurrence relations are adopted without detailed explanation,

$$
a_n^m = \begin{cases} \sqrt{\frac{(n-|m|+1)(n+|m|+1)}{(2n+1)(2n+3)}}, & \text{for } n \geq 0 \text{ and } -n \leq m \leq n, \\ 0, & \text{else}, \end{cases} \tag{4.8}
$$

$$
b_n^m = \begin{cases} \sqrt{\frac{(n-m-1)(n-m)}{(2n-1)(2n+1)}}, & \text{for } n \geq 0 \text{ and } 0 \leq m \leq n, \\ -\sqrt{\frac{(n-m-1)(n-m)}{(2n-1)(2n+1)}}, & \text{for } n \geq 0 \text{ and } -n \leq m \leq 0, \\ 0, & \text{else}. \end{cases} \tag{4.9}
$$

4.3.1 Rotation of spherical base solutions

Rotations can be decomposed into a combination of rotations around the z-axis and $\frac{\pi}{2}$-y-axis according to

$$\boldsymbol{r}' = \boldsymbol{Q}_z(\alpha_1 + \pi/2)\boldsymbol{Q}_{y\frac{\pi}{2}}\boldsymbol{Q}_z(\alpha_2 + \pi)\boldsymbol{Q}_{y\frac{\pi}{2}}\boldsymbol{Q}_z(\alpha_3 + \pi/2)\boldsymbol{r}. \tag{4.10}$$

This decomposition simplifies the transforms of the spherical base solutions. The two types of rotations are expressed below.

Transform for z-rotation:

$$c_{nm} = e^{im\theta_z} c'_{nm}. \tag{4.11}$$

Transform for $y\frac{\pi}{2}$-rotation:

$$c_{nm} = \sum_{n=0}^{N} \sum_{m'=-n}^{n} T_{n,n}^{m',m}(\boldsymbol{Q}_{y\frac{\pi}{2}})c'_{nm'}, \tag{4.12}$$

with the generating equation of the coefficients

$$T_{n-1,n-1}^{m',m+1}(\boldsymbol{Q}_{y\frac{\pi}{2}}) = \frac{\sqrt{2 - \delta_{m+1}}}{2b_n^m \sqrt{2 - \delta_m}} \Bigg\{ \sqrt{2 - \delta_{m'}} \Bigg\{ \frac{b_n^{m'-1}}{\sqrt{2 - \delta_{m'-1}}} T_{n,n}^{m'-1,m}(\boldsymbol{Q}_{y\frac{\pi}{2}}) $$
$$- \frac{b_n^{-m'-1}}{\sqrt{2 - \delta_{m'+1}}} T_{n,n}^{m'+1,m}(\boldsymbol{Q}_{y\frac{\pi}{2}}) \Bigg\} + 2a_{n-1}^{m'} T_{n,n}^{m',m}(\boldsymbol{Q}_{y\frac{\pi}{2}}) \Bigg\}, \tag{4.13}$$

using the initial values

$$T_{n,n}^{m',0} = \sqrt{\frac{4\pi}{2n + 1}} N_n^{|m'|} P_n^{|m'|}(0). \tag{4.14}$$

4.3.2 z-Translation of spherical base solutions

Performing translations of spherical base solutions is not as trivial as translations are in cartesian coordinates. It turns out that a single spherical base solution in the original coordinate system is an expansion of an infinite number of spherical base solutions in the dislocated coordinate system. Translations along the z-axis can be performed according to the following recurrence relations. An arbitrary translation vector $\boldsymbol{d}$ can easily be rotated onto the z-axis by a rotational transform $\boldsymbol{d} = \boldsymbol{Q}(\varphi_d, \vartheta_d, 0) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \|\boldsymbol{d}\|$ of the coordinate system.

$$c_{nm} = \sum_{n'=0}^{N'} \sum_{m'=-n}^{n} T_{n'n}^{m'm}(d_z) c_{n'm'}', \tag{4.15}$$

$$\begin{pmatrix} T_{n',n+1}^{m+1,m+1} \\ T_{n',n+1}^{m,m} \end{pmatrix} = \begin{pmatrix} \frac{1}{b_{n+1}^{-m-1}} \Big[-b_{n'+1}^{m} T_{n'+1,n}^{mm}(d_z) + b_{n'}^{-m-1} T_{n'-1,n}^{mm}(d_z) + b_n^m T_{n',n-1}^{m+1,m+1}(d_z) \Big] \\ \frac{1}{a_n^m} \Big[-a_{n'}^{m'} T_{n'+1,n}^{mm}(d_z) + a_{n'-1,n}^{m} T_{n'-1,n}^{mm}(d_z) + a_{n-1}^m T_{n',n-1}^{mm}(d_z) \Big] \end{pmatrix}, \tag{4.16}$$

using the initial values for d_z

$$T_{n'0}^{00}(d_z)\boldsymbol{I} = \sqrt{2n'+1} j_n(kd_z). \tag{4.17}$$

Chapter 5

Sources In Spherical Microphone Arrays

As we deal with musical instruments, we have no exact idea of how the wave spectrum is distributed. However, we may assume a set of distributions that are likely for certain characteristics of different instruments. For example, most brass instruments have a focused radiation in the direction of their cone, a flute exhibits a dipole-like radiation pattern because of two effective points of radiation and string instruments a more complex distribution since they have more parts radiating sound [Mey72].
For the evaluation of tracking algorithms it is essential to have sound sources of known positions. It is a difficult task to mount radiating sound-sources at a certain position in the surrounding microphone array, therefore we will use simulated sound-sources for this purpose.

5.1 Spherical multipole expansion

We can describe the radiation from bodies of finite extent with an expansion of multipoles [Wil99]. According to [GD04] multipoles are defined as derivatives of the Green function of the fundamental solution of the Helmholtz equation. Using the differentiation theorems for spherical elementary solutions, multipoles can be related to the spherical harmonics functions Y_n^m. Thus, an expansion of multipoles is similar to the spherical harmonics expansion written as [Wil99]

$$p(kr, \boldsymbol{\theta}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} c_{nm} h_n^{(1)}(kr) Y_n^m(\boldsymbol{\theta}). \tag{5.1}$$

In a manner, the above equation may be used to generate different multipole sources which are believed to resemble the behavior of musical instruments. Examples of such sources are given in section 5.3.

5.2 Higher orders due to shifted sources

The multipole coefficients of a source located at the origin of the coordinate system denotes as

$$M_n^m(\boldsymbol{r}) = h_n^{(1)}(kr)Y_n^m(\boldsymbol{\theta}). \tag{5.2}$$

With the recurrence relations given in section 4.3.2 a displacement of the multipole expansion can be performed with

$$M_n^m(\boldsymbol{r}) = \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} T_{n'n}^{m'm}(\boldsymbol{d})M_n'^m(\boldsymbol{r'}). \tag{5.3}$$

The translation of multipole coefficients again is a linear combination of an infinite sum of spherical harmonics. According to [GD04] the additional orders required for an accurate resolution of a shifted multipole is roughly estimated by

$$N_{analyze} \geq N_s + kd. \tag{5.4}$$

5.3 Examples of multipole expansions and their spectral distribution

In order to simulate the performance of acoustic centering algorithms, some types of sources shall be represented applicable as test set.
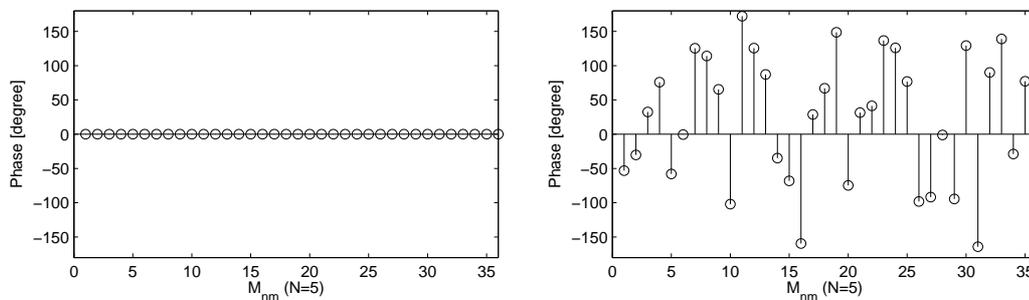

Figure 5.1: Comparison of phase relations in coefficient vector (left: constant phase, right: random phase)
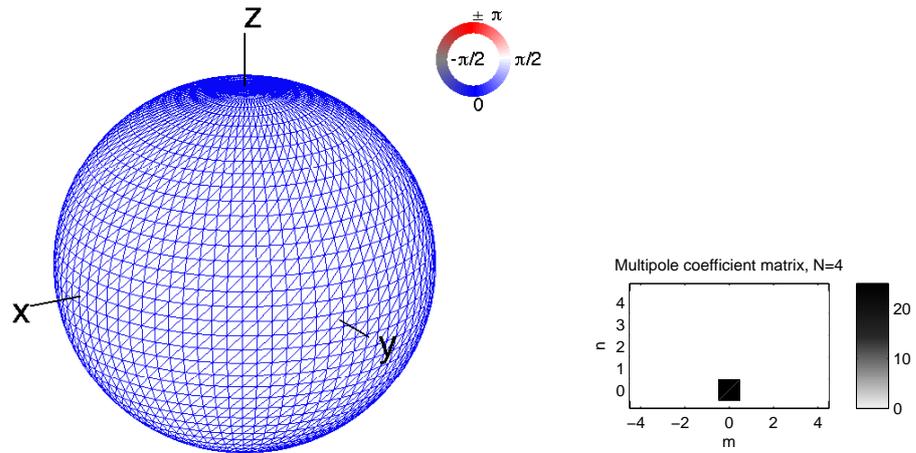
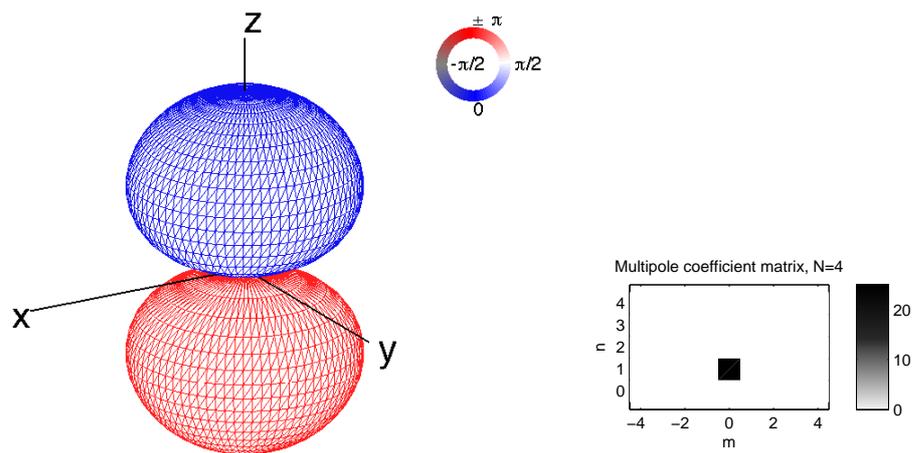Figure 5.2: Monopole radiation pattern and its spectral distribution

Figure 5.3: Dipole on z-axis, radiation pattern and its spectral distribution
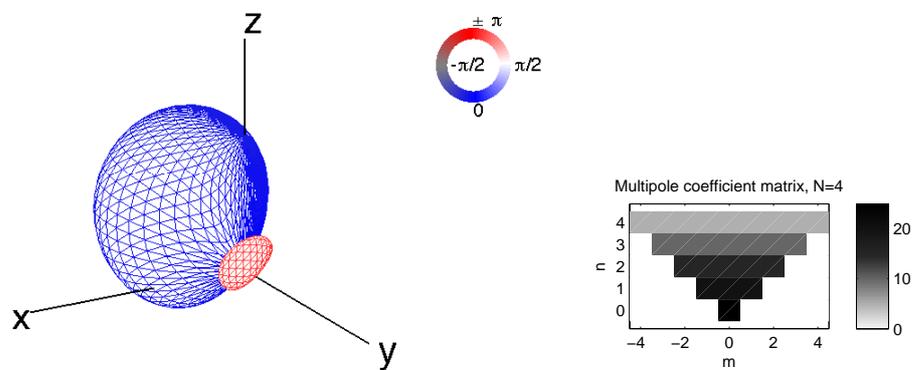

Figure 5.4: Example of a real-valued multipole radiation pattern with a distribution $\sim 1/n$ and its spectral distribution
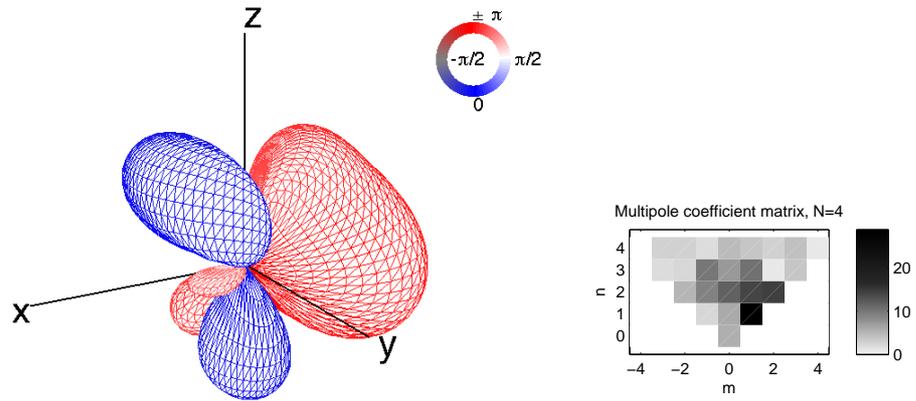
Figure 5.5: Example of a real-valued multipole radiation pattern with random coefficients and its spectral distribution
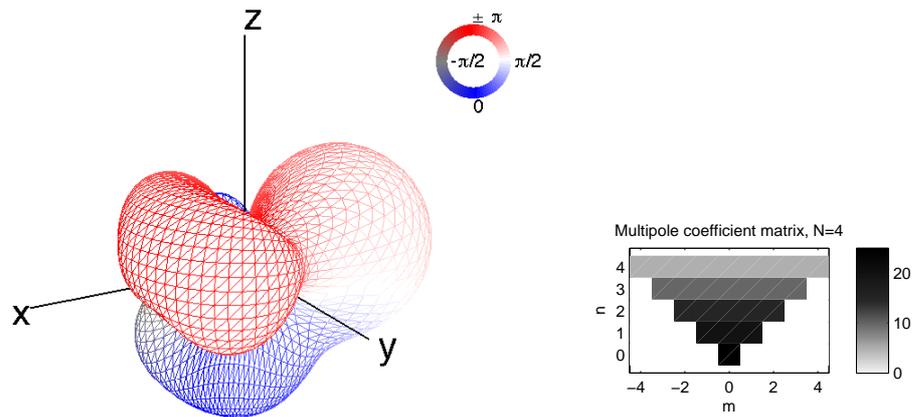

Figure 5.6: Example of a complex-valued multipole radiation pattern with a distribution $\sim 1/n$ and its spectral distribution
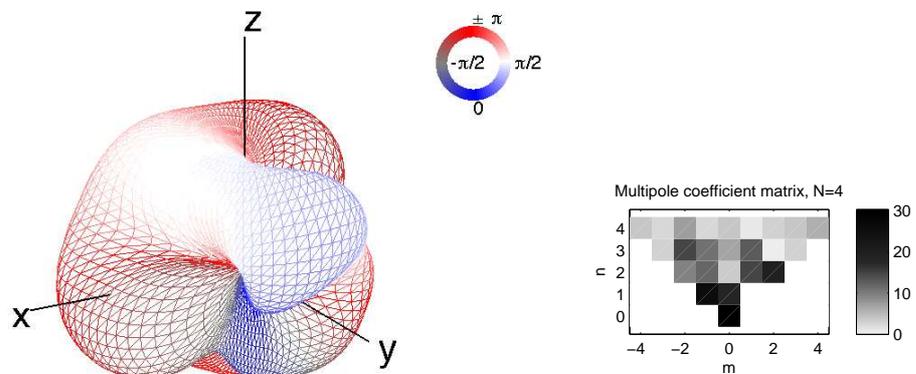

Figure 5.7: Example of a complex-valued multipole radiation pattern with random coefficients and its spectral distribution
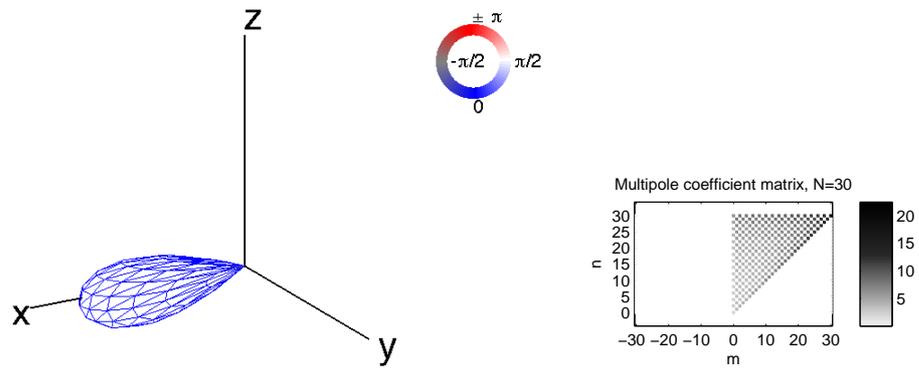
Figure 5.8: Example of a focusing multipole radiation pattern and its spectral distribution

Chapter 6

Rotational Tracking Of Sound Sources

This chapter deals with the practical tests of rotational tracking of moving sound sources. The assumption that musical instruments do not move fast allows to divide the recorded time-signal into multiple successive blocks of a constant time period. The Fourier coefficient $p^b(kr, \boldsymbol{\theta})$ for one block can be used for discrete spherical harmonic transform. A spherical correlation function is used to compare the spherical harmonics magnitude-patterns of two successive blocks. Rotational matching is achieved by rotating the older spherical harmonics magnitude-pattern and maximizing its spherical correlation to the newer pattern. The approach is first evaluated with a simulated complex-valued multipole source and finally an evaluation is demonstrated using real-world data from a recorded musical instrument.

6.1 Rotational matching

Rotational matching is done with a set of rotations obtained by the hypersphere sampling method described in section 4.1.5. At first a reference block needs to be found that can be used for comparison of its pattern to the patterns of subsequent blocks. Once calculated the estimated spherical harmonics decomposition of the reference block in all discrete rotations, the spherical harmonics decomposition of a subsequent block can be matched. Both spherical harmonics decompositions are compared by evaluating the correlation function

$$c(\boldsymbol{n}, \theta) = \frac{\boldsymbol{\gamma}_N^H \boldsymbol{\gamma}_N'(\boldsymbol{n}, \theta)}{\| \boldsymbol{\gamma}_N \| \| \boldsymbol{\gamma}_N' \|} \tag{6.1}$$

If the correlation function becomes maximal, the estimated rotation between both blocks has been found. This procedure is continued for every block of the recorded signal.
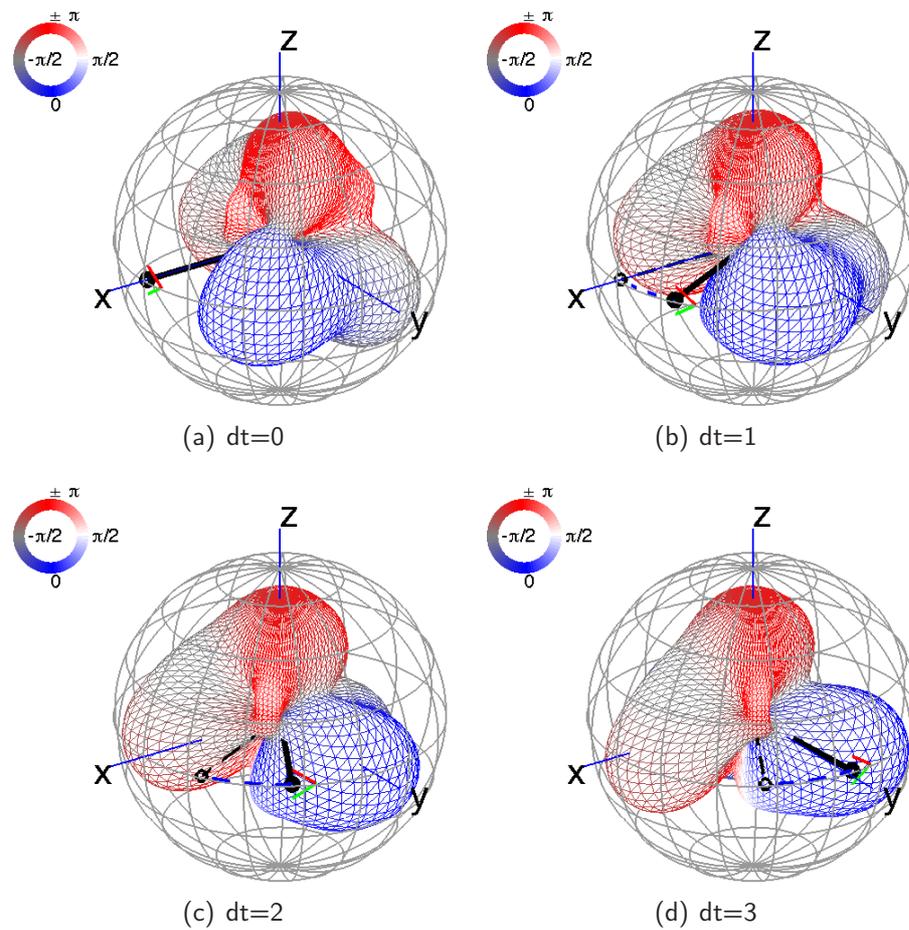
6.2   Simulation example

(a) dt=0

(b) dt=1

(c) dt=2

(d) dt=3

Figure 6.1: Example of rotational tracking, simulated complex valued multipole source rotating around z-axis ($f = 1000Hz$)

6.3   Experimental example
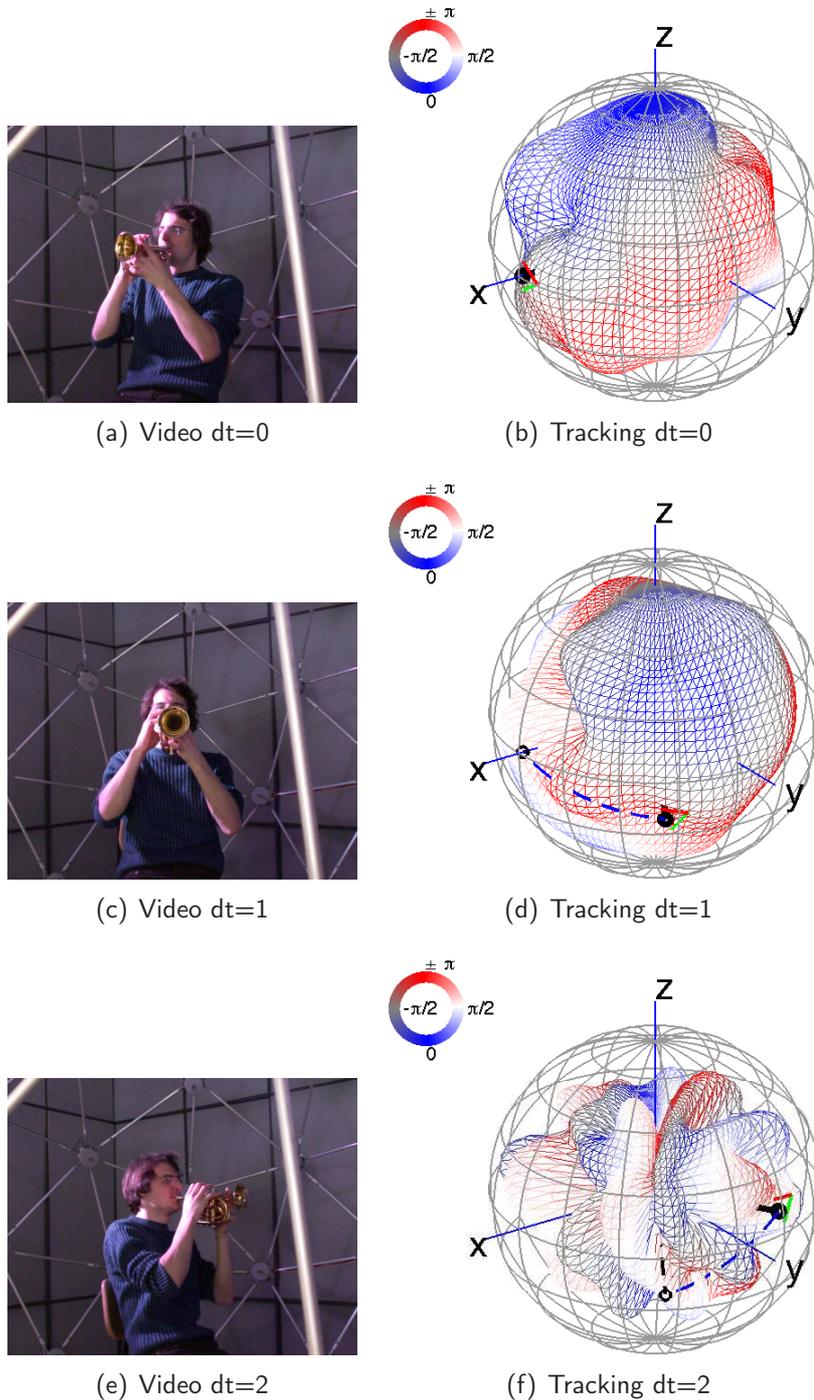

(a) Video dt=0

(b) Tracking dt=0

(c) Video dt=1

(d) Tracking dt=1

(e) Video dt=2

(f) Tracking dt=2

Figure 6.2: Example of rotational tracking, trumpet rotating around z-axis while playing Bb4, partial at $718Hz$

Chapter 7

Translational Tracking Of Sound Sources

Using the prototypical sources described in chapter 5, novel translational tracking approaches are being evaluated in this chapter for the static case.

7.1 Defining the acoustic center of musical instruments

As becomes clear in the previous chapter, radiating sound sources of finite extent can be expressed by a multipole expansion at the origin. The simplest form is an omnidirectional radiating source, i.e. a point source.

Defining an acoustic center can easily be done for loudspeakers radiating low frequencies where they become omnidirectional as long as the wavelength is considerably larger than the diameter of its baffle [Van07].

We may define the acoustic center of a radiating source as the position a at which higher order components become small. For a dislocated point source the acoustic center will then coincide with the point of its origin.

However, defining the acoustic center of musical instruments is not a simple task if not impossible. Musical instruments have lots of different vibrating parts which radiate sound to the air. We assume that the multipole expansion of a musical instrument may be quite complex.

7.2 Acoustic centering

Since a dislocated source outside the origin of the coordinate system can be represented by translation of its wave spectrum, we may transform back the acoustic center of a source by applying the so-called multipole-translations of the wave spectrum by $-a$. We will use the term *acoustic centering* for this procedure. Fig. 7.1 illustrates this

approach starting from the discrete pattern p_L of a dislocated source that is translated back to the origin of a surrounding microphone array.
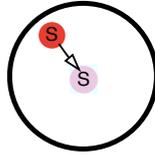

Figure 7.1: Scheme for acoustic centering

7.3 Proposed criteria

We will examine two different approaches to find the unknown acoustic center a within the surrounding microphone array. The first one takes into account the measured sound pressure and its phase relations. The second one directly implies the above-mentioned definition of the acoustic center. Cost functions will help us to evaluate the performance of these approaches.

7.3.1 Squared sum criterion J_{ssc}

A point source produces spherical waves radiating from its origin. We can see spheres of coherent phase for the sound pressure. Fig. 7.2 shows a monopole source being shifted out of the origin. The array will cut through different shells of coherent phase. Summing up the sound-pressure of all microphones can therefore lead to phase cancellations so the complex sum of the sound-pressures becomes smaller than in the original position. We assume the normalized sum of sound-pressures becomes maximal at the position a of the acoustic center.
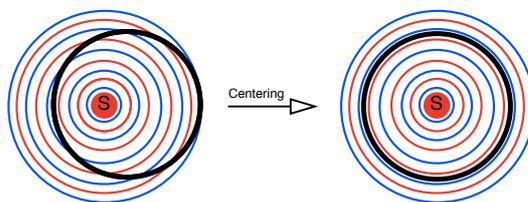

Figure 7.2: Different shells of coherent phase radiated from an omni-directional source (left: dislocated, right: centered)

For multipole expansions with real-valued coefficients that may have a common complex multiplier we can assume that the sum of the complex-squared sound pressure will become maximal at the acoustic center. As shown later, arbitrarily complex weighted multipoles may cause the criterion to fail in detecting a sensible acoustic center.

A cost function considering the sum of complex-squared sound pressure at all microphone positions can be written as

$$J_{ssc}(\boldsymbol{d}) = 1 - \left| \frac{\boldsymbol{p'}_L^T \boldsymbol{p'}_L}{\boldsymbol{p'}_L^H \boldsymbol{p'}_L} \right| = 1 - \left| \frac{\sum \boldsymbol{p'}_L^2}{\sum |\boldsymbol{p'}_L|^2} \right|. \tag{7.1}$$

7.3.2 Low order criterion J_{loc}

The coefficients c_{nm} of a wave spectrum express a weight for each order n and its degree m. Regarding the definition of the acoustic center, higher orders should be increasingly penalized in a cost function. Eq. (7.3) is a cost function in which each weight of the wave spectrum is multiplied by its order n. This corresponds to the concept of the actuarial expectation value and can be seen as the centre of mass of the order n. Its value should become minimal at the position $\boldsymbol{a}$ of the acoustic center.

$$w_n = n \tag{7.2}$$

$$J(\boldsymbol{d}) = \frac{\boldsymbol{c'}_N^H diag_N\{w_n\} \boldsymbol{c'}_N}{\boldsymbol{c'}_N^H \boldsymbol{c'}_N} \tag{7.3}$$

7.4 Examining the cost functions

In this section we will examine the behavior of the cost functions using simulated multipole sources presented in chapter 5 at different positions $\boldsymbol{a}$. The cost functions will be evaluated at different displacements $\boldsymbol{d}$. For a better readability only displacements along one axis of the coordinate system have been chosen. We would like to assume that the different structure of aliasing artifacts for other displacements are similar due to the uniform (however not regular) structure of the investigated spherical array geometry.

7.4.1 Translation comparison

This section presents two ways of obtaining an estimation of the spherical wave spectrum at a displaced origin.

Estimation of spherical wave spectrum at translated observation nodes As displacements are expressed as an addition in cartesian coordinates, it might look favorable to estimate the wave spectrum at translated cartesian coordinates.
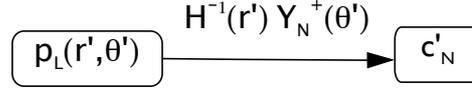
$$\boxed{p_L(r',\theta')} \xrightarrow{\quad H^{-1}(r')\ Y_N^{+}(\theta') \quad} \boxed{c'_N}$$

Figure 7.3: Scheme for translation in cartesian coordinates

The displacement of the cartesian coordinates affects the angular coordinates,

$$\mathring{r} = \begin{pmatrix} r' \\ \varphi' \\ \vartheta' \end{pmatrix} = \begin{pmatrix} \sqrt{(x+d_x)^2+(y+d_y)^2+(z+d_z)^2} \\ \arctan(\frac{(y+d_y)}{(x+d_x)}) \\ \arctan \frac{\sqrt{(x+d_x)^2+(y+d_y)^2}}{(z+d_z)} \end{pmatrix}, \tag{7.4}$$

and hyperinterpolation might become infeasible. In fact, the geometric distortion of the angular discretization nodes for hyperinterpolation no longer guarantees a good condition number. Thus we have to be careful with the matrix inversion.

Regularization using singular value decomposition as described in section 3.1.3 can be used to achieve a pseudo-inverse Y_N^{+}.

Translated spherical wave spectrum from estimate at original observation nodes   The second approach operates directly on the wave spectrum and preserves hyperinterpolation at its designated sampling nodes. It applies the multipole translations for the displacement d.

$$\boxed{p_L(r,\theta)} \xrightarrow{\quad H^{-1}(r)\ Y_N^{-1}(\theta) \quad} \boxed{c_N} \xrightarrow{\quad T_{n'n}^{m'm}(d) \quad} \boxed{c'_{N'}}$$
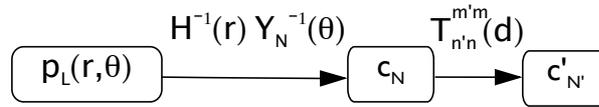
Figure 7.4: Scheme for translation method in the wave spectrum

In the following example singular-values s_i 20dB below s_0 have been cut off for the translation in spherical coordinates. The artifacts this regularization produces in the cost function is clearly observable for larger displacements d, cf. fig 7.5. The example shows a displacement along the z-axis. For this purpose the cost functions are evaluated step-wise for translations reaching from $\{-1.35...1.35\}m$ along the z-axis. As the source, a monopole was placed at the origin. For large displacements $d_z \to 1.35m$ the number of omitted singular values nearly approaches $(N+1)^2$. For this extreme case, the cost function loses its validity.

(a) J_{loc} with vs without regularization
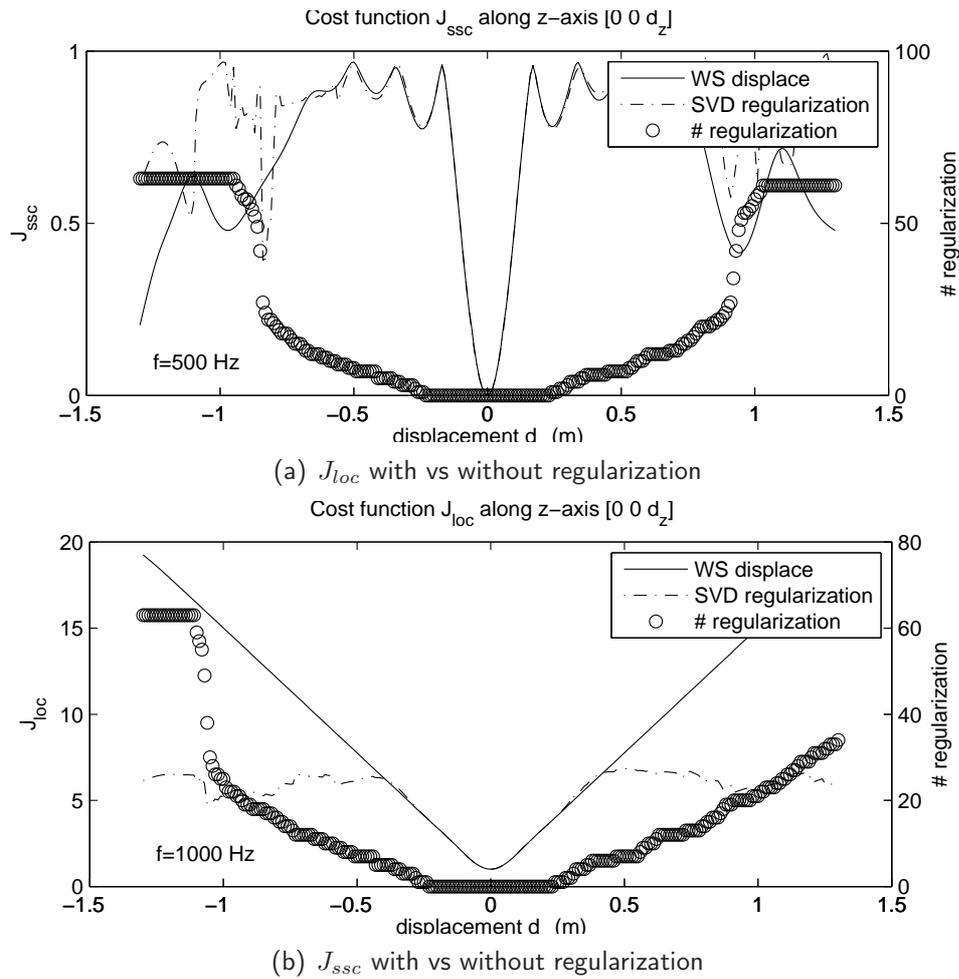

(b) J_{ssc} with vs without regularization

Figure 7.5: Effects on the cost function due to the regularization of the wave spectrum estimation

7.4.2   Truncation and aliasing effects

Truncation of acoustically centered coefficients

Obviously, the translation approach in the wave spectrum outperforms the translation of the cartesian coordinates. From section 4.3 we know that shifting in wave spectra is done by a linear combination of an infinite number of spherical base solutions. The coefficient vector c'_N therefore becomes infinitely large and must be truncated for numerical computations. The following example shows the effect of this truncation on the cost functions. If we truncate the coefficient vector c'_N at order N the translated wave spectrum c'_N lacks the components at higher orders generated through translation. This inaccuracy has much more influence on J_{loc} than on J_{ssc} because J_{loc} is mainly based on penalizing the higher order components. Thus, a large order N' is advisable the larger displacements and the higher the frequencies become according to eq. (5.4). The example in fig. 7.6 shows the results on the cost functions for a monopole source at the

origin.
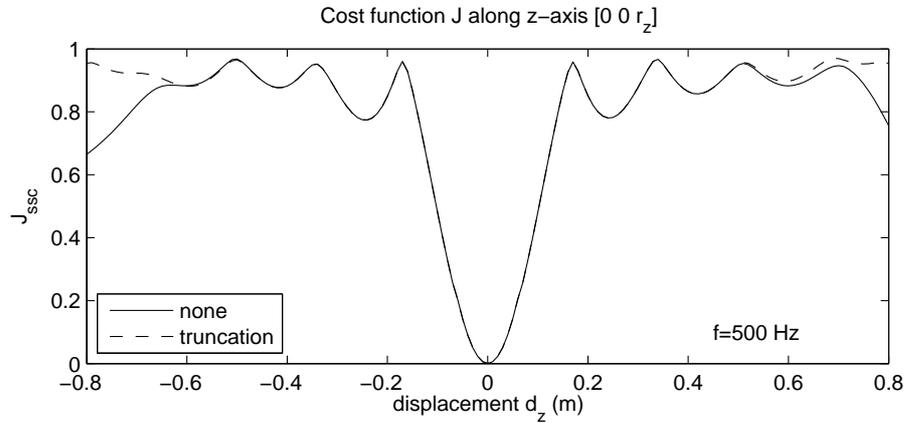
(a) J_{ssc} with vs. without truncation in displacement calculation
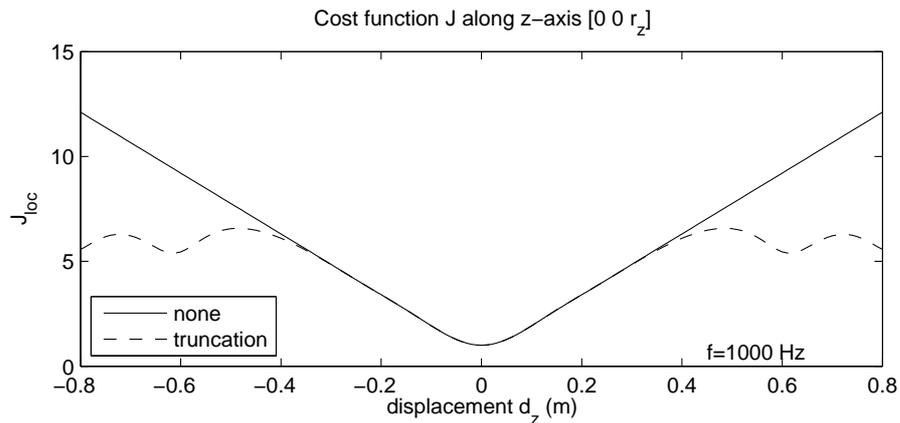
(b) J_{loc} with vs. without truncation in displacement calculation

Figure 7.6: Cost functions with vs. without truncation in the calculation of the displacement, $N' \gg N$

Truncation of source coefficients

An aspect we cannot get rid of in real recordings of the array is the spatial aliasing error in the discrete wave spectrum estimation. However, the aliasing error can be avoided in simulations in order to investigate the effect of aliasing on the cost functions. Aliasing emerges from higher order components ($N_s > N$) that cannot be correctly identified. In the following example, a dislocated monopole source has been used to illustrate the aliasing error. Since the coefficient vector of the dislocated source can also be truncated in the simulation, the effect of truncation is also visible in the example, cf. fig. 7.7.
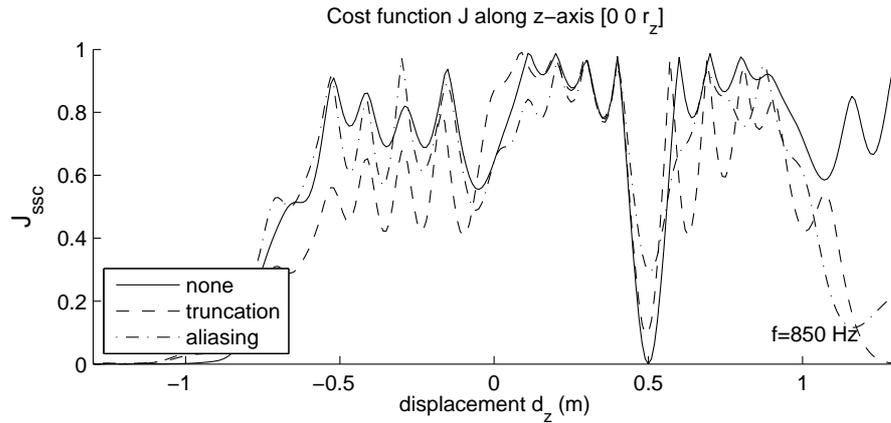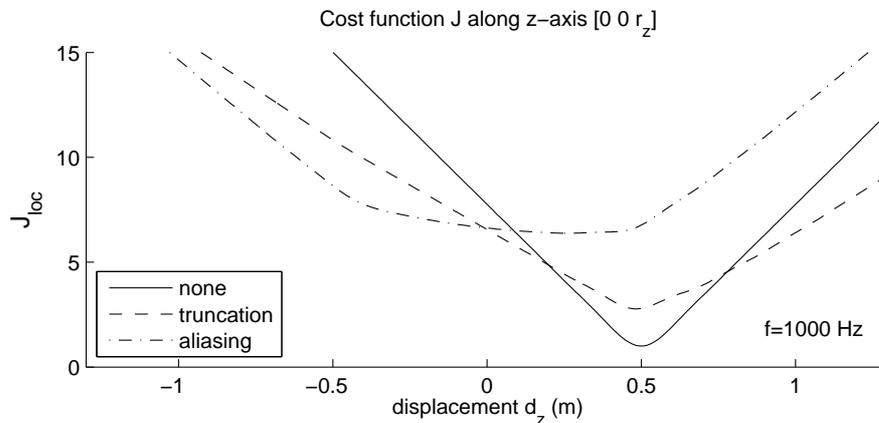
Cost function J along z-axis [0 0 r$_z$]

(a) J_{ssc} aliasing error of a dislocated source

Cost function J along z-axis [0 0 r$_z$]

(b) J_{loc} aliasing error of a dislocated source

Figure 7.7: Simulated dislocation of the source with $N_s \gg N$, truncation at N, and aliasing effect

7.4.3 Cost functions evaluated for different sources along z-axis

Now that we have seen different effects of errors influencing the cost functions we can take a closer look at the behavior for different sources located at the origin of the array. Displacements have been calculated ranging up to a radius of $r = 1.3m$ which is nearly the whole diameter of the array frame. If the translation approaches the array boundary we sometimes observe that the sound-pressure based cost function J_{ssc} becomes minimal. This may occur if a sampling node comes close to the near-field of the sound-source (see Fig. 2.2).
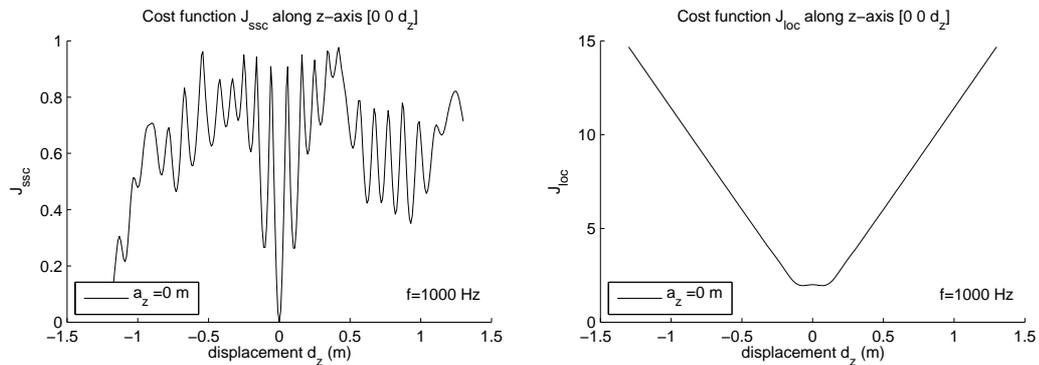
Monopole source
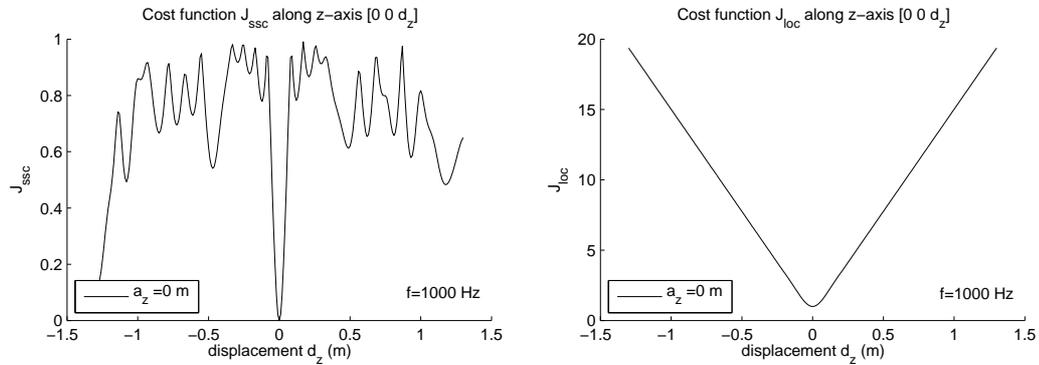
(a) J_{ssc} cost function for a monopole source   (b) J_{loc} cost function for a monopole source

Figure 7.8: For a monopole source the J_{ssc} function becomes very sharp at the origin with a lot of wiggles whereas the J_{loc} remains convex over the whole displacement range. On the other hand the J_{loc} criterion seems to be shallower as the J_{ssc} criterion.

Dipole source

(a) J_{ssc} cost function for a dipole source along z-axis   (b) J_{loc} cost function for a dipole source along z-axis

Figure 7.9: In both cost functions one can guess the dipole-like radiation of the sound-source. J_{ssc} becomes minimal at three positions of which one is located at the origin, whereas J_{loc} has two minima, both at a small displacement. It can be assumed that both plots show the positions where two monopoles with an inverse phase may yield a dipole-like radiation pattern. In the case of J_{ssc} the minimum at the origin should vanish of we evaluate the sum of p_L instead of the complex-squared sum of p_L. This may be beneficial for the purpose of monopole-source synthesis.
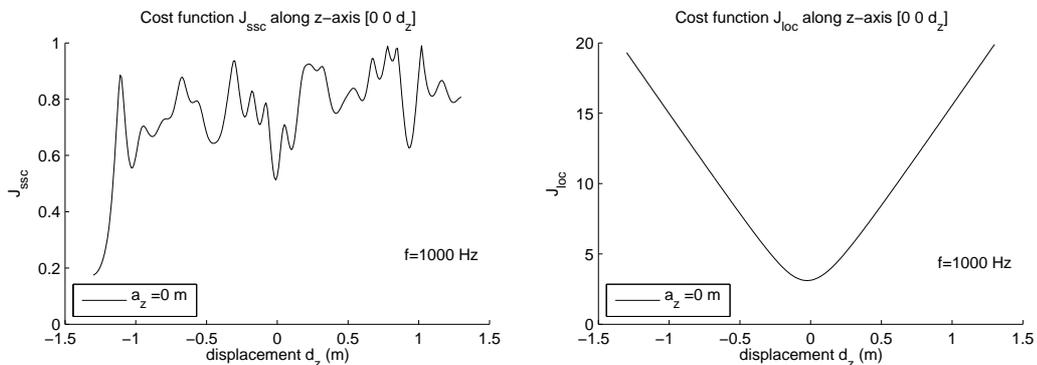
Real-valued multipole source
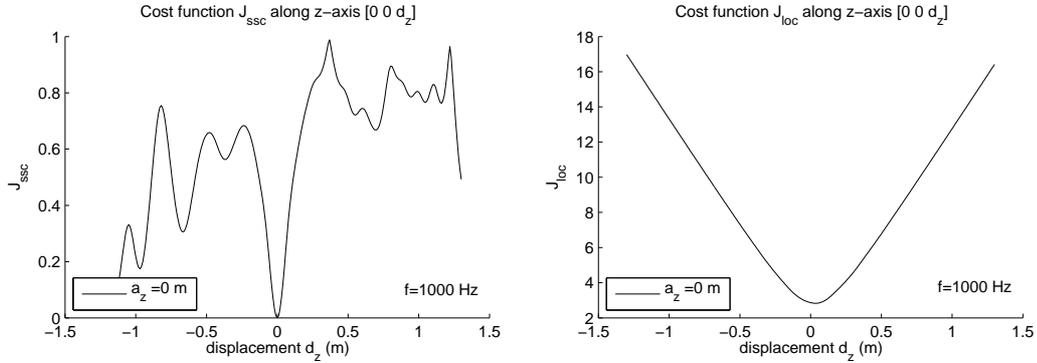
(a) J_{ssc} cost function for real-valued multipole, $N_s = 4$

(b) J_{loc} cost function for real-valued multipole, $N_s = 4$

Figure 7.10: For real-valued multipoles both criteria work pretty well and become minimal at the origin. A small displacement of the J_{loc} function can be guessed that is assumed to arise because of a directionally radiating source.

Complex-valued multipole source

(a) J_{ssc} cost function for complex-valued multipole $N_s = 4$

(b) J_{loc} cost function for complex-valued multipole $N_s = 4$

Figure 7.11: For a complex-valued multipole with random phase distribution we see the J_{ssc} fail, whereas the J_{loc} criteria still has a convex behavior and the minimum seems to be located exactly at the origin.

Cost functions for different frequencies

When calculating the cost functions at higher frequencies of the sound source we can see that in both cases the curves become steeper, cf. fig. 7.12
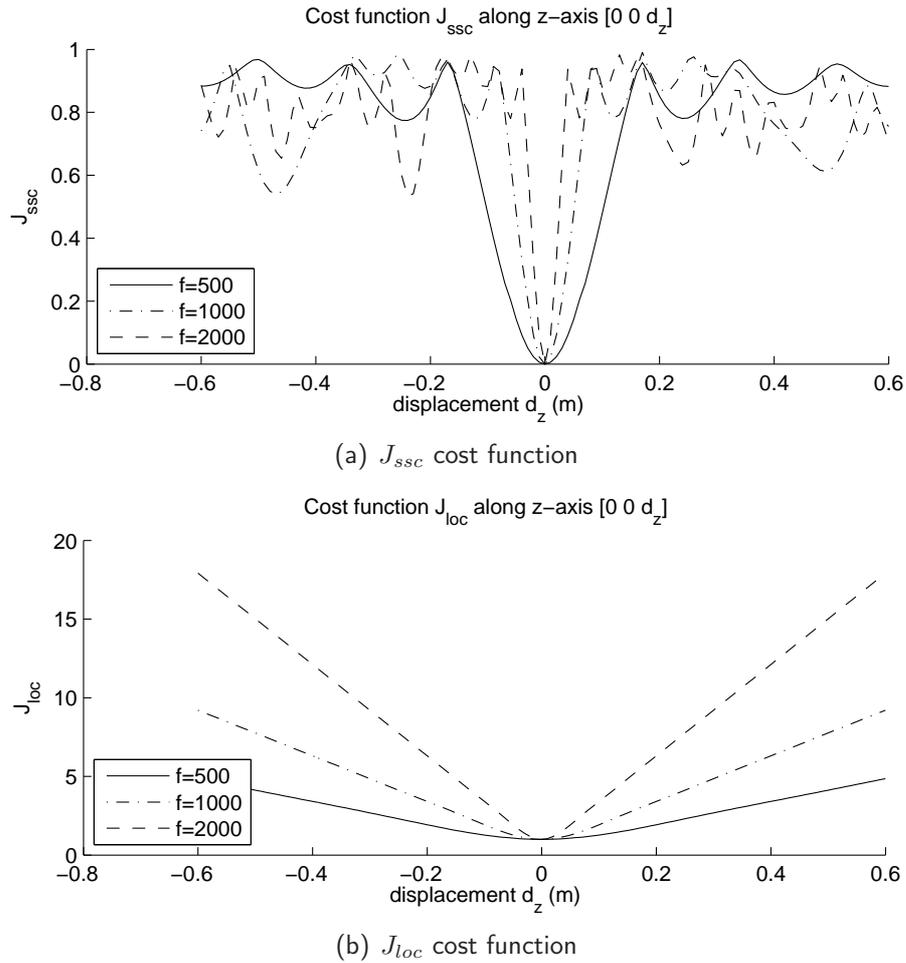
Diploma thesis, Daniel Deboy

Cost function J$_{ssc}$ along z–axis [0 0 d$_z$]

(a) J_{ssc} cost function

Cost function J$_{loc}$ along z–axis [0 0 d$_z$]

(b) J_{loc} cost function

Figure 7.12: Cost functions at different frequencies (centered monopole source)

Cost functions for different positions a_z

The fig. 7.13, and 7.14 respectively, show cost functions for two different frequencies of a monopole source positioned at different centers a. Since spatial aliasing depends on kR_a, we can assume that both criteria fail due to aliasing effects at certain displacements depending on the frequency. For $f = 500Hz$ the criteria begin to fail between 0.5 and 0.7 meters, whereas at $f = 1000Hz$ the radial boundary lies below 0.5 meter.
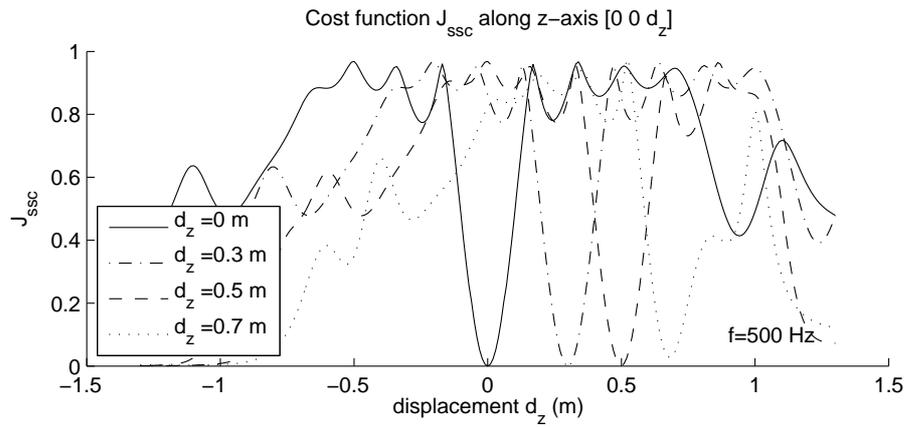
(a) J_{ssc} cost function
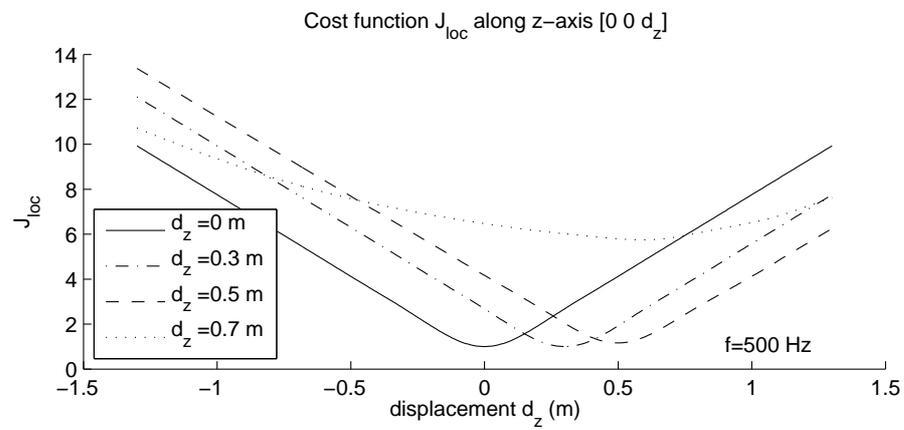
(b) J_{loc} cost function

Figure 7.13: Cost functions for different source positions a_z (centered monopole source at 500 Hz)
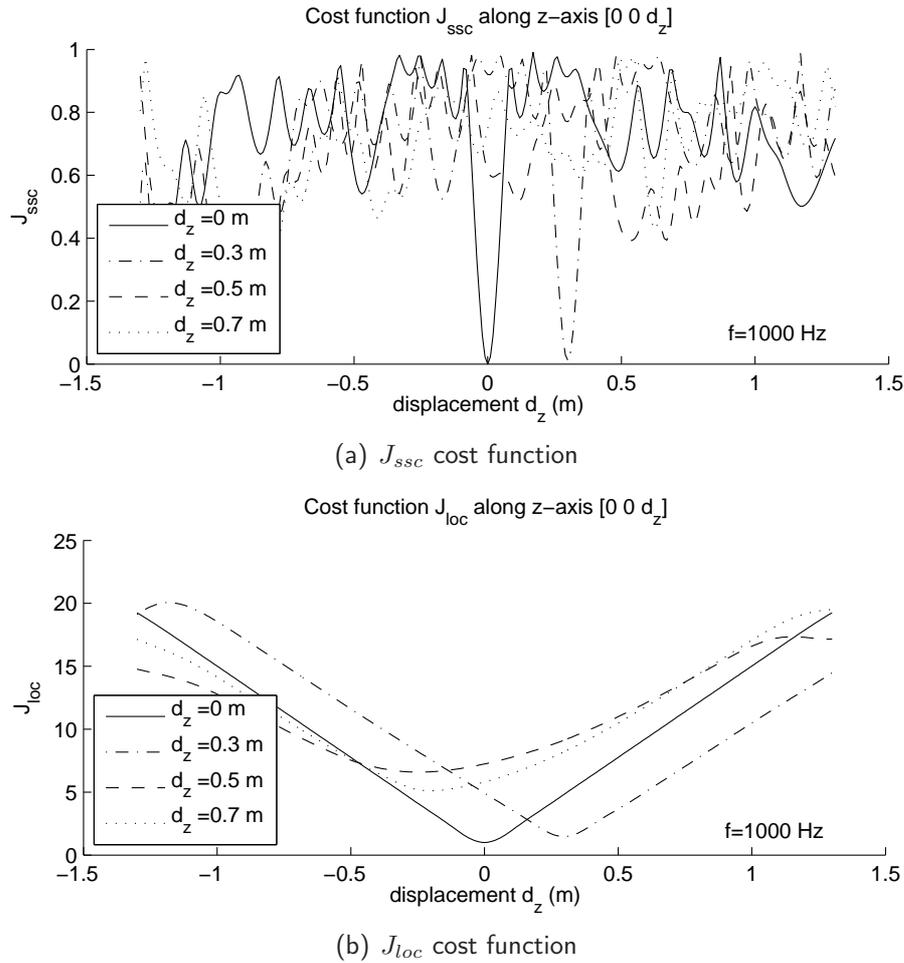
(a) J_{ssc} cost function

(b) J_{loc} cost function

Figure 7.14: Cost functions for different source positions a_z (centered monopole source at 1000 Hz)

7.4.4 Localization maps

Evaluating the cost functions at displacements along other axes than the z-axis can be done by rotating the displacement vector d onto the z-axis. This rotation needs to be computed on the wave spectrum in advance of the translation.

Using this decomposition, a three dimensional localization map can be derived with the formula for z-translations. For numerical computation we need to sample the sphere within the array with a volume sampling method. Suitable sampling methods are introduced in section 4.2.

A slice through the sphere can be done at the x-y, x-z and y-z plane (or arbitrary) to illustrate localization maps in two dimensions. For this purpose the computation can be reduced to a linearly sampled plane (for example $d_x = \{-r...0...r\}$ and $d_y = \{-r...0...r\}$). Fig. 7.15 shows an example of a localization map in the x-y plane for a monopole source at the origin at a frequency of 1000 Hz.

Note that the illustrated localization maps depict small values in dark and big values of the cost functions in bright pixels. Dark spots of the cost function therefore indicate the estimated position of the acoustic center. Localization maps for examples of simulated multipoles at the origin can be found in appendix B.

(a) x-y view of J_{loc} function
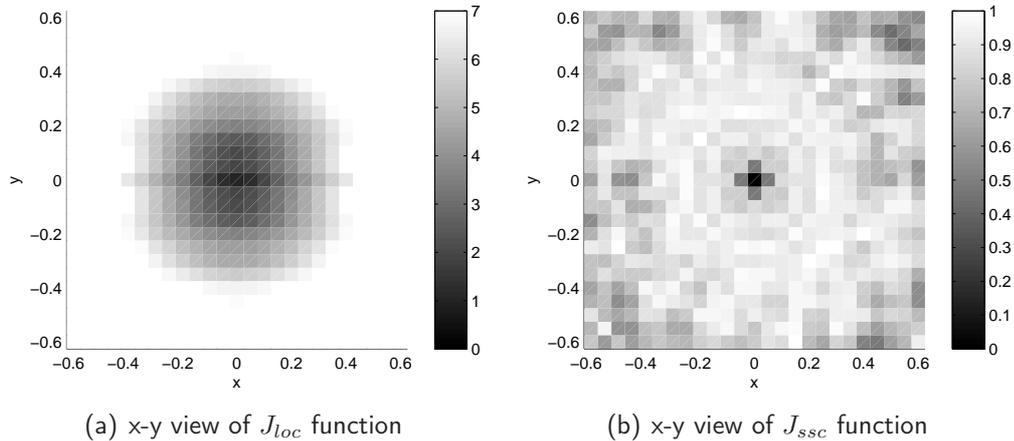

(b) x-y view of J_{ssc} function

Figure 7.15: Example localization map - centered monopole source, slice at $z = 0$

7.5 Optimization and accuracy analysis

Optimization

Essentially, assuming the cost function indicates correct positions, the goal of acoustic centering is to find a displacement $\boldsymbol{d}$ where the cost function has its global minimum. This displacement can be seen as the acoustic center ($\boldsymbol{d}_{min} = -\boldsymbol{a}$ for J_{loc} or $J_{ssc} \rightarrow min$). An optimization by a simple search algorithm evaluates the cost function in a given radial bound using discrete sampling of the volume. However, sampling nodes of a volume do not necessarily hit the global minimum of the cost functions. Especially in the case of J_{ssc} the narrow minimum may be overseen by the optimization. Therefore, a sophisticated unconstrained nonlinear optimization using *fminsearch* from the optimization toolbox in MATLAB is probably more accurate [Mat10a]. The *fminsearch* function uses the simplex search method.

Accuracy

The aliasing artifacts of a dislocated source mark bounds the optimization may accurately work within. It depends on $k\mathrm{R}_a$, and since the radius is a fixed construction parameter of the array, it in fact only depends on the frequency. Furthermore, they are assumed to depend on the wave spectrum of the source. Higher orders tend to produce more spatial aliasing inherently. Thus, another bound is given by eq. (5.4) as $N_{analyze}$ should not exceed N = 7. This bound depends on the amount of translation and the frequency.

We examine the accuracy for monopole, real-valued ($0/180°$ phase) and complex-valued (random phase for each Y_n^m) sources of order $N_s = 4$ at different dislocations and frequencies.

Error measure

For the analysis of the accuracy the absolute error of the displacement vector

$$v_{error} = |\boldsymbol{a} - \boldsymbol{d}_{min}| \tag{7.5}$$

is plotted in fig. 7.16 - fig. 7.22.

7.5.1 Optimization using discrete volume sampling methods

The following example shows the absolute error for a monopole source with a dislocation of $a_z = \{-0.7...0.7\}$ using a simple optimization by discrete search based on cubic tightest sphere packing volume sampling with approximately 10000 sampling nodes. In fig. 7.16, the J_{ssc} criteria shows large errors even in the central region of the array. The global minimum of the cost function is not reached for all dislocations due to the fixed sampling nodes that do not hit the global minimum and indicate erroneous positions instead. It is assumed that a higher resolution of sampling nodes has an increased accuracy but requires much more computational effort. Discrete search using J_{loc} seems to yield oscillating errors, but delivers much better results due to its shallower convex steep. A radial bound may be set at approximately $0.5m$.
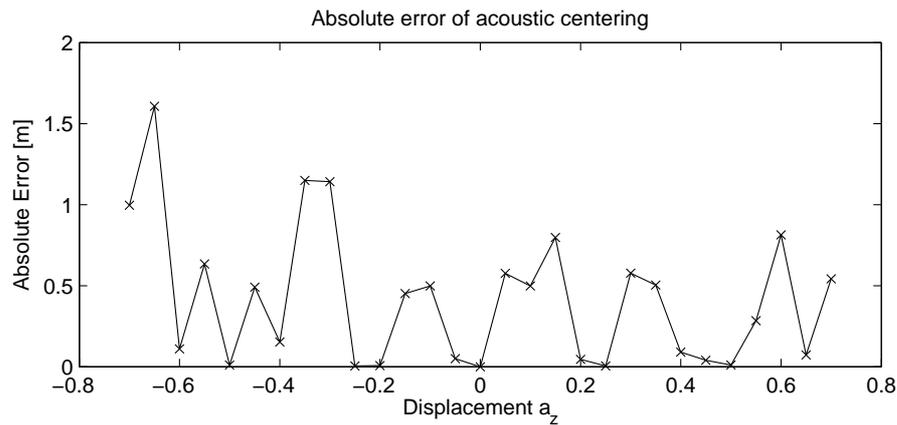
(a) Absolute Error of centering result for Monopole source J_{ssc}
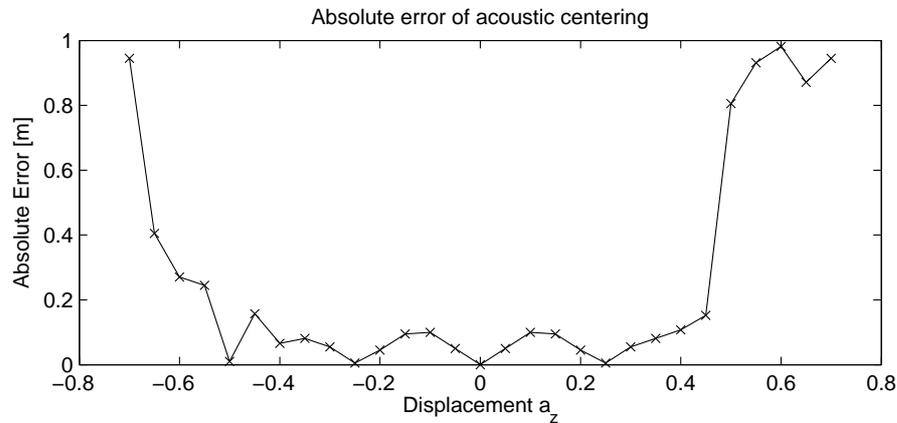
(b) Absolute Error of centering result for Monopole source J_{loc}

Figure 7.16: Centering error for a monopole source using cubic tightest sphere packing (10000 pts), $f = 1000Hz$.

7.5.2 Optimization using simplex search method

The simplex algorithm obviously yields a good estimate of the center of a dislocated source, as long as it stays within the radial bound of approximately $0.3m$. Then small distortions of the cost functions due to aliasing effects, as seen in fig. 7.13 may evoke errors. At a certain dislocation, the error strongly increases, this may be the bound where aliasing renders the cost function useless, see fig. 7.14. In the case of a complex-valued multipole source, the J_{ssc} criterion fails as seen in fig. 7.11. The initial values for the optimization algorithm have been linearly varied along the z-axis to avoid the optimization detecting local minima instead of the global minimum.
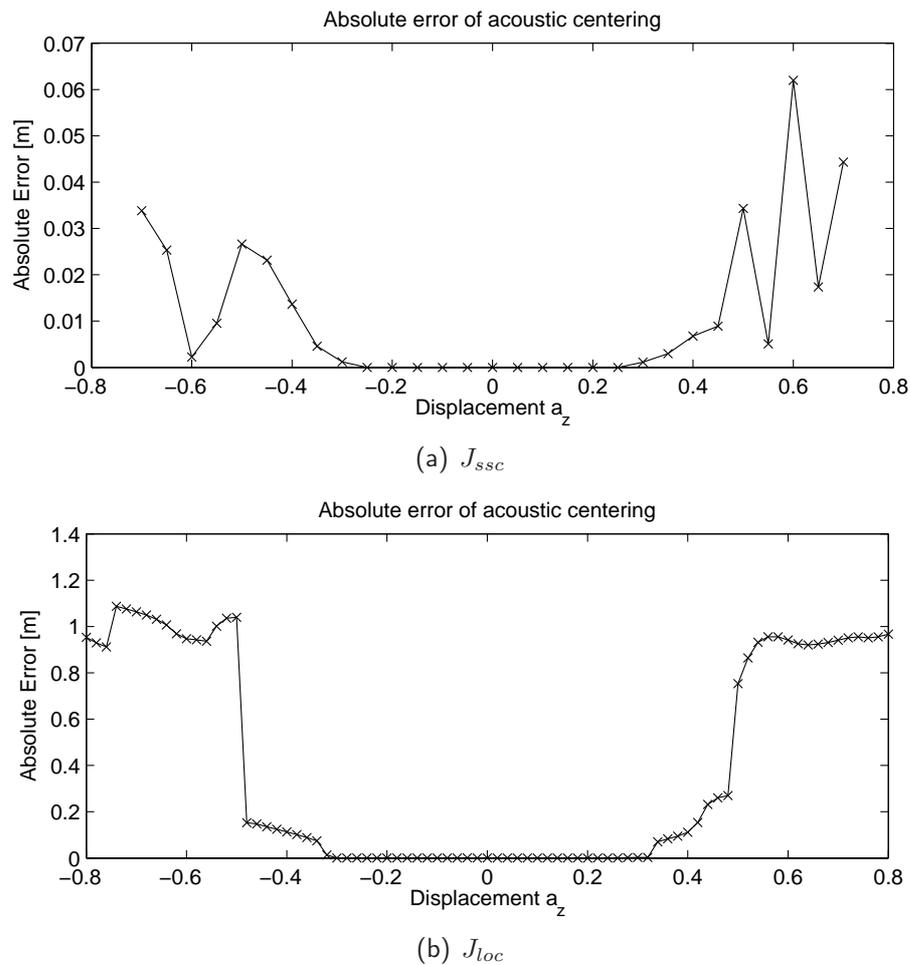

(a) J_{ssc}

(b) J_{loc}

Figure 7.17: Absolute centering error for a monopole source using fminsearch optimization, $f = 1000Hz$. Both criteria perform well in the case of a monopole source. A radial bound of confidence can be estimated.
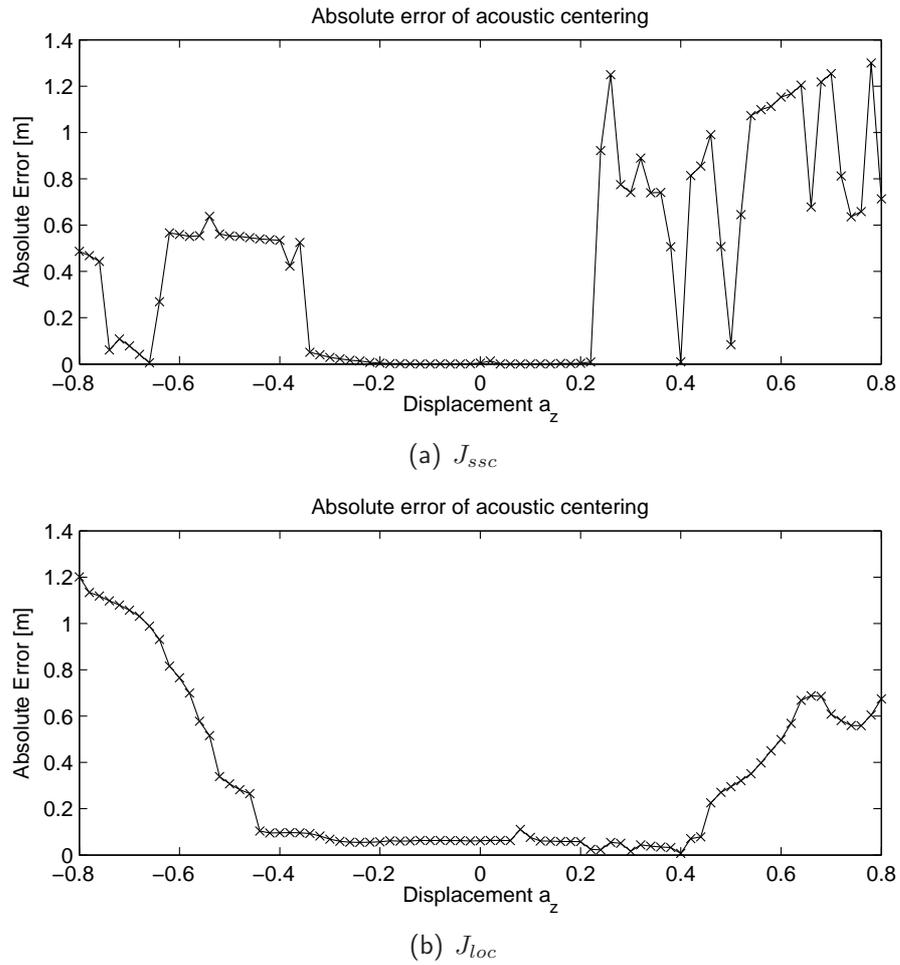
(a) J_{ssc}

(b) J_{loc}

Figure 7.18: Absolute centering error for a real-valued source using fminsearch optimization, $f = 1000Hz$. In case of J_{ssc}, the criterion fails at a smaller bound than with a monopole source. An error bias can be observed when using the J_{loc} cost function, maybe due to the directional sound radiation.
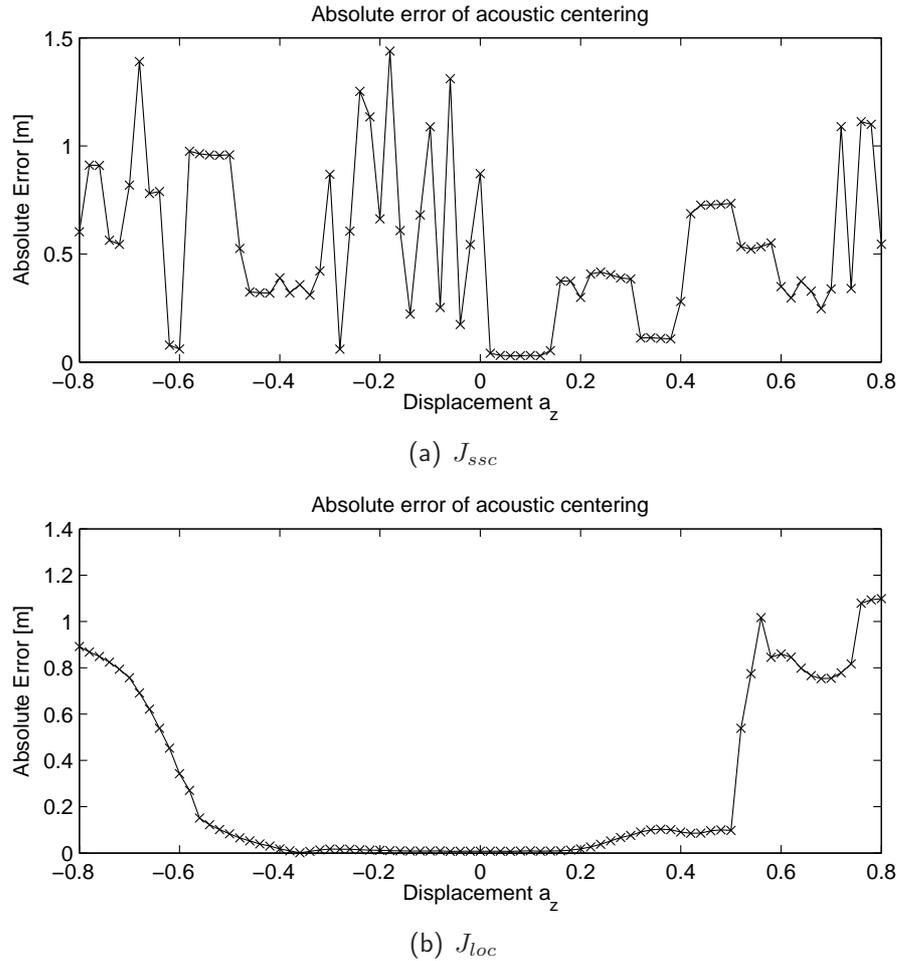
(a) J_{ssc}

(b) J_{loc}

Figure 7.19: Absolute centering error for a complex-valued multipole source, $f = 1000Hz$. The J_{ssc} criterion fails throughout all shifts except for a small range. A statistical analysis is assumed to show random results since the phase relations are randomly chosen. The J_{loc} crierium still performs well in a certain radial bound.

Frequency bounds for optimization of J_{loc}

Due to the robust performance of J_{loc}, the dependency of the error on frequency is only regarded for J_{loc}. As seen in fig. 7.17-7.19, a radial bound for a certain frequency can be estimated from the results of the optimization error. In the following examples the absolute error has been calculated for different frequencies at a certain dislocation to obtain an estimate on the frequency range within the optimization may find the correct acoustic center. The results have been found to correspond to the theoretical bound given by eq. (5.4) with $N = 7$. At low frequencies a small error occurs. This may be due to numerical issues that arise when the product $j_n h_n$ is calculated at the translation of the wave spectrum, with $j_n \ll$ and $h_n \gg$.
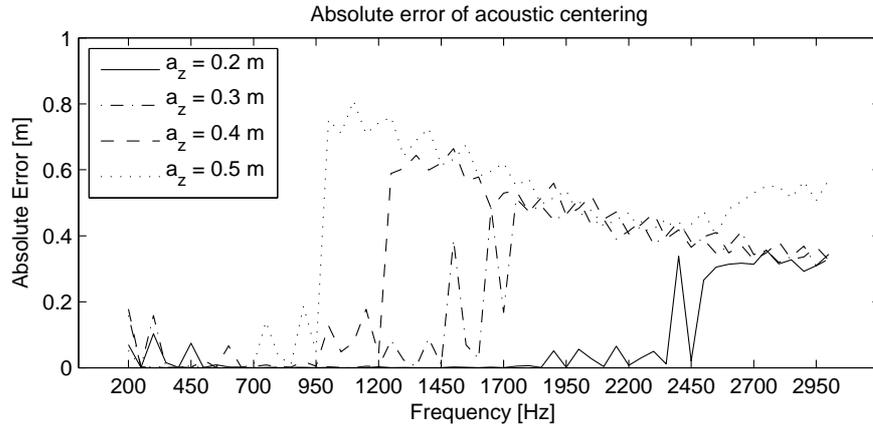
Figure 7.20: Centering error for a monopole source. The upper frequency bound corresponds to the values examined by eq. 5.4 (1893Hz for $a_z = 0.2m$, 1263Hz for $a_z = 0.3m$, 947Hz for $a_z = 0.4m$ and 758Hz for $a_z = 0.5m$)
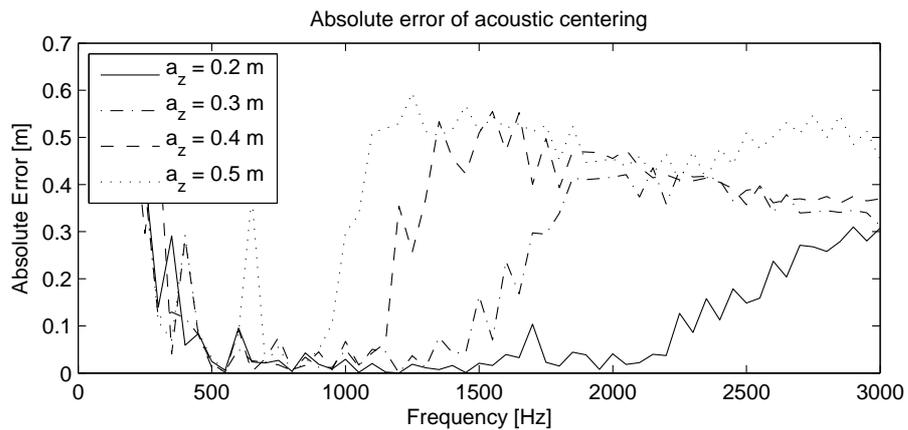

Figure 7.21: Centering error for a real-valued multipole source of order $N_{mult} = 4$. Corresponding to eq. 5.4, the frequency bound is assumed to fall due to higher-order components of the source. However, the strong increase of the error stays at the bounds of the above plot
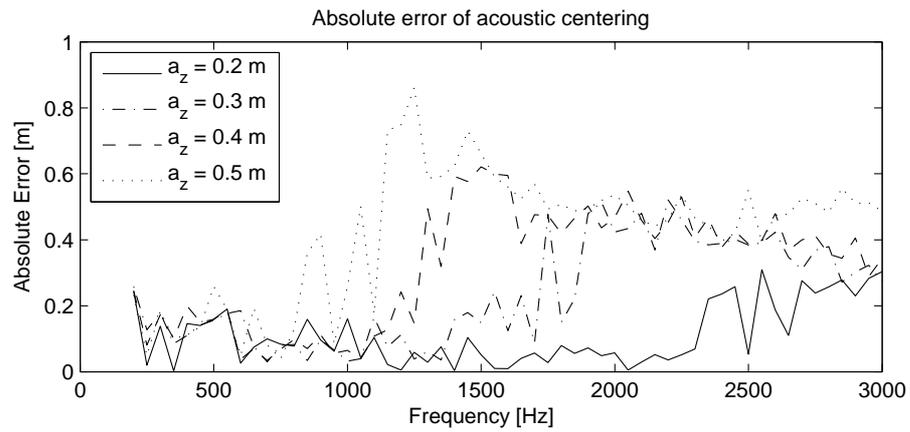
Figure 7.22: Centering error for a complex-valued multipole source of order $N_{mult} = 4$. The result is similar to the case of a real-valued multipole as expected. However, the error seems to be more arbitrary

7.6   Experimental evaluation

The following simulation examples have been made by using the unconstrained nonlinear optimization algorithm to determine $\boldsymbol{d}_{min}$. As before, the optimizations are based on optimization using *fminsearch* with the J_{loc} criterion.

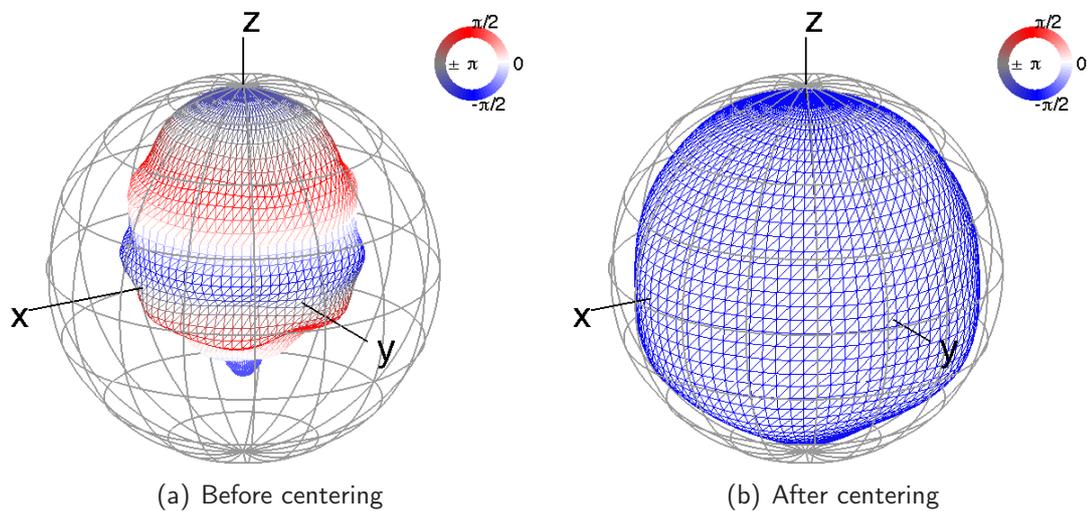### 7.6.1   Radiation patterns with vs. without acoustic centering

Shifted monopole

(a) Before centering   (b) After centering

Figure 7.23: Radiation pattern comparison (monopole source dislocated at $a_z = 0.3m$, $f = 1000Hz$). The centering by $-\boldsymbol{d}_{min}$ yields a coherent phase distribution on the surface.
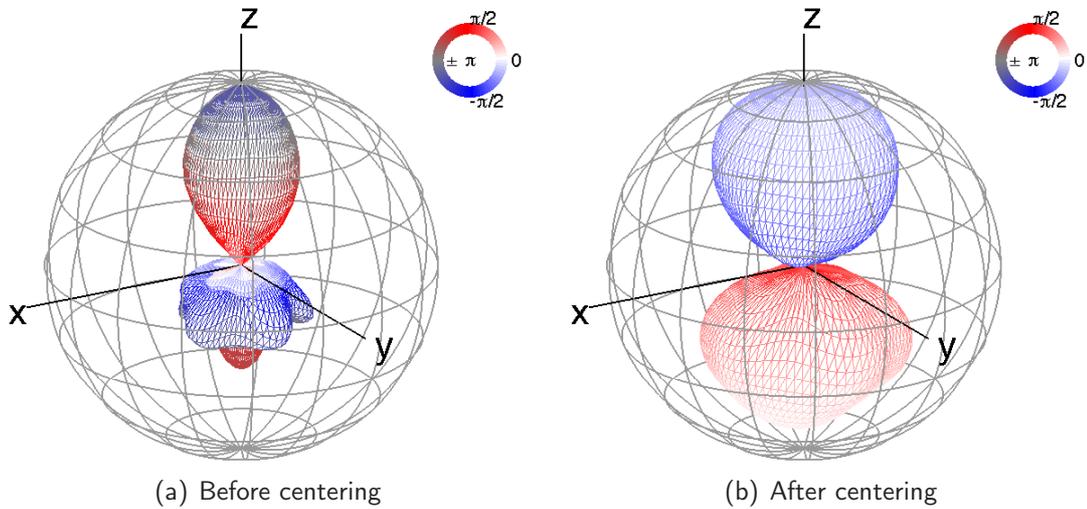
Shifted dipole

(a) Before centering   (b) After centering

Figure 7.24: Radiation pattern comparison (dipole source dislocated at $a_z = 0.3m$, $f = 1000Hz$). The centering by $\boldsymbol{d}_{min}$ yields the dipole-like phase distribution.

Directional source
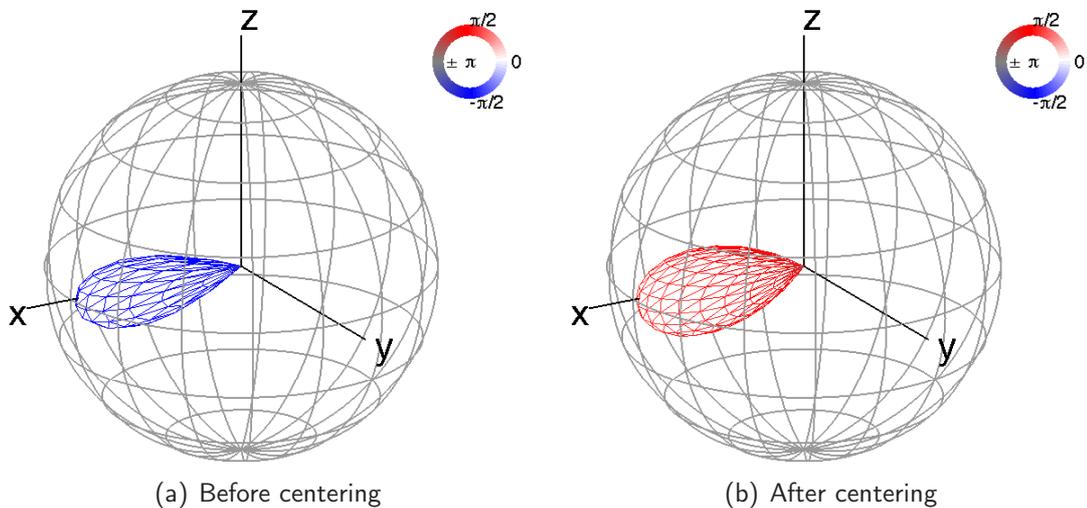

(a) Before centering   (b) After centering

Figure 7.25: Radiation pattern comparison (directional source at the origin, $f = 1000Hz$). A directional source may be centered by the optimization algorithm even if the source has not been dislocated.

7.6.2   Case study on musical instruments

To perform a test in the real-world environment, several recordings of different musical-instruments are evaluated. The optimization algorithm is used to determine the assumed

acoustic centers of different partials of the note C4 (262 Hz). The frequencies have been estimated in the total power spectrum of all 64 channels of the microphone array. The localization maps are sliced at the supposed acoustic center estimated by the simplex algorithm using J_{loc}.

On the localization maps we can see in general that the first partial may be covered by room modes and will not be reliably centered.

Bassflute

The dipole-like radiation pattern of the bass flute can be recognized in the localization map having in mind the one of a simulated dipole source (see appendix B). It is prominent in the map of the second partial.
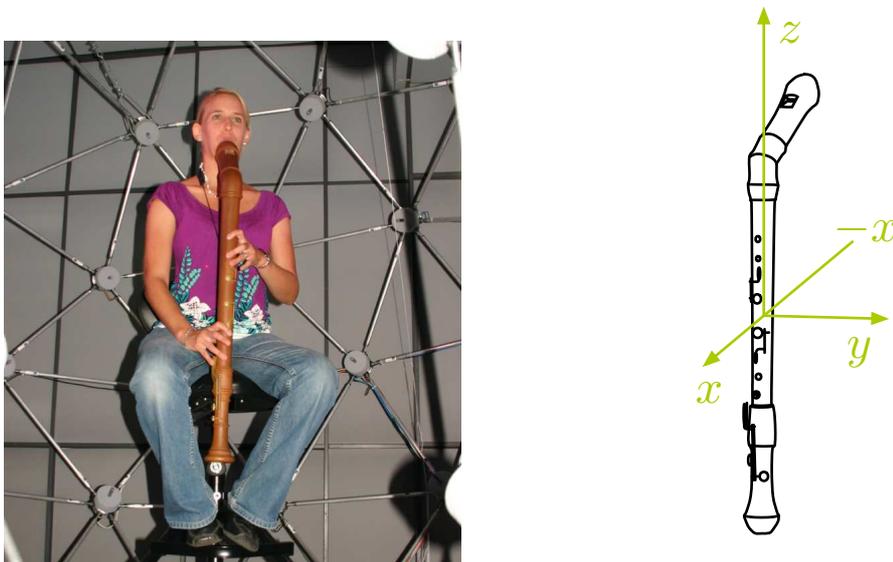

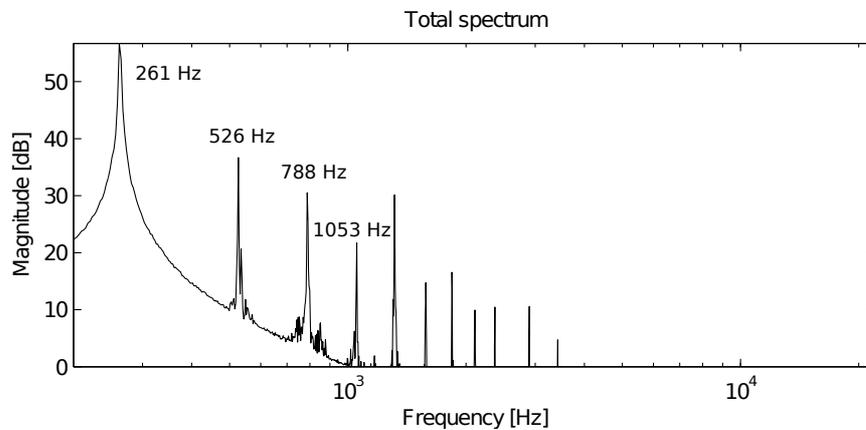Figure 7.26: Bass flute player and position of the instrument [Hoh09]

Figure 7.27: Spectrum and partials of bass flute playing C4
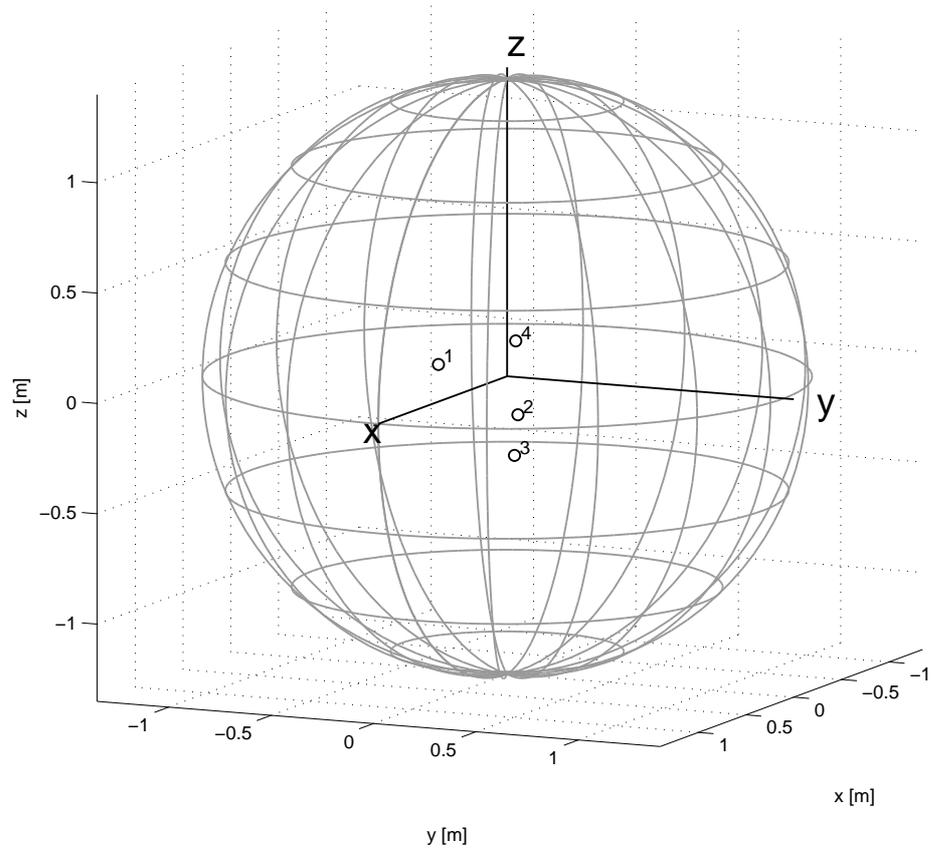
Figure 7.28: Bass flute C4, position of partials 1 to 4.

(a) x-y view of J_{loc}

(b) x-y view of J_{ssc}

(c) x-z view of J_{loc}

(d) x-z view of J_{ssc}

(e) y-z view of J_{loc}

(f) y-z view of J_{ssc}
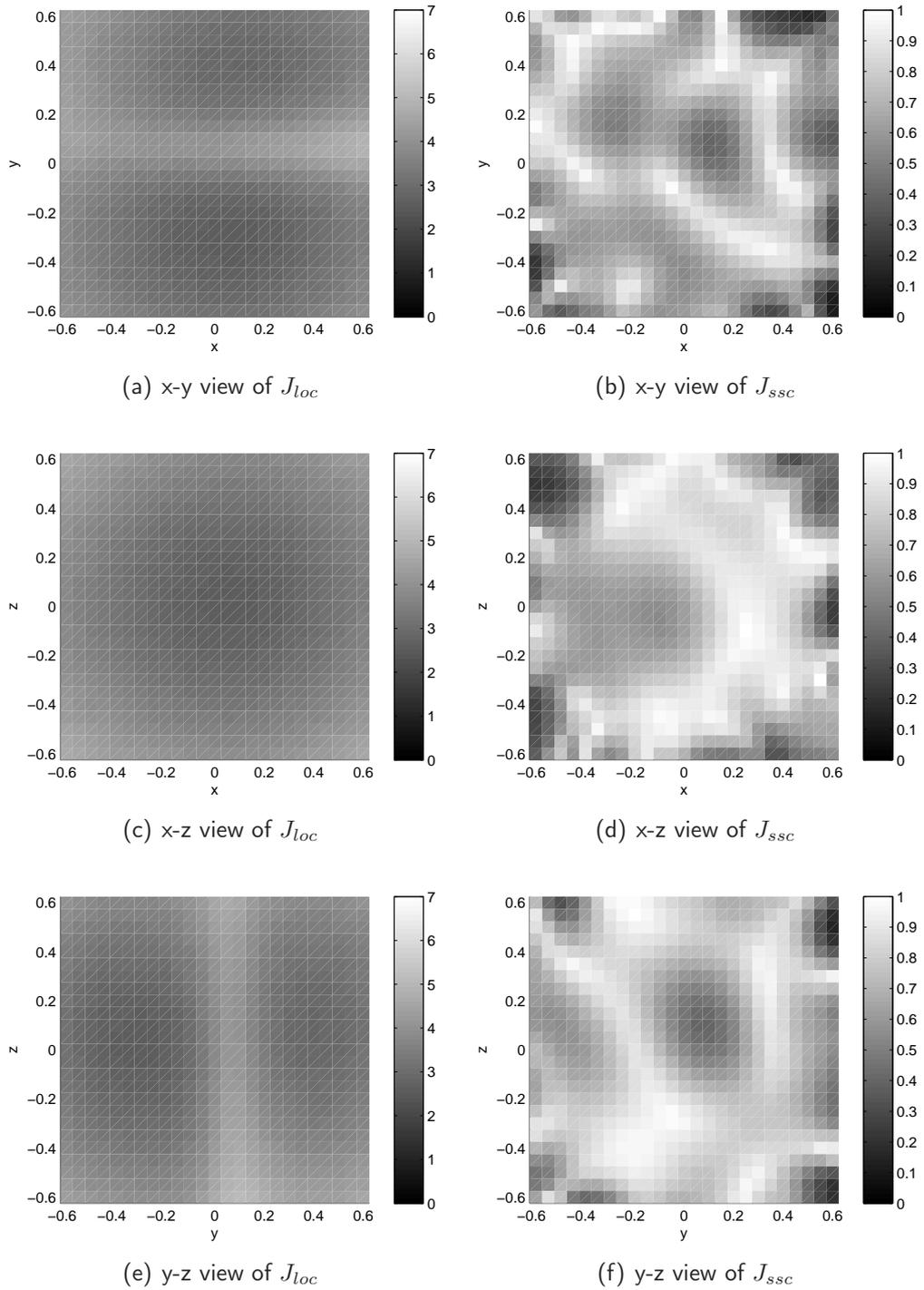
Figure 7.29: Bassflute, C4, partial at 262 Hz, slices at $(0.07, -0.30, 0.04)$m

(a) x-y view of J_{loc}

(b) x-y view of J_{ssc}

(c) x-z view of J_{loc}

(d) x-z view of J_{ssc}

(e) y-z view of J_{loc}

(f) y-z view of J_{ssc}
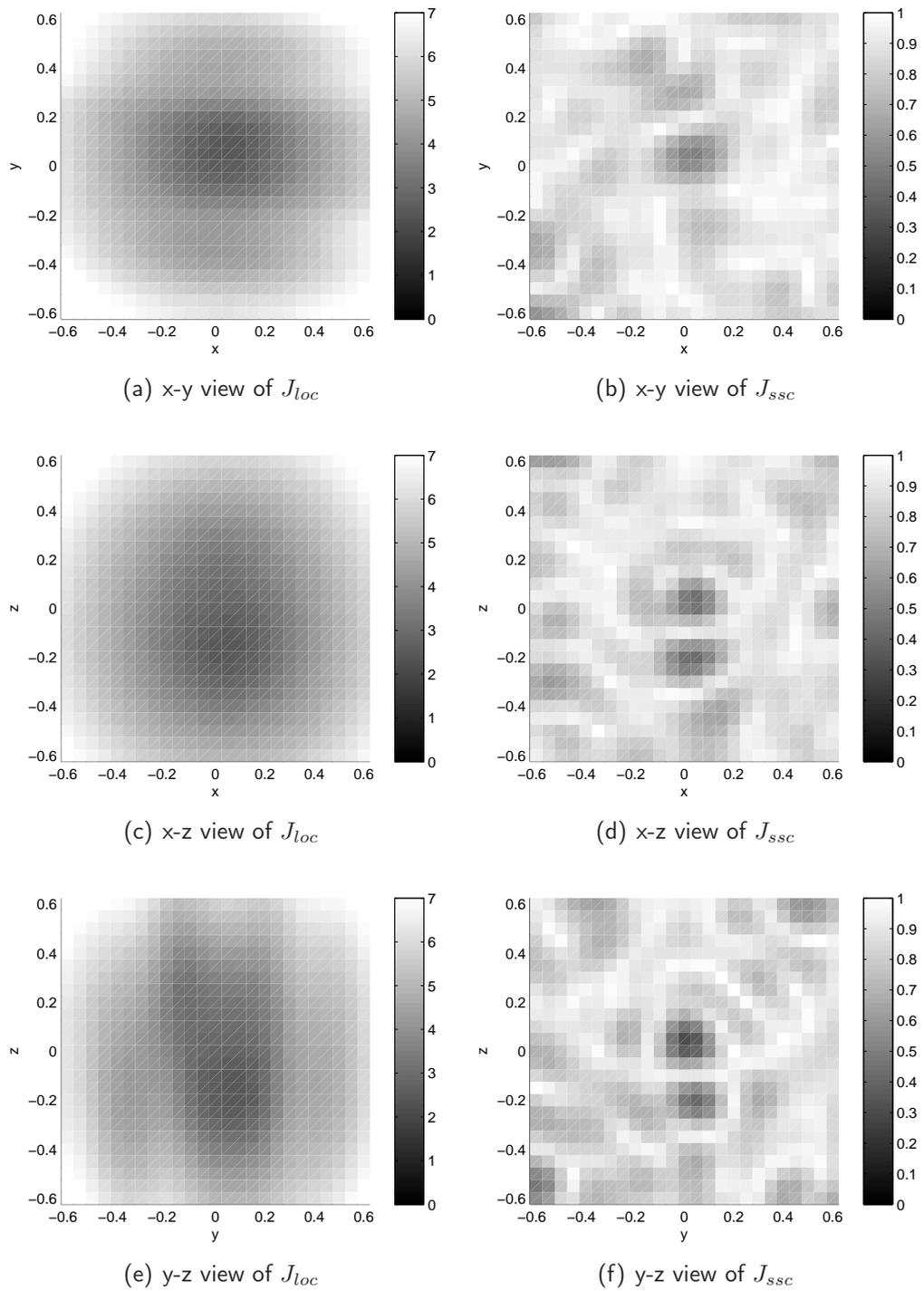
Figure 7.30: Bassflute, C4, partial at 526 Hz, slices at $(0.04, 0.07, -0.16)$m

(a) x-y view of J_{loc}

(b) x-y view of J_{ssc}

(c) x-z view of J_{loc}

(d) x-z view of J_{ssc}

(e) y-z view of J_{loc}

(f) y-z view of J_{ssc}
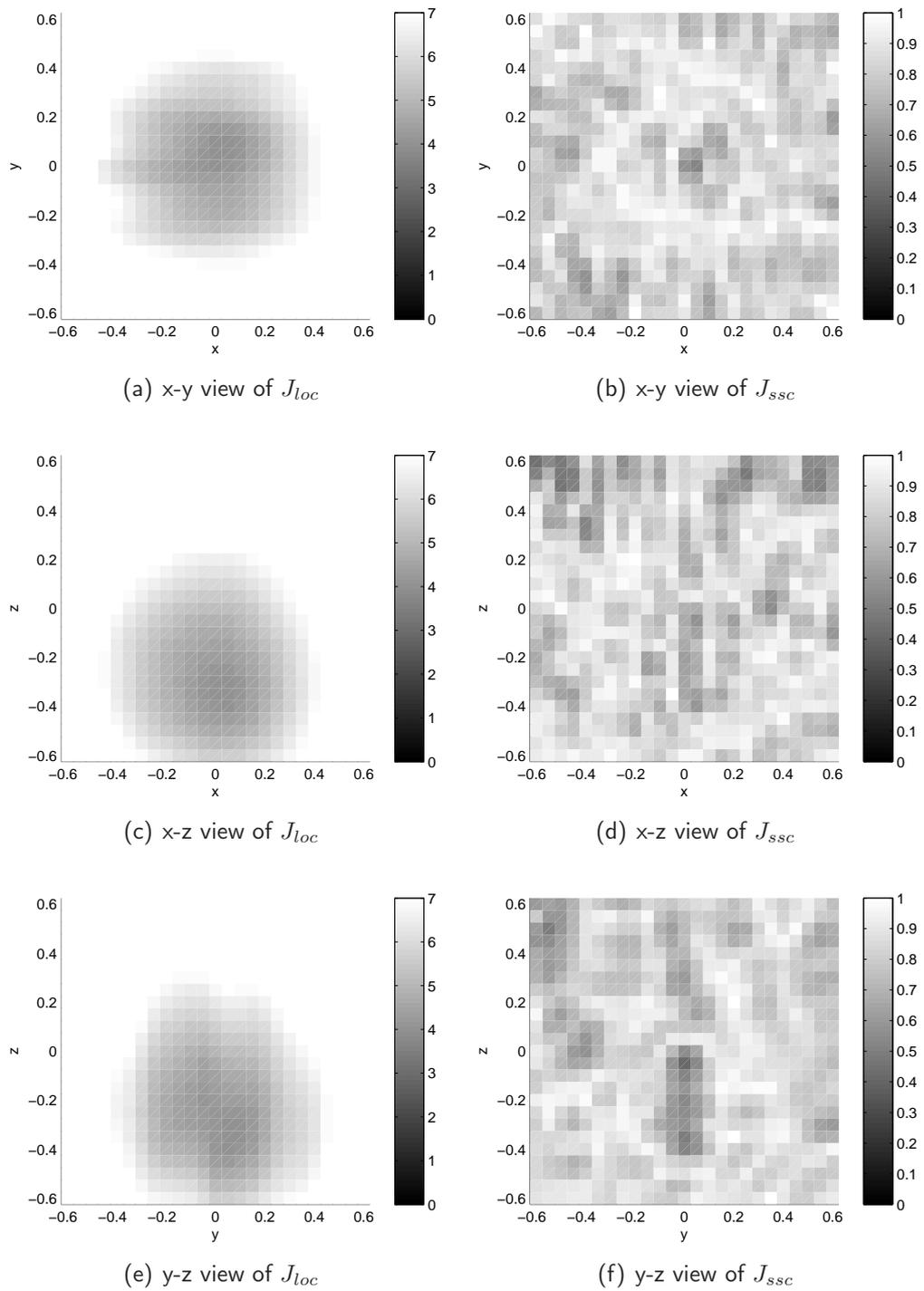
Figure 7.31: Bassflute, C4, partial at 788 Hz, slices at $(0.03, 0.05, -0.35)$m

(a) x-y view of J_{loc}

(b) x-y view of J_{ssc}

(c) x-z view of J_{loc}

(d) x-z view of J_{ssc}

(e) y-z view of J_{loc}

(f) y-z view of J_{ssc}

Figure 7.32: Bassflute, C4, partial at 1053 Hz, slices at $(0.11, 0.09, 0.18)$m

Violoncello

Figure 7.33: Violoncello player and position of the instrument [Hoh09]

Figure 7.34: Spectrum and partials of violoncello playing C4
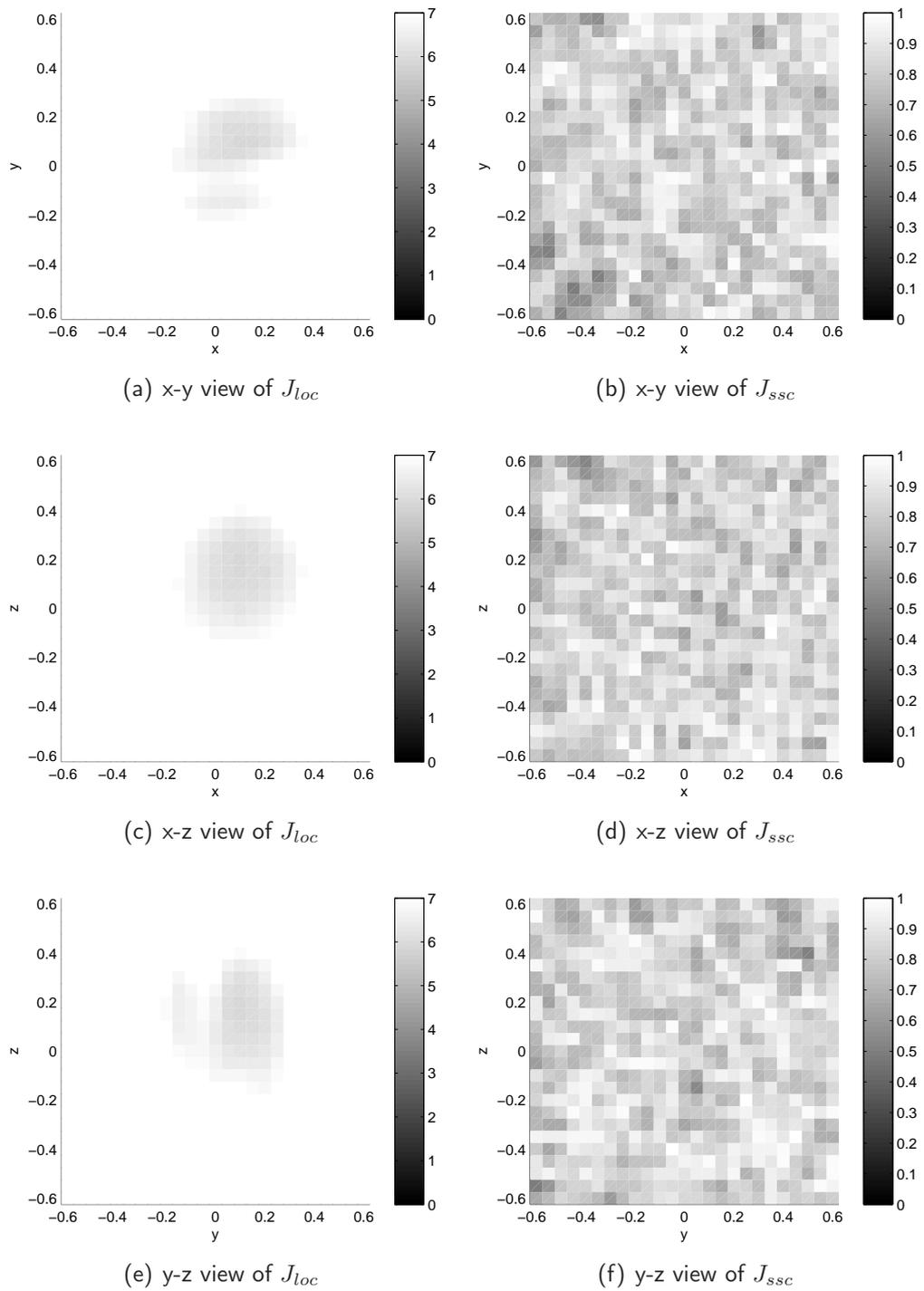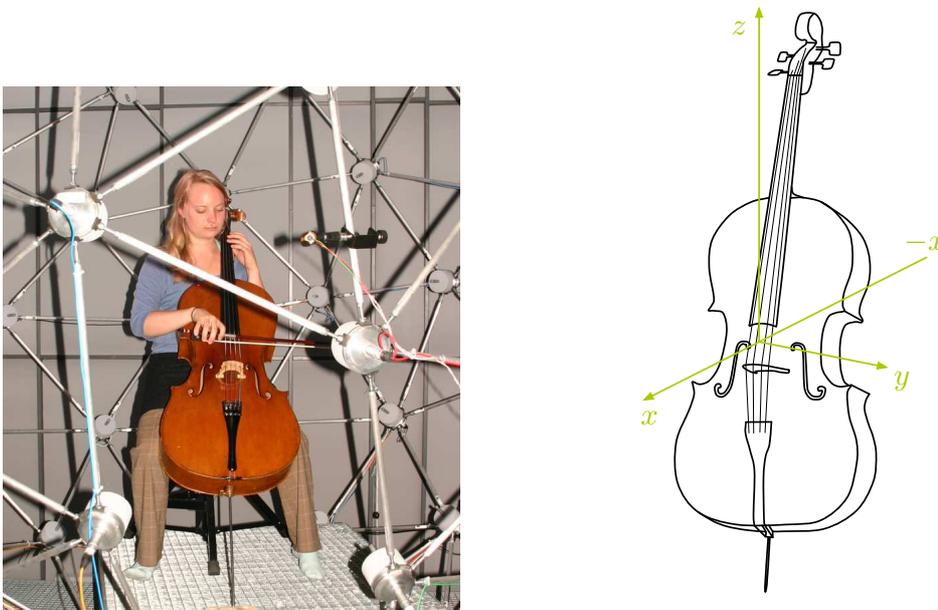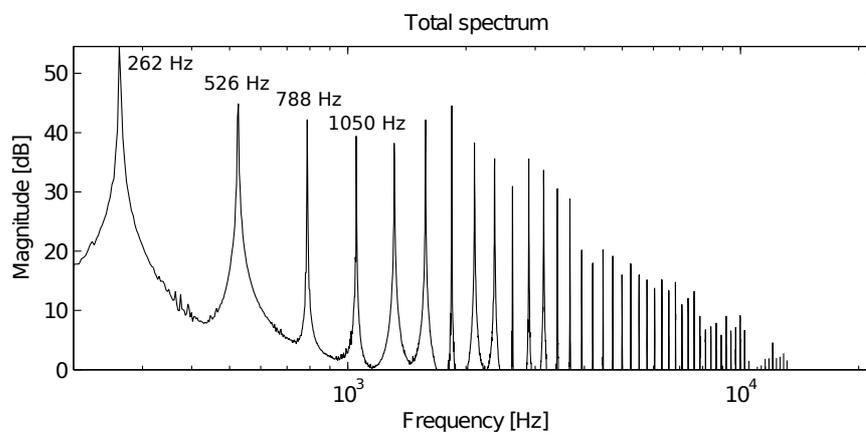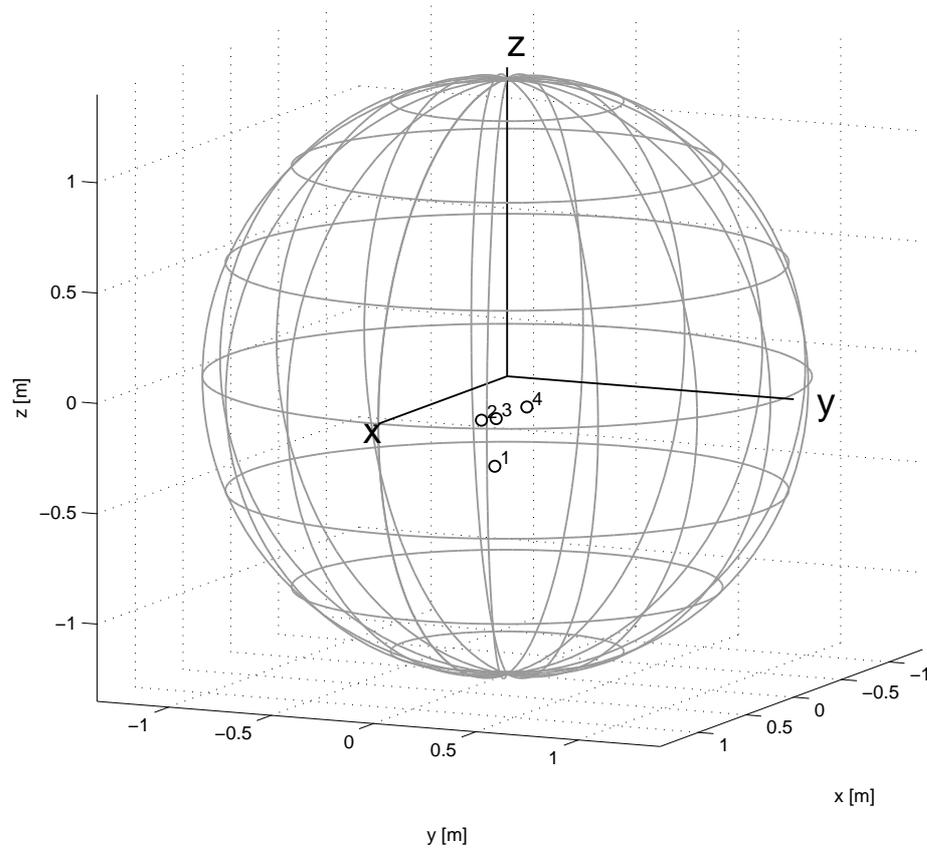
Figure 7.35: Violoncello C4, positions of partials 1-4

(a) x-y view of J_{loc}

(b) x-y view of J_{ssc}

(c) x-z view of J_{loc}

(d) x-z view of J_{ssc}

(e) y-z view of J_{loc}

(f) y-z view of J_{ssc}

Figure 7.36: Cello, C4, partial at 262 Hz, slices at $(-0.02, -0.07, -0.42)$m

(a) x-y view of J_{loc}

(b) x-y view of J_{ssc}

(c) x-z view of J_{loc}

(d) x-z view of J_{ssc}

(e) y-z view of J_{loc}

(f) y-z view of J_{ssc}
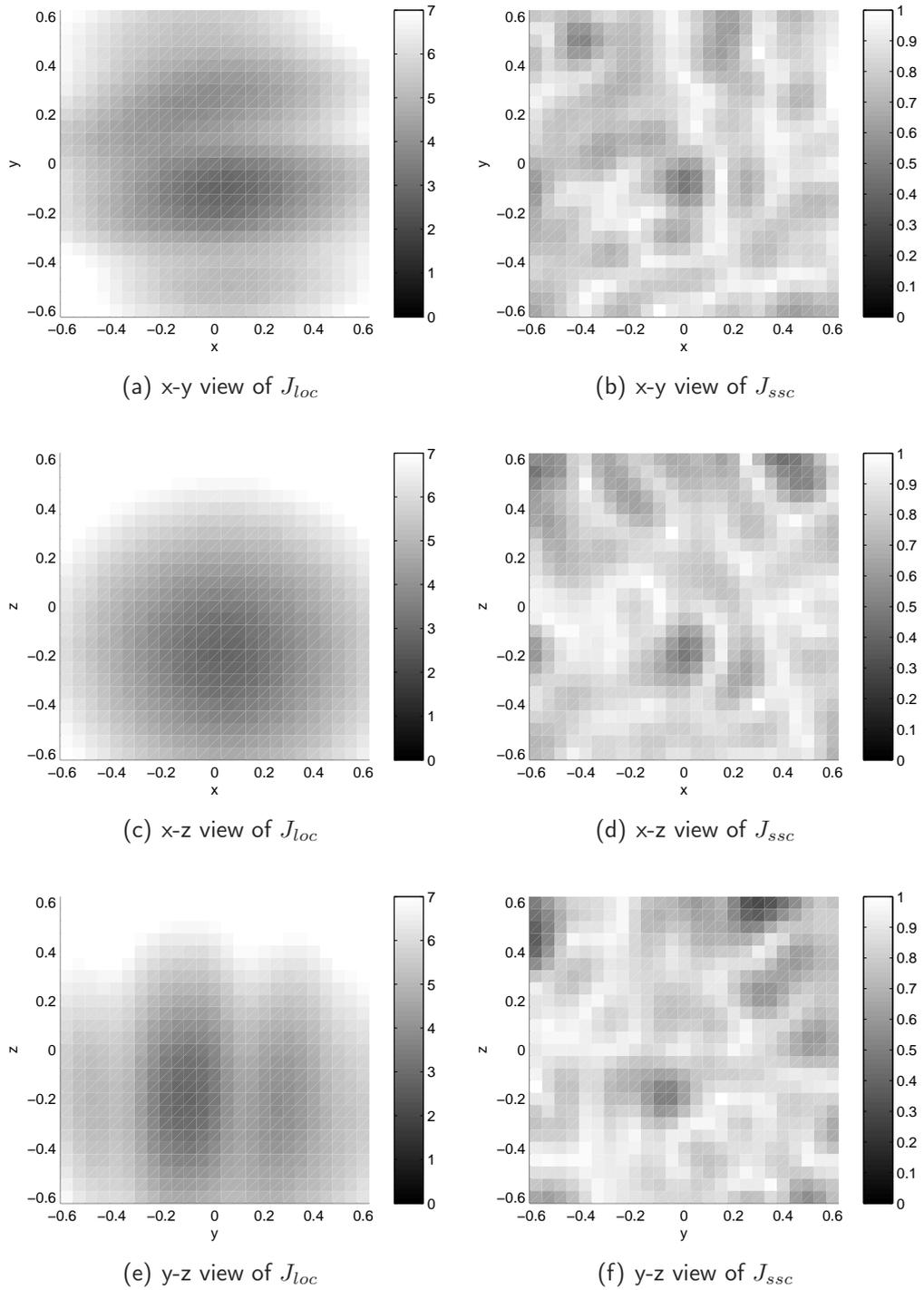
Figure 7.37: Cello, C4, partial at 526 Hz, sliced at $(0.07, -0.11, -0.20)$m

(a) x-y view of J_{loc}

(b) x-y view of J_{ssc}

(c) x-z view of J_{loc}

(d) x-z view of J_{ssc}

(e) y-z view of J_{loc}

(f) y-z view of J_{ssc}
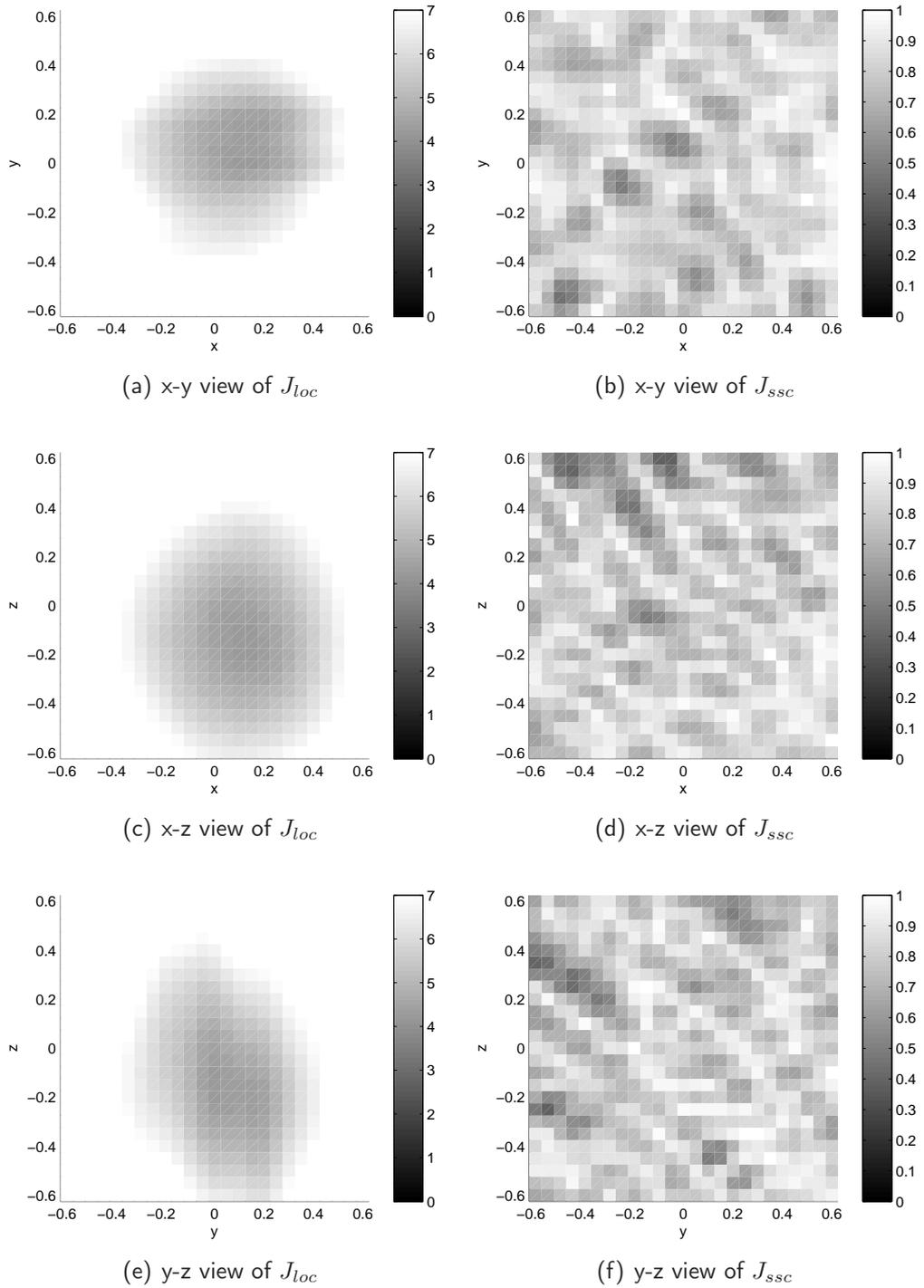
Figure 7.38: Cello, C4, partial at 788 Hz, sliced at $(0.13, -0.01, -0.17)$m

(a) x-y view of J_{loc}

(b) x-y view of J_{ssc}

(c) x-z view of J_{loc}

(d) x-z view of J_{ssc}

(e) y-z view of J_{loc}

(f) y-z view of J_{ssc}

Figure 7.39: Cello, C4, partial at 1050 Hz, sliced at $(-0.04, 0.08, -0.14)$m

Trombone

Figure 7.40: Trombone player and position of the instrument [Hoh09]
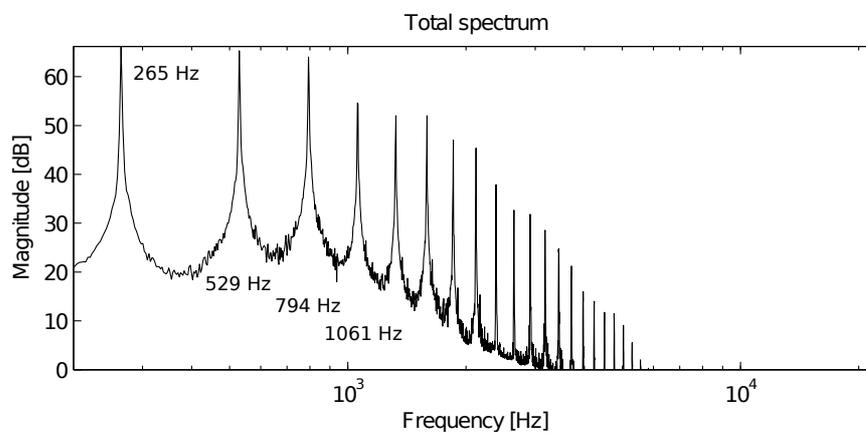
Figure 7.41: Spectrum and partials of trombone playing C4
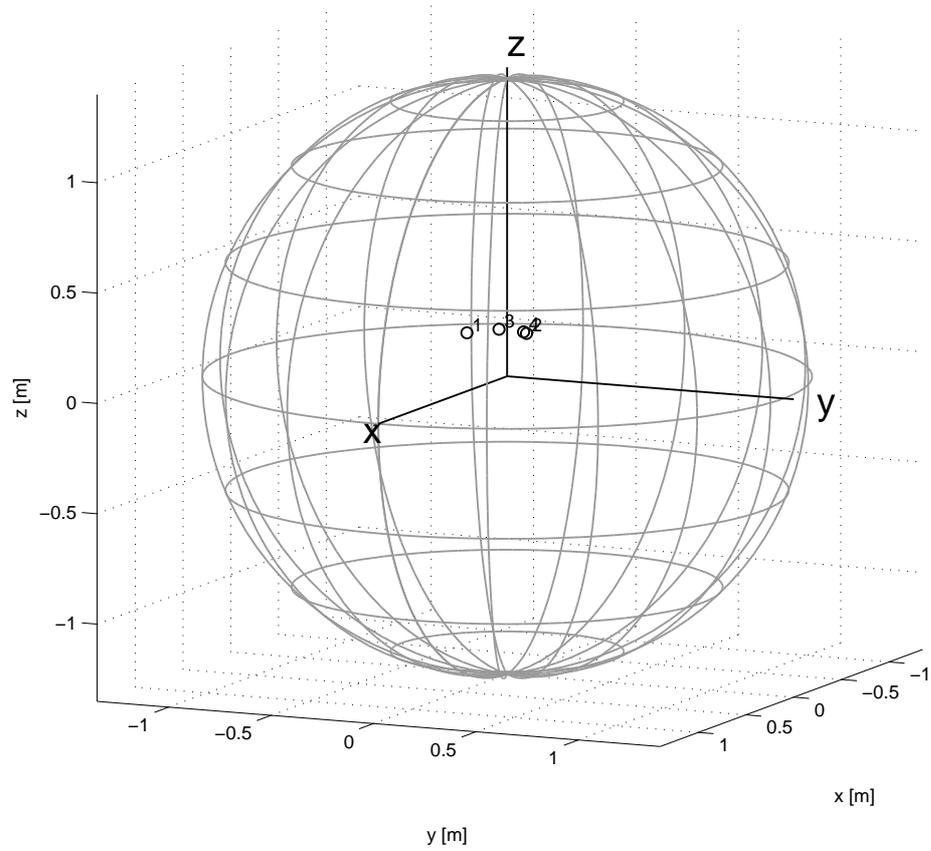
Figure 7.42: Trombone C4, positions of partials 1 to 4.
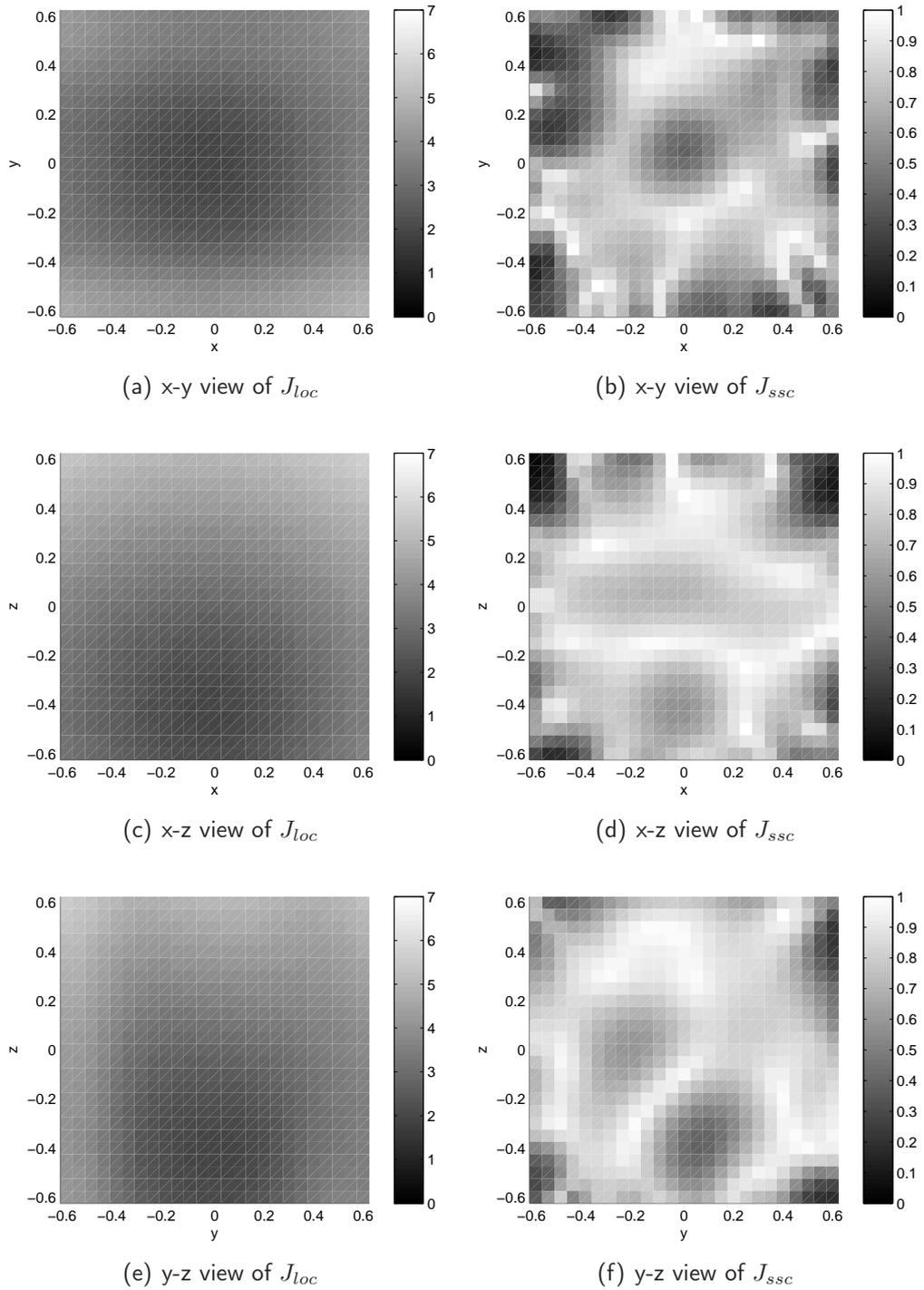
(a) x-y view of J_{loc}
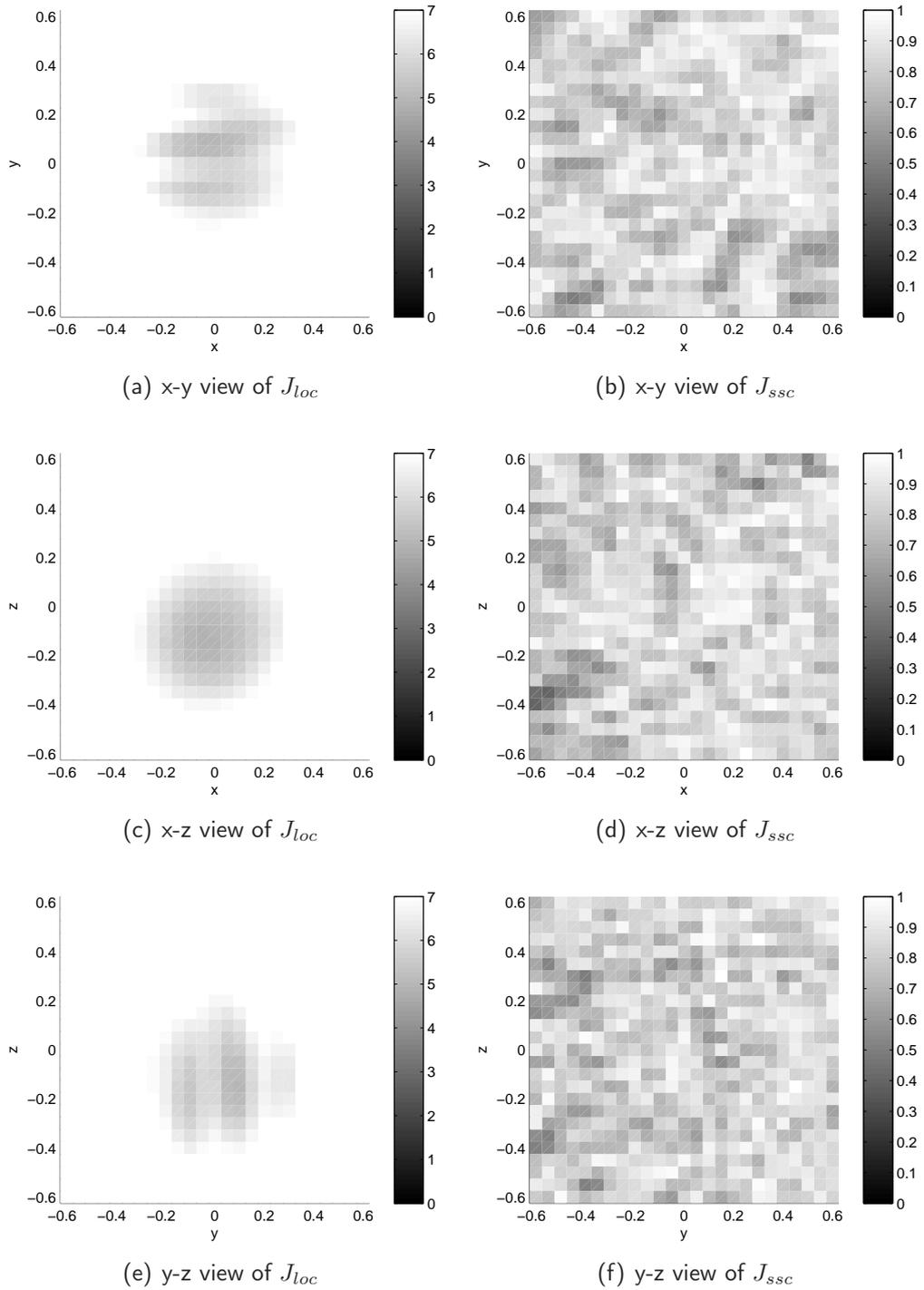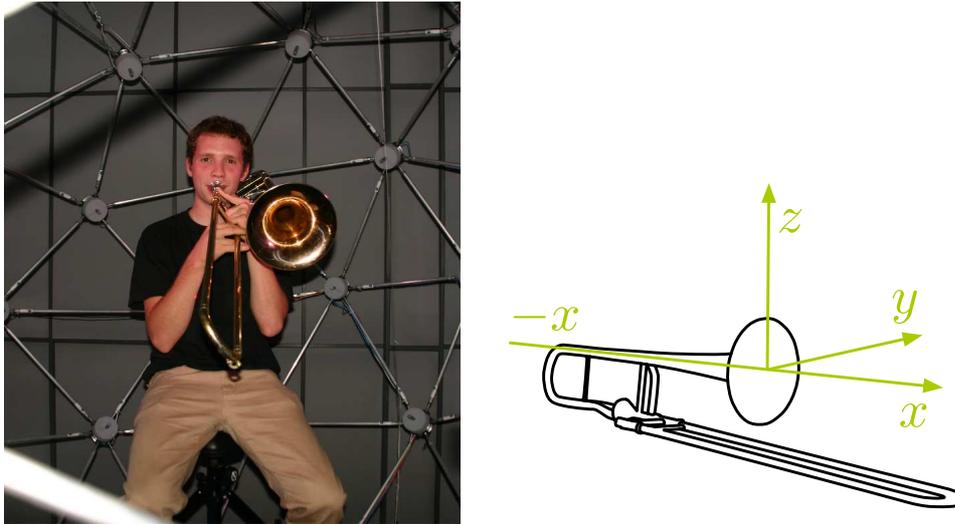
(b) x-y view of J_{ssc}

(c) x-z view of J_{loc}

(d) x-z view of J_{ssc}

(e) y-z view of J_{loc}

(f) y-z view of J_{ssc}
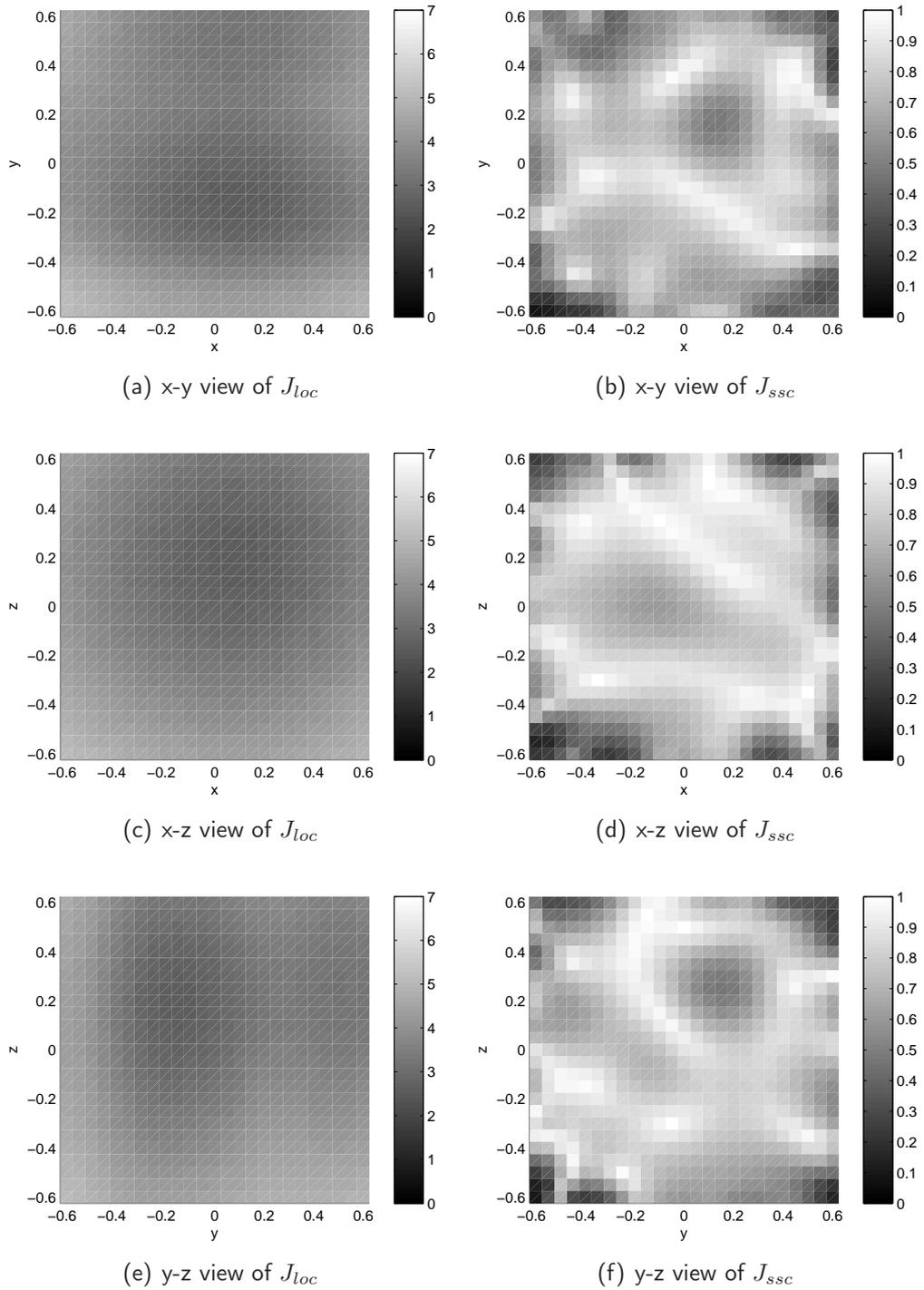
Figure 7.43: Trombone, C4, partial at 265 Hz, sliced at $(0.09, -0.15, 0.20)$m

(a) x-y view of J_{loc}

(b) x-y view of J_{ssc}

(c) x-z view of J_{loc}

(d) x-z view of J_{ssc}

(e) y-z view of J_{loc}

(f) y-z view of J_{ssc}
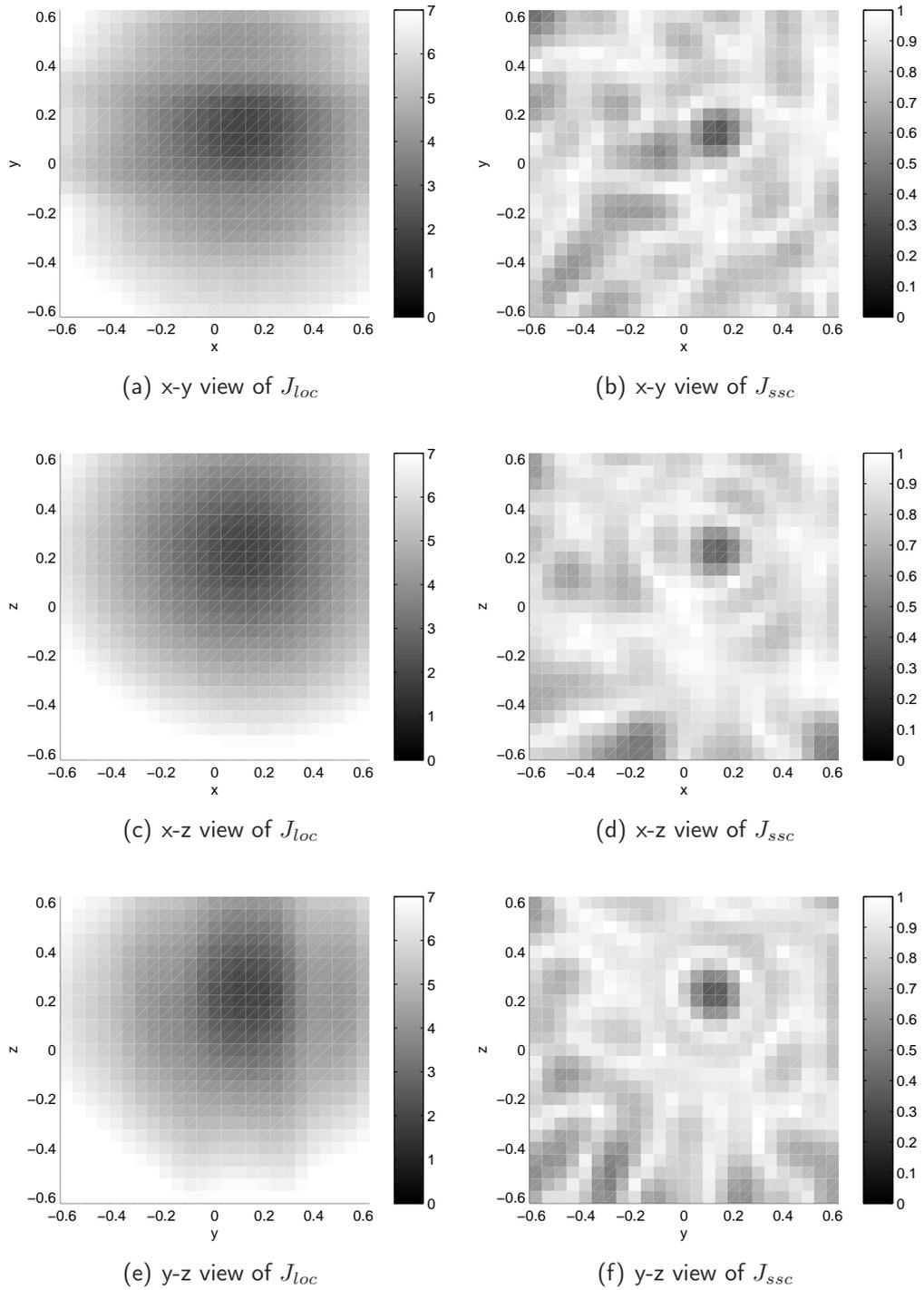
Figure 7.44: Trombone, C4, partial at 529 Hz, sliced at $(0.13, 0.16, 0.23)$m

(a) x-y view of J_{loc}

(b) x-y view of J_{ssc}

(c) x-z view of J_{loc}

(d) x-z view of J_{ssc}

(e) y-z view of J_{loc}

(f) y-z view of J_{ssc}
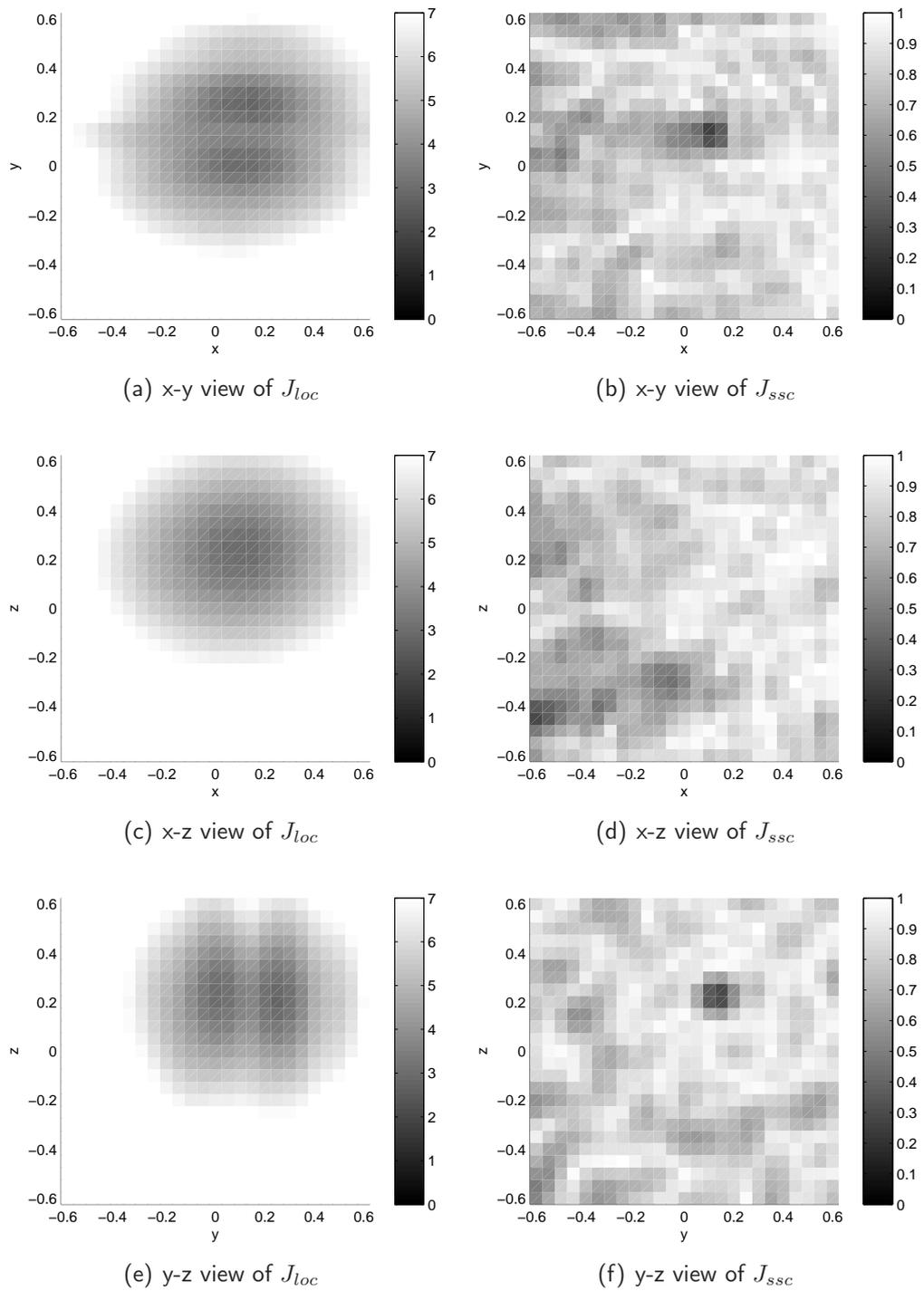
Figure 7.45: Trombone, C4, partial at 794 Hz, sliced at $(0.11, 0.01, 0.23)$m

(a) x-y view of J_{loc}

(b) x-y view of J_{ssc}

(c) x-z view of J_{loc}

(d) x-z view of J_{ssc}

(e) y-z view of J_{loc}

(f) y-z view of J_{ssc}

Figure 7.46: Trombone, C4, partial at 1061 Hz, sliced at $(0.12, 0.14, 0.23)$m

7.6.3   "Tropfenlautsprecher" - Loudspeaker in small enclosure

In order to evaluate the optimization algorithm in terms of its accuracy a loudspeaker in a small enclosure was positioned at several places within the array, cf. picture 7.48. The relative distance between two locations was determined with a measuring tape. The impulse response has been recorded using the exponential sweep method. Reflections of the reverberant room are cut in the post processing of the analysis. With the sm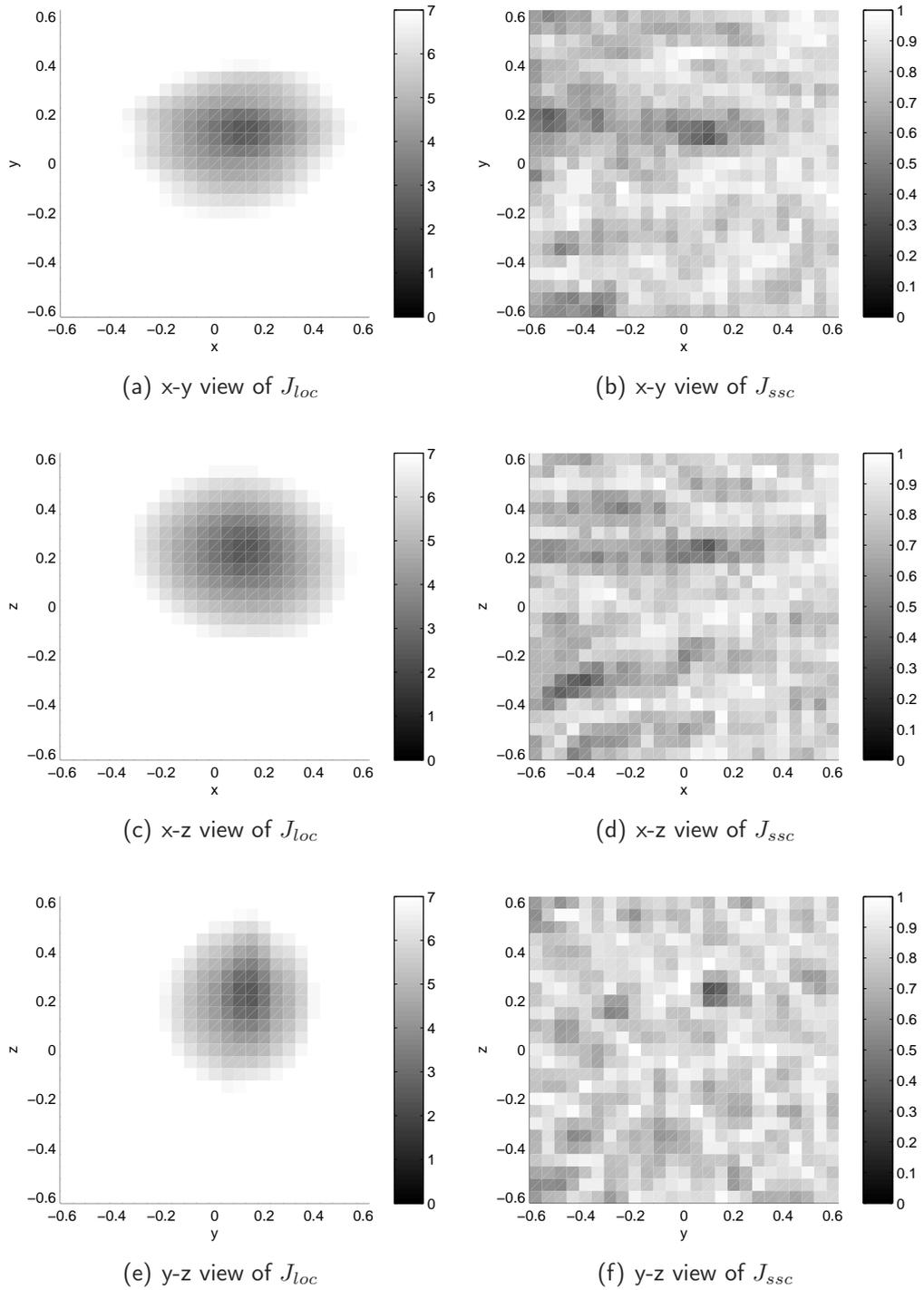all enclosure of the loudspeaker it should radiate as a point source and we should be able to track the acoustic center of the speaker. Within a bound of $0.50m$ we see a well performing localization of the acoustic center. All positions have been calibrated to the first one which was meant to be at the origin of the array.
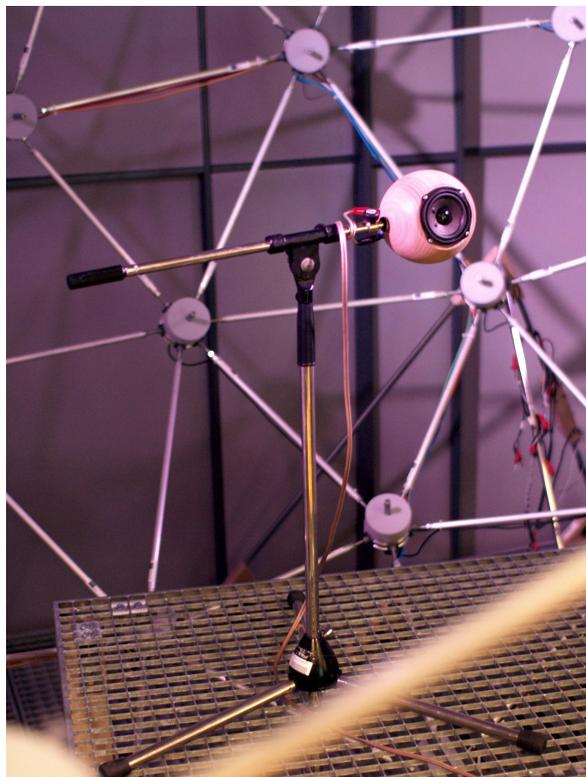

Figure 7.47: Picture of the "Tropfenlautsprecher"

Position	$\boldsymbol{a}([a_x a_y a_z]cm)$			$-d_{min}([d_x d_y d_z]cm)$			$v_{error}(cm)$
1	0	0	0	0	0	0	$N.A.$
2	0	0	5	1	0	4	1.41
3	0	0	10	1	1	10	1.41
4	0	0	20	1	0	20	1
5	0	0	30	2	2	29	3
6	0	0	40	1	0	38	2.23
7	0	0	50	1	0	45	5.01
8	0	0	60	2	1	54	6.4
9	20	0	0	18	1	0	2.23
10	20	0	20	19	1	22	2.45
11	−20	0	20	−20	2	21	2.23

Table 7.1: Positions of the loudspeaker
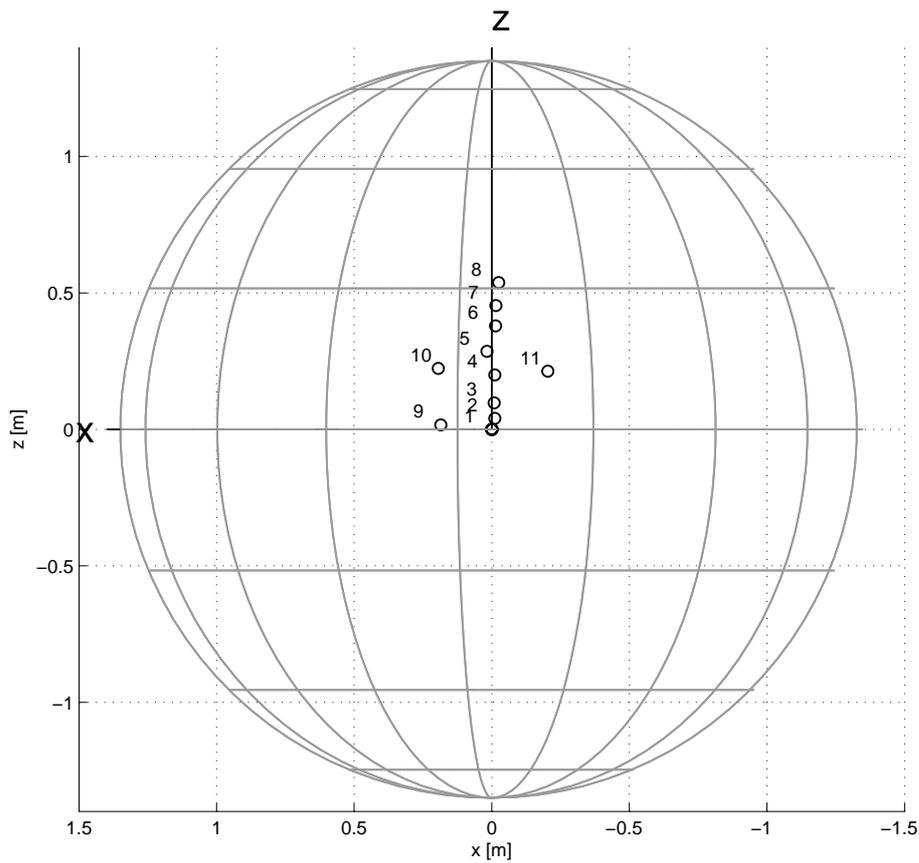
Figure 7.48: Loudspeaker positions found with centering (side-view), exponential sweep measurement, centering frequency 600 Hz

7.7   Tangential Intensity Vector Sum $I_{tg,sum}$

Another concept which may yield a measure for the shift of a source within a surrounding spherical microphone array is to derive only the tangential parts of the acoustic intensity at each sampling node. The sum of the tangential intensity vector may indicate the direction in which the source has been shifted.

The tangential acoustic intensity is given by

$$\boldsymbol{I}_{tg} = \int_{\mathbb{S}} p^* \boldsymbol{v}_{tg} d\boldsymbol{\theta}, \tag{7.6}$$

and the tangential derivative of the sound-pressure is proportional to the tangential velocity vector. Thus, the intensity can be obtained from the sound-pressure (the impedance is neglected) and a tangential component can be written

$$\int_{\mathbb{S}} p^* \nabla_{tg} p d\boldsymbol{\theta} = \sum_{nm} \sum_{n'm'} \boldsymbol{c}_n^{m*} \mu_{n'}^{m'} \int_{\mathbb{S}} Y_n^m(\boldsymbol{\theta}) Y_{n'}^{m'}(\boldsymbol{\theta}) d\boldsymbol{\theta} = \sum_{nm} \boldsymbol{c}_n^{m*} \mu_n^m. \tag{7.7}$$

Due to an unidentified mistake in a lengthy derivation, the approach could not be finished and evaluated. Related investigations remain subject to future work.

Chapter 8

Conclusion

This thesis has investigated new methods of tracking displaced sources in spherical surrounding microphone arrays. An in-depth study based on newly proposed cost functions has been performed to obtain a profound understanding of the capacity of algorithms that detect dislocated sound sources within the array.

The presented simulations show that a cost function penalizing higher-order wave-spectral components is robust and outperforms a measure based on interference of a complex-squared sum of sound-pressures. A case study on musical instruments showed that even dealing with complex radiation patterns, reasonable results can be achieved.

Moreover, discretization of rotations has been investigated for the purpose of rotational tracking. This work demonstrates that robust results can be obtained when using such an approach on absolute sound-pressures even without acoustic centering.

The task of acoustic centering in surrounding spherical microphone arrays is important for a more compact description of radiation patterns from musical instruments. Moreover, compact spherical loudspeaker arrays with low resolution benefit from the simplified radiation patterns obtained by acoustic centering. The performance of the developed centering algorithm depends on the spatial aliasing error. As expected, limitations that are evident in the analytic descriptions have been shown to predominate the simulations as well.

Future work should investigate ways of making the robust centering methods computationally efficient so that they can be applied in realtime applications. One proposition, the derivation of the tangential intensity vector as a prospectively robust approach is subject to future research.

Note that the herein presented results are not determined to obtain highly accurate estimations of the location of one or more acoustically effective sources of vibration. For that purpose, other measures instead of the herein presented centers for a multipole decomposition should be discussed as well.

For example, the less robust interference based measure might be considered again preferable, if the goal is to accurately locate superposed individual source positions in a spherical sound field, cf. monopole source synthesis [Gir96].

Appendix A

Rotational Calibration Of The Array

This chapter describes a concept how to calibrate the spherical coordinates of the analysis to a set of absolute points of the array. To achieve this at least three known directions are needed and a measurement of their absolute coordinates in the analysis grid. These coordinates can be obtained by directing a focusing sound-source to the known absolute directions at the array and extracting the sound pressure maxima at the analysis sphere grid. We have to take into account that these measured positions will all be superimposed by an error vector. The method will yield the mean of the error vectors and gives us a 3×3 rotation matrix which rotates the measured positions to the known absolute array positions

$$\boldsymbol{x}_{2i} = \boldsymbol{Q}_{cal}\boldsymbol{x}_{1i}. \tag{A.1}$$

The method has been adopted from optical tracking systems [Kwo98].

Least-Squares Approach

$$\frac{1}{P}\sum_{i=1}^{P}(\boldsymbol{x}_{2i} - \boldsymbol{Q}_{cal}\boldsymbol{x}_{1i}) \Rightarrow min \tag{A.2}$$

$$\frac{1}{P}\sum_{i=1}^{P}\boldsymbol{x}_{2i}^{t}\boldsymbol{Q}_{cal}\boldsymbol{x}_{1i} \Rightarrow max \tag{A.3}$$

$$\frac{1}{P}\sum_{i=1}^{P}\boldsymbol{x}_{2i}^{t}\boldsymbol{Q}_{cal}\boldsymbol{x}_{1i} = tr\left(\boldsymbol{Q}_{cal}^{t}\frac{1}{P}\sum_{i=1}^{P}\boldsymbol{x}_{2i}\boldsymbol{x}_{1i}^{t}\right) \tag{A.4}$$

$$C = \frac{1}{P} \sum_{i=1}^{P} \boldsymbol{x}_{2i} \boldsymbol{x}_{1i}^{t} \tag{A.5}$$

SVD Decomposition

Singular Value Decomposition of C:

$$C = \boldsymbol{U}\boldsymbol{W}\boldsymbol{V}^{t} \tag{A.6}$$

Trace

$$tr(\boldsymbol{Q}_{cal}^{t}\boldsymbol{C}) \Rightarrow max \tag{A.7}$$

$$tr(\boldsymbol{Q}_{cal}^{t}\boldsymbol{U}\boldsymbol{W}\boldsymbol{V}^{t}) = tr(\boldsymbol{V}^{t}\boldsymbol{Q}_{cal}^{t}\boldsymbol{U}\boldsymbol{W}) \tag{A.8}$$
$$= tr((\boldsymbol{V}^{t}\boldsymbol{Q}_{cal}^{t}\boldsymbol{U})\boldsymbol{W}) \tag{A.9}$$
$$= tr(\boldsymbol{C}'\boldsymbol{W}) \Rightarrow max \tag{A.10}$$

$$C' = \boldsymbol{V}^{t}\boldsymbol{Q}_{cal}^{t}\boldsymbol{U} \tag{A.11}$$

Since matrix $\boldsymbol{W}$ is a diagonal matrix and $\boldsymbol{C}$ is orthogonal since it is composed from three orthogonal matrices the trace $tr(\boldsymbol{C}'\boldsymbol{W})$ reaches its maximum if $\boldsymbol{C}'$ is the Identity matrix:

$$\boldsymbol{V}^{t}\boldsymbol{T}^{t}\boldsymbol{U} = \boldsymbol{I} \tag{A.12}$$

Deriving a calibrating rotation matrix

$$\boldsymbol{Q}_{cal} = \boldsymbol{U}\boldsymbol{V}^{t} \tag{A.13}$$

$$\boldsymbol{Q}_{cal} = \boldsymbol{U} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & det(\boldsymbol{U}\boldsymbol{V}^{t}) \end{pmatrix} \boldsymbol{V}^{t} \tag{A.14}$$

(a) x-y view of original positions

(b) x-y view of resulting positions

(c) x-z view of original positions

(d) x-z view of resulting positions

(e) y-z view of original positions

(f) y-z view of resulting positions
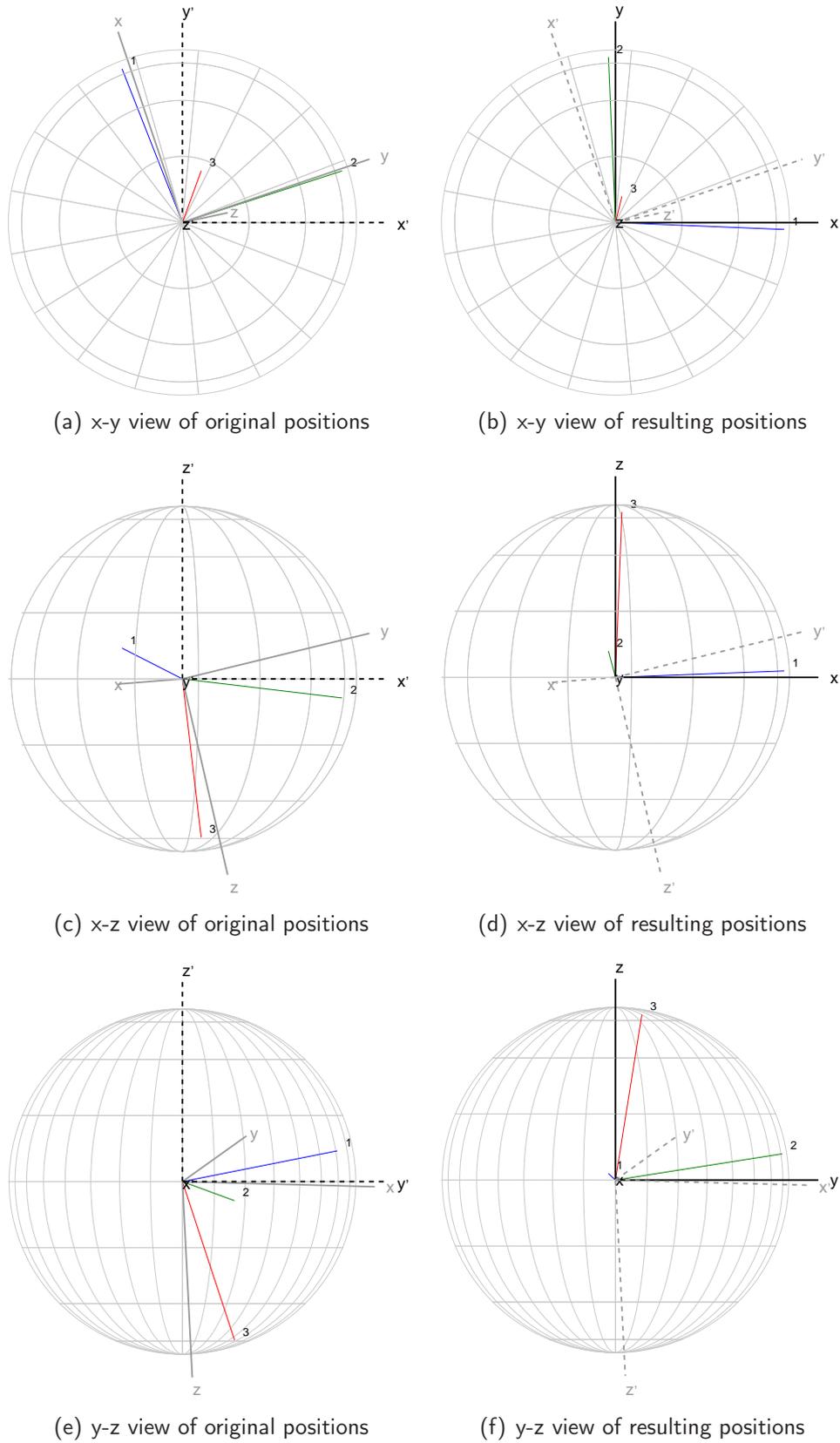
Figure A.1: Example calibration with three peak positions

(a) x-y view of original positions

(b) x-y view of resulting positions

(c) x-z view of original positions

(d) x-z view of resulting positions

(e) y-z view of original positions

(f) y-z view of resulting positions

Figure A.2: Example calibration with icosahedron (20 loudspeakers)

Appendix B

Localization Maps Of Simulated Sources

monopole at origin

(a) x-y view of J_{loc}

(b) x-y view of J_{ssc}

(c) x-z view of J_{loc}

(d) x-z view of J_{ssc}

(e) y-z view of J_{loc}
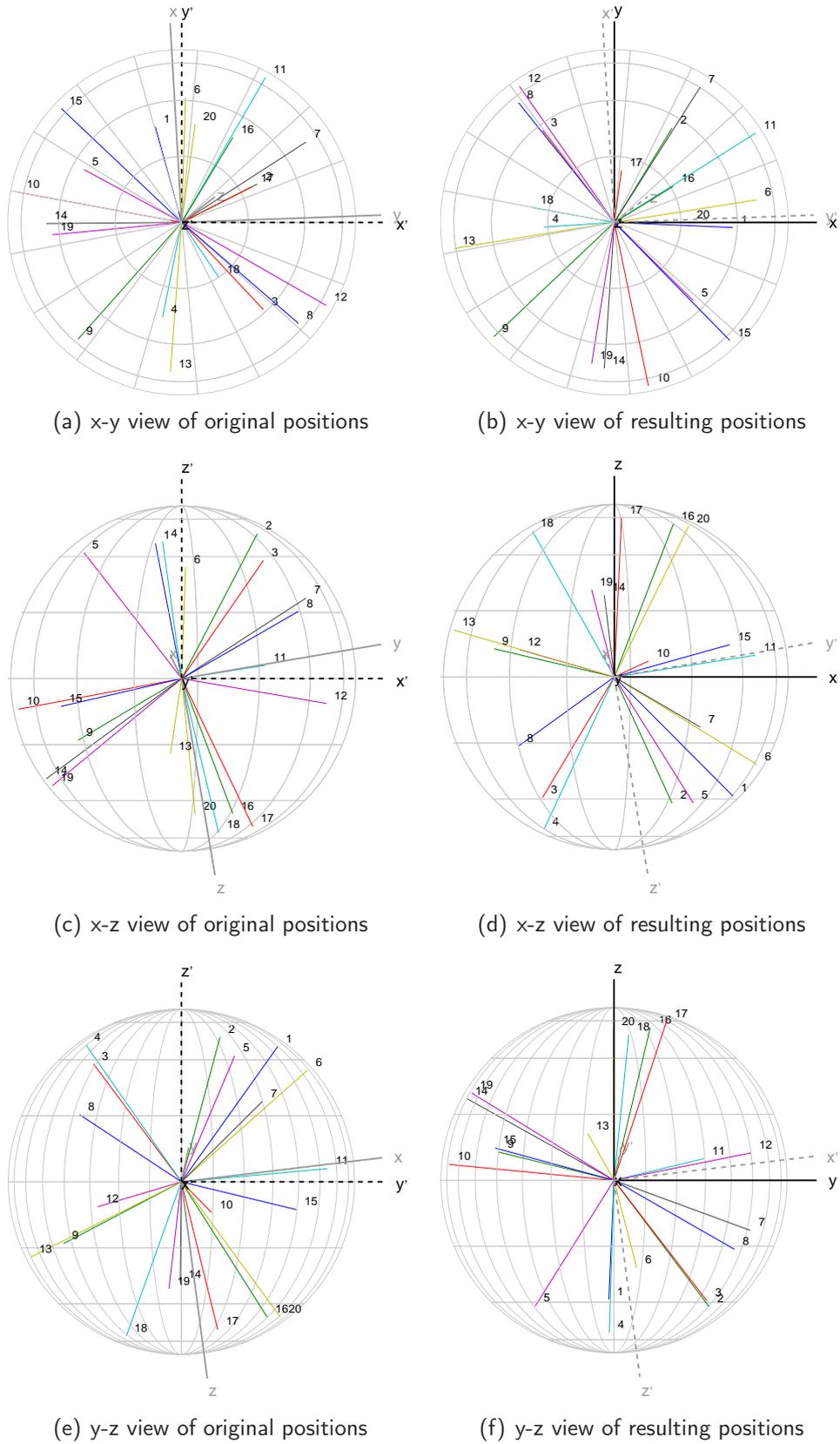
(f) y-z view of J_{ssc}
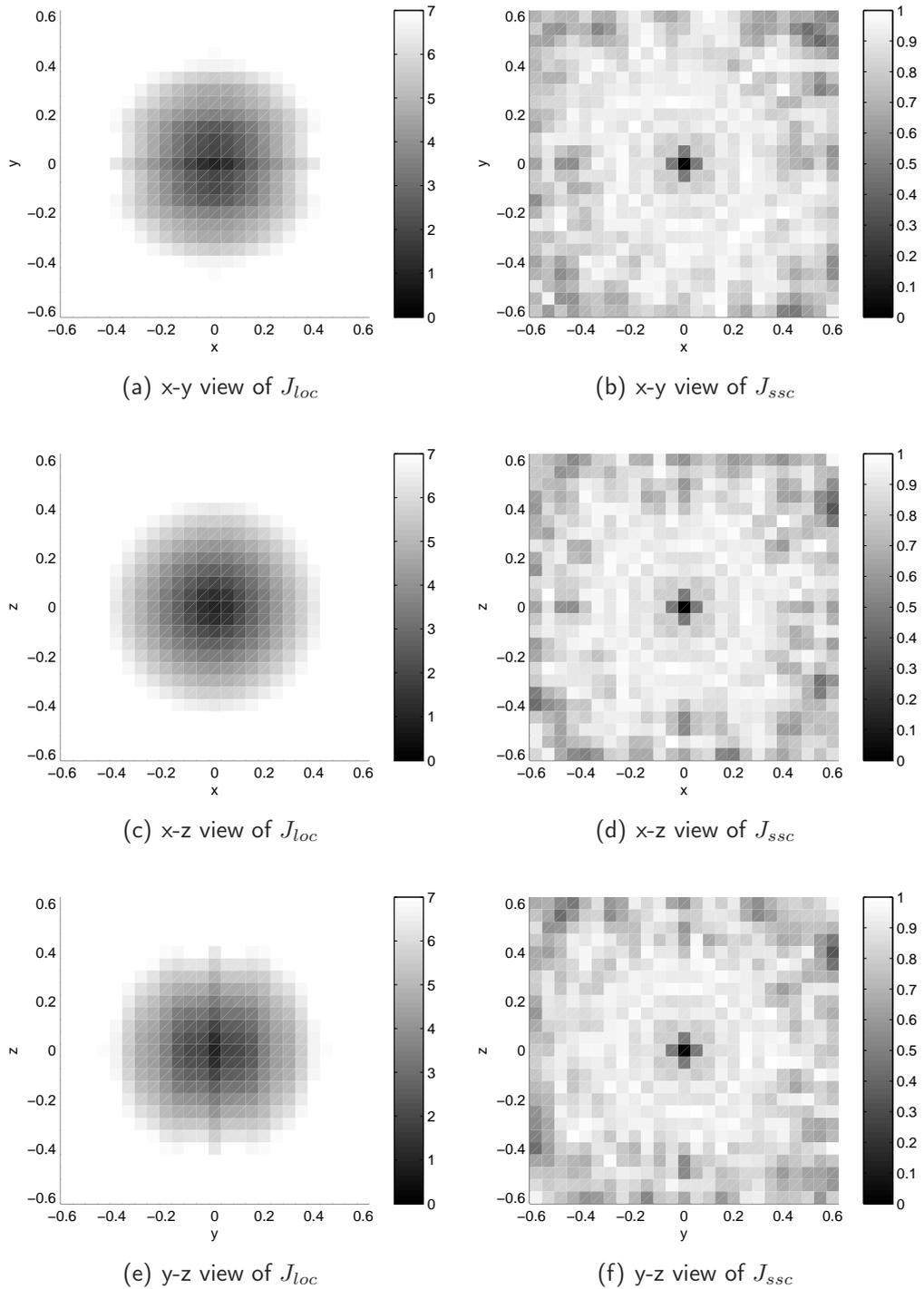
Figure B.1: monopole source, centered

Diploma thesis, Daniel Deboy

Dipole at origin

(a) x-y view of J_{loc}

(b) x-y view of J_{ssc}

(c) x-z view of J_{loc}

(d) x-z view of J_{ssc}

(e) y-z view of J_{loc}

(f) y-z view of J_{ssc}
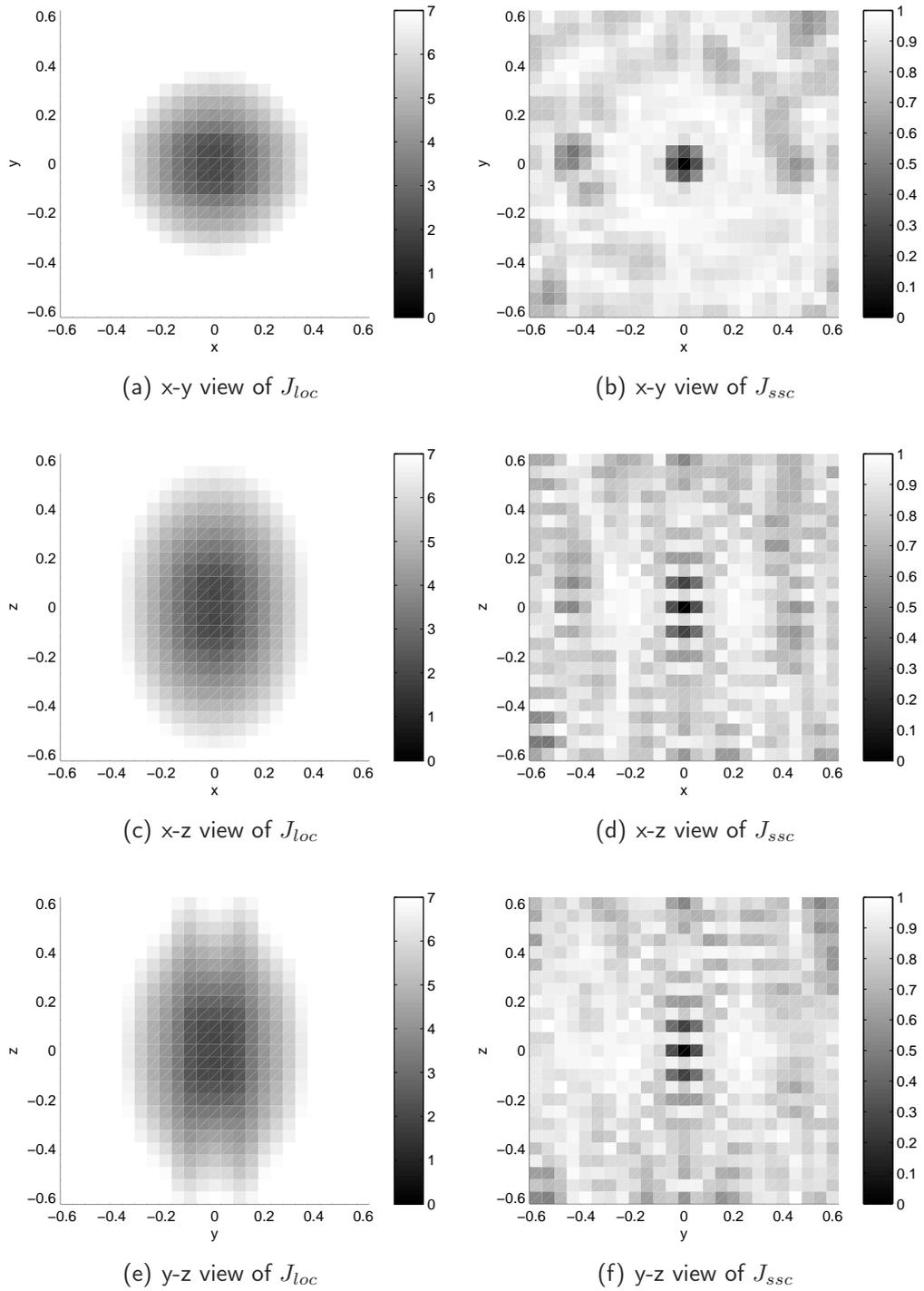
Figure B.2: dipole source on z-axis, centered

Real-valued multipole expansion

(a) x-y view of J_{loc}

(b) x-y view of J_{ssc}

(c) x-z view of J_{loc}

(d) x-z view of J_{ssc}

(e) y-z view of J_{loc}

(f) y-z view of J_{ssc}

Figure B.3: Multipole source with real-valued coefficients, centered

Complex-valued multipole expansion with random phase

(a) x-y view of J_{loc}
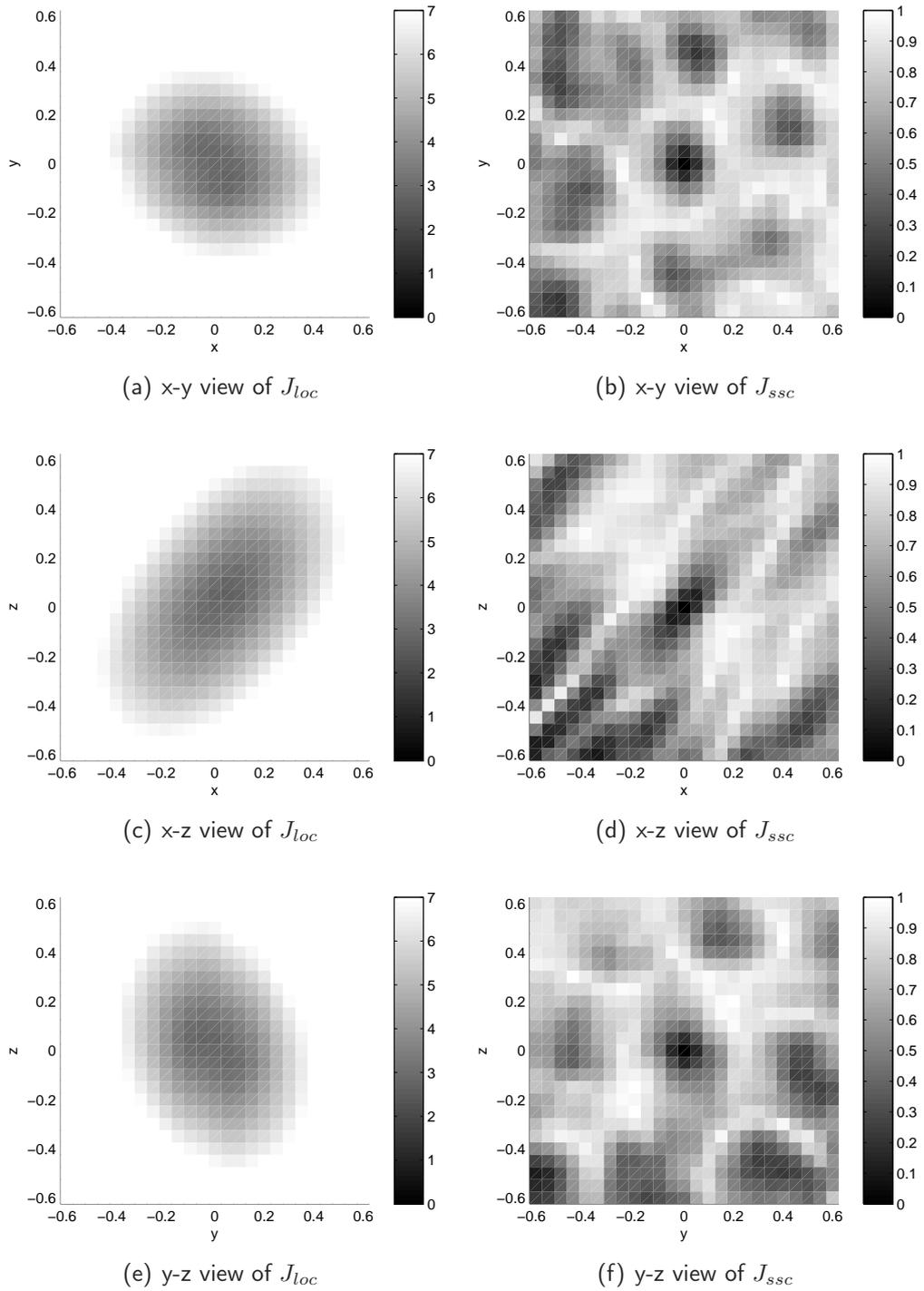
(b) x-y view of J_{ssc}

(c) x-z view of J_{loc}

(d) x-z view of J_{ssc}

(e) y-z view of J_{loc}

(f) y-z view of J_{ssc}
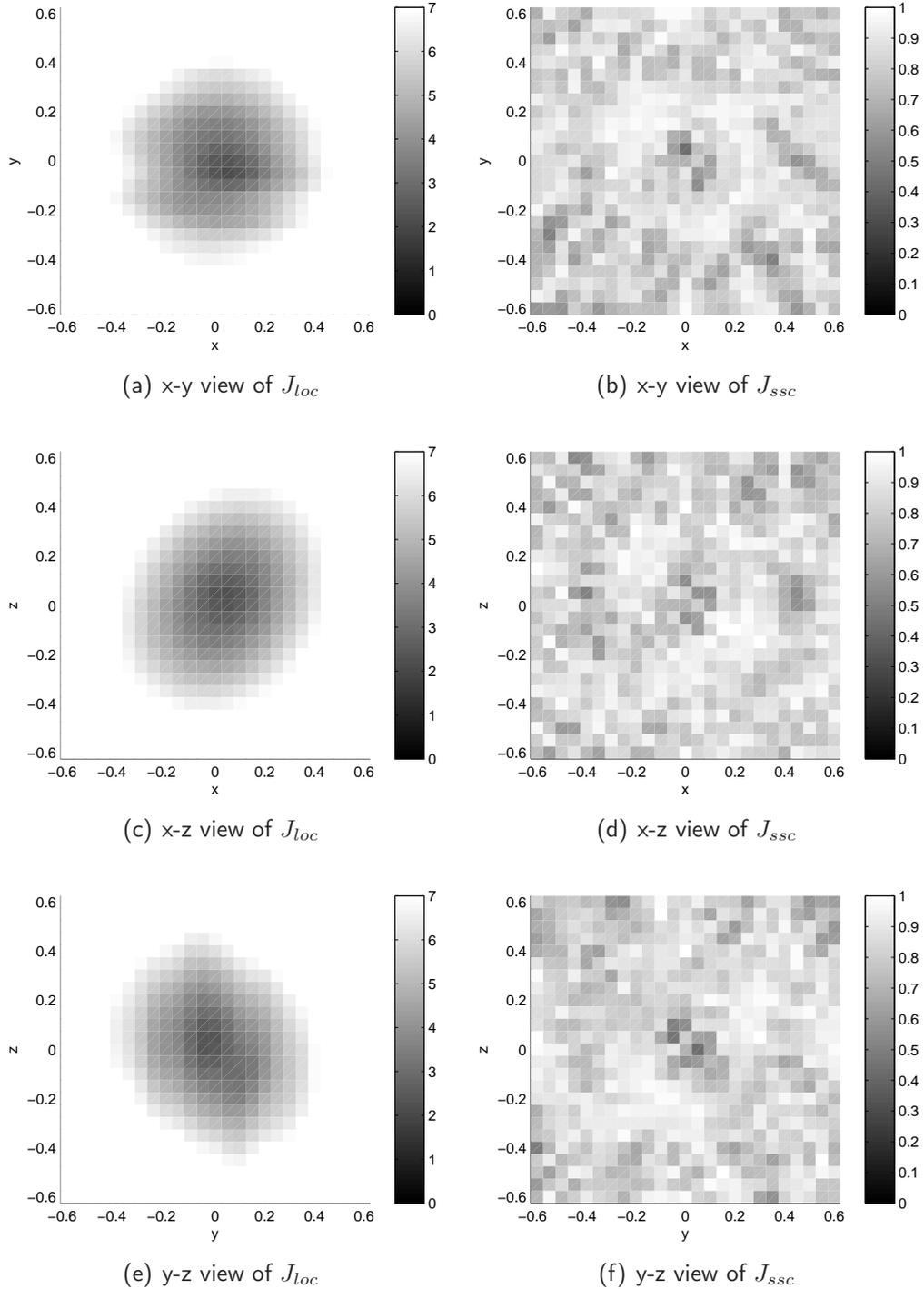
Figure B.4: Multipole source with complex-valued coefficients, centered

Bibliography

[Bög03] S. Böge, "Lineare Algebra I," 2003, Lecture Notes.

[GD04] N. A. Gumerov and R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*, 1st ed. Elsevier, 2004, no. 0-08-044371-0.

[Gir96] F. Giron, "Investigations about the directivity of sound sources," Ph.D. dissertation, Ruhr-Universität Bochum, 1996.

[Hoh09] F. Hohl, "Kugelmikrofonarray zur Abstrahlungsvermessung von Musikinstrumenten," Master's thesis, Institut für Elektronische Musik und Akustik, Universität für Musik und darstellende Kunst, Graz, Österreich, November 2009.

[Kwo98] Y. Kwon, "Computation of the rotation matrix," 1998. [Online]. Available: http://www.kwon3d.com/theory/jkinem/rotmat.html

[Leo06] P. Leopardi, "Distributing points on the sphere: Partitions, separation, quadrature and energy," Ph.D. dissertation, School of Mathematics and Statistics, Departement of Applied Mathematics, University of New South Wales, November 2006.

[Leo10] ——. (2010) Equal area partitioning toolbox for MATLAB. [Online]. Available: http://eqsp.sourceforge.net/

[Mat10a] T. MathWorks. (2010) fminsearch. [Online]. Available: http://www.mathworks.com/access/helpdesk/help/techdoc/ref/fminsearch.html

[Mat10b] W. MathWorld. (2010) Sphere packing. [Online]. Available: http://mathworld.wolfram.com/SpherePacking.html

[Mey72] J. Meyer, *Akustik und Musikalische Aufführungspraxis*, 1st ed. Das Musikinstrument, Frankfurt/Main, 1972.

[Mit07] J. C. Mitchell, "Discrete uniform sampling of rotation groups using orthogonal images," *Departments of Mathematics and Biochemistry*, 2007.

[Ple09] P. Plessas, "Rigid Sphere Microphone Arrays for Spatial Recording and Holography," Master's thesis, Graz University of Technology, Austria Institute of Electronic Music and Acoustics, University of Music and Performing Arts Graz Arts Graz, Austria CNMAT Center for New Music and Audio Technologies, University of California, Berkeley, November 2009.

[Van07] J. Vanderkooy, "Applications of the acoustic centre," *122nd AES Convention*, 2007.

[WA80] G. Weinreich and E. B. Arnold, "Method for measuring acoustic radiation fields," *Journal of the Acoustical Society of America*, vol. 68, August 1980.

[Wei10] E. Weisstein. (2010) Euler angles. [Online]. Available: http://mathworld.wolfram.com/EulerAngles.html

[Wik10] Wikipedia. (2010) N-spheres. [Online]. Available: http://en.wikipedia.org/wiki/N-sphere

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

[WS01] R. S. Womersley and I. H. Sloan, "How good can polynomial interpolation on the sphere be?" April 2001.

[Zot09] F. Zotter, "Analysis and Synthesis of Sound-Radiation with Spherical Arrays," Ph.D. dissertation, Insitute of Electronic Music and Acoustics, University of Music and Performing Arts Graz, Austria, 2009.