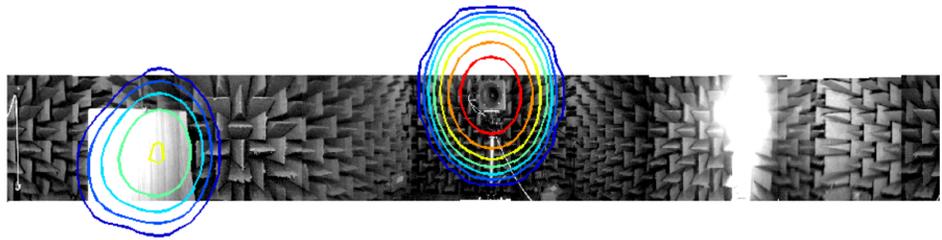


LOCALIZATION OF SOUND SOURCES USING 3D MICROPHONE ARRAY

Master of Science in Engineering Thesis



by M.Sc.Eng. Student
Svend Oscar Petersen

Supervisor: Peter Møller Juhl
University of Southern Denmark - September 2004

The title page shows the result of a Spherical Harmonics Beamforming in reflection free room with sound source and reflecting surface at Odense University College of Engineering.

Table of Contents

Table of Contents	iv
Abstract	vii
Acknowledgements	viii
Preface	ix
Introduction	1
1 Wave field in spherical coordinates	3
1.1 Introduction	3
1.2 Spherical Coordinates	4
1.3 Wave equation in spherical coordinates	4
1.4 Spherical Harmonics	6
1.5 Sound field on sphere	7
1.6 Simulation of sound field	12
1.7 Chapter summary	17
2 Spherical Harmonics Beamforming	18
2.1 Introduction	18
2.2 Spherical Harmonics Beamformer response	20
2.3 Properties of the SHB response	23
2.4 Focusing radius	25
2.5 Chapter summary	27
3 Numerical Integration on the Sphere	29
3.1 Introduction	30
3.2 Reproducing kernel	30

3.3	Spherical t-designs	33
3.4	Beamformer response using cubature rule	36
3.5	Error evaluation	37
3.6	Chapter summary	42
4	Analysis, optimization and implementation	43
4.1	Surface of the sphere	43
4.2	Stability	45
4.3	Optimization	49
4.4	Geometrical properties of selected array	56
4.5	Practical implementation of SHB	57
4.6	Chapter summary	60
5	SHB versus Delay and Sum Beamforming	61
5.1	Introduction	61
5.2	Delay and Sum Beamforming	61
5.3	Comparing Performances	66
5.4	Chapter summary	69
6	Tests of Beamformer	70
6.1	Introduction	70
6.2	Test and Measurement conditions	71
6.3	Maximum Sidelobe Level	72
6.4	Direction of the mainlobe	75
6.5	Radius dependance	77
6.6	Two sound sources	77
6.7	Full performance of beamformer	79
6.8	Test with reflecting wall	82
6.9	Chapter summary	84
7	Conclusion and Perspectives	85
7.1	Conclusion	85
7.2	Perspectives	87
	Appendices	88
A	Tests	88
A.1	Hardware	88
A.2	Test Setups	89
A.3	Test A - Single Sinus source	91

A.4	Test B - Two single frequency sources at different frequencies	93
A.5	Test C - Test Two sources at same frequency	96
A.6	Test D - Single Sinus source at different radii	98
A.7	Test E - Single Sinus source at different angles	100
A.8	Test F - White noise sound source	102
A.9	Test G - White noise and sinus source	105
A.10	Test H - Reflecting plate	107
B	Symbols, units and abbreviations	110
B.1	Symbols	111
B.2	Units	113
B.3	Abbreviations	113
	Bibliography	114

Abstract

This thesis investigates localization of sound sources by the use of 3D-microphone array. A beamforming method named *Spherical Harmonics Beamforming* (SHB) has been developed. The method is based on decomposing the acoustical wave field, measured on the surface of a sphere, into spherical harmonics, and thereafter extracting the angular part of the result in order to determine the direction of the wave field. The SHB-method has been compared with the conventional Delay and Sum Beamforming method, and showed improvements both in the directional gain and in the resolution.

The theoretical investigations in this thesis involves the decomposition of wave fields into spherical harmonics and Bessel functions, used both for simulations of acoustical wave fields and in the development of the SHB-method. Numerical integration on the sphere is examined in order to determine the decomposition constants of the SHB-response. Furthermore, this thesis involves optimization by error minimization, stability evaluations and an implementation method using FFT.

A test model of the SHB has been constructed, and tested to verify the theoretical performance of the method.

Acknowledgements

I would like to thank Brüel & Kjær for their interest in this project and for providing me with test equipment as well as the construction of the test model. I would especially like to thank Lars Alkestrup for helping with the construction of the test model and Claus Blaabjerg who helped me to get familiar with the test equipment.

I would furthermore like to thank Peter Møller Juhl who has been a great supervisor throughout the entire project, with brilliant guidance and helpful discussions.

This project completes my studies as Civil-Engineer in Physics and Technologies at Odense University College of Engineering and the University of Southern Denmark. I would therefore like to thank all the committed and inspiring teachers who have made my studies interesting and motivating, but also my fellow students Lykke, Hansen, Dr. Banke and Frue for their friendship.

Finally my thanks go to my wife Michele for her unwavering support.

Preface

This master of science in engineering thesis is the product of a masters project dealing with localization of sound sources using 3D-microphone arrays, and in particular an array of microphones placed on a sphere.

Before the project started, I highly prioritized working with both theoretical considerations, computer simulations as well as practical verifications through tests. This has all been achieved in this project with great help from Brüel & Kjær, with whom this project has been made in cooperation with.

Computer simulations have during the entire process of the project been a major tool for testing and verification of all theory. All simulation programs are made in MATLAB, and therefore most of the figures in this thesis are exported figures from MATLAB. The programs are not explained in this thesis, but all theory behind the simulations is. The theory is verified by illustrations from simulations.

In this thesis, \bar{x} denotes the complex conjugate of x , \vec{x} denotes a vector and $\overline{\overline{X}}$ denotes a matrix.

A prior knowledge of acoustics, signal processing and mathematics can be helpful in understanding this thesis, but is not a prerequisite.

This thesis is meant to be in colour, as many of the illustrations and graphs only can be interpreted by the colour. An electronic copy of this thesis can be found on my personal web page: www.oscarnpetersen.dk.

Svend Oscar Petersen
M.Sc.Eng. Student
September 1st 2004

Introduction

Microphone arrays are used for determining the direction of a present sound field or the location from which it originates from. The array and the following *array signal processing* is also called a *beamformer* and the technique is called *beamforming*. Beamforming can be described as a kind of *spatial filtering*. The beamformer can be steered in a particular direction/area in order to enhance signals from that direction/area and suppress signals from all other directions/areas [5].

Many conventional beamformers are based on *planner arrays*, where all the microphones are placed in the same two dimensional plane. The disadvantage of this type of beamformer is that it can not distinguish if the signals are approaching the array from the front or the back. The planner array beamformer works fine in situations where the array is placed outside the body from which the sound source originates, but has a major disadvantage inside confined enclosures, such as car cabins or aeroplane cabin.

A way to handle this problem is to place the microphones in a three dimensional array instead of a two dimensional planner array. In this project the position of the microphones are limited to be on a sphere with constant radius, a . Furthermore, the possibility of developing the sound field into spherical harmonics has been investigated.

An overview of the contents of each chapter in this thesis follows:

Chapter 1 starts this thesis with an explanation of why the spherical coordinate system has been chosen, and of how to describe the acoustical wave field, in spherical coordinates.

Chapter 2 explains the theory of beamforming and the properties that characterizes a beamformer response. Hereafter follows the development of the Spherical Harmonics Beamforming method (SHB), and the properties of the method.

Chapter 3 investigates how to perform numerical spherical integrations in order to perform the SHB developed in chapter 2.

Chapter 4 is a broad chapter, which analyzes the stability of the SHB, followed by an optimization of an array with 64 microphones, and finally an implementation method of the SHB technique.

Chapter 5 compares the performance of the SHB-method with the conventional Delay & Sum Method (DSB).

Chapter 6 outlines and discusses the results from tests made on the SHB test model.

Chapter 7 gives an overall conclusion of what has been achieved in this thesis, as well as the perspectives to be used in further studies of sound source localization using 3D-microphones arrays.

Appendix A explains in more detail about the various tests performed on the SHB test model as well as showing the results. Appendix B includes lists of symbols, units and abbreviations used in this thesis. Finally, an index of important words and terms can be found on the last pages.

Chapter 1

Wave field in spherical coordinates

In this chapter, a detailed explanation of how to express an acoustical wave field in spherical coordinates, is given, both at any point in space, and in particular on the surface of a sphere.

1.1 Introduction

In beamforming, it is common to mix a contour plot of the response of the microphone array, with a picture of the surrounding environment in which the beamformer operates. This makes it easy to determine where the sound originates from, instead of dealing with a set of coordinates.

The picture of the surroundings is the response of the optical, and describes the intensity and wavelength of the light that approaches the lens from a particular angle of incidence. To get a picture of the surroundings, all around the microphone array, in all possible directions in the space, it is necessary to combine the picture from more wide-angle lens cameras.

In the same way as the camera, the response of the microphone array describes both the intensity and wavelength (or frequency), but for the sound field instead of the light. To combine these two graphical responses, a common coordinate system should be chosen. The spherical coordinate system is the most appropriate for describing the angle of incidence of both the light and the sound.

1.2 Spherical Coordinates

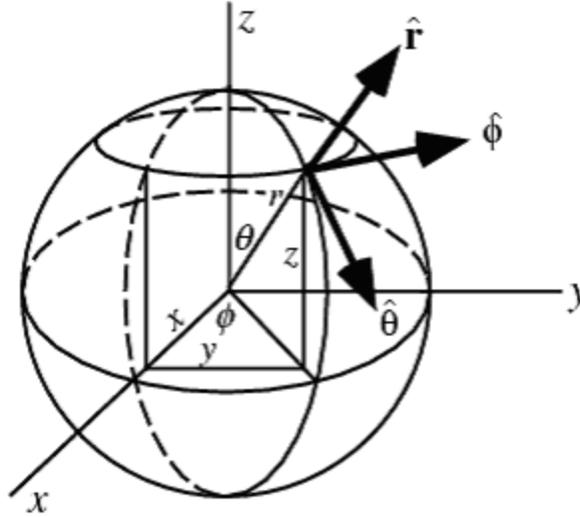


Figure 1.1: Definition of spherical coordinates.

The *spherical coordinate system* has three coordinates (θ, ϕ, r) . θ describes the *elevation*, ϕ the *azimuth* and r the radius from *origo*. A point in the spherical coordinate system can be found in the *cartesian coordinate system* $\{x, y, z\}$ as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = r \cdot \begin{bmatrix} \cos(\phi) \sin(\theta) \\ \sin(\phi) \sin(\theta) \\ \cos(\theta) \end{bmatrix}. \quad (1.2.1)$$

The definition of both the cartesian and the spherical axes is shown in figure 1.1.

1.3 Wave equation in spherical coordinates

The *wave equation* used to describe the wave field in acoustics, is in *spherical coordinates*[14] [13]

$$\frac{1}{R} \frac{\partial^2 \Phi}{\partial r^2} + \frac{2}{r} \frac{\partial \Phi}{\partial r} + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Phi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \Phi}{\partial \phi^2} = \frac{1}{c^2} \frac{\partial^2 \Phi}{\partial t^2}, \quad (1.3.1)$$

where the solution, Φ , is the *velocity potential* of the acoustical wave field. In this project the solution to the wave equation is not discussed, but just shown here. The details for this can be found in [1] and [14]. The solution to equation 1.3.1 can be written, by the separation of the variables, as a product of four functions

$$\Phi(r, \theta, \phi, t) = R(r) U(\theta) V(\phi) T(t). \quad (1.3.2)$$

The first function, $R(r)$, depends solely on the radius, r , and the solution can be described by using *spherical Bessel functions*. $U(\theta)$ and $V(\phi)$ depend on the angles, θ and ϕ , and the solution to these can be described by using spherical harmonics. The *time dependency* is separated to the function $T(t)$. In this thesis, time harmonic signals are assumed, defined as

$$T(t) = e^{-i\omega t}, \quad (1.3.3)$$

where ω is the *angular frequency* related to the frequency as $\omega = 2\pi f$.

The velocity potential can be used to find the *sound pressure*, p , as

$$p = \rho_0 \frac{\partial \Phi}{\partial t}, \quad (1.3.4)$$

where ρ_0 is the density of the media. The *particle velocity*, \vec{u} , can be found as the gradient of the velocity potential, as

$$\vec{u} = -\nabla \Phi. \quad (1.3.5)$$

1.4 Spherical Harmonics

The *spherical harmonics* play a major role in this thesis, and especially for the angular part of the velocity potential, and are therefore discussed in detail in this section.

Any complex function, $f \in S^2$, can be described by a linear combination of spherical harmonics (S^2 is the space defined on a sphere of constant radius). The spherical harmonics act as a set of basis functions.

$$f(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^n A_n^m Y_n^m(\theta, \phi) \quad (1.4.1)$$

A_n^m are a set of *decomposition constants*, and Y_n^m are the spherical harmonics defined as

$$Y_n^m(\theta, \phi) = \xi_n^m P_n^m(\cos \theta) \cdot e^{im\phi} \quad (1.4.2)$$

with

$$\xi_n^m = \sqrt{\frac{2n+1}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} \quad (1.4.3)$$

$P_n^m(x)$ are the so called *associated Legendre polynomials*, where n denotes the degree, and m the order ($n = \{0, 1, 2, \dots\}$ and $m = \{-n, -n+1, \dots, 0, 1, \dots, n-1, n\}$). The constant ξ_n^m is a normalization constant made to satisfy

$$\int_0^{2\pi} \int_0^\pi Y_n^m(\theta, \phi) \overline{Y_{n'}^{m'}}(\theta, \phi) \sin \theta d\theta d\phi = \delta_{nn'} \delta_{mm'}. \quad (1.4.4)$$

The top bar in $\overline{Y_n^m}$ denotes the complex conjugate of Y_n^m , and δ_{nm} is the mathematical *Kronecker delta*. Equation 1.4.4 is a very important property of the spherical harmonics, which means that the spherical harmonics are a set of orthonormal basis functions.

Equation 1.4.4, can also be used to determine the constants, A_n^m in 1.4.1, by multiplying both sides of 1.4.1 with the complex conjugate of the spherical harmonics, and then integrating over S^2 , we get

$$\begin{aligned}
\int_0^{2\pi} \int_0^\pi f(\theta, \phi) \overline{Y_{n'}^{m'}}(\theta, \phi) \sin \theta d\theta d\phi &= \int_0^{2\pi} \int_0^\pi \sum_{n=0}^{\infty} \sum_{m=-n}^n A_n^m Y_n^m(\theta, \phi) \overline{Y_{n'}^{m'}}(\theta, \phi) \sin \theta d\theta d\phi \\
\Downarrow \\
\int_0^{2\pi} \int_0^\pi f(\theta, \phi) \overline{Y_{n'}^{m'}}(\theta, \phi) \sin \theta d\theta d\phi &= \sum_{n=0}^{\infty} \sum_{m=-n}^n A_n^m \delta_{nn'} \delta_{mm'} \\
\Downarrow \\
A_n^m &= \int_0^{2\pi} \int_0^\pi f(\theta, \phi) \overline{Y_n^m}(\theta, \phi) \sin \theta d\theta d\phi.
\end{aligned} \tag{1.4.5}$$

It is then possible to describe any function in the spherical coordinate system by the use of spherical harmonics. The decomposition constants are found by the double integral in equation 1.4.5.

1.5 Sound field on sphere

In this section, we look into how the sound field on a sphere can be described by spherical harmonics and *spherical Bessel functions*. We look at two specific situations; first a *transparent sphere*, where the sound wave can propagate undisturbed through the sphere, and next a *hard sphere*, where the incident sound wave is scattered at the surface of the sphere. Furthermore, we look at two different incident waves; the plane wave and the spherical wave originating from a point outside the sphere. The *time dependency*, $T(t)$, is omitted in the expression of the velocity potential, Φ .

$$\Phi = \Phi(r, \theta, \phi) = R(r) U(\theta) V(\phi) \tag{1.5.1}$$

For a plane incident wave approaching the acoustically transparent sphere from the angles θ_0 and ϕ_0 the *velocity potential* at (r, ϕ, θ) is [1] [14]

$$\Phi = \Phi_i = \sum_{n=0}^{\infty} i^n \cdot [j_n(kr)] \cdot \sum_{m=-n}^n \overline{Y_n^m}(\theta_0, \phi_0) \cdot Y_n^m(\theta, \phi). \quad (1.5.2)$$

The radial part is described by the *spherical Bessel functions* of first kind, j_n . Generally, the radial part is described by a combination of spherical Bessel functions of first kind and *spherical Neumann functions*, n_n , but the spherical Neumann approaches infinity at origo [1]. Inclusion of the spherical Neumann functions will then represent a point source at origo. In this case, where the sphere is acoustically transparent, the value of the velocity potential is finite at origo, which is why it is sufficient to use only spherical Bessel functions of the first kind. The center of the sphere is placed at origo. k is the wave number, related to the frequency, f , as

$$k = \frac{\omega}{c} = \frac{2\pi f}{c}, \quad (1.5.3)$$

where c is the propagation speed of sound and ω is the *angular frequency*.

The *spherical Bessel* and *spherical Neumann* functions, j_n and n_n , are given by the normal *cylindrical Bessel* and *cylindrical Neumann* functions, J_n and N_n , as

$$j_n(z) = \sqrt{\frac{\pi}{2z}} \cdot J_{n+1/2}(z) \quad (1.5.4)$$

$$n_n(z) = \sqrt{\frac{\pi}{2z}} \cdot N_{n+1/2}(z).$$

J_n is also called the Bessel function of the first kind, and the Neumann function N_n is called the Bessel function of the second kind. Equation 1.5.2 can be used to describe the velocity potential at any point in the room. To calculate the field on the surface of an acoustically transparent sphere, choose $r = a$.

In the case of a spherical incident wave originating from a point, at radius r_0 , outside the sphere of radius, a , the velocity potential at (r, ϕ, θ) is

$$\Phi = \Phi_i = i \sum_{n=0}^{\infty} h_n^{(1)}(kr_>) \cdot [j_n(kr_<)] \cdot \sum_{m=-n}^n \overline{Y_n^m}(\theta_0, \phi_0) \cdot Y_n^m(\theta, \phi), \quad (1.5.5)$$

where the radii, $r_>$ and $r_<$, depends on the position of r and r_0

$$r_> = r \wedge r_< = r_0 \quad \text{for } r \geq r_0 \quad (1.5.6)$$

$$r_> = r_0 \wedge r_< = r \quad \text{for } r < r_0.$$

To find the wave field on the sphere of radius a , originating from the point source at radius r_0 , equation 1.5.7 would be

$$\Phi = \Phi_i = i \sum_{n=0}^{\infty} h_n^{(1)}(kr_0) \cdot [j_n(ka)] \cdot \sum_{m=-n}^n \overline{Y_n^m}(\theta_0, \phi_0) \cdot Y_n^m(\theta, \phi). \quad (1.5.7)$$

$h_n^{(1)}$ is the *spherical Hankel function* of the first kind, defined as

$$h_n^{(1)}(z) = j_n(z) + i \cdot n_n(z). \quad (1.5.8)$$

If the surface of the sphere is hard, so that the incident wave field is scattered, the total sound field will be a combination of both the incoming field, Φ_i , and the scattered field, Φ_{sc} .

$$\Phi = \Phi_i + \Phi_s \quad (1.5.9)$$

For the case where the incident wave is plane, the total wave field is [1]

$$\Phi = \Phi_i + \Phi_s = \sum_{n=0}^{\infty} i^n \cdot [j_n(kr) - a'_n \cdot h_n^{(1)}(kr)] \cdot \sum_{m=-n}^n \overline{Y_n^m}(\theta_0, \phi_0) \cdot Y_n^m(\theta, \phi). \quad (1.5.10)$$

Note that the radial part now contains the spherical Neumann functions, which represent the scattered wave. We do not now have the singularity problem at $r = 0$, since the shell of the sphere is hard, and we only look at the wave field outside the scattering surface. The acoustical center of the scattered wave is at the center of the sphere.

The constant a'_n in 1.5.10 is found by looking at the *boundary condition* on the surface of the sphere. We know that the particle velocity at the surface of a hard rigid body is zero. Using this, and the fact that the particle velocity is proportional to the first derivative of the velocity potential with respect to r (see equation 1.3.5), we get

$$\begin{aligned}
\frac{\partial \Phi}{\partial r} \Big|_{r=a} &= 0 \\
\Downarrow \\
\frac{\partial j_n(kr)}{\partial r} \Big|_{r=a} - a'_n \cdot \frac{\partial h_n^{(1)}(kr)}{\partial r} \Big|_{r=a} &= 0 \\
\Updownarrow \\
j'_n(ka) - a'_n h_n^{(1)'}(ka) &= 0 \\
\Updownarrow \\
a'_n &= \frac{j'_n(ka)}{h_n^{(1)'}(ka)}.
\end{aligned} \tag{1.5.11}$$

Finally, in the case of a spherical incident wave, the total wave field will be [1] [14]

$$\Phi = i \cdot \sum_{n=0}^{\infty} h_n^{(1)}(kr_0) [j_n(kr) - a'_n \cdot h_n^{(1)}(kr)] \cdot \sum_{m=-n}^n \overline{Y_n^m}(\theta_0, \phi_0) \cdot Y_n^m(\theta, \phi). \tag{1.5.12}$$

Summing up on these four situations, we can write the velocity potential from equation 1.5.2, 1.5.7, 1.5.10 and 1.5.12 as a general expression (equation 1.5.13).

$$\boxed{\Phi(r, \theta, \phi) = \sum_{n=0}^{\infty} R_n(r, r_0) \cdot \sum_{m=-n}^n \overline{Y_n^m}(\theta_0, \phi_0) \cdot Y_n^m(\theta, \phi)} \tag{1.5.13}$$

The radial part, R_n , contains the information that decides whether the sphere is hard or transparent, and if the incident wave is plane or spherical. Table 1.1 shows R_n in these four situations.

R_n	Hard	Transparent
Plane wave	$i^n \cdot [j_n(ka) - a'_n \cdot h_n^{(1)}(ka)]$	$i^n \cdot j_n(ka)$
Spherical wave	$i \cdot h_n^{(1)}(kr_0)[j_n(ka) - a'_n \cdot h_n^{(1)}(ka)]$	$i \cdot h_n^{(1)}(kr_0) \cdot j_n(ka)$

Table 1.1: Radial part, R_n , of the velocity potential of sound field in spherical coordinates.

1.6 Simulation of sound field

The formula in equation 1.5.13, using the appropriate radial part, R_n , from tabel 1.1, can be used to calculate the sound field at any point in the room, both outside the hard rigid sphere, as well as inside and outside the transparent sphere. The transparent sphere is equivalent to having no sphere present at all.

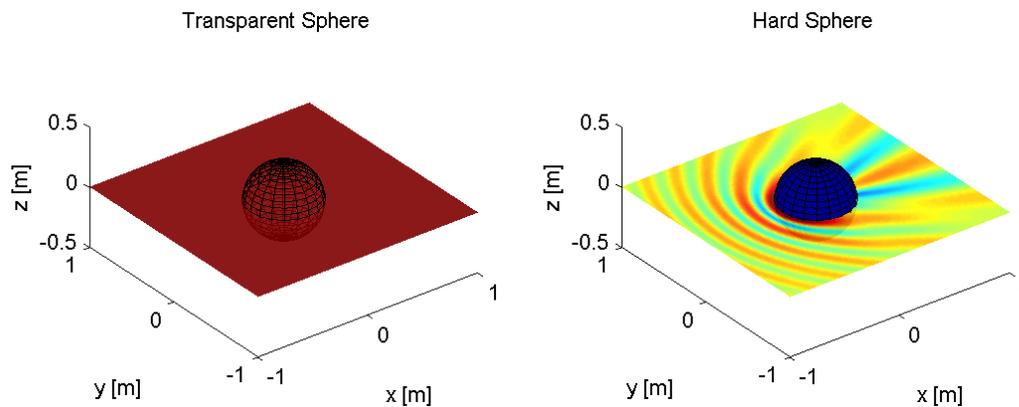


Figure 1.2: Incident plane wave travelling in the positive x-direction. Left: Absolute value of the sound field travelling undisturbed through a transparent sphere. Right: The same field scattered on a hard rigid sphere. Note the interference of the incident and the scattered field on the right. ($ka = 5$, and radius of sphere $a = 0.3$)

Having a plane incident wave travelling in the positive x direction, the field in the xy-plane ($z = 0$) is simulated using MATLAB, to the degree, $N = 100$, of spherical harmonics. This is illustrated in figure 1.2 using a sphere of radius $a = 0.3m$ (seen at the center of the figure), and at the frequency $ka = 5$. Using higher degrees of spherical harmonics does not give a noticeable difference in the plotting. The colour in the xy-plane at $z = 0$ represents the absolute value of the velocity potential, which is proportional to the absolute value of the sound pressure (see equation 1.3.4). The left picture of figure 1.2, shows the case of a transparent sphere, where the plane sound wave travels undisturbed through the sphere (the overall red colour indicates that the intensity of the pressure field is constant). The right picture shows the situation of a hard rigid scattering sphere. The incoming plane wave interferes with the scattered

wave, resulting in an interference pattern where the intensity varies throughout the plane.

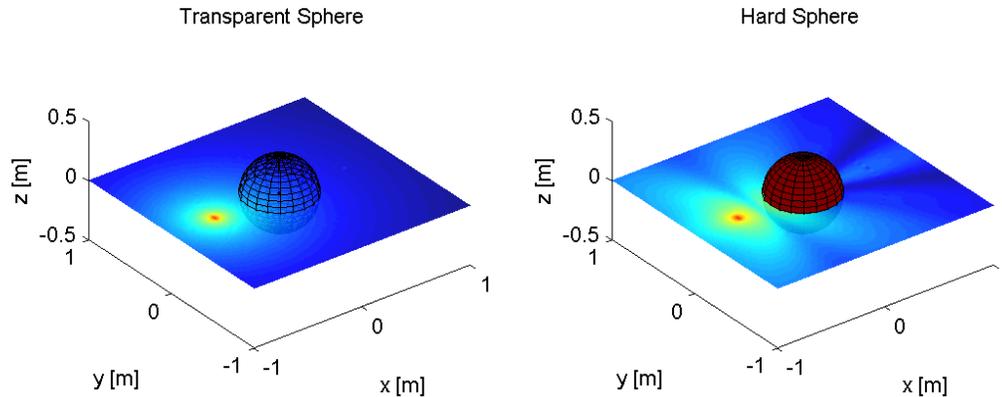


Figure 1.3: Incident spherical wave from point source placed at $(x, y, z) = (-0.6, 0, 0)$. The development into spherical harmonics is made to the degree $N = 50$. Left shows the situation with the transparent sphere, and right the hard rigid sphere. ($ka = 5$, and radius of sphere $a = 0.3$)

Figure 1.3 shows another situation, where the incident wave is spherical and originates from a point source located at -0.60 m on the x -axis. Again the field is simulated at the same plane, for both the transparent sphere (left), and for the scattering sphere (right). The scaling of the colour is logarithmical, because the value of the sound pressure at the center of the point source is approaching infinity. Otherwise, it will not be possible to see the interference pattern for the scattering sphere, and the xy -plane would be completely blue, except for a red dot at the point source location.

Equation 1.5.13 can also be used to calculate the wave field on the surface of a sphere. Figure 1.4 shows the relative amplitude distribution of the sound pressure over a sphere, for a plane wave travelling in the positive x -direction, at $ka = 1$. What cannot be seen in figure 1.4, is that the pressure rises at the back of the sphere as expected (see [11]). The sound pressure at the black equator line in figure 1.4, is shown in figure 1.5. Here it is also shown how the amplitude of the pressure converges; at $N = 3$, it is very close to the exact solution. To calculate the sound pressure at positions outside the sphere, the maximum degree of spherical harmonics necessary to

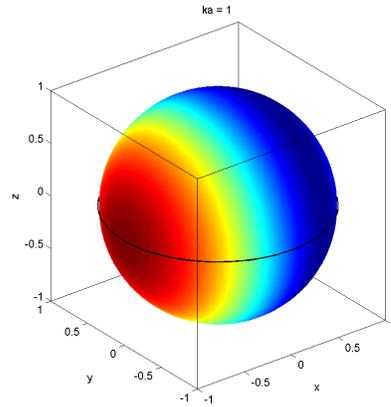


Figure 1.4: Sound field on scattering sphere at $ka = 1$.

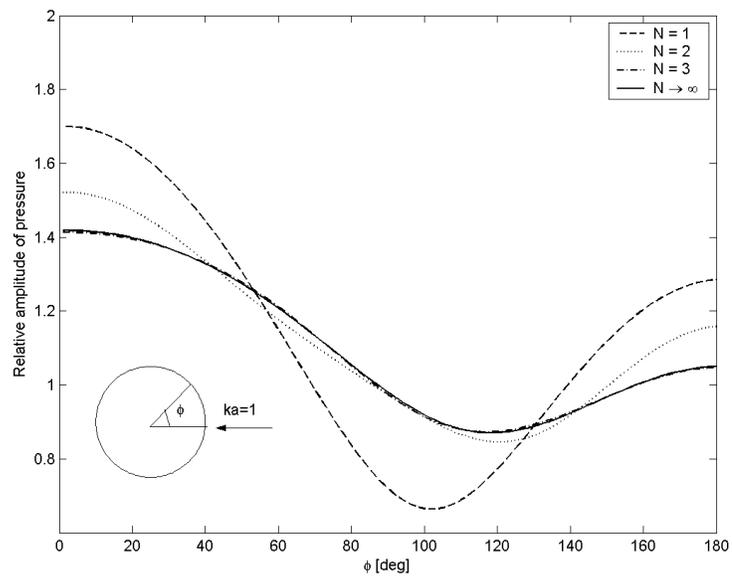


Figure 1.5: Relative amplitude of pressure on the surface of the scattering sphere at $ka = 1$.

get a result as close as possible to the exact solution, depends on the radius r . Figure 1.6 illustrates this at $ka = 1$; for $N = 5$, the error is already noticeable at $r = 3a$, for $N = 10$, the error is noticeable at around $r = 7a$ and for $N = 15$, it is possible to calculate the wave field at distances around $r < 10a$. So, when including higher degrees of spherical harmonics, it is possible to calculate the wave field at greater distances from the sphere.

The error of the wave field depends also on the frequency, which results in the need for higher degree, N , at higher frequency. To illustrate this, an expression for how the amplitude of the wave field converges as a function of the frequency, ka , and the degree, N , has been defined (equation 1.6.1). This function is based on the situation shown in figure 1.5; the amplitude of the pressure at the equator, for a plane incident wave approaching from $(\theta_0, \phi_0) = (\pi/2, 0)[rad]$.

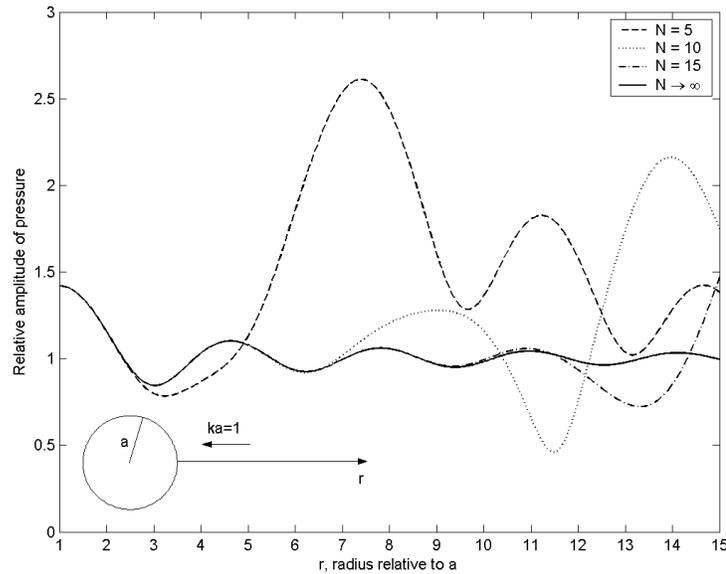


Figure 1.6: Relative amplitude of pressure at radius relative to the radius of the sphere for $ka = 1$. At distances further away from the sphere, higher degree of spherical harmonics is required.

$$\mu_N(ka) \equiv \frac{1}{2\pi} \int_0^{2\pi} \frac{|\Phi_N(\frac{\pi}{2}, \phi) - \Phi_{N-1}(\frac{\pi}{2}, \phi)|}{|\Phi_\infty(\frac{\pi}{2}, \phi)|} d\phi \quad (1.6.1)$$

Using equation 1.6.1, we get an idea of how much each step from $N - 1$ to N contributes, $|\Phi_N(\frac{\pi}{2}, \phi) - \Phi_{N-1}(\frac{\pi}{2}, \phi)|$, to the final amplitude of the wave field, $|\Phi_\infty(\frac{\pi}{2}, \phi)|$, in average around the equator of the sphere.

N	$\mu_N(ka = 0.1)$	$\mu_N(ka = 1)$	$\mu_N(ka = 10)$
1	0.096158	0.76179	0.24145
2	0.0027986	0.23824	0.31493
3	5.0284e-005	0.043073	0.37429
4	6.5479e-007	0.0056282	0.43028
5	6.7271e-009	0.00058077	0.48631
6	5.72e-011	4.9789e-005	0.54908
7	4.1478e-013	3.6146e-006	0.61556
8	2.66e-015	2.258e-007	0.682
9	1.4888e-017	1.2735e-008	0.7336
10	7.5699e-020	6.4321e-010	0.68891
20	1.1275e-044	9.87e-025	3.908e-005
30	1.0282e-071	8.9248e-042	5.1973e-012

Table 1.2: $\mu_N(ka)$ at different values of ka , and different values of N , showing that the Φ_N converges faster at low frequencies.

If the contribution from each step converges, then Φ_N must converge towards the exact solution when $N \rightarrow \infty$. This is calculated numerically from 360 discrete points around the equator, as in figure 1.4 and 1.5, and Φ_{50} is used for Φ_∞ (higher N did not show any changes on the first five decimals of μ_N). The results are listed in table 1.2.

When developing a wave field into spherical harmonics and spherical Bessel functions, there exist two major factors that determine how high a degree, N , is needed to get an acceptable result; namely r and ka . Generally, a higher N is needed when:

- The radius for the point of calculation increases.

- The frequency increases.

1.7 Chapter summary

A discussion of the acoustical wave field in spherical harmonics is made, leading to a general expression of the wave field 1.5.13, using spherical harmonics and spherical Bessel functions. The expression can be used in both the situations of having either a plane incoming wave field or a spherical incoming wave field, by choosing the appropriate radial part from table 1.1. In table 1.1, it is furthermore possible to choose between a scattering sphere or a transparent sphere.

In section 1.6, some examples of simulations are made to get an idea of how high degree, N , of the spherical harmonics and order of the spherical Bessel functions, are needed to calculate a suitable wave field at a given point using 1.5.13. Generally, if the frequency or the distance increases, higher degree, N , is needed.

Chapter 2

Spherical Harmonics Beamforming

In this chapter, the concept of beamforming and the parameters characterizing the performance of a beamformer response are discussed. This is followed by the development of a method to perform beamforming by using spherical harmonics, and this method is named *Spherical Harmonics Beamforming*(SHB).

The work in this chapter has been inspired by the work of P. Thushara D. Abhayapala who has described the theory of modal analysis of beamforming [5]. He has used it for radial transformations from near field into far field beamforming. Among other interesting studies on this subject is the work of Hulseboe, Vries and Bourdillat who have used a circular array of microphones to decompose the sound field into cylindrical harmonics for a recreation of the sound field in the horizontal plane [6].

2.1 Introduction

In the concept of beamforming, there exists three stages. The first stage is the measurement or detection of the sound field performed by the hardware; i.e. the microphone array. The next stage is the computation of the measured signals. In this stage, the algorithm steers the microphone array in order to amplify the signal from a given direction, and suppress the signals from other directions. This is also called *spatial filtering*. This computational stage could also include a filtering in the frequency domain (i.e. low-pass filtering, FFT etc.). The last stage is the read out, usually done as a contour plot, mixed with a picture of the surroundings in which the

microphone array is placed. The picture is taken by a camera placed in connection with the array.

The *beamformer response* is a general expression for the output of the beamforming algorithm, and there exists two types; *beam pattern* and *steered response*. The difference between them is [10]:

Beam pattern The microphone array is steered in one direction, while the angle of the incoming wave field is varied.

Steered response The angle of the incoming wave field is set, and the steering direction is varied.

These two beamformer responses can appear very similar, but can differ in cases where the array is good at steering in one direction. In this thesis, the beamformer response is denoted as $b(\theta, \phi, \theta_0, \phi_0)$, where (θ, ϕ) is the steering direction and (θ_0, ϕ_0) is the angle of the incident wave field.

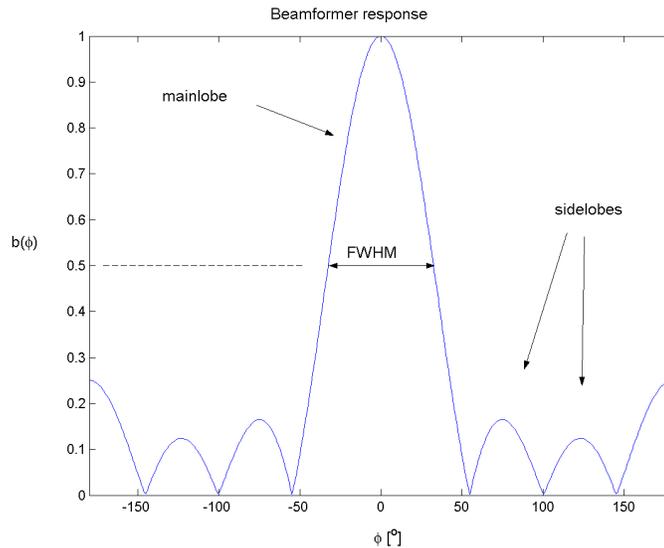


Figure 2.1: Beamformer response in one dimension with mainlobe and sidelobes

The beamformer response is characterized by having a *mainlobe* in the (θ_0, ϕ_0) direction, and *sidelobes* in other directions, see figure 2.1. The performance of the beamformer is described by its *resolution* and the attenuation of the sidelobes. The resolution is determined by the width of the mainlobe, which should be as little as possible, measured by the *Full Width Half Max* (FWHM). This value determines the minimum angle between two incoming waves, where it is possible to distinguish them from each other. The attenuation of the sidelobes is described by the level of the biggest sidelobe, the *Maximum Sidelobe Level* (MSL), relative to the mainlobe. The term *directional gain* (DG), is also used and is related to MSL as: $DG = -MSL[dB]$ (see [2], [3] and [10]).

2.2 Spherical Harmonics Beamformer response

The ideal beamformer response from a spherical microphone array, is a function having a maximum value proportional to the field strength when looking in the direction of an incoming wave field, and zero when looking in all other directions. A function with this property is the *delta function* and is denoted by $\delta(x - x_0)$. The ideal beamformer response could then be expressed as

$$b_{ideal}(\theta, \phi, \theta_0, \phi_0) = \delta(\theta - \theta_0) \delta(\phi - \phi_0). \tag{2.2.1}$$

The angles (θ, ϕ) denote the steering direction, and (θ_0, ϕ_0) denote the direction in which the incoming wave approaches the microphone array from. If we decompose the ideal beamformer response 2.2.1 into spherical harmonics, using equation 1.4.1 and 1.4.5, we get

$$b_{ideal}(\theta, \phi, \theta_0, \phi_0) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \alpha_{ideal}^{n,m} Y_n^m(\theta, \phi) \tag{2.2.2}$$

$$\alpha_{ideal}^{n,m} = \int_0^{2\pi} \int_0^\pi \delta(\theta - \theta_0) \delta(\phi - \phi_0) \overline{Y_n^m}(\theta, \phi) \sin \theta d\theta d\phi.$$

The integral property of the delta function results in the following decomposition constants

$$\alpha_{ideal}^{n,m} = \overline{Y_n^m}(\theta_0, \phi_0). \tag{2.2.3}$$

The ideal beamformer response decomposed into N^{th} degree of spherical harmonics can now be written as

$$\begin{aligned} b_N(\theta, \phi, \theta_0, \phi_0) &= \sum_{n=0}^N \sum_{m=-n}^n \overline{Y_n^m}(\theta_0, \phi_0) \cdot Y_n^m(\theta, \phi) \\ &\rightarrow b_{ideal}(\theta, \phi, \theta_0, \phi_0) \qquad \text{for } N \rightarrow \infty. \end{aligned} \tag{2.2.4}$$

When decomposing the beamformer response into higher degrees of spherical harmonics, it approaches the delta function.

The next problem is then to retrieve this from the wave field on a sphere of radius, a . The wave field in spherical harmonics (as described in equation 1.5.13) looks very similar to the the ideal beamformer response (equation 2.2.4) apart from the radial part, R_n .

Proposition 2.2.1. *The ideal beamformer response decomposed to N^{th} degree of spherical harmonics, can be obtained from the wave field, Φ_a , on a sphere of constant radius a , as*

$$b_N(\theta, \phi, \theta_0, \phi_0) = \sum_{n=0}^N \sum_{m=-n}^n \frac{A_n^m}{R_n(a, r_0)} \cdot Y_n^m(\theta, \phi), \tag{2.2.5}$$

where

$$A_n^m = \int_0^{2\pi} \int_0^\pi \Phi_a(\theta, \phi, \theta_0, \phi_0) \overline{Y_n^m}(\theta, \phi) \sin \theta d\theta d\phi. \tag{2.2.6}$$

Proof. Inserting the wave field Φ_a , on the sphere of radius a , from equation 1.5.13 into equation 2.2.6, the constants A_n^m are

$$\begin{aligned}
 A_n^m &= \int_0^{2\pi} \int_0^\pi \left[\sum_{n'=0}^\infty R_{n'}(a, r_0) \sum_{m'=-n'}^{n'} \overline{Y_{n'}^{m'}}(\theta_0, \phi_0) \cdot Y_{n'}^{m'}(\theta, \phi) \right] \overline{Y_n^m}(\theta, \phi) \sin \theta d\theta d\phi \\
 &\Downarrow \\
 A_n^m &= \sum_{n'=0}^\infty \sum_{m'=-n'}^{n'} R_{n'}(a, r_0) \cdot \overline{Y_{n'}^{m'}}(\theta_0, \phi_0) \cdot \int_0^{2\pi} \int_0^\pi Y_{n'}^{m'}(\theta, \phi) \cdot \overline{Y_n^m}(\theta, \phi) \cdot \sin \theta d\theta d\phi.
 \end{aligned}
 \tag{2.2.7}$$

From equation 1.4.4, we know that the double integral in 2.2.7 equals one when $n' = n$ and $m' = m$, and zero otherwise. Equation 2.2.7 can then be reduced to

$$A_n^m = R_n(a, r_0) \cdot \overline{Y_n^m}(\theta_0, \phi_0).
 \tag{2.2.8}$$

Inserting this into equation 2.2.5, we get

$$\begin{aligned}
 b_N(\theta, \phi, \theta_0, \phi_0) &= \sum_{n=0}^N \sum_{m=-n}^n \frac{R_n(a, r_0) \cdot \overline{Y_n^m}(\theta_0, \phi_0)}{R_n(a, r_0)} \cdot Y_n^m(\theta, \phi) \\
 &= \sum_{n=0}^N \sum_{m=-n}^n \overline{Y_n^m}(\theta_0, \phi_0) \cdot Y_n^m(\theta, \phi),
 \end{aligned}
 \tag{2.2.9}$$

which is the same as equation 2.2.4. □

Equation 2.2.5 and 2.2.6 are the basic equations needed to obtain a SHB-response from a spherical microphone array, if the acoustical wave field is known in the entire space S^2 on the sphere of radius a . In the division by the radial part, $R_n(a, r_0)$, in equation 2.2.5, the known radius of the sphere, a , is needed, as well as the distance to the acoustical center, r_0 , for the incoming spherical wave. This distance is not necessarily known and can be considered as a *focusing radius* (the consequences of this will be discussed later).

2.3 Properties of the SHB response

The ideal beamformer response, found using equation 2.2.4 at a limited degree of spherical harmonics, N , is characterized by having one mainlobe in the (θ_0, ϕ_0) direction and N sidelobes symmetrically around the mainlobe. This can be seen in figure 2.2.

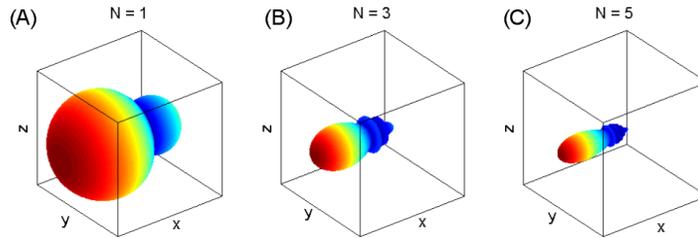


Figure 2.2: Ideal beamformer response, $b(\theta, \phi)$, with $(\theta_0, \phi_0) = (\pi/2, 0)$. (A) $N=1$, (B) $N=3$ and (C) $N=5$.

The ideal SHB-response is rotational symmetric around the axis in the direction of the incoming wave (θ_0, ϕ_0) . The symmetry itself is however independent of the direction (θ_0, ϕ_0) . This is illustrated in figure 2.3. The performance of the SHB response is therefore the same in all directions of the 4π -room, also called *omnidirectional*.

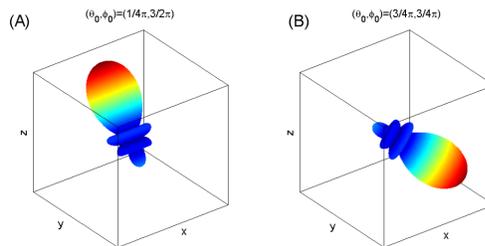


Figure 2.3: Ideal beamformer response, $b(\theta, \phi)$, with different values of (θ_0, ϕ_0) .

When using higher degrees of spherical harmonics, the width of the mainlobe decreases, which results in a better resolution (see figure 2.4), and the attenuation of the sidelobes is larger, especially for the far sidelobes. The first sidelobe (nearest the

mainlobe) does not decrease significantly when N reaches a value of 6 or more, but it moves closer to the mainlobe (note that $b_N(\phi)$ for $N = 20$ in figure 2.4 is closer to the *delta function* than for $N = 5$ or $N = 1$).

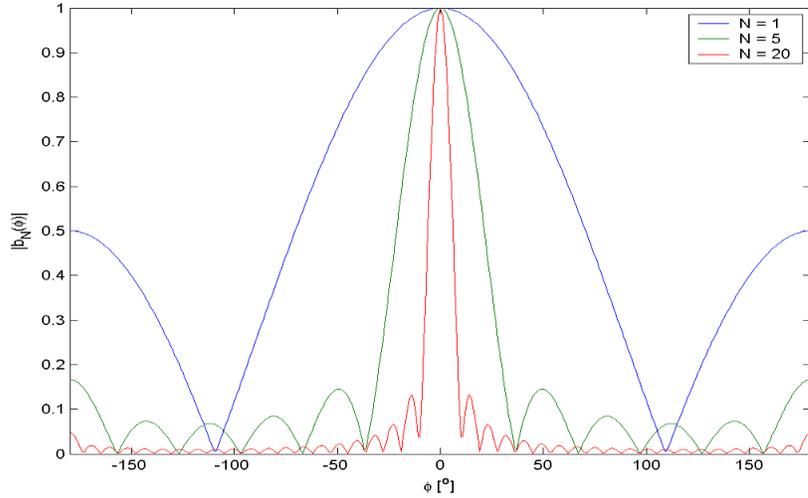


Figure 2.4: Absolute value of ideal spherical harmonics beamformer response, developed to different degrees of N .

The most significant sidelobe at low degree of spherical harmonics, $N \leq 6$, is the *backlobe*. (The backlobe is the sidelobe in the opposite direction of the mainlobe). At higher degrees, $N > 6$, the attenuation of the *backlobe* increases further, and the first sidelobe becomes the most significant, with an MSL of -17 dB, see figure 2.5. Figure 2.5 also shows the resolution measured as the FWHM of the mainlobe. The FWHM decreases rapidly at low N , from 143° at $N = 1$ to under 45° at $N = 5$.

In general, both the resolution and the attenuation of sidelobes increase when using higher degrees of spherical harmonics, but most significantly for $N \leq 6$.

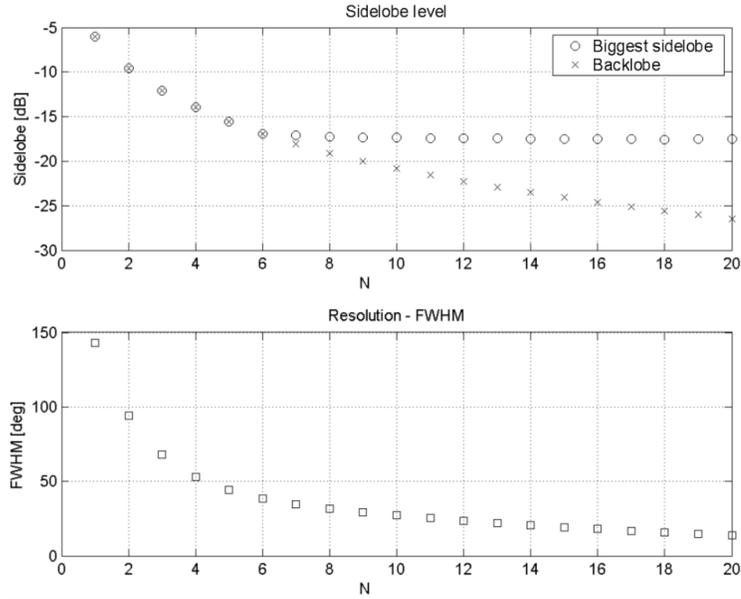


Figure 2.5: Sidelobe attenuation and resolution of the ideal beamformer response of degree N .

2.4 Focusing radius

When operating with the SHB, the distance to the acoustical center of the sound source is usually not known, but is required when calculating the beamformer response, b_N , (see equation 2.2.5). In this section, we examine the consequences of choosing a *focusing radius*, that is not the same as the actual radius.

In equation 2.2.9, we anticipated that the radius, r_0 , was known. In the case where we do not know from what distance the incoming wave originated from, we would have to guess the radius, and this radius we call the focusing radius, denoted r_{foc} . Using this focusing radius, we get the following beamformer response

$$b_N(\theta, \phi, \theta_0, \phi_0) = \sum_{n=0}^N \sum_{m=-n}^n \frac{R_n(a, r_0)}{R_n(a, r_{foc})} \cdot \overline{Y_n^m}(\theta_0, \phi_0) \cdot Y_n^m(\theta, \phi). \quad (2.4.1)$$

This means, that the different degrees, n , of the beamformer response gets weighted incorrectly, since the radial part, R_n , does not cancel out in equation 2.4.1. Figure

2.6 shows what happens to $|b_N|$, for $N = 5$ at low frequencies, $ka = 1$.

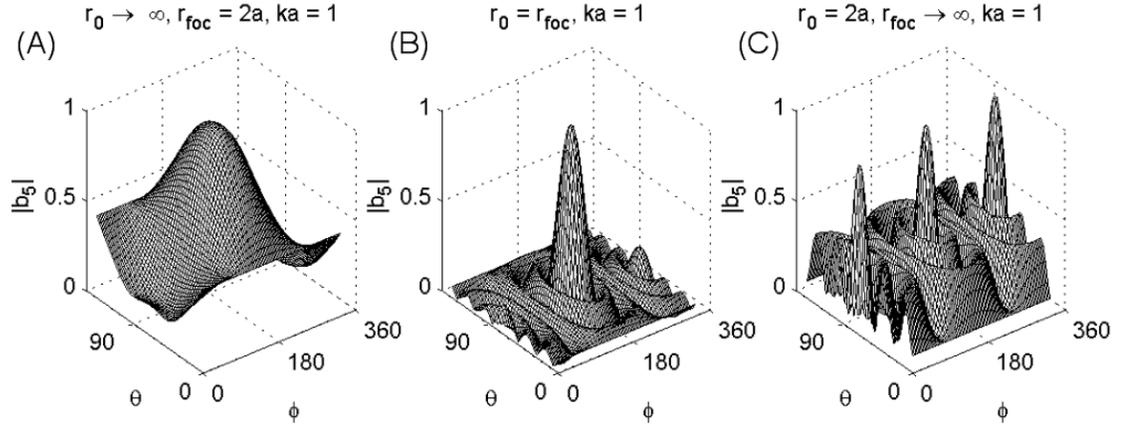


Figure 2.6: Consequences of focusing radius at low frequency, $ka = 1$. (A) $r_{\text{foc}} = 2a$ and plane wave approaching. (B) $r_{\text{foc}} = r_0$ the right focusing radius. (C) point source at $r_0 = 2a$ and radius focused for plane waves.

If the incoming wave is plane and we focus the beamformer at the distance two times the radius of the sphere, $r_{\text{foc}} = 2a$, we get a response where the mainlobe is wider and the sidelobes are lower than expected (see figure 2.6 A). In the opposite case, where the incoming wave originates from a distance two times the radius, and the beamformer is focussed at high radius, we get a response where the mainlobe is narrower but the sidelobes are higher than expected (see figure 2.6 C). Figure 2.6 B shows the right focusing radius. The problem is worst at low frequencies. Figure 2.7 shows the same situations as figure 2.6 but at a higher frequency, $ka = 5$.

The result of a wrong focusing radius is that if the focusing radius is too low, the beamformer response will have worse resolution, but a larger attenuation of the sidelobes. If the focusing radius is too high, then the resolution is improved but at the cost of lower attenuation of the sidelobes.

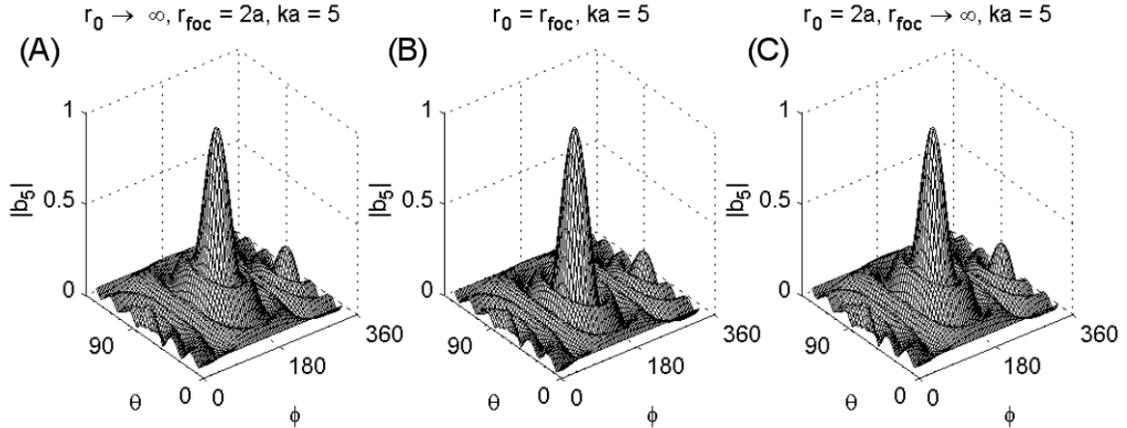


Figure 2.7: Consequences of focusing radius at high frequency, $ka = 5$. (A) $r_{foc} = 2a$ and plane wave approaching. (B) $r_{foc} = r_0$ the right focusing radius. (C) point source at $r_0 = 2a$ and radius focused for plane waves.

2.5 Chapter summary

To sum up this chapter, a list of design steps to achieve the beamformer response in equation 2.2.5, of the developed Spherical Harmonics Beamforming-method, is presented. Furthermore, it gives the reader an overview of the main equations in this chapter.

1. Decide what type of sphere to use. Hard scattering sphere or an transparent sphere.
2. Determine the sound field on the sphere (or measure using microphones).
3. Calculate the decomposition constants of equation 2.2.6, using the sound field.
4. Insert these constants in equation 2.2.5, to determine the beamformer response.
5. Set the focusing radius to determine the radial part, R_n , of the beamformer response.

The properties of the SHB-method have been investigated, and the performance showed to be omnidirectional. Furthermore, the investigations showed that the MSL

and the FWHM of the beamformer response both decrease (i.e. improved directional gain and resolution), when higher N is included.

The next challenge is step 2 on the above list, determining the sound field on the sphere from a limited amount of discrete points on the sphere. This is discussed in the next chapter.

Chapter 3

Numerical Integration on the Sphere

One of the major problems in this project is to perform the surface integral for determining the decomposition constant of the beamformer response in spherical harmonics, 2.2.5, and achieve a result as close as possible to the ideal constants 2.2.3. The integral needed to be solved is

$$A_n^m = \int_0^{2\pi} \int_0^\pi \Phi_a(\theta, \phi) \overline{Y_n^m}(\theta, \phi) \sin \theta d\theta d\phi. \quad (3.0.1)$$

In the practical implementation, it is impossible to measure the continuous variation of the sound field on the spherical surface of radius a . It is only possible to measure the sound field at discrete points by the use of microphones. This leads to the investigation of numerical integration of spherical functions on the sphere. An expression of the error on the decomposition constants, caused by using the selected numerical integration method, is found. This chapter uses the shorter notation $x_0 = (\theta_0, \phi_0)$ for the angle of the incoming wave field, and $x_q = (\theta_q, \phi_q)$ for the coordinates of the points on the sphere representing the microphone positions.

3.1 Introduction

The most common and simple method for performing numerical integration is the *cubature rule* [15] [8]

$$F = \sum_{q=1}^Q w_q f(x_q), \quad (3.1.1)$$

where the function needed to be integrated is f . However, the values of f are only known at the discrete set of Q points, $\vec{x} = \{x_1, x_2, \dots, x_Q\}$. w_q is the *cubature weight* at x_q . Using this numerical integration method, the decomposition constants for the beamformer response can be found as

$$\alpha_n^m(x_0) = \frac{A_n^m(x_0)}{R_n(r_0, a)} = \frac{1}{R_n(r_0, a)} \sum_{q=1}^Q w_q \Phi_a(x_0, x_q) \overline{Y_n^m(x_q)}. \quad (3.1.2)$$

Firstly, a method for determining the cubature weights by using a reproducing kernel is discussed. Then in section 3.3 the work of Hardin and Sloane, who have worked on finding optimal designs for spherical integration on a sphere, by using equal cubature weights, is outlined.

3.2 Reproducing kernel

The cubature weights can be found by solving the following matrix equation [15]

$$\overline{\overline{G}} \vec{w} = \vec{e}, \quad (3.2.1)$$

where \vec{w} is a vector containing the cubature weights, \vec{e} is a vector of 1's and $\overline{\overline{G}}$ is the *reproducing kernel matrix*. The elements of the *reproducing kernel matrix*, are found from the orthonormal set of basis functions, which in this case are the spherical harmonics, as [15]

$$G_N(x, y) = \sum_{n=0}^N \sum_{m=-n}^n Y_n^m(x) \overline{Y_n^m(y)}, \quad (3.2.2)$$

where $x, y \in S^2$. The property of the reproducing kernel can be seen in the following proposition.

Proposition 3.2.1. *Having a function $p \in \mathbb{P}_N(S^2)$, where \mathbb{P}_N is the space of spherical polynomials of degree at most N , the reproducing kernel has the property*

$$(p(x), G_N(x, y))_x = p(y), \quad (3.2.3)$$

where $(p, q)_x$ denotes the inner product, with x as the integration variable, defined as

$$(p, q)_x \equiv \int_{S^2} p(x) \overline{q(x)} ds(x). \quad (3.2.4)$$

Proof. In equation 3.2.2 we saw the definition of the reproducing kernel, and by inserting that into equation 3.2.3 we get

$$\begin{aligned} (p(x), G_N(x, y))_x &= \int_{S^2} p(x) \sum_{n=0}^N \sum_{m=-n}^n \overline{Y_n^m(x)} Y_n^m(y) ds(x) \\ &= \sum_{n=0}^N \sum_{m=-n}^n \left(\int_{S^2} p(x) \overline{Y_n^m(x)} ds(x) \right) Y_n^m(y). \end{aligned} \quad (3.2.5)$$

If we decompose p into spherical harmonics, we only need to include spherical harmonics at most degree N because $p \in \mathbb{P}_N(S^2)$. From equation 1.4.5 we get the following decomposition constants

$$A_n^m = \int_{S^2} p(x) \overline{Y_n^m(x)} ds(x). \quad (3.2.6)$$

and by using 1.4.1 we can retrieve p at a given point, y , as

$$\begin{aligned}
p(y) &= \sum_n^N \sum_{m=-n}^n A_n^m Y_n^m(y) \\
&= \sum_n^N \sum_{m=-n}^n \int_{S^2} p(x) \overline{Y_n^m(x)} ds(x) Y_n^m(y) \\
&= (p(x), G_N(x, y))_x.
\end{aligned} \tag{3.2.7}$$

□

Given a fundamental system, $\vec{x} = \{x_1, \dots, x_{d_N}\}$, where d_N is the number of points, it is possible to find $p(x)$ at these fundamental points, \vec{x} , as

$$p(x_i) = (p(x), G_N(x, x_i)), \quad i = 1, \dots, d_N, \tag{3.2.8}$$

and $G_N(x, x_i)$ is called the *reproducing kernel basis*, also denoted as $g_i(x) = G_N(x, x_i)$. This still requires that p is known at any point on S^2 , and this is why the *discrete inner product* is introduced by using the cubature rule 3.1.1

$$(p(x), q(x)) = \sum_{j=1}^{d_N} w_j p(x_j) q(x_j), \tag{3.2.9}$$

and by using this in equation 3.2.8, we get

$$p(x_i) = \sum_{j=1}^{d_N} w_j p(x_j) G_N(x_j, x_i) \quad x_i, x_j = 1, \dots, d_N. \tag{3.2.10}$$

This can be written as a set of equations, and by introducing $g_{i,j} = G_N(x_i, x_j)$ we get the following matrix equation

$$\begin{bmatrix} p(x_1) \\ \vdots \\ p(x_i) \\ \vdots \\ p(x_{d_N}) \end{bmatrix} = \begin{bmatrix} p(x_1)g_{1,1} & \cdots & p(x_j)g_{j,1} & \cdots & p(x_{d_N})g_{d_N,1} \\ \vdots & \ddots & \vdots & & \vdots \\ p(x_1)g_{1,i} & \cdots & p(x_j)g_{j,i} & \cdots & p(x_{d_N})g_{d_N,i} \\ \vdots & & \vdots & \ddots & \vdots \\ p(x_1)g_{1,d_N} & \cdots & p(x_j)g_{j,d_N} & \cdots & p(x_{d_N})g_{d_N,d_N} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_j \\ \vdots \\ w_{d_N} \end{bmatrix}. \tag{3.2.11}$$

As in one dimension where two points are required to find a polynomial of first degree, three points to find a polynomial of second degree, and $(N + 1)$ points to find a polynomial of N^{th} degree, the number of points required in two dimensions to find a polynomial of N^{th} degree is $d_N = (N + 1)^2$ (see, for example [15]). In order to find $p(x) \in \mathbb{P}_N$ exact, we need $d_N = (N + 1)^2$ number of points. Solving this in the case where $p(x) = 1$ in the entire space, $x \in S^2$, we get the equation as in 3.2.1

$$\begin{bmatrix} g_{1,1} & \cdots & g_{j,1} & \cdots & g_{d_N,1} \\ \vdots & \ddots & \vdots & & \vdots \\ g_{1,i} & \cdots & g_{j,i} & \cdots & g_{d_N,i} \\ \vdots & & \vdots & \ddots & \vdots \\ g_{1,d_N} & \cdots & g_{j,d_N} & \cdots & g_{d_N,d_N} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_j \\ \vdots \\ w_{d_N} \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} \Leftrightarrow \overline{G}\overline{w} = \overline{e}. \quad (3.2.12)$$

We now have a tool to find the cubature weights of equation 3.1.1 from a set of discrete basis points, $\overline{x} = \{x_1, \dots, x_{d_N}\}$, on the surface S^2 . Having found the cubature weights, \overline{w} , we can calculate the numerical integration of a function, $f \in \mathbb{P}_N$, on the sphere exactly, if the number of points satisfy $d_N = (N + 1)^2$. This method is optimized for integrating *spherical polynomials*, but the spherical polynomials are closely related to the spherical harmonics used as the basis functions.

3.3 Spherical t-designs

Another simpler way of using the cubature rule for numerical integration, is by choosing equal cubature weights, also called the *quasi-Monte Carlo cubature rule*. Equation 3.1.1 can then be written as [15]

$$F = \frac{1}{Q} \sum_{q=1}^Q f(x_q). \quad (3.3.1)$$

One of the advantages of using quasi-Monte Carlo, is that we are not restricted to use a quadratic number of points, as in the case of using the reproducing kernel. In this thesis, the work of Hardin and Sloane [7], has been a major inspiration for the use of this method. Hardin and Sloane have investigated different spherical designs with high order of symmetry, in order to perform the numerical integration exactly, including as high as possible degree of spherical polynomials. A spherical design constructed by a set of Q points, $\vec{x} = \{x_1, \dots, x_Q\}$ on the unit sphere S^2 forms a *spherical t -design* if the identity

$$\int_{S^2} f(x) ds(x) = \frac{1}{Q} \sum_{q=1}^Q f(x_q) \quad (3.3.2)$$

holds for all polynomials f of degree $\leq t$ [7]. (In this section, t is not the symbol for time, but symbolizes the degree of spherical polynomials.)

Q	$\tau(Q)$	Symmetry order	Symmetry description
24	7	24	improved snub cube
25	5	10	$10^2 + 5$
36	8	12	3 snub tetrahedra
48	9	24	2 snub cubes
60	10	12	5 snub tetrahedra
64	10	12	$12^5 + 4$

Table 3.1: Properties of selected spherical t-designs. τ denotes the highest degree of spherical polynomials possible for exact integration.

The highest possible degree of *spherical polynomials* that satisfy equation 3.3.2 from a specific set of Q points, is denoted $\tau(Q)$. Hardin and Sloane have searched for systems consisting of 1 point up to 100 points, which have the highest $\tau(Q)$, and they have listed the results with the highest order of symmetry (see [7]). Table 3.1 lists some of the results that show very high τ from a relatively low number of points Q . For comparison, $Q = 25$ is also listed to show that a higher number of points, does not necessarily result in a higher $\tau(Q)$. It is possible to perform an exact integration up to a degree of spherical polynomials at most 7 from only 24 points, but with just

one more point, it is only possible to perform an exact integration to a degree at most 5.

In the practical design of the spherical beamformer, and for the directivity of the beamformer response, a high order of symmetry would be an advantage. For example, a symmetry order of 12 would mean that one basis point can by the use of *symmetry operators* be mapped at 12 different symmetrical positions. Therefore, to construct the system of 60 points with the symmetry order of 12, all we need are 5 basis points, and the rest of the points can be determined by a set of symmetry operators.

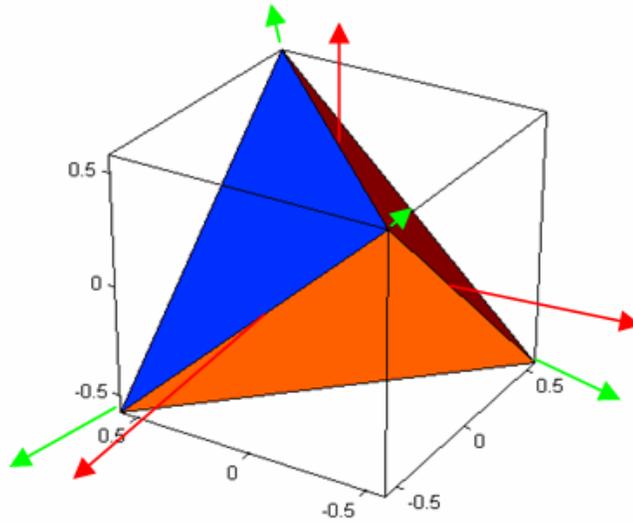


Figure 3.1: Symmetry of the tetrahedra structure. Red arrows are C_2 rotational symmetry axis, and the green arrows are C_3 rotational symmetry axis.

Figure 3.1 shows the symmetry axes of the *tetrahedra structure*, which is equivalent to the symmetry of $Q = 60$ in table 3.1. The green arrows represent the C_3 symmetry axis, and the red arrows represent the C_2 symmetry axis. A C_n axis is a *rotational symmetry axis*, with the property that if the whole system is rotated $\frac{360}{n}$ degrees around the symmetry axis, the new system will appear to be the same as before the symmetric operation. [4]

In order to test the reproducing kernel, described in last section, equation 3.2.1

has been solved for all t-designs consisting of an quadratic number of points. In all results the cubature weights, w , where equal, as expected.

3.4 Beamformer response using cubature rule

Use of the cubature rule, for numerical integration on the sphere, results in the following decomposition constants for the beamformer response

$$\alpha_{cub}^{n,m} = \frac{A_n^m}{R_n(a, r_0)} = \frac{1}{R_n(a, r_0)} \sum_q^Q w_q \Phi_a(x_0, x_q) \overline{Y_n^m(x_q)}, \quad (3.4.1)$$

where the cubature weights can be found from equation 3.2.1, by using the reproducing kernel. The cubature weights are equal if the system is a spherical t-design.

Using the cubature rule results in the following beamformer response (see equation 2.2.5)

$$b_N(x) = \sum_{n=0}^N \sum_{m=-n}^n \frac{1}{R_n(a, r_0)} \sum_q^Q w_q \Phi_a(x_0, x_q) \overline{Y_n^m(x_q)} \cdot Y_n^m(x)$$

$$\Downarrow$$

$$b_N(x) = \sum_q^Q \left[\sum_{n=0}^N \frac{1}{R_n(a, r_0)} \sum_{m=-n}^n w_q \overline{Y_n^m(x_q)} \cdot Y_n^m(x) \right] \Phi_a(x_0, x_q).$$

(3.4.2)

In the framed version of equation 3.4.2, it can be seen that the beamformer response, $b_N(x)$, steered in the direction, x , can be found by summing up the wave field, $\Phi_a(x_0, x_q)$, at the Q microphone positions, x_q , which are each multiplied by the factor in the square brackets of equation 3.4.2. The factor multiplied at each microphone signal is independent of the direction of the incoming field, but still depends on the radius from where the incoming field originates, r_0 , as well as the steering direction, x .

3.5 Error evaluation

Now that we have an expression for finding the decomposition constants of the beamformer response from a discrete set points on the sphere, we are able to evaluate an error. The error to be evaluated can be found by comparing the decomposition constants determined numerically by 3.4.1, with the ideal constants defined to be 2.2.3

$$\alpha_{ideal}^{n,m} = \overline{Y_n^m}(\theta_0, \phi_0). \quad (3.5.1)$$

If it is possible to recreate the ideal decomposition constants, then it is possible to get the ideal beamformer response. The decomposition constants depend on the angle of incidence, x_0 , therefore

$$\varepsilon^{n,m}(x_0) = \alpha_{cub}^{n,m}(x_0) - \alpha_{ideal}^{n,m}(x_0). \quad (3.5.2)$$

By finding the mean square error through an integration over x_0 , we get an expression for the error that is independent of the angle of incidence, as

$$\begin{aligned} mse_n^m &= \frac{1}{4\pi} \int_{S^2} |\varepsilon^{n,m}(x_0)|^2 ds(x_0) \\ &\Downarrow \\ mse_n^m &= \frac{1}{4\pi} \int_{S^2} |\alpha_{cub}^{n,m}(x_0) - \alpha_{ideal}^{n,m}(x_0)|^2 ds(x_0). \end{aligned} \quad (3.5.3)$$

In this way, it is possible to optimize a system so that the general performance increases in all directions. mse_n^m still depends on the degree and order of spherical harmonics as well as the frequency. Equation 3.5.3 can be written as

$$mse_n^m = \frac{1}{4\pi} \int_{S^2} \left(\underbrace{\alpha_{cub}^{n,m} \overline{\alpha_{cub}^{n,m}}}_A + \underbrace{\alpha_{ideal}^{n,m} \overline{\alpha_{ideal}^{n,m}}}_B - \underbrace{\alpha_{cub}^{n,m} \overline{\alpha_{ideal}^{n,m}}}_C - \underbrace{\overline{\alpha_{cub}^{n,m}} \alpha_{ideal}^{n,m}}_D \right) ds(x_0), \quad (3.5.4)$$

resulting in four separate integrals, with integrands denoted by A, B, C and D.

Integral of A: This first term involves the integral of constants found by the cubature rule integral. Inserting these constants from equation 3.4.1 we get

$$\begin{aligned} \int_{S^2} \alpha_{cub}^{n,m} \overline{\alpha_{cub}^{n,m}} ds(x_0) &= \int_{S^2} \frac{1}{|R_n(r_0,a)|^2} \sum_{q1} w_{q1} \cdot \overline{\Phi}(x_0, x_{q1}) \cdot Y_n^m(x_{q1}) \\ &\cdot \sum_{q2} w_{q2} \cdot \Phi(x_0, x_{q2}) \cdot \overline{Y_n^m}(x_{q2}) ds(x_0). \end{aligned} \quad (3.5.5)$$

Inserting the wave field on the sphere, Φ , from equation 1.5.13 we get

$$\begin{aligned} \int_{S^2} \alpha_{cub}^{n,m} \overline{\alpha_{cub}^{n,m}} ds(x_0) &= \int_{S^2} \frac{1}{|R_n(r_0,a)|^2} \sum_{q1} \left[w_{q1} Y_n^m(x_{q1}) \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} \overline{R_{n'} Y_{n'}^{m'}(x_0)} \overline{Y_{n'}^{m'}(x_{q1})} \right] \\ &\cdot \sum_{q2} \left[w_{q2} \overline{Y_n^m}(x_{q2}) \sum_{n''=0}^{\infty} \sum_{m''=-n''}^{n''} R_{n''} \overline{Y_{n''}^{m''}(x_0)} Y_{n''}^{m''}(x_{q2}) \right] ds(x_0) \end{aligned} \quad (3.5.6)$$

The only terms dependant on the integration variable, $s(x_0)$ are $Y_{n'}^{m'}(x_0)$ and $\overline{Y_{n''}^{m''}(x_0)}$. The orthogonality of the spherical harmonics (see 1.4.4) results in $n'' = n'$ and $m'' = m'$, therefore

$$\begin{aligned} \int_{S^2} \alpha_{cub}^{n,m} \overline{\alpha_{cub}^{n,m}} ds(x_0) &= \frac{1}{|R_n(r_0,a)|^2} \sum_{q1} \sum_{q2} w_{q1} w_{q2} Y_n^m(x_{q1}) \overline{Y_n^m}(x_{q2}) \\ &\cdot \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} |R_{n'}|^2 \overline{Y_{n'}^{m'}(x_{q1})} Y_{n'}^{m'}(x_{q2}) \end{aligned} \quad (3.5.7)$$

Integral of B: The second term, that only involves the ideal decomposition constants, is more simple

$$\int_{S^2} \alpha_{ideal}^{n,m} \overline{\alpha_{ideal}^{n,m}} ds(x_0) = \int_{S^2} |Y_n^m(x_0)|^2 ds(x_0) = 1. \quad (3.5.8)$$

Integral of C: This third term involves both the numerically computed and ideal decomposition constants. We have

$$\begin{aligned}
\int_{S^2} \alpha_{cub}^{n,m} \overline{\alpha_{ideal}^{n,m}} ds(x_0) &= \int_{S^2} \frac{1}{R_n(r_0,a)} \sum_q w_q \Phi(x_0, x_q) \overline{Y_n^m(x_q)} Y_n^m(x_0) ds(x_0) \\
&= \int_{S^2} \frac{1}{R_n(r_0,a)} \sum_q w_q \left(\sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} R_{n'}(a, r_0) \overline{Y_{n'}^{m'}(x_0)} Y_{n'}^{m'}(x_q) \right) \overline{Y_n^m(x_q)} Y_n^m(x_0) ds(x_0) \\
&= \frac{1}{R_n(r_0,a)} \sum_q w_q \overline{Y_n^m(x_q)} \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} R_{n'}(a, r_0) \int_{S^2} \overline{Y_{n'}^{m'}(x_0)} Y_n^m(x_0) ds(x_0) \cdot Y_{n'}^{m'}(x_q).
\end{aligned} \tag{3.5.9}$$

Again by using the orthogonality properties of the spherical harmonics, resulting in $n' = n$ and $m' = m$, it is easy to see that:

$$\begin{aligned}
\int_{S^2} \alpha_{cub}^{n,m} \overline{\alpha_{ideal}^{n,m}} ds(x_0) &= \frac{1}{R_n(r_0,a)} \sum_q w_q \overline{Y_n^m(x_q)} R_n(a, r_0) Y_n^m(x_q) \\
&= \sum_q w_q \overline{Y_n^m(x_q)} Y_n^m(x_q)
\end{aligned} \tag{3.5.10}$$

Integral of D: This last term D, can easily be shown to be the same as term C.

$$\int_{S^2} \overline{\alpha_{cub}^{n,m}} \alpha_{ideal}^{n,m} ds(x_0) = \sum_q w_q \overline{Y_n^m(x_q)} Y_n^m(x_q) \tag{3.5.11}$$

In summing up these four terms, the mse_n^m can be rewritten as

$$\begin{aligned}
4\pi \cdot mse_n^m &= 1 - 2 \sum_q w_q \overline{Y_n^m(x_q)} Y_n^m(x_q) \\
&+ \frac{1}{|R_n(r_0,a)|^2} \sum_{q1} \sum_{q2} w_{q1} w_{q2} Y_n^m(x_{q1}) \overline{Y_n^m(x_{q2})} \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} |R_{n'}|^2 \overline{Y_{n'}^{m'}(x_{q1})} Y_{n'}^{m'}(x_{q2}).
\end{aligned} \tag{3.5.12}$$

Equation 3.5.12 can be written much shorter, by the use of vector and matrix notation, as

$$mse_n^m = \frac{1}{4\pi} \left[1 - 2\vec{w}^T \vec{v} + \frac{1}{|R_n(r_0, a)|^2} \vec{w}^T \overline{\overline{M}} \vec{w} \right], \quad (3.5.13)$$

where w is a vector of length Q containing the cubature weights, w_q , and \vec{v} is a vector, also of length Q , with elements defined as

$$v_q = |Y_n^m(x_q)|^2, \quad (3.5.14)$$

and M is a matrix of dimensions ($Q \times Q$), with elements defined as

$$M_{q1, q2} = Y_n^m(x_{q1}) \overline{\overline{Y_n^m}}(x_{q2}) \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} |R_{n'}(r_0, a)|^2 \overline{\overline{Y_{n'}^{m'}}}(x_{q1}) Y_{n'}^{m'}(x_{q2}). \quad (3.5.15)$$

Using equation 3.5.13, it is possible to calculate an error, mse_n^m , from the coordinates of the Q microphone positions on the sphere, $\vec{x} = \{x_1, x_2, \dots, x_q, \dots, x_Q\}$. The error, mse_n^m , is furthermore independent of the angle of incidence of a present wave field. However, the mse_n^m is still dependant on the frequency through the *Bessel functions* in R_n . mse_n^m also depends on all the spherical harmonics of all degrees and orders through $\overline{\overline{M}}$. (see equation 3.5.15).

An example of the mse_n^m for the t-design previously described having 64 points on the sphere (table 3.1), can be seen in figure 3.2, at $ka = 1$. Note that the error on the decomposition constants of the same degree, n , does not vary much. The error increases rapidly when the order increases. The effect on the actual beamformer response can be seen in figure 3.3, which shows the same system as shown in figure 3.2. For $N = 5$, the error is not noticeable on the beamformer response, but for $N = 6$ the error becomes significant.

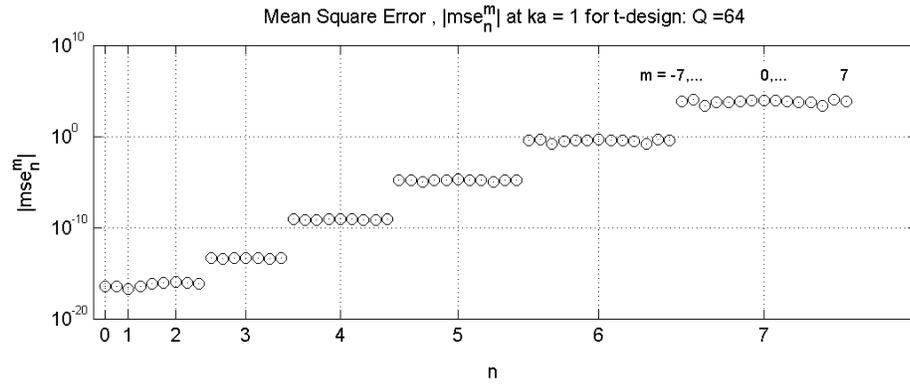


Figure 3.2: $|mse_n^m|$ of the decomposition constants, $\alpha_{cub}^{n,m}$, for the t-design having $Q = 64$ at $ka = 1$. The degree n is denoted on the horizontal axis, having the order m ranging from $-n$ to n , plotted symmetrically around n . (i.e. see $n = 7$)

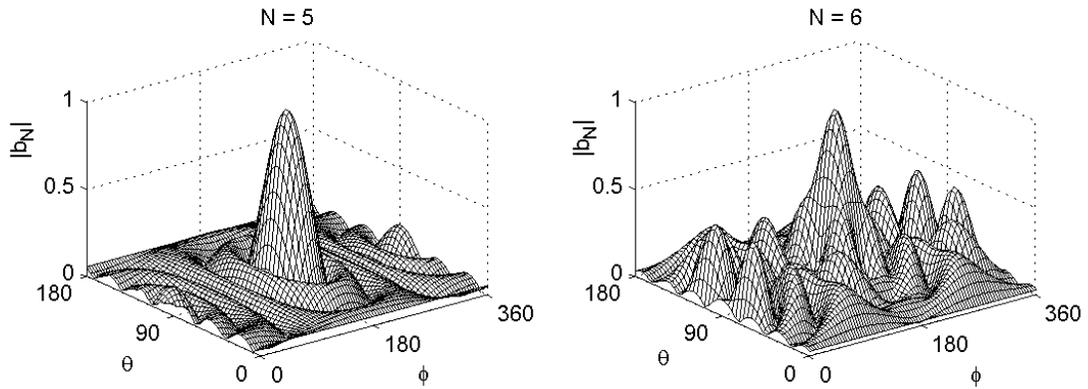


Figure 3.3: $|b_N|$ for the t-design with $Q=64$. Left shows a beamformer response with no noticeable error at $N = 5$. At $N = 6$ on the right, the error is too large to get the correct beamformer response.

3.6 Chapter summary

In this chapter, it has been shown how to perform the spherical integral used to determine the decomposition constants of the SHB-response. Two numerical methods are described, both based on the cubature rule. The first method is used to determine the cubature weights by the use of the reproducing kernel. This method requires a quadratic number of integration points. The second method uses equal cubature weights, and the positions of the integration points determine how high a degree of spherical polynomials is possible. The designs of point positions found by Hardin and Sloane are described, and they are also characterized by having high order of symmetry.

The new SHB response using the cubature rule is shown (equation 3.4.2), followed by an expression determining the error of the decomposition constants (equation 3.5.13), and can be used for optimization of an array.

Chapter 4

Analysis, optimization and implementation

An analysis of the Spherical Harmonics Beamforming method is made in this chapter, which leads to an optimization of the microphone array. The analysis includes a selection of the optimal surface of the beamformer sphere, as well as the stability of the system when phase and amplitude errors on the microphones occur. After the optimization follows a description of the geometrical properties of the selected optimized array. At the end, a practical implementation method of the SHB is given, since the microphones are not capable of measuring the velocity potential, as presumed in the theoretical version of the SHB.

4.1 Surface of the sphere

Theoretically, it is possible to use both types of spheres; a hard scattering sphere with the front of the microphones aligned with the surface of the sphere, or an acoustically transparent sphere with the microphones set in a lattice. The only requirement is that the appropriate radial part, R_n , from table 1.1 is chosen when calculating the decomposition constants (equation 3.4.1).

One disadvantage of using the hard sphere, is that the scattering surface "colours" the wave field. The transparent sphere allows the wave field to travel undisturbed through the sphere, at least at low frequencies (at higher frequencies the microphones

and the lattice-structure can cause reflections and diffraction).

In the equation, for determining the beamformer response, the division by the radial part, $R_n(a, r_0)$, occurs (see equation 3.4.2). This requires that $R_n(a, r_0)$ is never zero. This only holds for the hard sphere, where R_n is a combination of spherical Bessel and Neumann functions. For the transparent sphere, R_n is described by spherical Bessel functions only. In figure 4.1 the two types of the radial part, R_n , are shown.

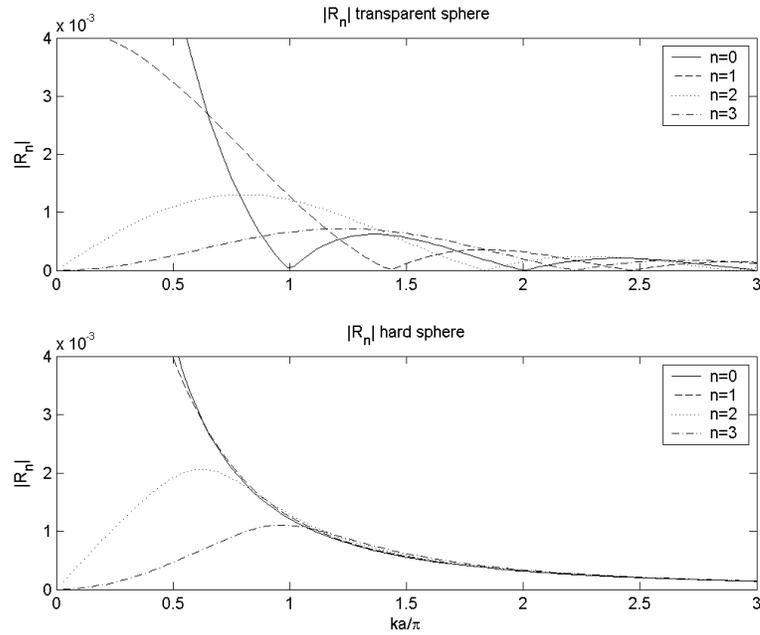


Figure 4.1: $|R_n|$ as a function of ka/π , at different values of n , for the transparent sphere (top) and the hard sphere (bottom).

The zeros of the radial part for the transparent sphere would result in great uncertainty for the beamformer response at certain frequencies (i.e. $ka = \pi$). This is why the hard sphere is chosen as the best solution.

4.2 Stability

In this section, an investigation of the stability is made to evaluate how the beamformer response reacts to an error at different frequencies and degrees, N , of spherical harmonics. This is done by simulating a random error at each microphone, at different values of N , and at different values of the frequency, ka . Firstly, the consequences of a *phase error* is investigated.

The simulations are made by applying a *random normally distributed phase error* to each microphone, and then seeing how it affects the beamformer response. Applying the same error to all microphones would not result in any error at the normalized absolute value of the beamformer response, since it is the internal phase and amplitude difference between the microphones that results in the response. The velocity potential, with the applied phase error, is

$$\Phi_{ph,q} = \Phi_q e^{i\sigma_{ph}\gamma_q}, \quad (4.2.1)$$

where σ_{ph} is the *standard deviation* of the normally distributed phase errors. $\gamma_{ph} = \{\gamma_{ph,1}, \dots, \gamma_{ph,Q}\}$ is a set of random normally distributed numbers, with a standard deviation of 1. The phase error at microphone number q , is then $\sigma_{ph}\gamma_{ph,q}$.

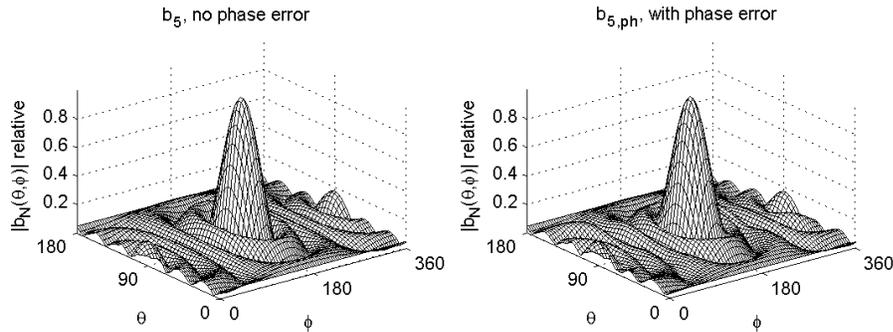


Figure 4.2: Absolute value of beamformer response of t-design having $Q=64$ at $ka=4$ and $N=5$. Left: No phase error attached. Right: Phase error resulting in an error of the beamformer response of $\epsilon = 0.01$.

In order to evaluate the error of the beamformer response, we need to define an expression of the error of the beamformer response. The mean absolute error on the beamformer response from the phase error, relative to the maximum of the response without any phase error, is defined as

$$\epsilon_{ph}(b_N) \equiv \frac{1}{4\pi} \int_{S^2} \frac{|b_N(x) - b_{N,ph}(x)|}{|b_{N,max}(x)|} dS(x), \quad (4.2.2)$$

where $b_N(x)$ is the beamformer response with no phase error attached, and $b_{N,ph}$ is with the phase error. Allowing a maximum error on the beamformer response of $\epsilon_{ph}(b_N) = 0.01$, the maximum allowable phase error as a function of ka and N , can be simulated. Figure 4.2 shows an example where the error on the beamformer response is $\epsilon_{ph}(b_N) = 0.01$. The effect of the error is slightly noticeable on the sidelobes (right picture), but the response can still be used to determine the direction of the incoming wave.

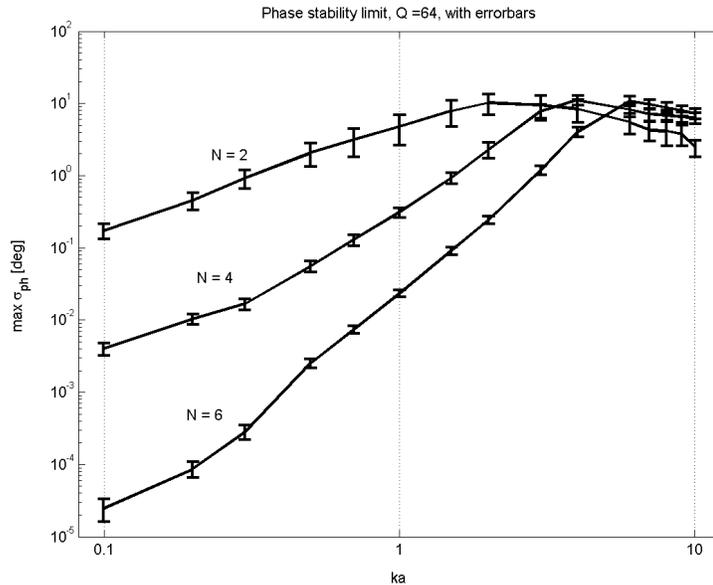


Figure 4.3: Limit of phase stability, where $\epsilon_{ph} = 0.01$ for t-design having $Q=64$, at $N = 2, 4, 6$. The error bars show the standard deviation of the calculated graphs made from 30 simulations, having random normally distributed phase error.

By using the limit for the error of the beamformer response of $\epsilon_{ph}(b_N) = 0.01$, a random phase error $\sigma_{ph}\gamma_q$ has been simulated in MATLAB, where the standard deviation σ_{ph} , has been decreased until $\epsilon_{ph}(b_N) = 0.01$ is achieved. This is done for $N = \{1, 2, 3, 4, 5, 6\}$ at the following frequencies, $ka = \{0.1, 0.2, 0.3, 0.5, 0.7, 1, 1.5, 2, 3, 4, 6, 7, 8, 9, 10\}$ by using the same set of Q random errors $\sigma_{ph}\gamma_{ph} = \sigma_{ph}\{\gamma_{ph,1}, \dots, \gamma_{ph,Q}\}$. This is simulated 30 times by using a new set of random normally distributed phase errors. The average of σ_{ph} from these 30 simulations are illustrated in figure 4.3, including the error bars showing the standard deviation of the results from the 30 simulations. The figure illustrates that an phase error of around 8° is acceptable at $ka = 2$ for $N = 2$, but not for $N = 4$ or 6 , which demand a phase error of less than 2° and 0.2° respectively (see figure 4.3).

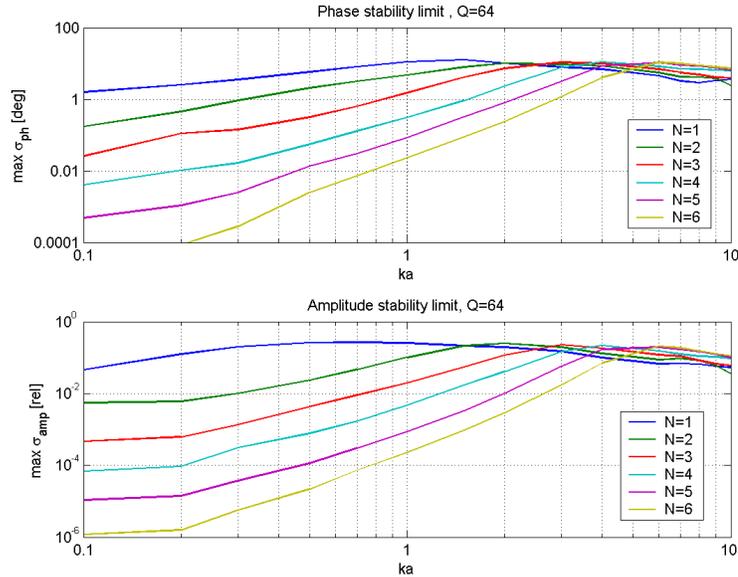


Figure 4.4: Stability of t-design having $Q = 64$. Maximum allowed standard deviation of random normally distributed error, for the phase error (top) and the relative amplitude error (bottom). (All calculations are the mean value from 30 simulations, and the standard deviations from this are shown in figure 4.3)

The same simulations have been made to evaluate the effect of an error in the amplitude of each microphone. The error $\sigma_{amp}\gamma_q$ is again assumed to be random

normally distributed, having the standard deviation σ_{amp} . The velocity potential at microphone q , with an *amplitude error* attached, is defined as

$$\Phi_{amp,q} \equiv \Phi_q(1 + \sigma_{amp}\gamma_q). \quad (4.2.3)$$

Figure 4.4 shows the results for both the phase error and the amplitude error, for the t-design having $Q = 64$, at $N = 1$ to $N = 6$.

The same simulations have also been made for systems with a lower number of points. Figure 4.5, shows the the results of maximum allowable σ_{ph} from the three t-designs having 24, 48 and 64 points, for $N=3$. Note that a higher number of points result in a higher allowable random normally distributed phase error.

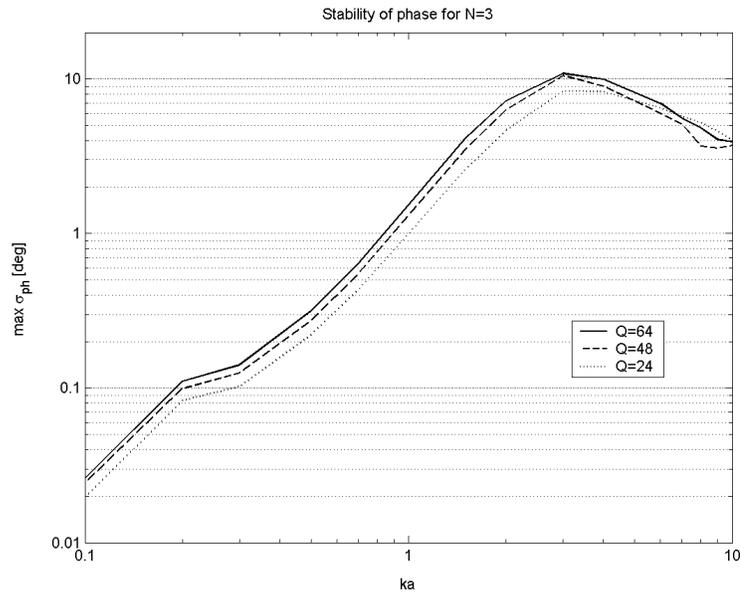


Figure 4.5: Maximum allowable σ_{ph} , at degree $N = 3$, for the three different t-designs having 24, 48 and 64 points respectively.

The analysis of the stability generally shows that the systems are most stable at frequencies from $ka = 2$ to $ka = 6$ (dependant on N). The higher the N , the more the stability falls at lower frequencies. It will therefore probably be a problem to retrieve the higher degrees, N , of the beamformer response, b_N , at low frequencies

(i.e. at $ka = 1$ and $N = 6$ the maximum allowed standard deviation of the random normally distributed amplitude error should be lower than 0.03%. See figure 4.4).

4.3 Optimization

To find the best system of microphones for performing Spherical Harmonics Beamforming, it is necessary to look at what parameters need to be optimized. The available parameters are:

1. Radius of sphere, a
2. Frequency, f
3. Maximum degree of spherical harmonics to include, N
4. Number of microphones, Q
5. Microphone positions, $x = \{x_1, \dots, x_q, \dots, x_Q\}$

The radius of the sphere, a , and the frequency, f , are connected, since the beamformer response calculations are made as a function of $ka = \frac{2\pi f}{c}a$; if the radius of the sphere is increased, then the frequency range is moved relatively down. The radius of the sphere can therefore be chosen to get the desired frequency range. There will of course exist some practical limitations of the radius (e.g. the size of the microphones sets a lower limit). In further optimization analysis, ka will be used as a value for the frequency and radius.

The higher degree of spherical harmonics, N , (and thereby higher order of Bessel functions) included in the beamformer response, results in a better resolution, (i.e. smaller FWHM) as well as larger attenuation of sidelobes. The highest possible N is therefore desired. Furthermore, N is connected to the number of microphones, Q , since in general higher Q results in higher N .

Q is not the only parameter determining N . The final parameters to be tuned are the positions of the microphones. Having Q microphones gives us $2Q$ parameters,

since the position of each microphone is described by two coordinates, $x_q = \{\phi_q, \theta_q\}$. The optimization can be performed by minimizing the errors of the decomposition constants, mse_n^m , up to the desired degree at most N , as described by equation 3.5.13. For a desired beamformer response of degree N , there exist $(N+1)^2$ errors to minimize (since $n = \{0, 1, 2, \dots, N\}$ and $m = \{-n, \dots, n\}$). This is why the following expression is introduced, in order to find the average error of the decomposition constants of degree n and order m up to a maximum degree of N .

$$\nu(ka, N) \equiv \frac{1}{(N+1)^2} \sum_n^N \sum_{m=-n}^m |mse_n^m(ka)| \quad (4.3.1)$$

In this way, there is only one parameter to minimize instead of $(N+1)^2$.

The minimization in this project is done by moving the points on the sphere (the microphone positions), until the nearest local minimum of $\nu(ka, N)$ is found. $\nu(ka, N)$ depends also on the positions of the points through mse_n^m (see section 3.5). The nearest local minimum of $\nu(ka, N)$ has been found, when the displacement of any single point 10^{-10} [rad] in any direction (on the surface of the sphere) did not result in a lower value of $\nu(ka, N)$.

Finding the nearest local minimum of $\nu(ka, N)$ at a particular ka and N , does not necessarily mean that the global minimum is found. This is why the following strategies have been considered and tried out.

1. Minimize $\nu(ka, N)$ from a random start guess of the microphone positions (possibly by picking the best result of many different start guesses).
2. Use one of the already optimized t-designs.
3. Minimize $\nu(ka, N)$ and use existing t-designs as a start guess and keep the cubature weights equal.
4. Minimize $\nu(ka, N)$ and use existing t-designs as a start guess, and have a square number of microphones, Q , that allows unequal cubature weights, w , through the reproducing kernel.

5. as suggestion 3 and 4 but keeping the symmetry of the existing t-design.

The problem of using the first suggestion (having a random start guess of the microphone positions) is that, $\nu(N)$ has many local minima. When the microphones are moved, the system will fall into the minimum nearest the starting guess, which is not necessarily the global minimum. Using this approach requires a lot of different starting guesses as well as a lot of computational time, before finding the global minimum or any minimum close to the global minimum. Another disadvantage of this strategy, is that no symmetry of the point positions exist, resulting in a possible variation of the performance in different directions. In the test of this strategy, the point positions appeared to spread regularly out over the surface of the sphere when minimizing $\nu(ka, N)$.

The second suggestion is using one of the existing t-designs optimized (by Hardin and Sloane) for exact integration of spherical polynomials of degree t . The function to be integrated is the product between the velocity potential, Φ_a , and the spherical harmonics. The velocity potential, Φ , is described by a combination of spherical harmonics and spherical Bessel and Neumann functions that are frequency dependent. The existing t-designs are not necessarily the optimal setup to find the decomposition constant up to a degree of t (initial simulation shows more likely up to a degree of $t/2$). Figure 4.6 shows the performance of the t-designs having $Q = 60$ and $Q = 64$, by plotting $\nu(N, ka)$ as a function of ka at different values of N . Note that $\nu(ka, N) < 0.001$ for $N \leq 5$, for frequencies $ka < 6$. Something else interesting is that the t-design which has $Q = 60$ has lower $\nu(ka, 6)$ than for $Q = 64$. In other words, the existing t-design of 60 points is better at creating a beamformer response of degree $N = 6$ than for the t-design having 64 points, and in a much wider frequency range. Furthermore, both of these t-designs showed much lower $\nu(ka, N)$ than any designs found by minimizing $\nu(ka, N)$ from a random starting guess of the microphone positions.

The third suggestion, is to try to optimize the existing t-designs while keeping equal weights in the numerical integration. Since the t-designs are already optimized by using equal weights, this strategy did not show any improvement, neither at low

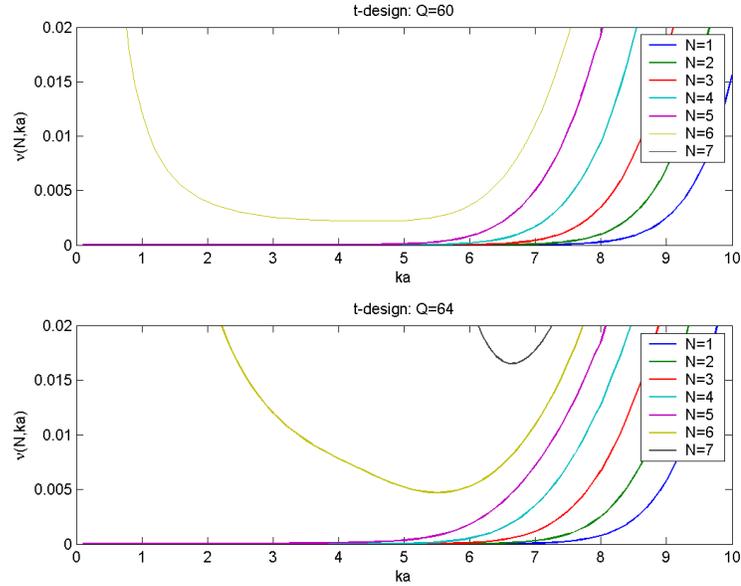


Figure 4.6: $\nu(ka, N)$ for t-design having $Q = 60$ (top) and $Q = 64$ (bottom).

($ka = 1$) nor high ($ka = 7$) frequencies.

The fourth suggestion, is again to use the existing t-designs and optimize them by allowing the cubature weights to be unequal. This is done by using the reproducing kernel to calculate new cubature weights after each move of the points, and requires the number of points to be quadratic (see section 3.3). This strategy showed a slight improvement of $\nu(ka, N)$ around the optimizing frequency, but a worsening at other frequencies.

The last tried approach is using the existing t-designs as a start guess, and then finding the nearest local minimum of $\nu(ka, N)$, while still keeping the symmetry. By minimizing $\nu(ka, N)$ of the t-design with 60 points while keeping the symmetry, the weights would still have to be kept equal (since the number of points is not quadratic, see section 3.3). Since this system is already optimized having equal weights, it is not possible to minimize $\nu(ka, N)$ further.

For the t-design with 64 points, unequal weights can be allowed, since 64 is a quadratic number, and are found by the use of the reproducing kernel (see section 3.2). This optimization is tried out at different frequencies, in order to extend the frequency range at $N = 6$.

The values of $\nu(N, ka)$ before the optimization of the t-design with 64 points, is seen as the yellow line in the bottom picture in figure 4.6 and the blue line in figure 4.7. The results after the optimization at $ka = 1$ and $ka = 7$ can also be seen in figure 4.7. Optimizing at the high frequency of $ka = 7$, showed a slight improvement at high frequency by having a lower $\nu(ka, N)$. Optimizing at the lower frequency $ka = 1$ showed a significant improvement at low frequencies, and a slight improvement at high frequencies. Optimizing at lower frequencies than $ka = 1$, did not show any improvement. In the stability analysis, the system showed anyway to be very unstable at low frequencies, allowing a maximum standard deviation of a normally distributed phase error of 0.01° at $ka = 1$ (see figure 4.4).

To see what effect this optimization has on the actual beamformer response, b_N , at degree $N = 6$, the beamformer response has been simulated both before and after the optimization, and the result can be seen in figure 4.9.

point number	x	y	z
1	-0.32134051642889	0.91871757250132	-0.22956109095073
2	0.11889351497826	0.98736875238214	-0.10472477698924
3	-0.53994094605482	-0.82075188837576	-0.18662827358430
4	-0.12503345547114	0.86604541844656	0.48407847318453
5	0.69237632636429	0.65300799683460	0.30691298239134
6	0.57735026918963	0.57735026918963	-0.57735026918963

Table 4.1: Coordinates of the 6 basis points of the optimized system of points with tetrahedra symmetry.

Optimization of higher N than 6 has been tried, but without any significant success for systems of 64 points or less. The best improvement after the optimization is made using the existing t-design with 64 points as a start guess. The optimization is made by moving the microphones, while still keeping the tetrahedra symmetry. This is done until the nearest local minimum of $\nu(ka, N)$ at $ka = 1$ and $N = 6$ is found, while allowing unequal weights in the numerical integration by using the reproducing

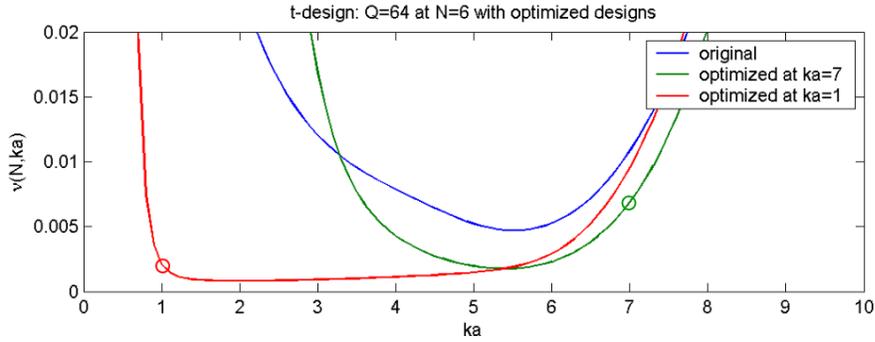


Figure 4.7: $\nu(ka, N)$ for t-design having $Q = 64$ at $N = 6$, optimized at the frequencies $ka = 1$ and $ka = 7$. 'o' indicates the optimization frequency.

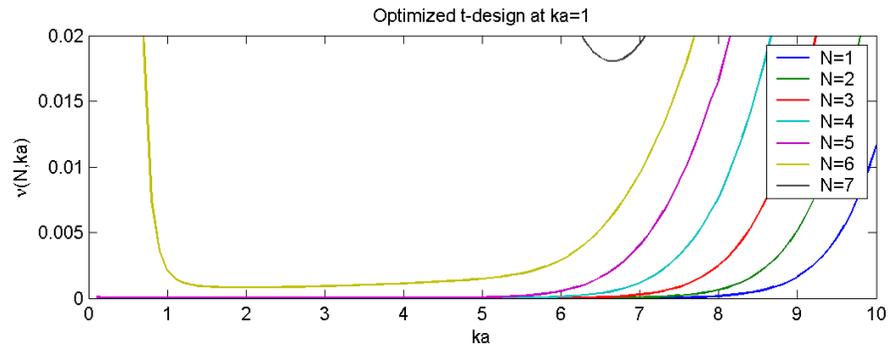


Figure 4.8: t-design with $Q = 64$ optimized at $ka = 1$. Figure shows $\nu(ka, N)$ at all $N = \{1, 2, \dots, 7\}$

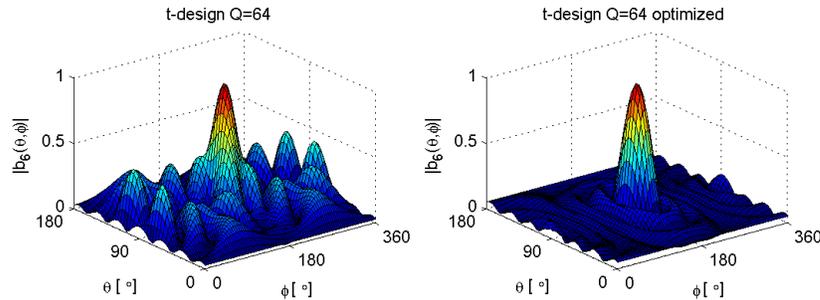


Figure 4.9: Resulting beamformer response at $ka = 1$. Left: existing t-design having $Q=64$ points. Right: Same t-design optimized at $ka = 1$ for $N = 6$.

kernel. The coordinates of the 6 basis points are listed in table 4.1, and the rest of the 64 points can be found by using the symmetry operators, C_3 and C_2 , described in section 3.3. The first 5 points will each result in 11 new points. Point number 6 is placed on a C_3 symmetry axis, which means that the symmetry operators will result in 3 new points. The total number of points will therefore be: $Q = 5 \times 12 + 1 \times 4 = 64$.

The minimum, average and maximum values of the cubature weights, belonging to the 64 optimized points, found by solving the matrix equation, $Gw = e$ (see section 3.2), are listed in table 4.2. The sum of the cubature weights, represents the integration of a function on the sphere with a constant value of 1, and should always equal 4π .

w_{min}	w_{avg}	w_{max}	$\sum w$
0.1271	0.1963	0.2262	$12.5664 \simeq 4\pi$

Table 4.2: The minimum, average, maximum and sum of the 64 cubature weights (first four decimals shown).

4.4 Geometrical properties of selected array

In this section, we investigate the geometrical properties of the optimized array.

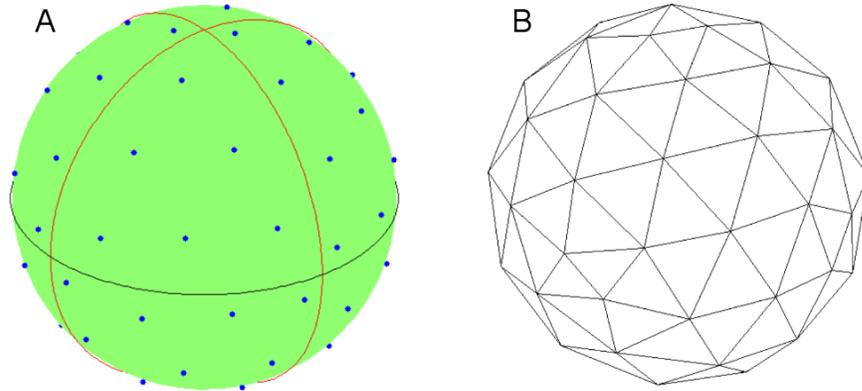


Figure 4.10: A) Optimized 64 points on the sphere. The black equator line parts the sphere in two identical half spheres. The red line parts the sphere in the same identical half spheres because of the symmetry. B) The delaunay triangulation of the points.

Figure 4.10 shows the position and symmetry of the optimized 64 points on the sphere, as well as the *delaunay triangulation*¹. The tetrahedral symmetry of the points makes it possible to part the sphere in two identical half spheres at the equator ($\theta = 90^\circ$). The two vertical red lines in figure 4.10 A, result in the same two half spheres. The symmetry order of 12, makes it furthermore possible to part the sphere in 12 identical shells with 5 points on each, with the last four points placed at corners of the 12 shells.

The upper frequency limit of around $ka = 7$ (see figure 4.8) can be explained from the delaunay triangulation. In the conventional Delay and Sum beamforming, the upper frequency limit is determined by the microphone spacing, where the distance between them may not exceed the half of the wavelength, λ , of the incoming wave (see [10] and [3]). A frequency limit of $ka = 7$ results in the following minimum wavelength

¹Given a set of data points, the Delaunay triangulation is a set of lines connecting each point to its natural neighbours.

$$\begin{aligned}
ka &= \frac{2\pi}{\lambda}a = 7 \\
&\Downarrow \\
\lambda &= \frac{2\pi}{7}a = 0.8976 \cdot a,
\end{aligned} \tag{4.4.1}$$

allowing a maximum spacing of the microphones of $d_{(ka=7)} = \frac{\lambda}{2} = 0.449 \cdot a$. Table 4.3 lists the minimum, average and maximum distances from the delaunay triangulation. (Note how close d_{avg} is to $d_{(ka=7)}$.)

d_{min}	d_{avg}	d_{max}
$0.3169 \cdot a$	$0.4752 \cdot a$	$0.5446 \cdot a$

Table 4.3: The minimum, average and maximum distances of the delaunay triangulation of the optimized 64 points.

4.5 Practical implementation of SHB

The SHB-response was in chapter 3 expressed as a weighted sum of the velocity potentials, $\Phi_a(x_0, x_q)$, at the microphone positions, $x_q = (\theta_q, phi_q)$, on the sphere of radius a (see equation 3.4.2). Equation 3.4.2 can also be written as

$$b_N(x) = \sum_q^Q W_q(N, a, r_0, x_q, x, \omega) \Phi_a(x_0, x_q), \tag{4.5.1}$$

where the weights, W_q are found as

$$W_q(N, a, r_0, x_q, x, \omega) = \left[\sum_{n=0}^N \frac{1}{R_n(a, r_0)} \sum_{m=-n}^n w_q \overline{Y_n^m}(x_q) \cdot Y_n^m(x) \right]. \tag{4.5.2}$$

The velocity potential, $\Phi_a(x_0, x)$, holds information about the phase of the wave field. However, the microphone can not measure the velocity potential. The microphone measures the real value of the sound pressure, related to the velocity potential as (see equation 1.3.4)

$$p(t) = \text{Re} \left\{ \rho_0 \frac{\partial \Phi}{\partial t} \right\}. \quad (4.5.3)$$

Throughout this thesis the implicit time factor, $e^{i\omega t}$ is assumed, resulting in

$$p(t) = \text{Re} \{ i\omega \rho_0 \Phi \}. \quad (4.5.4)$$

In the practical implementation, the phase information of the signals can be obtained through a *Fast Fourier Transformation* FFT. The FFT is a fast implementation of the *Discrete Fourier Transformation*, DFT. The L -point DFT can, from L samples of the signal $p(t)$, be found as

$$P(\omega_\eta) = \sum_{l=0}^{L-1} p(t_l) e^{-i\omega_\eta l}, \quad (4.5.5)$$

where $p(t_l)$ is a sample of the sound pressure, at time t_l , determined by the sampling frequency as $t_l = \frac{l}{f_s}$. L is the number of samples, and $l = 0, \dots, L-1$, denotes the sample number. The *angular frequency*, ω_η is related to the real frequency as $\omega_\eta = 2\pi f_\eta$, where $\eta = \{0, 1, \dots, \frac{L}{2}\}$ denotes the number of the possible discrete frequencies from the L -point DFT. The *frequency resolution* is therefore determined by the number of samples, L , and the *sampling frequency*, f_s , as $\Delta f = \frac{f_s}{L}$. The discrete frequencies are found as $f_\eta = \frac{\eta f_s}{L}$. To perform the FFT, the number of samples must agree $L = 2^\varrho$ where $\varrho \in \mathbb{N}$ (i.e. $L = 2, 4, 16, 32, \dots, 1024, \dots$). [12]

The FFT output, $P_q(\omega_\eta)$, made on the signal from microphone number q , holds information about the amplitude and the phase of the sound pressure at the microphone and at the specific frequency ω_η . To perform the SHB, it is the internal phase and amplitude difference that determines the absolute value of the beamformer response. $P(\omega_\eta)$ can therefore be used instead of the velocity potential when calculating the beamformer response. By using $P(\omega_\eta)$, it is furthermore possible to select the frequency of interest.

Figure 4.11 shows the signal flow in the practical implementation of the SHB method. The microphones on the right measure the sound field, which is sampled and sent through the FFT routine, whereafter the actual SHB routine starts. The SHB is performed by first multiplying the weights, W_q , from equation 4.5.3 and then adding all the results, ending up with the beamformer response, b .

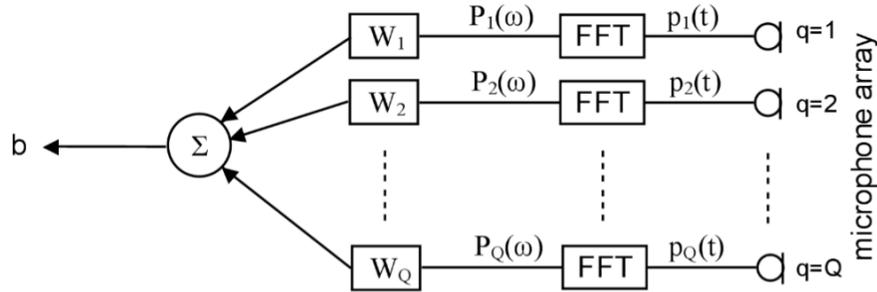


Figure 4.11: Signal flow diagram of the implementation method using FFT.

Figure 4.12 illustrates an example from a simulation of two incoming plane waves at different frequencies and different angles of incidence. The signals at each microphone of the array, are sampled at the sampling frequency $f_s = 8kHz$. The FFT of $L = 1024$ samples from each microphone are made, after which the spherical harmonics beamforming is executed. The left picture shows the SHB method performed on the output of the FFT's for $\eta = 256$, equivalent to the frequency $f_\eta = \frac{256f_s}{L} = 2000Hz$, and the right picture for $\eta = 261$, equivalent to the frequency $f_\eta = \frac{261f_s}{L} = 2031.25Hz$.

The output of the FFT can also be used to determine the intensity of the sound field at a given frequency. It has not been examined closer in this thesis.

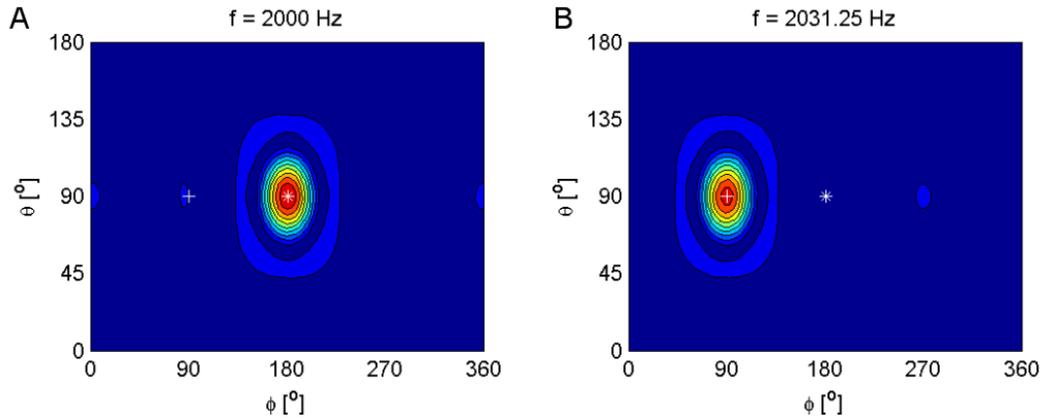


Figure 4.12: Simulations of SHB on signals after the FFT. '+' indicates the direction in which a signal of frequency 2030 Hz approaches from, and '*' indicates the direction in which a signal of 2000 Hz approaches the array from. (A) SHB performed at $f_{\eta=256} = 2000Hz$. (B) SHB performed $f_{\eta=261} = 2031.25Hz$

4.6 Chapter summary

In this chapter, an analysis of the beamformer properties has been investigated, leading to the optimization of a spherical beamformer based on 64 microphones placed in alignment with the surface of a hard sphere. The surface was chosen to be hard instead of transparent in order to avoid division by zero in calculation of the beamformer response. The stability was investigated, showing to be most stable at around $ka = 5$ for $N = 6$, and increasingly unstable at lower frequencies, see figure 4.4. The final design, chosen after the optimization, keeps the tetrahedra symmetry from the t-design of 64 points. The design is capable of getting a beamformer response of degree 6 at frequencies between $ka = 1$ and $ka = 6$ (See figure 4.8). The investigation of the geometrical properties showed similarity between the distances of the microphones and the upper frequency limit of around $ka = 7$. Finally, a method for implementing the SHB method, by using FFT, was shown, and simulation using the method was made. Furthermore, the method made it possible to distinguish signals of different frequencies.

Chapter 5

SHB versus Delay and Sum Beamforming

In this chapter, the performance of the beamformer technique using spherical harmonics (SHB), is compared with the more conventional *Delay and Sum Beamforming* (DSB) technique. This is done in order to show the advantages of using SHB.

5.1 Introduction

The DSB method is the oldest and simplest method for determining the direction of an incoming wave field and is still commonly used. The DSB-method is therefore the most obvious to compare the SHB-method with.

5.2 Delay and Sum Beamforming

When an incoming plane wave, with the wave number vector $\vec{k} = -k\vec{\kappa}$, approaches an array of Q microphones placed at locations r_q ($q = 1, 2, \dots, Q$), from the direction $\vec{\kappa}$, each microphone will measure the same pressure signals, but with an internal phase difference, because of their spatial positions. By applying appropriate delays, $\Delta_q(\vec{\kappa})$, to the signals from each microphone, p_q , and then adding the results, it is possible to enhance signals from a specific direction and suppress signals from other directions.

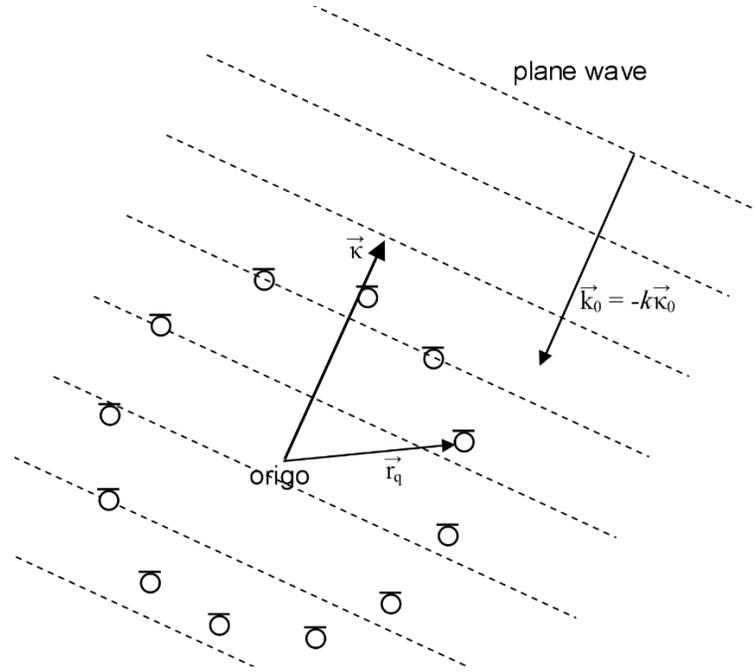


Figure 5.1: Illustration of circular microphone array steered in the correct direction from where a plane wave propagates from.

This is called Delay and Sum Beamforming (DSB), and the beamformer response is [9]

$$z(\vec{\kappa}, t) = \sum_{q=1}^Q p_q(t - \Delta_q(\vec{\kappa})), \quad (5.2.1)$$

where $p_q(t)$ is the sound pressure measured at microphone number q in the time domain. The time delays can be found from geometrical consideration (see figure 5.1 or [9])

$$\Delta_q(\vec{\kappa}) = \frac{\vec{\kappa} \cdot \vec{r}_q}{c}, \quad (5.2.2)$$

where \vec{r}_q is the position of microphone number q , and c is still the propagation speed of sound.

In the frequency domain, the beamformer response is a Fourier transformation of 5.2.1

$$Z(\vec{\kappa}, \omega) = \sum_{q=1}^Q P_q(\omega) e^{-i\omega\Delta_q(\vec{\kappa})}, \quad (5.2.3)$$

where ω is the temporal *angular frequency*. $P_q(\omega)$ is the frequency domain version of the measured signal at microphone number q , (in practice this could be the result of a FFT of p_q). $\vec{k} = -k\vec{\kappa}$ is the wave number vector for the expected plane wave incident from the steering direction κ . The actual wave number vector of the incident wave is denoted $\vec{k}_0 = -k\vec{\kappa}_0$, and the pressure measured at microphone q would then be

$$P_q(\omega) = P_0 e^{-i\vec{k}_0 \cdot \vec{r}_q}, \quad (5.2.4)$$

where P_0 is the amplitude of the wave field. By inserting $P(\omega)$ and the time delay 5.2.2 into equation 5.2.3, we get the following frequency domain version of the beamformer response for plane waves

$$Z(\vec{\kappa}, \omega) = P_0 \sum_{q=1}^Q e^{i(\vec{k} - \vec{k}_0) \cdot \vec{r}_q}, \quad (5.2.5)$$

In the near field situation, where the sound source is close to the array, we would have to consider spherical waves instead of plane waves. Since the direction of the propagating wave now varies throughout the array, there would have to be found a steering direction vector for each microphone. So, instead of one steering direction vector, $\vec{\kappa}$, we have Q different steering direction vectors, found as $\vec{\kappa}_q$ all directed at the same point. This point is called the steering point, and is the point where the sound source is expected to originate from (see figure 5.2). The amplitude of the pressure measured at microphone number q is inverse proportional to the distance between the microphone and the sound source, resulting in the pressure

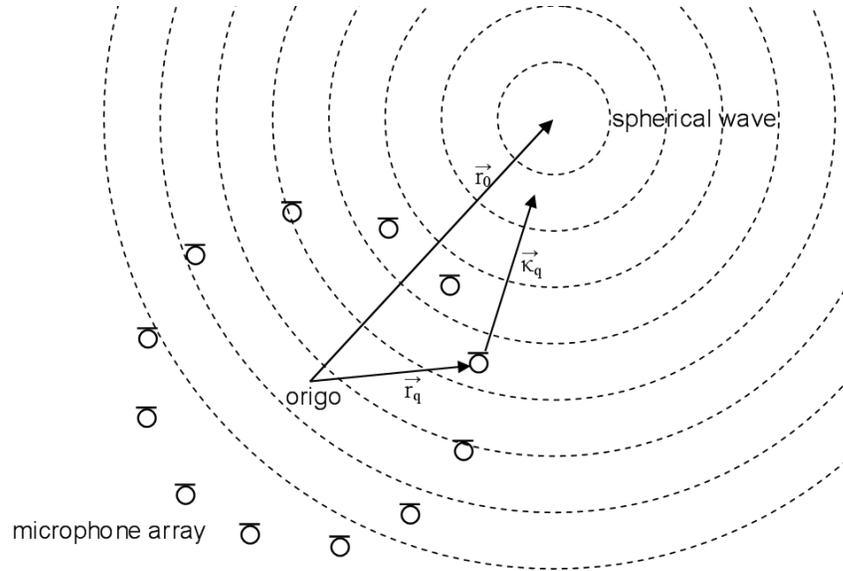


Figure 5.2: Illustration of microphone array steered in the correct direction from where an spherical wave propagates from.

$$P_q(\omega) = \frac{P_0}{|\vec{r}_0 - \vec{r}_q|} e^{-i\vec{k}_{0,q} \cdot \vec{r}_q}, \quad (5.2.6)$$

This results in the following beamformer response for spherical waves

$$Z(\vec{k}, \omega) = P_0 \sum_{q=1}^Q \frac{1}{|\vec{r}_0 - \vec{r}_q|} e^{i(\vec{k}_q - \vec{k}_{0,q}) \cdot \vec{r}_q}. \quad (5.2.7)$$

The DSB-technique can be applied to any microphone array that is acoustically transparent, as it does not account for any reflections. By using the DSB-method on the same microphone array as optimized for spherical harmonics in chapter 5 using 64 microphones, but without the hard scattering surface, we can simulate the DSB-response. The beamformer response is calculated by using equation 5.2.7 in near field situations and equation 5.2.5 in far field situations. Figure 5.3 shows a simulation of the beamformer response at the frequency $ka = 3$, where the source is a point source located at $(\theta, \phi) = (90^\circ, 180^\circ)$ at a distance $r_0 = 15a$.

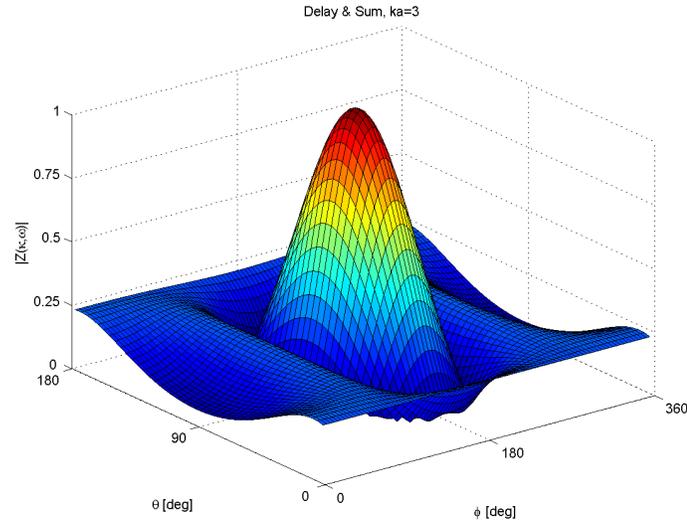


Figure 5.3: Steered response for the Delay and Sum beamformer, using a spherical array of 64 microphones.

The beamformer response using the DSB method, is also characterized by having a mainlobe and sidelobes, as seen for the SHB using spherical harmonics. The resolution of the DSB, measured by the FWHM of the mainlobe, depends on the aperture of the array, and the wavelength of the incoming wave. The aperture is the spatial extension of the array. In other words, this means that the resolution depends on the frequency, which results in a wide mainlobe at low frequencies and a narrow mainlobe at high frequencies. The upper frequency limit is determined by the distance between the microphones and thereby in general the number of microphones. [10]

5.3 Comparing Performances

To emphasize the advantage of using the SHB-method over the performance of the DSB-method, these two techniques have been simulated using the same array of 64 microphones. The surface of the sphere is hard for the SHB, and acoustically transparent for the DSB method.

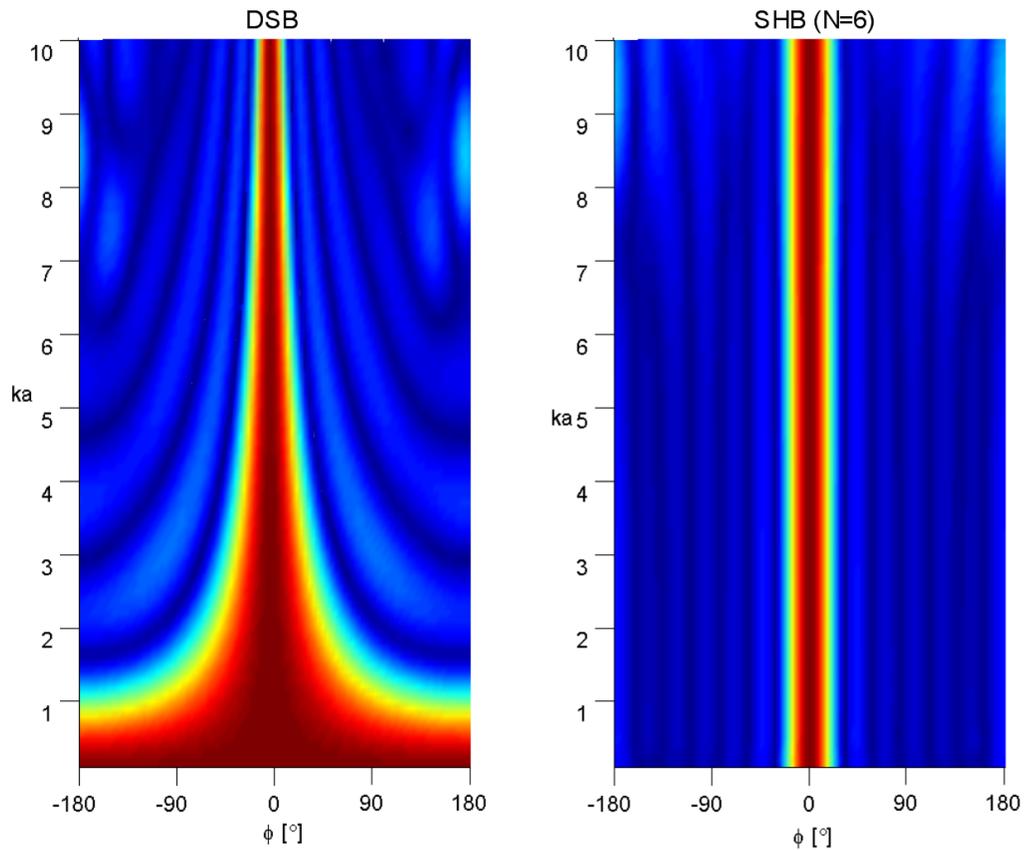


Figure 5.4: Beamformer response sweep for a wave field incident from $(\theta, \phi) = (90^\circ, 0^\circ)$, steered in directions along the equator. Left: the delay and sum beamformer. Right: Spherical harmonics beamformer using $N = 6$.

Figure 5.4 shows the results of the simulation as a frequency sweep from $ka = 0.1$ to $ka = 10$ on the vertical axis. The incoming wave is simulated coming from $(\theta, \phi) = (90^\circ, 0^\circ)$ and the beamformers are steered in directions along the equator, ranging

from -180 to 180 degrees. The red area is the mainlobe of the beamformer response, and the sidelobes are seen as the light blue area on each side of the mainlobe.

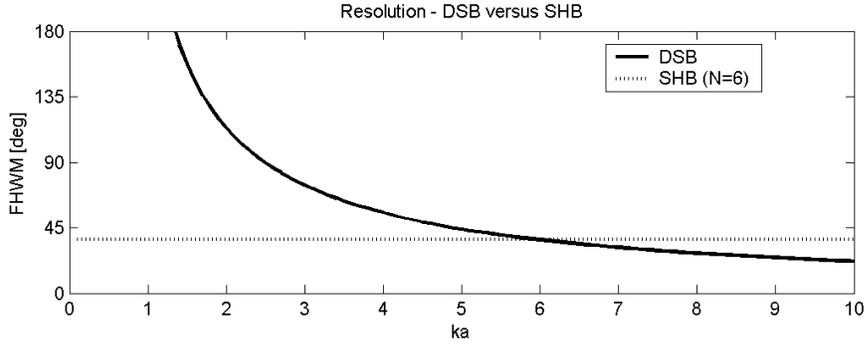


Figure 5.5: FWHM of the beamformer response as a function of the frequency, for both the DSB and SHB-method at $N = 6$.

The SHB at $N = 6$ shows the same width of the mainlobe at all frequencies, ($FWHM \approx 36.6^\circ$). The DSB has a very wide mainlobe at low frequencies, which makes it difficult to determine the direction of the incoming field (see also figure 5.5). At higher frequencies, $ka > 6$, the DSB shows a narrower mainlobe than the SHB. At frequencies higher than $ka = 7$, errors on the sidelobes occur for both beamforming techniques because of the limited amount of microphones, $Q = 64$. This upper frequency limit can be increased for both techniques by using a higher number of microphones. It will not improve the resolution of the DSB, but will make it possible to use higher N , and thereby achieve a better resolution for the SHB-method. The lower MSL using SHB, can be seen in figure 5.6 showing the gain of the sidelobes. The biggest sidelobe level using the DSB-method is around -12 dB in the frequency range $ka = 3$ to $ka = 7$. At low frequencies $ka < 2$ the mainlobe of the DSB-method is so wide (see figure 5.4) that there exists no actual sidelobes.

Using the SHB-method the biggest sidelobe is about 4 dB lower than for the DSB-method. The usable frequency range in which the MSL is lower than -15 dB for the SHB-method is $ka = 1$ to $ka = 7$ which is equivalent to about 2.7 octaves.

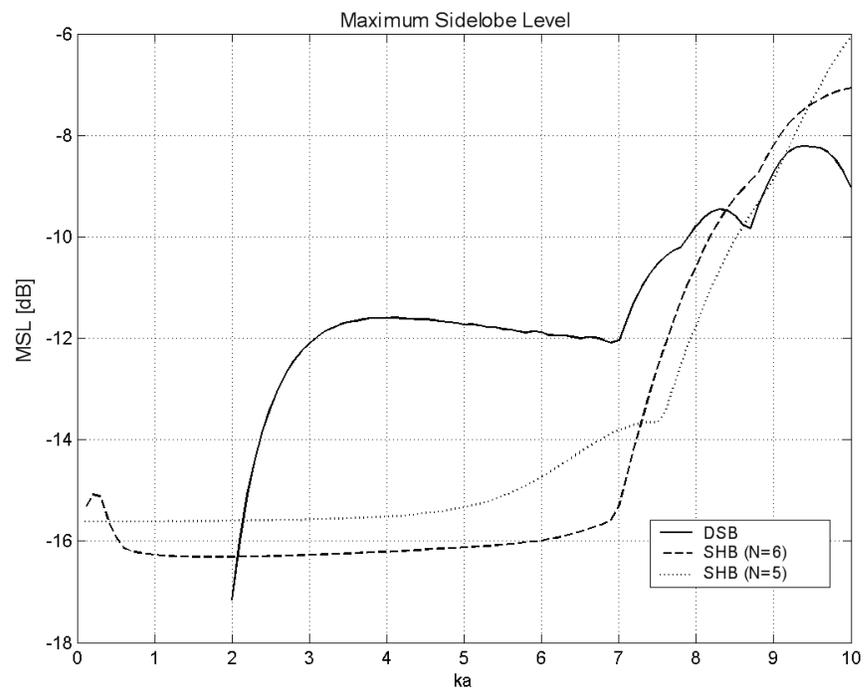


Figure 5.6: Maximum Sidelobe Level for both the DSB-method and the SHB-method

5.4 Chapter summary

The beamforming method developed in this thesis (SHB), has in this chapter been compared with the conventional delay and sum beamforming method (DSB), and has showed significant improvement of the attenuation of the sidelobes on the beamformer response. The DSB-method has a MSL of approximately -12 dB in the frequency range from $ka = 3$ to $ka = 7$, and the SHB-method has an MSL of less than -15 dB in a frequency range of around 2.7 octaves (from $ka = 1$ to $ka = 7$). Furthermore, the resolution showed an improvement at frequencies up to $ka = 6.0$ (i.e. FWHM for DSB at $ka = 1$ is higher than 180° , and only 36.6° using SHB).

The FWHM of the DSB is smaller than for the SHB at degree $N = 6$ for frequencies higher than $ka = 6.0$. The solution to this could be to use DSB at high frequencies, or to include higher N in the SHB-method, though this results in less attenuation of the sidelobes. Both methods showed a constant MSL at frequencies up to $ka = 7$, (-12 dB for DSB and -16 dB for SHB), so at frequencies between $ka = 6$ and $ka = 7$, the user could for example choose between having good resolution using DSB or good attenuation of sidelobes using SHB.

Chapter 6

Tests of Beamformer

In order to verify the theoretical performance of the Spherical Harmonics Beamformer (SHB), a model has been built using the optimized array of 64 microphones described in chapter 4. The model has been tested in various situations described in detail in appendix A. The results of the tests are discussed in this chapter and compared to the simulations based on theory.

6.1 Introduction

The test model of the SHB is constructed with a shell in hard plastic and with a radius of 0.140 metres. The radius of the sphere was chosen in order to fit in all the 64 microphones, cables and plugs. The chosen radius determines the frequency range. Choosing a bigger radius would push the frequency range down, and a smaller radius would push the frequency range up. In figure 5.6 the Maximum Sidelobe Level was simulated, and shows that the beamformer is capable of keeping a good attenuation of the sidelobes, $MSL < -15dB$, using $N = 6$ up to around $ka \simeq 7$, which is equivalent to a frequency of approximately 2.74 kHz, when the radius of the beamformer sphere is 0.140 meter.



Figure 6.1: The test model of the Spherical Harmonics Beamformer, having 64 microphones.

6.2 Test and Measurement conditions

The beamformer has been tested in an acoustical reflection free room at Odense University College of Engineering. The room is free of reflections down to 100 Hz and up to more than 20 kHz. The SHB been tested using a 5.5" loudspeaker in a closed cabinet as a sound source. The tests have been carried out in the following situations:

- Test A - Single frequency source.
- Test B - Two sources at different frequencies.
- Test C - Two sources at same frequencies.
- Test D - Single sinus source at different radii.
- Test E - Single sinus source at different angular position.
- Test F - White noise source.

- Test G - White noise source and sinus source.
- Test H - Single sinus source and reflecting plate.

A more detailed description of the tests, and the instruments, can be found in appendix A.

The samplings frequency of 16384 Hz, and low pass filter cutting off at 6.4 kHz, have been used in order to test the beamformer at frequencies higher than the performance limit at 2.74 kHz. The sampling frequency is a standard in the B&K front-end hardware and PULSE software used in the tests, as well as the low-pass filter.

6.3 Maximum Sidelobe Level

Figure 6.2 shows absolute value of the beamformer response from both the measured data (left) and a simulation (right), for a sound source placed at the distance $r_0 = 2.5m$ and at the angular position $(\theta_0, \phi_0) = (90^\circ, 180^\circ)$. The beamformer response from both the test and the simulation are very alike, having a maximum difference of 0.059 and an average difference of 0.0094 times the normalized value, 1, in the mainlobe direction.

The results at other frequencies are shown in figure 6.3 as a contour plot, at degree $N = 6$ (For other values of N see figure A.3). The colour represents the beamformer response level, relative to the level at the angle of the incident spherical wave coming from $(\theta_0, \phi_0) = (90^\circ, 0^\circ)$ at distance $r_0 = 2.5m$. The system shows to be stable from around 1.2 kHz and up to around 2.8 kHz. In section 4.2 both the phase and amplitude stability was investigated. At $ka = 3$ (equivalent to $f = 1173Hz$ for $a = 0.14$) the maximum allowed standard deviation of random normally distributed phase errors, was found to be around 1° , and the maximum allowed standard deviation of random normally distributed relative amplitude errors was found to be around 0.02 (see figure 4.4). This is the cause of the instability of the system at low frequencies,

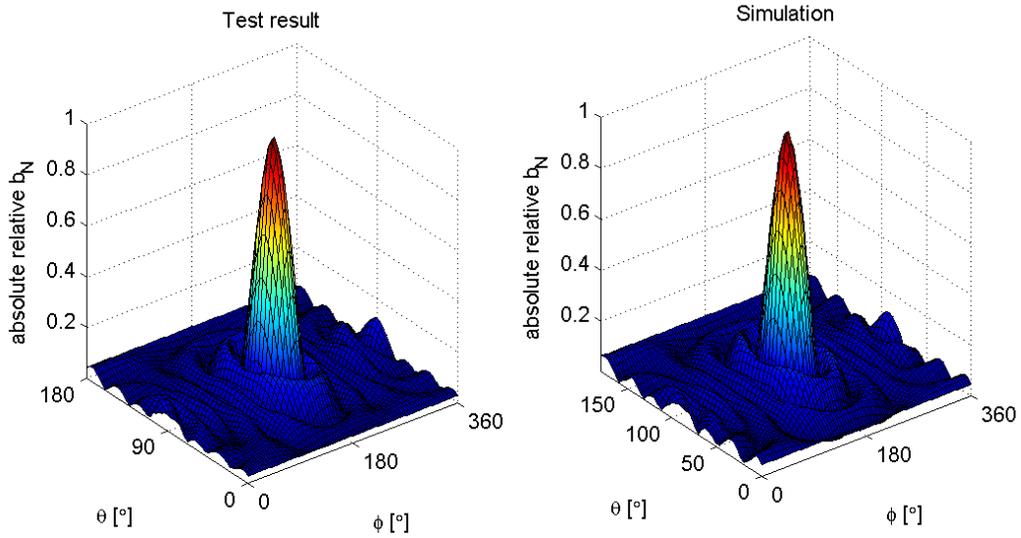


Figure 6.2: Absolute value of the beamformer response from 2000 Hz sound source placed at $(r_0, \theta_0, \phi_0) = (2.5m, 90^\circ, 180^\circ)$. Left: Results from measured data. Right: Computer simulation.

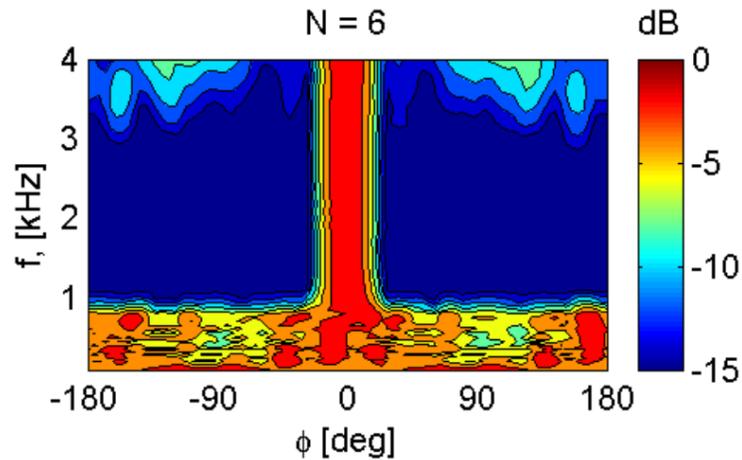


Figure 6.3: Contour plot of the beamformer response level from test results, having the frequency on the vertical axis and the angle ϕ on the horizontal axis. Incident wave from $(\theta_0, \phi_0) = (90^\circ, 0^\circ)$.

and could be improved by a better calibration of the microphones. At high frequencies, $f > 2.8kHz$, the error mostly effects the sidelobe level, and not the direction of the mainlobe (e.g. see figure 6.3 at $f = 4kHz$).

In figure 6.4 showing the Maximum Sidelobe Level (MSL), from the same test as in figure 6.3, is compared to the MSL simulated on the same array. For frequencies $f > 1200Hz$ the tested MSL differs less than 1 dB from the simulated MSL. The errors at high frequencies, $f > 2700Hz$, are also present in the simulations, and the frequency limit is related to the array. So, allowing a little deviation of the MSL from the simulated values; ie. $MSL < -15dB$ for $N = 6$ will result in a frequency range from 1200 Hz to 2770 Hz. In table 6.1 the frequency range for other values of N are shown for defined allowed limits of MSL.

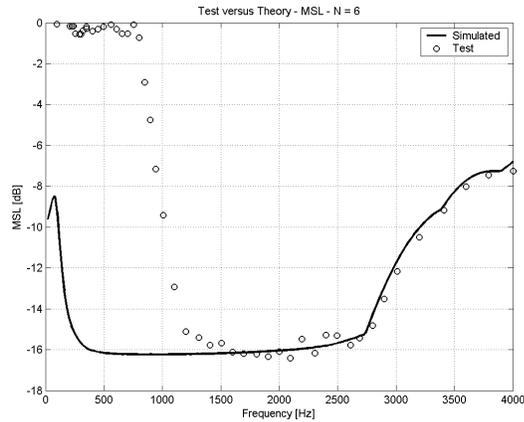


Figure 6.4: Maximum Sidelobe Level (MSL), of incident wave originating from $(r_0, \theta_0, \phi_0) = (2.5m, 90^\circ, 180^\circ)$. Simulated: '-', and from test: 'o'.

The SHB was also tested using a white noise sound source (Test F appendix A). White noise is characterized by containing all frequencies in the range from 20 Hz - 20 kHz, and by the use of the Fast Fourier Transformation, the white noise signal can be used to test the beamformer performance at all frequencies at once. Figure 6.5 shows the Maximum Sidelobe Level found from a 20 second measurement, at the discrete frequencies from the 1024-point FFT $f_n = \{16, 32, \dots, 5008\}[Hz]$, at the degrees $N = 1, 2, \dots, 6$, where the white noise source is placed in direction $(\theta_0, \phi_0) =$

N	ideal MSL	FWHM	allowed MSL	$f_{min}[Hz](ka)$	$f_{max}[Hz](ka)$
1	-6.02 dB	141.0°	<-5 dB	130 Hz (0.33)	>4000 Hz (10.22)
2	-9.54 dB	88.5°	<-8 dB	200 Hz (0.51)	3800 Hz (9.7)
3	-12.04 dB	65.0°	<-11 dB	390 Hz (1.00)	3500 Hz (9.0)
4	-13.979 dB	51.6°	<-13 dB	610 Hz (1.56)	2850 Hz (7.3)
5	-15.563 dB	42.8°	<-14 dB	870 Hz (2.22)	2770 Hz (7.1)
6	-16.902 dB	36.6°	<-15 dB	1200 Hz (3.06)	2770 Hz (7.1)

Table 6.1: Frequency range of the beamformer at degree N, at allowed limits of MSL, based on results from test A appendix A. The ideal MSL and the resolution, FWHM, are also listed.

(90°, 0°). The frequency ranges found from test A are listed in table 6.1 with the defined allowed MSL. (See section A.8 for more results on this test, including the beamformer response frequency sweep in the ϕ -direction.)

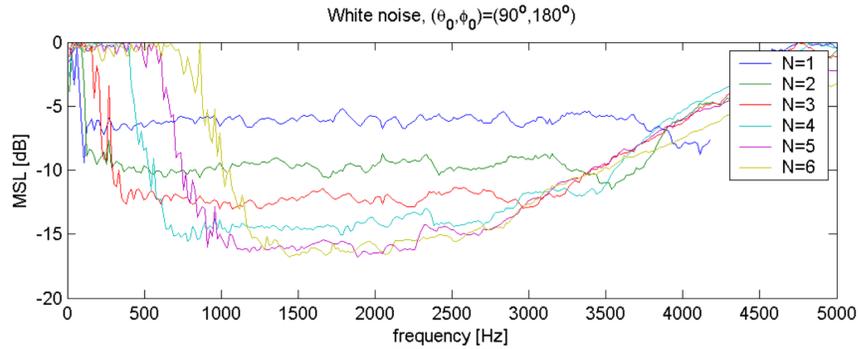


Figure 6.5: Maximum Sidelobe Level of beamformer response, from white noise sound source with angular position $(\theta_0, \phi_0) = (90^\circ, 0^\circ)$, as a function of the frequency, and at different degree $N = 1, \dots, 6$.

6.4 Direction of the mainlobe

The analysis of the results so far have been related to how good the beamformer is to suppress the sidelobes. The direction of the mainlobe is also important to investigate. The mainlobe is supposed to point in the direction from where the incoming wave approaches the array from. One of the problems involved with the testing of this, is

to determine the actual position, especially the angular position, of the sound source relative to the beamformer array. This has been achieved with a precision of $\pm 5^\circ$, and is relatively accurate in comparison to the beamformer resolution of 36.6° at $N = 6$. The result of this test can be found in appendix A, section A.7. The direction of the mainlobe showed to differ less than 4° from the direction of sound source in all tests.

The tests using white noise sound sources also showed good accuracy in the mainlobe direction, which can be seen on the beamformer response frequency sweep from test F appendix A. Figure 6.6 shows the frequency sweep at $N = 6$, and the direction of the mainlobe is in the correct direction, $\phi = 0^\circ$, from around 1.0 kHz and up. Even in the high frequency area, $f > 3kHz$, where the sidelobe level is increased, the direction of the mainlobe is still pointed towards the sound source.

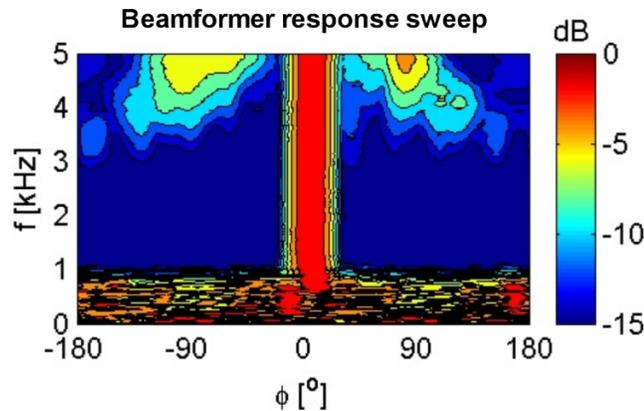


Figure 6.6: Beamformer response level sweep, at $N = 6$ from white noise source.

The beamformer responses showed in Figure 6.7 are from test F, using a white noise sound source. The figure shows that the SHB works on white noise signals, having the mainlobe of the beamformer response pointing in the right direction, both at the low frequency, $f = 1190Hz$, and at the high frequency, $f = 2800Hz$. Both results have a $MSL < -15dB$.

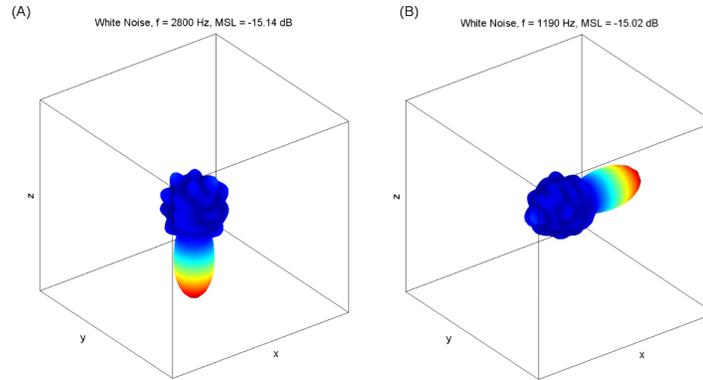


Figure 6.7: (A) White noise source placed at $\theta = 180^\circ$ (southpole), and SHB performed at $f = 2800Hz$. MSL = -15.14 dB. (B) White noise source placed at $(\theta, \phi) = (90^\circ, 180^\circ)$ (negative x-direction), and SHB performed at $f = 1190Hz$. MSL = -15.02 dB.

6.5 Radius dependance

Analysis of the effect of the distance of the sound source has also been carried out (see section A.6). The tests showed that it is possible to get a beamformer response of degree $N = 6$ at lower frequencies (i.e. $f = 800Hz$) if the source is close to the beamformer array (i.e. $r_0 = 0.25m$). See figure A.8.

6.6 Two sound sources

The resolution has been tested by having two sound sources. If the frequency of the two signals are different, then the FFT will separate the two signals as shown in test B. The frequency resolution in the FFT determines how close two frequencies can be. In this test, 1024 samples at a samplings frequency of 16384 Hz, results in frequency resolution of $\delta f = 16Hz$. If the frequency of the two sources are the same, they will appear in the same beamformer response. This is tested in test C, where the internal angle between the two sources is 52° , seen from origo. At the low frequency, $f = 624Hz$, the maximum degree of spherical harmonics possible to retrieve is $N = 4$, having the resolution, $FWHM = 51.6^\circ$. This resolution is not

good enough to distinguish the two sources. At the higher frequency, $f = 2432\text{Hz}$, the degree $N = 6$ having the resolution $FHWM = 36.6^\circ$ is possible, and now the two sources can easily be distinguished from each other (see figure ??).

A disadvantage of having two discrete sources of the same frequency, is that the sidelobes from one source can disturb the beamformer response of the other source, resulting in a displacement of the mainlobes. This can also be seen in figure ?? at the bottom, where the beamformer response is calibrated for the source on the right, resulting in a small displacement of the mainlobe for the source on the left. The blue rings above and below the middle of the two sources, are caused by the interference of the sidelobes.

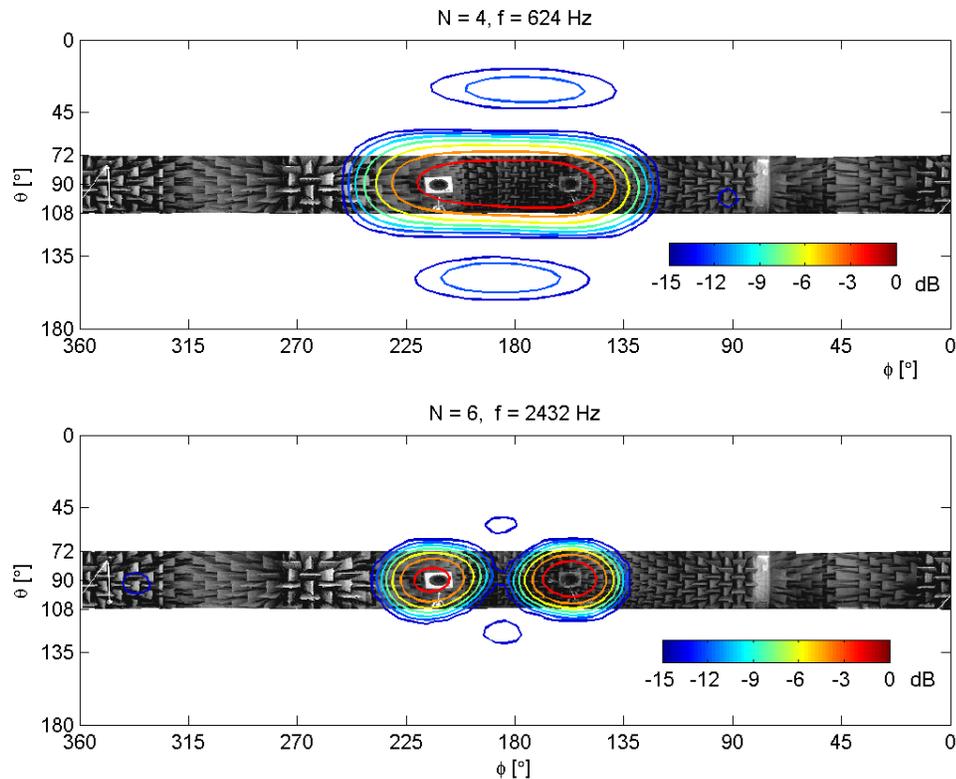


Figure 6.8: Beamformer response level in combination with picture. Top: At $f = 624\text{Hz}$ it is only possible to get the result of degree $N = 4$. Bottom: At $f = 2432\text{Hz}$ a result of degree $N = 6$ is possible.

6.7 Full performance of beamformer

In the test using white noise as a sound source, the full performance of the beamformer at all frequencies up to 6.4 kHz (the cut-off frequency of the low pass filter) can be evaluated, at any degree N . The measurements from test F having the sound source positioned at $(r_0, \theta_0, \phi_0) = (1.00m, 90^\circ, 180^\circ)$, are used, and the MSL is found at all the discrete frequencies from the FFT, $f_n = 16, 32, 48, \dots, 6400[Hz]$, and for $N = 1, 2, \dots, 20$.

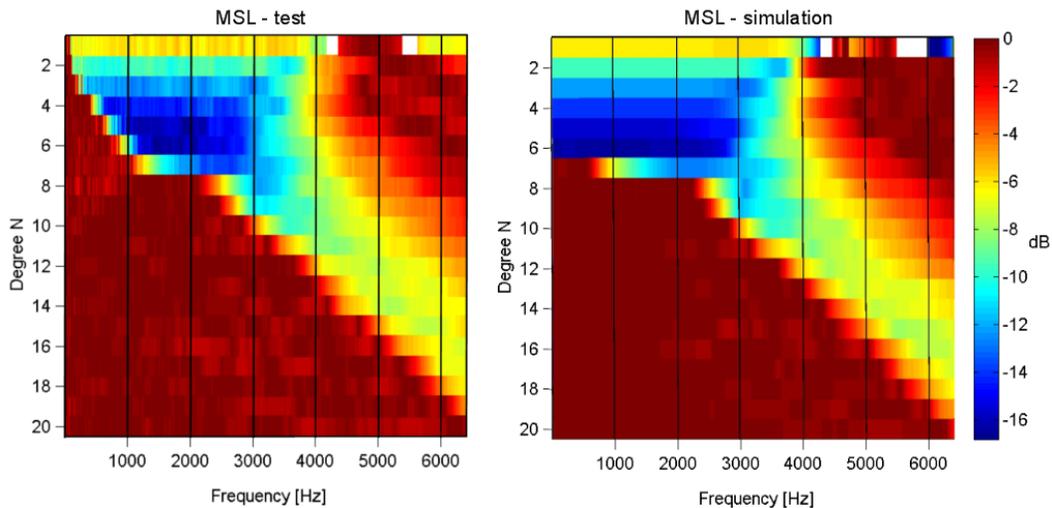


Figure 6.9: Maximum sidelobe level of incident wave from sound source at $(r_0, \theta_0, \phi_0) = (1.00m, 90^\circ, 180^\circ)$, for $N = \{1, 2, \dots, 20\}$. Left: test result. Right: simulation.

Figure 6.9 shows the MSL (represented in colour) from the measurement on the left, and from the simulation at the same conditions on the right. The performance of the beamformer in the test situation is very similar to the simulations, only differing significantly at low frequencies for $N \leq 6$. The figure illustrates that it is possible to get a beamformer response at higher frequencies than the 2.7 kHz, by using higher degree, N , and still achieve a MSL of around -8 dB (green colour). This is shown for $N = 14$ at $f = 5300Hz$ in figure 6.10, again both from measured data, and from a simulation. The resolution at $N = 14$ is $FWHM = 17^\circ$. Note how similar the test and the simulations are, which confirms that the simulations are correct, and

that the error of the sidelobes are entirely due to the array geometry (A beamformer response without errors would have N sidelobes, seen as concentric circles around the mainlobe).

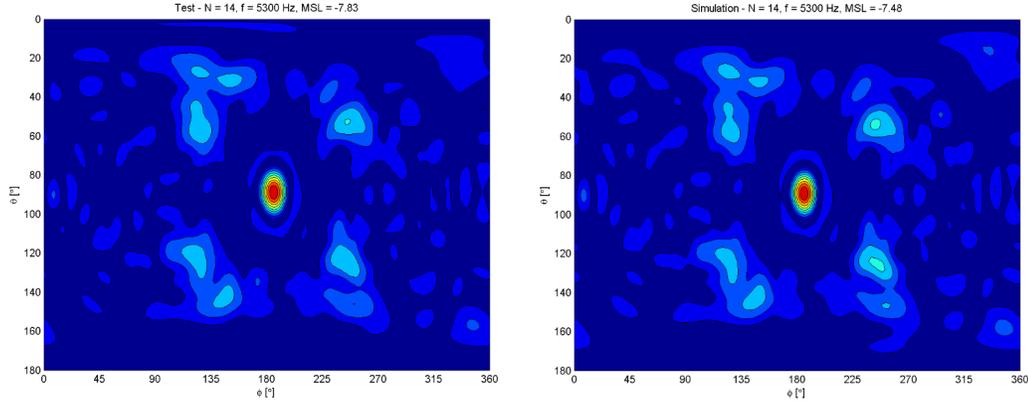


Figure 6.10: Absolute value of beamformer response, from sound source at position $(r_0, \theta_0, \phi_0) = (1.00m, 90^\circ, 180^\circ)$, and frequency $f = 5300$, and the degree $N = 14$.

By using the beamformer response of the degree N , showing the best performance at a specific frequency, it is possible to get an extended frequency range. This is done in figure 6.11 from $f = 16Hz$ to $f = 6400$, by using the beamformer responses of degree $N = \{1, 2, \dots, 15\}$. In the frequency range, $f = [150Hz, 6400Hz]$, the MSL is less than -7 dB. For all measurements and simulations made at high frequencies, $f > 2700Hz$, the attenuation of the sidelobes around the mainlobe and in the opposite direction of the mainlobe showed to be good, $MSL < -15dB$.

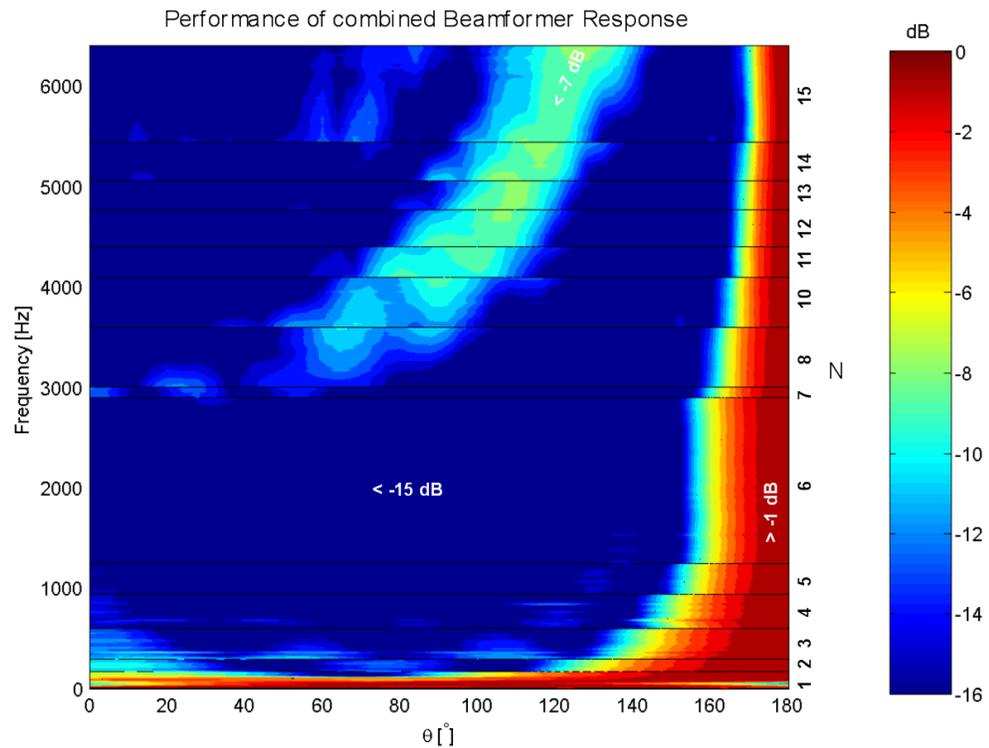


Figure 6.11: Combined beamformer response level sweep of measured data from white noise sound source placed at $\theta = 180^\circ$ at a distance of 1.00 m. The beamformer response showing the best performance at a given frequency is used. In the darkest blue areas, the highest level of sidelobes is less than -15 dB, and in the green area at high frequencies, the highest level of sidelobes is less than -7 dB.

6.8 Test with reflecting wall

In a confined space such as a car cabin, or similar spaces where the spherical beamformer would be operating, the existence of reflecting surfaces would be common. The last test (Test H appendix A) is made to see how the beamformer would respond to a situation having a sinus source and a reflecting surface. The reflecting surface is placed in such a way that it reflects the incoming wave from the sound source directly towards the beamformer sphere. The sound wave travelling directly from the sound source to the center of the sphere, travels 1.00 meter, and the wave being reflected by the reflecting surface, travels a total path of 2.26 meter. This results in the following intensity level of the reflected wave compared to the intensity level of the direct wave¹

$$L_{refl} = 20 \log \left(\frac{1m}{2.26m} \right) = -7.1dB. \quad (6.8.1)$$

The result of the beamformer response level is shown in figure 6.12. As expected, the largest mainlobe showed to be in the direction of the sound source, and a secondary mainlobe appeared in the direction of the reflecting surface. The level of the secondary mainlobe is -5.89 dB, and was expected to be -7.1 dB (see equation 6.8.1).

¹The amplitude of the sound pressure of a spherical sound wave is inverse proportional to the distance from its center[14]

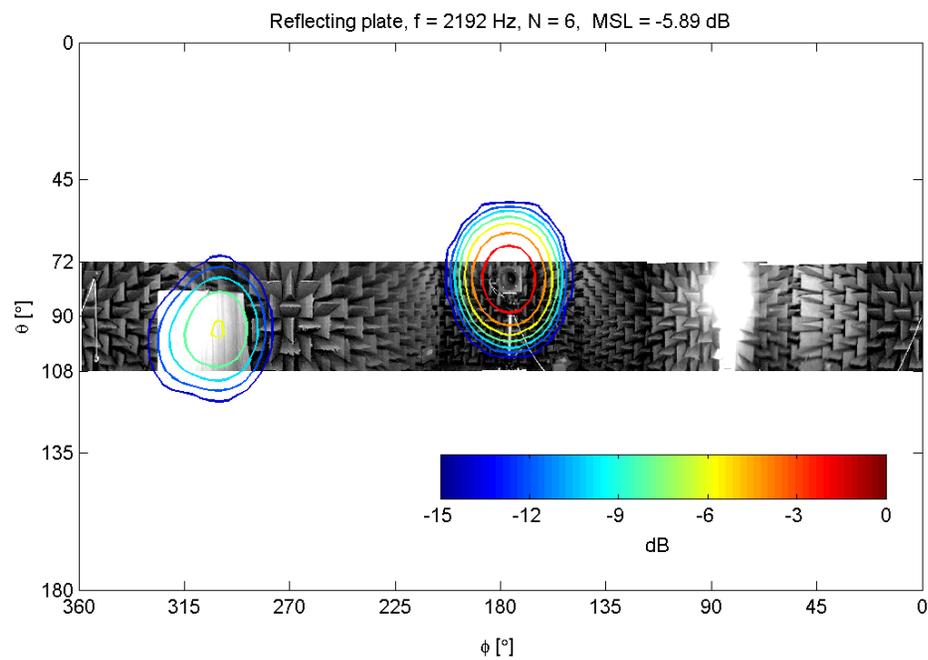


Figure 6.12: Combination of contour plot of the beamformer response level, $20\log\left(\frac{|b(\theta,\phi)|}{|b(\theta,\phi)|_{max}}\right)$, and a picture of the surrounding environment around the equator. The bright white light at $\phi = 85^\circ$ is the sunshine leaking into the room and confirms that we did get a little sunshine this summer of 2004 while the measurements were performed.

6.9 Chapter summary

In all the tests, the SHB technique showed to work on the Fast Fourier Transformed data of the measurements, making it possible to detect signals at a desired frequency.

The results of spherical harmonics beamforming, performed on the measured data from the test model, showed to be similar to the simulated results from the frequency of around 1200 Hz and up for $N = 6$. The Maximum Sidelobe Level in the test only differed by less than 1 dB from the simulated values, and the direction of the mainlobe was less than 4° from the actual direction of the sound source. The error on beamformer response at lower frequencies, $f < 1200\text{Hz}$, for $N = 6$, is due to the sensitivity towards the internal phase and amplitude errors of the microphones. This instability could be improved by better calibrations of the microphones.

The beamforming at $N = 6$ was possible in the frequency range $f = [1200\text{Hz}, 2770\text{Hz}]$ having a MSL lower than -15 dB. By combining beamformer responses at other degrees N it is possible to extend the frequency range. This is done by using the b_N which shows the best performance at a specific frequency, and it is thereby possible to perform the beamforming down to 130 Hz with $N = 1$, and up to 6400 Hz with $N = 15$, with a MSL of less than -7 dB (see figure 6.11).

Chapter 7

Conclusion and Perspectives

In this chapter the conclusions of the project are drawn, followed by reflections to be used for further studies.

7.1 Conclusion

A new method for sound source location using 3D-microphone arrays, the Spherical Harmonics Beamforming method (SHB), has been presented.

The sound field described in spherical coordinates has been examined, and can be described by the use of spherical harmonics and spherical Bessel functions. Simulations of the sound field were made in order to show how high degree of spherical harmonics and order of spherical Bessel functions are needed in various situations.

The ideal function for describing the direction of an incoming wave field, the delta function, was decomposed into spherical harmonics, and compared to the sound field on a sphere. The examinations led to a method of extracting the angular part of the sound field on a sphere, which required spherical integration on the sphere. This led to the closer study of numerical integration on the sphere, due to the fact that the sound field can only be determined at the positions of the microphones on the sphere. Two methods, both using the cubature rule, were examined; a method to determine the cubature weights, and spherical t-designs optimized for spherical integration by using equal cubature weights. The examination also led to an expression of the error,

to be used for optimization.

An analysis of the SHB method has been made, resulting in the choice of a hard shell as the optimal surface of the beamformer sphere. In order to construct a test model of the SHB-array, an optimization of an array of 64 microphones has been made. The optimization was made by using the existing t-design optimized by Hardin and Sloane, with a symmetry order of 12, but combined with the use of the reproducing kernel resulting in unequal cubature weights. The array is optimized to be capable of performing SHB at a degree at most 6, in the frequency range from $ka = 1$ to $ka = 7$. Simulation of phase and amplitude errors on the microphone signals, showed that the beamformer response is very sensitive at low frequencies and higher N , and most stable at $ka = 6$ for $N = 6$.

The performance of the SHB method on the optimized array has been compared with the performance of the Delay and Sum Beamforming (DSB) method on the same array, but without the hard shell. The resolution of the DSB-method shows to be frequency dependant, having a wide mainlobe at low frequencies ($FWHM > 180^\circ$ at $ka < 1.4$), and a narrower mainlobe at high frequencies, with a Maximum Sidelobe Level (MSL) of around -12 dB. The SHB-method showed to have a constant resolution in the frequency range from $ka = 1$ to $ka = 7$ with a $FWHM = 36.6^\circ$. Simulations of Maximum Sidelobe Level of SHB showed to be $MSL < -15dB$ in a frequency range of around 2.7 octaves.

Results from measurements, made on the test model, showed to agree with all simulations except at low frequencies. The beamformer response of degree 6 was only possible down to a frequency of around 1200 Hz ($\sim ka = 3.07$). This error is caused by the phase and amplitude errors on the microphone signals, as expected in the stability evaluations.

All in all, the method for localizing a sound source, by decomposing the sound field into spherical harmonics, has been a success. Improvements and suggestions of further studies are presented in the next section.

7.2 Perspectives

The optimization of the microphone positions could be investigated further in order to find a design with a wider frequency range and perhaps higher N (i.e. better resolution). This can be done by running further computer optimizations for finding the global (or a lower local) minimum of the errors on the decomposition constants of the SHB-response. The optimizations could be done to achieve a higher upper frequency limit, by optimizing at a lower N , if the resolution is less important than the frequency range.

Calibration of the microphones should be investigated, in order to reduce the error on the beamformer response at low frequencies, which is caused by the sensitivity towards errors on the measured signals from the microphones. Perhaps an automated calibration method could be developed by measuring known signals from a known sound source placed at various positions in the room.

The microphone positions in this project are limited to a sphere, but other ways of placing the microphones could be investigated, and still by using spherical harmonics. Perhaps another set of microphones placed at a distance from the surface of the sphere help to achieve a higher N or wider frequency range, could be interesting.

Appendix A

Tests

This appendix explains the tests performed on the SHB test model, and shows the results of the tests.

A.1 Hardware

The following hardware has been used to perform the tests:

Transducer

- Spherical beamformer having 64 microphones mounted on a hard sphere in alignment with the surface. The radius of the sphere is 0.140 m.
- 64 x B&K microphones type 4935.

Detection

- Enclosure B&K, type 3560D
- 5 x 12 channel module type 3038B
- 1 x 5 channel controller module type 7537

- Laptop computer, 1500 MHz centrino processor

Sound source and Signal generator

- 5.5" Loudspeaker (Vifa P13WH-00 08) in 1.49 L, closed cabinet.
- B&K Sine-Random Generator Type 1024

A.2 Test Setups

All tests have been carried out in the reflection free room, at Odense University College of Engineering. The room is acoustically reflection free in the frequency range from around 100 Hz and up to more than 20 kHz. The general setup for all the experiments is seen in figure A.1. All measurements are performed at a samplings frequency of 16384 Hz, and filtered by the built-in low pass filter cutting off at 6.4 kHz.

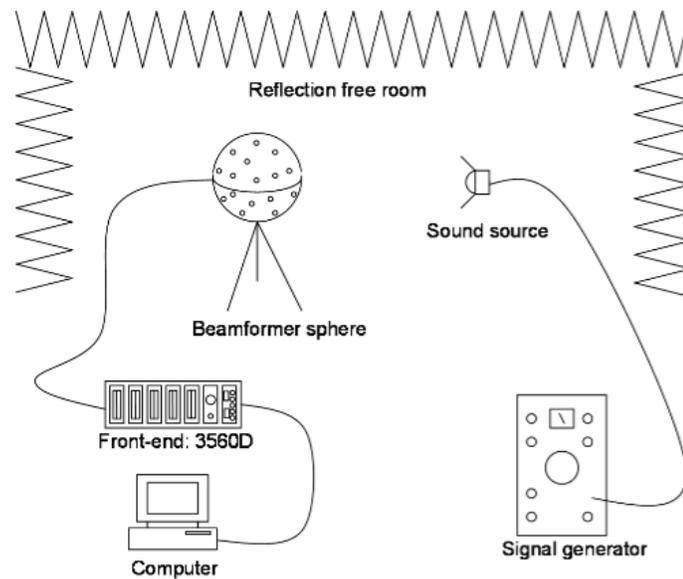


Figure A.1: General setup for experiments with beamformer sphere in acoustically reflection free room.

The following tests have been performed:

- Test A - Single frequency source.
- Test B - Two sources at different frequencies.
- Test C - Two sources at same frequency.
- Test D - Single sinus source at different radii.
- Test E - Single sinus source at different angular positions.
- Test F - White noise source.
- Test G - White noise source and sinus source.
- Test H - Single sinus source and reflecting plate.

Further descriptions of the tests and the results are to be found in the following sections:

A.3 Test A - Single Sinus source

Purpose: To investigate how the beamformer responds to a sound field originating from a single frequency sound source, and to see how great a degree of spherical harmonics it is possible to achieve at different frequencies.

Setup: See figure A.1.

Test conditions: The sound source is placed at the angles $(\theta_0, \phi_0) = (90^\circ, 180^\circ)$ and at the distance $r_0 = 2.50m$ from the center of the sphere. The measurements were performed at the following frequencies each at a duration of one second: $f = \{100, 200, 220, 240, 260, 280, 300, 340, 360, 380, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 3000, 3200, 3400, 3600, 3800, 4000\} [Hz]$.

Results: Figure A.2 shows the beamformer response generated from the measured data at frequency $f = 2000Hz$, developed to the degree of spherical harmonics $N = 6$.

Figure A.3 shows the result in [dB] at different degree N. Each figure shows $20\log\left(\frac{|b(90^\circ, \phi)|}{|b(90^\circ, 180^\circ)|}\right)$, at the frequencies defined in the test conditions.

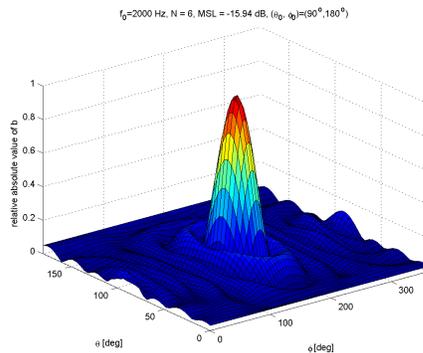


Figure A.2: Absolute value of the beamformer response at $N = 6$ generated from measured data. Sound source is placed at distance $r_0 = 2.50m$ and at angles $(\theta_0, \phi_0) = (90^\circ, 180^\circ)$, emitting an sinusoidal signal at frequency $f = 2000Hz$.

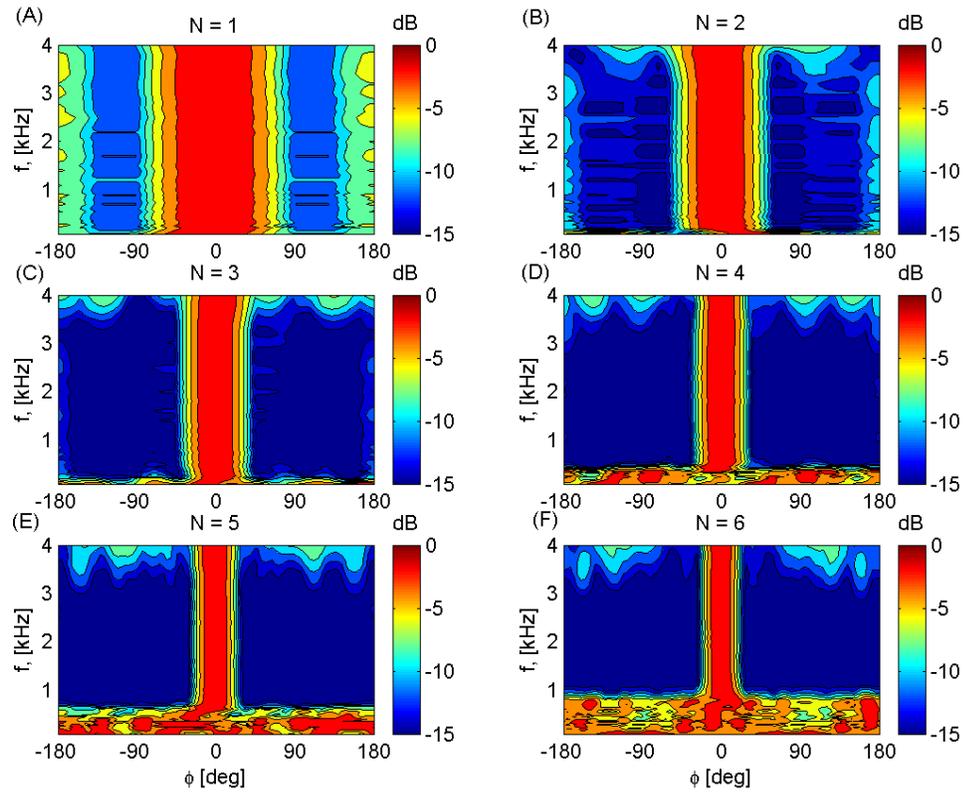


Figure A.3: Sweep of beamformer response level from sound source with single frequency signal. The frequency is marked on the vertical axis, and the azimuth angle with 0 in the direction of the source on the horizontal axis. SHB performed at $N = 1$ to $N = 6$ shown at plot A to plot F respectively.

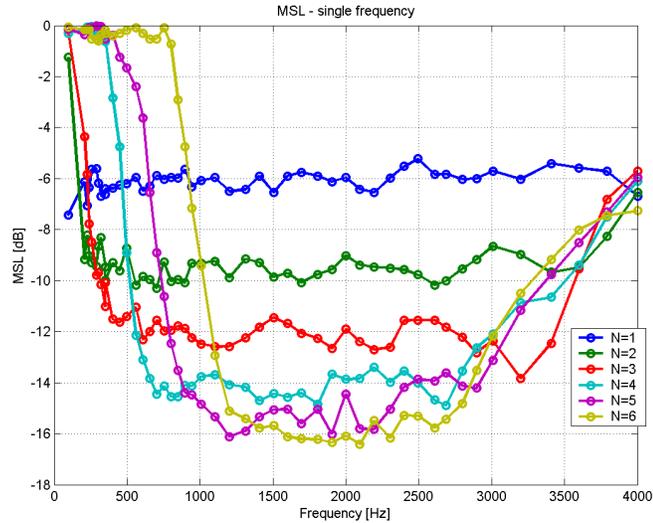


Figure A.4: Maximum Sidelobe Level of SHB-response at $N = \{1, 2, \dots, 6\}$, from tests with single frequency sound source.

A.4 Test B - Two single frequency sources at different frequencies

Purpose: To investigate how two sound sources at different frequencies affect each others beamformer response.

Setup: As in figure A.1, but with the use of two signal generators and two loudspeakers as sound sources. Both sound sources are placed at distance $r_0 = 1.00m$. The following angular positions of the sources have been used. S1: $(\theta_0, \phi_0) = (90^\circ, 135^\circ)$ and S2: $(\theta_0, \phi_0) = (90^\circ, 180^\circ)$

Test conditions: Three tests have been performed on test situation one:

- In the first measurement, source one is present with a frequency of 1520 Hz and source two with a frequency of 2000 Hz.

- In the second measurement, the second source is turned off, and source one is present alone at $f = 1520Hz$.
- In the last measurement source two is present alone at $f = 2000Hz$

Results: The results from test one in figure A.5 (A) and (B) show the absolute value of the beamformer response where both signals are present. The SHB is performed on the FFT-output having $f_\eta = 1520$ (A) and $f_\eta = 2000$ (B). The results shown on figure (C) and (D) are from the test where the respectively signals are present alone. Pictures (E) and (F) show the difference of the above results, and the difference shows in both cases to be less than 0.003 times the maximum response, equivalent to an error of less than -50 dB.

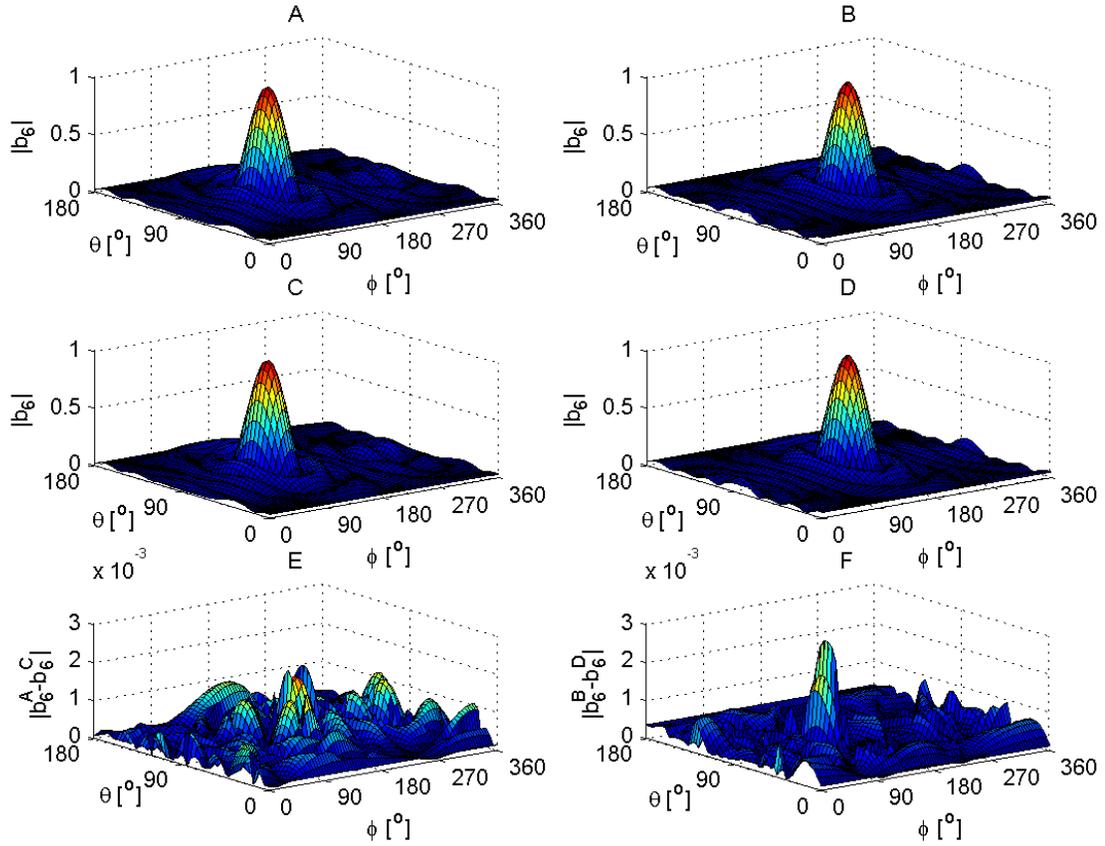


Figure A.5: A) Beamformer response focused at $f_{focus} = 1520$ having both sources present. B) $f_{focus} = 2000$ having both sources present. C) $f_{focus} = 1520$ source one alone. D) $f_{focus} = 1520$ source two alone. E) shows difference between (A) and (C). F) shows difference between (B) and (D).

A.5 Test C - Test Two sources at same frequency

Purpose: To investigate the importance of the resolution in order to distinguish two signals at same frequency.

Setup: As in figure A.1, but with the use of two signal generators and two loudspeakers as sound sources. Both sound sources are placed at distance $r_0 = 1.00m$. The following angular positions of the sources have been used. S1: $(\theta_0, \phi_0) = (90^\circ, 210^\circ)$ and S2: $(\theta_0, \phi_0) = (90^\circ, 158^\circ)$

Test conditions: The results of the test have been made to combine the beamformer response level with a picture of the surrounding environment, in order to see if it is possible to distinguish two discrete sources at same frequency. Their internal angle seen from the beamformer is 52° . The test has been made at two frequencies; at low frequency $f = 624Hz$ and at high frequency $f = 2432Hz$.

Results: Figure A.6 shows the results. At the low frequency $f = 624Hz$, it is only possible to get a beamformer response of degree $N = 4$, where it is difficult to distinguish the sources as two separate sources with an internal angle of 52° . At the higher frequency $f = 2432Hz$ the response of degree $N = 6$ is possible, as it is then easier to distinguish

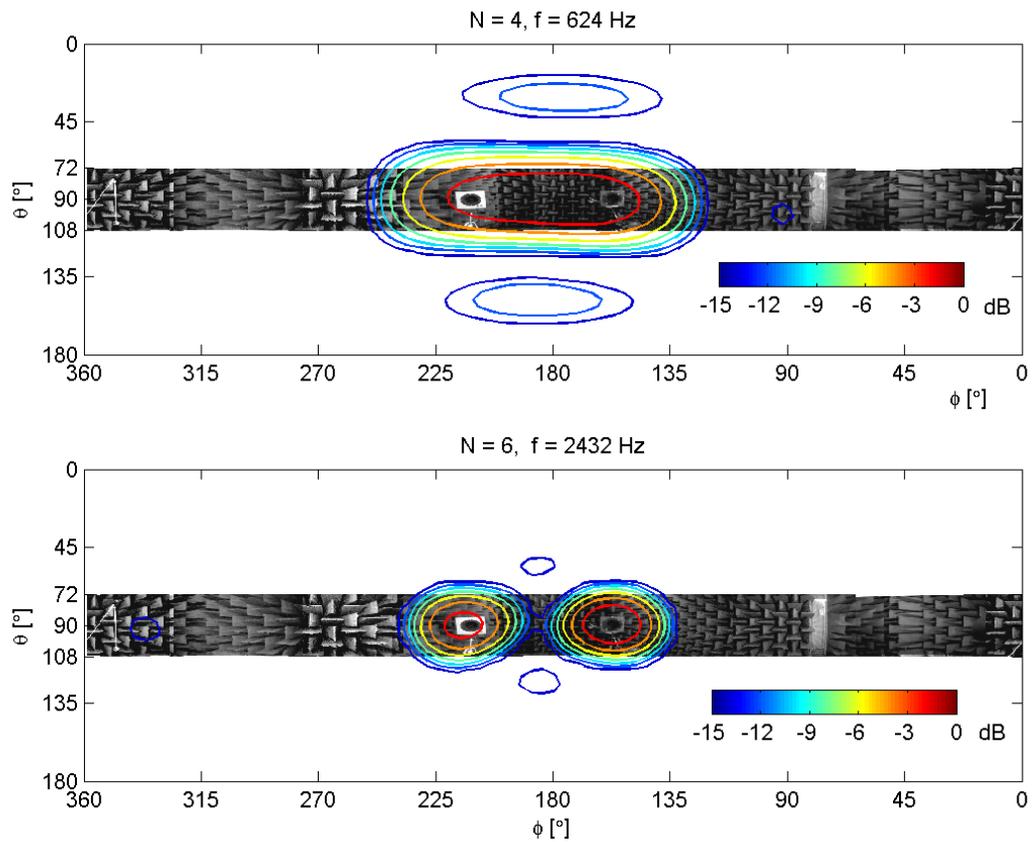


Figure A.6: Beamformer response level in combination with picture. Top: At $f = 624\text{Hz}$ it is only possible to get the result of degree $N = 4$. Bottom: At $f = 2432\text{Hz}$ a result of degree $N = 6$ is possible.

A.6 Test D - Single Sinus source at different radii

Purpose: To investigate how the distance to the source affects the beamformer response.

Setup: See figure A.1.

Test conditions: The sound source is placed at the angles $(\theta_0, \phi_0) = (90^\circ, 180^\circ)$. The measurements are made at the distances $r_0 = \{0.25m, 0.32m\}$ and at the frequencies $f = \{800Hz, 2000Hz\}$.

Results: Figure A.7 shows the absolute value of the beamformer response at frequency $f = 2000Hz$. At both plots the beamformer is focused at the expected distances of 0.25 m and 3.2 m respectively. In both situations the beamformer is capable of detecting the source to a degree of $N = 6$. The Maximum Sidelobe Level is slightly higher $MSL = -14.8dB$ at the shorter distance than at the higher distance $MSL = 16.6dB$. The ideal MSL is simulated to be 16.1 dB.

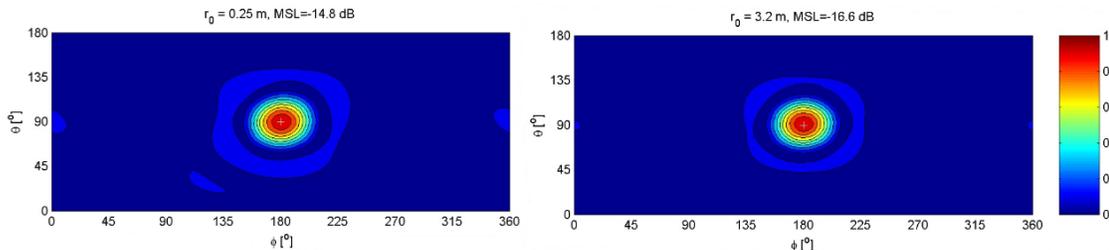


Figure A.7: Absolute value of beamformer response, $|b_N|$, at degree $N = 6$, from two test situations at frequency $f = 2000Hz$. Left: Source is placed at distance $r_0 = 0.25m$, and the Maximum Sidelobe Level (MSL) shows to be -14.8 dB Right: Source is placed at distance $r_0 = 3.2m$, and MSL is here -16.8 dB. White '+' indicates the direction of the source.

Figure A.8 shows the same situation, but at frequency $f = 800Hz$. The left picture is again at the close distance of $r_0 = 0.25m$ and the picture on the right has, $r_0 = 3.2m$. The system is expected to be unstable at this low frequency for $N = 6$,

but for $r_0 = 0.25$ the beamformer shows to be capable of detecting the direction of the sound source, but with some error. The MSL showed to be around -12.1 dB. The beamformer was incapable of detecting the direction of the source at the higher distance (see right picture of figure A.8).

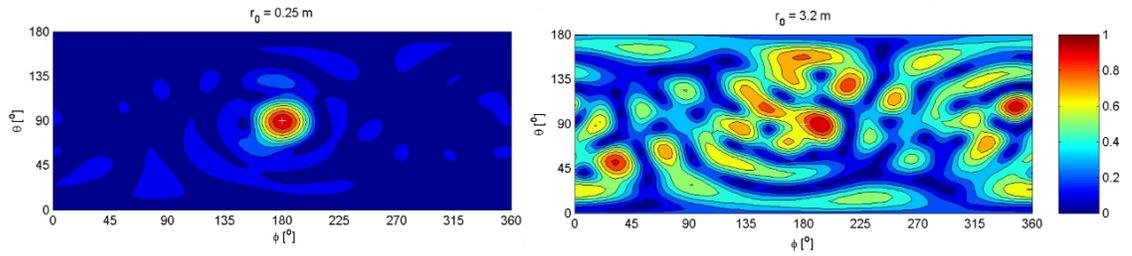


Figure A.8: Absolute value of beamformer response, $|b_N|$, at degree $N = 6$, from two test situations at frequency $f = 800Hz$. Left: Source is placed at distance $r_0 = 0.25m$. Right: Source is placed at distance $r_0 = 3.2m$. White '+' indicates the direction of the source.

A.7 Test E - Single Sinus source at different angles

Purpose: To investigate if the beamformer response shows maximum value in the expected direction (i.e. the direction in which the sound source is placed).

Setup: See figure A.1.

Test conditions: The distance of the sound source is in all tests $r_0 = 1.00m$, and the frequency is $f = 2000Hz$. The position of the sound source is determined by the use of a meter band, resulting in some uncertainty of the angle position. If the position of the sound source is determined with an uncertainty of $\pm 2cm$ at a distance $r_0 = 1m$, it will result in the uncertainty of the angle to be:

$$\Delta\vartheta_{source} = \pm \frac{0.02[m]}{2\pi \cdot 1.00[m]} \cdot 360^\circ = 1.15^\circ \quad (\text{A.7.1})$$

The center of the beamformer is set to be the center of the coordinate system, but the angular position of the sphere is also considered to have some uncertainty. This uncertainty is evaluated to be less than 3° . This results in the uncertainty of the position of the sound source in relation to the spherical beamformer to be less than $\pm 5^\circ$.

Results: Figure A.9 illustrates the results of the beamformer response from 9 different measurements. The angular position of the sound source is stated above each graph, and is illustrated on the graph with a white '+'-sign. In all results the maximum of the beamformer response was within 4° of the expected angle of incidence.

A.7. TEST E - SINGLE SINUS SOURCE AT DIFFERENT ANGLES 501

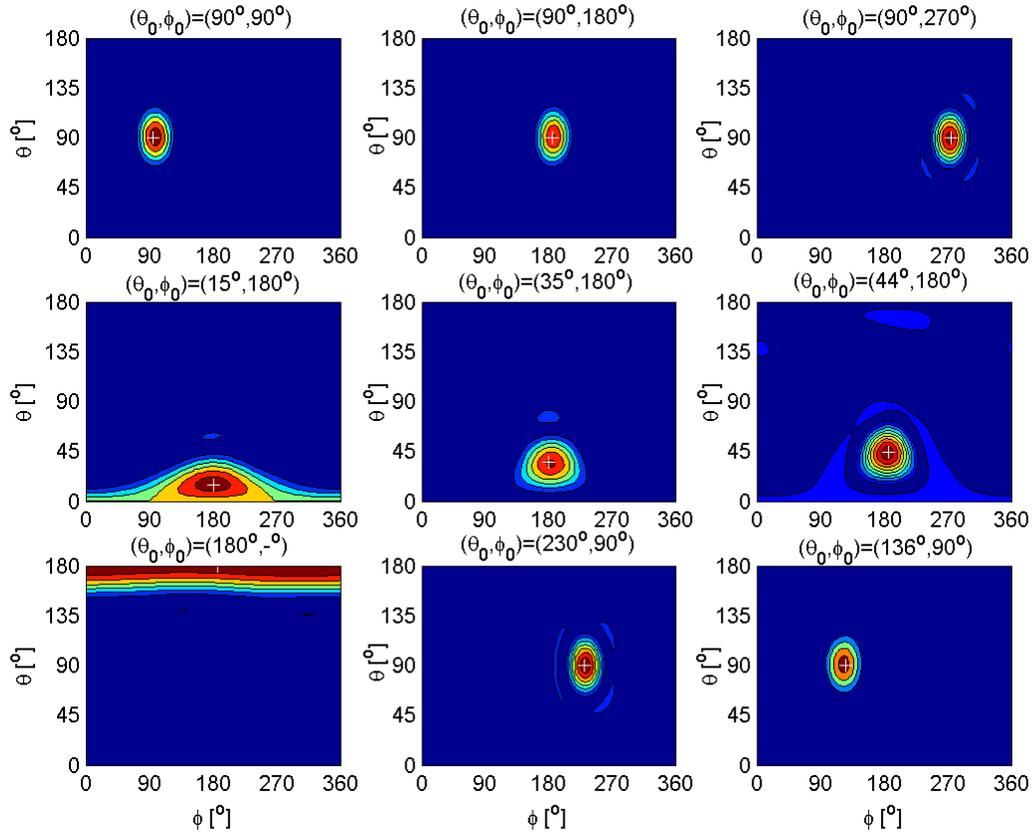


Figure A.9: Absolute value of beamformer response, b_N , for different angles of the incident wave, stated above each plot. Degree of spherical harmonics: $N = 6$, frequency: $f = 2000Hz$.

A.8 Test F - White noise sound source

Purpose: To investigate how the beamformer will respond to a white noise signal.

Setup: See figure A.1.

Test conditions: The duration of the measurements in this test is 20 seconds, and the samplings frequency is $f_s = 16384Hz$. The signal is split up into 320 pieces of 1024 samples. Each piece is Fast Fourier Transformed, and then the 320 resulting FFT's are averaged before the beamformer response is calculated. Two test cases are carried out, both having the noise source at the distance $r_0 = 1.00m$. In the first test the angular position of the sound source is $(\theta_0, \phi_0) = (90^\circ, 0^\circ)$, and in the second test the angular position is $\theta_0 = 180^\circ$ (*the south pole*).

Results: Figure A.10 shows the result of the beamforming made on the 512 discrete results of the FFT's, with a frequency resolution of, $\Delta f = f_s/1024 = 16Hz$. The frequency is represented at the vertical axis. The horizontal axis shows the azimuth angle ϕ , and the colour of the colour plot represents the maximum value of the absolute value of the beamformer response at the given frequency and given angle ϕ , relative to the maximum value at all angles in dB. The six different pictures represent the different results by using the maximum degree of spherical harmonics, from $N = 1$ to $N = 6$ respectively.

Figure A.11 shows the MSL (Maximum Sidelobe Level) of the same results as figure A.10. The MSL is found as the biggest local minimum relative to the value at maximum of the mainlobe, positioned at $(\theta_0, \phi_0) = (90^\circ, 0^\circ)$.

Figure A.12 and A.13 shows the same kind of results as plotted in figure A.10 and A.11, but for the situation where the sound source is placed at the direction of the south pole.

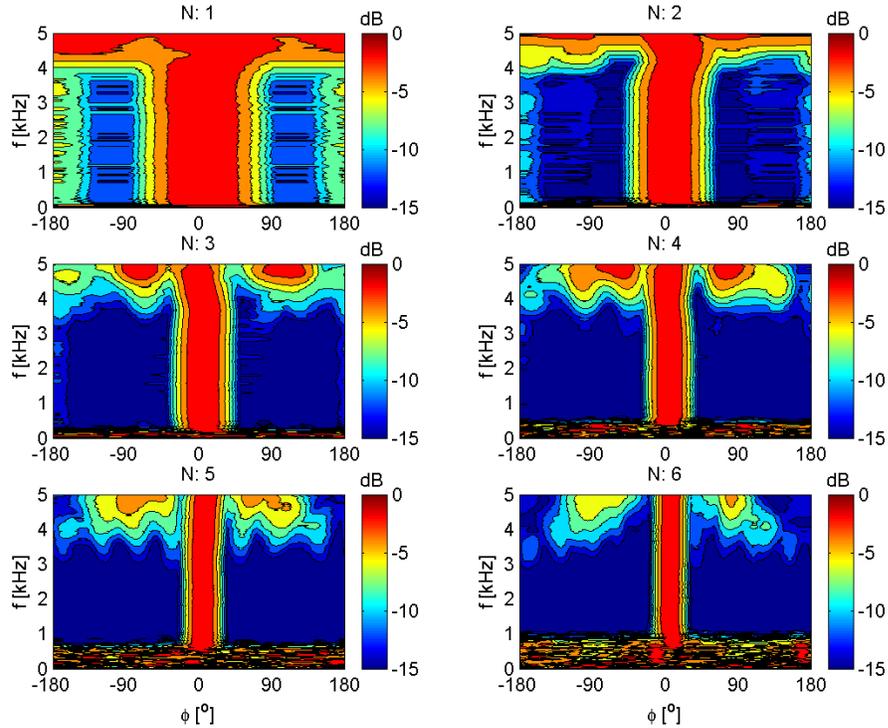


Figure A.10: Relative beamformer response level, as a function of the angle and the frequency, using development of spherical harmonics to a degree of max $N = 1$ to $N = 6$ respectively. The angular position of the sound source is $(\theta_0, \phi_0) = (90^\circ, 0^\circ)$. In the dark blue areas the beamformer response level is lower than -15 dB relatively to the highest value.

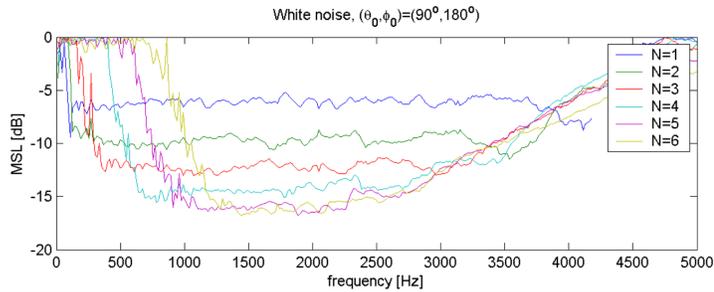


Figure A.11: Maximum Sidelobe Level of beamformer response, from white noise sound source with angular position $(\theta_0, \phi_0) = (90^\circ, 0^\circ)$, as a function of the frequency, and at different degree $N = 1, \dots, 6$.

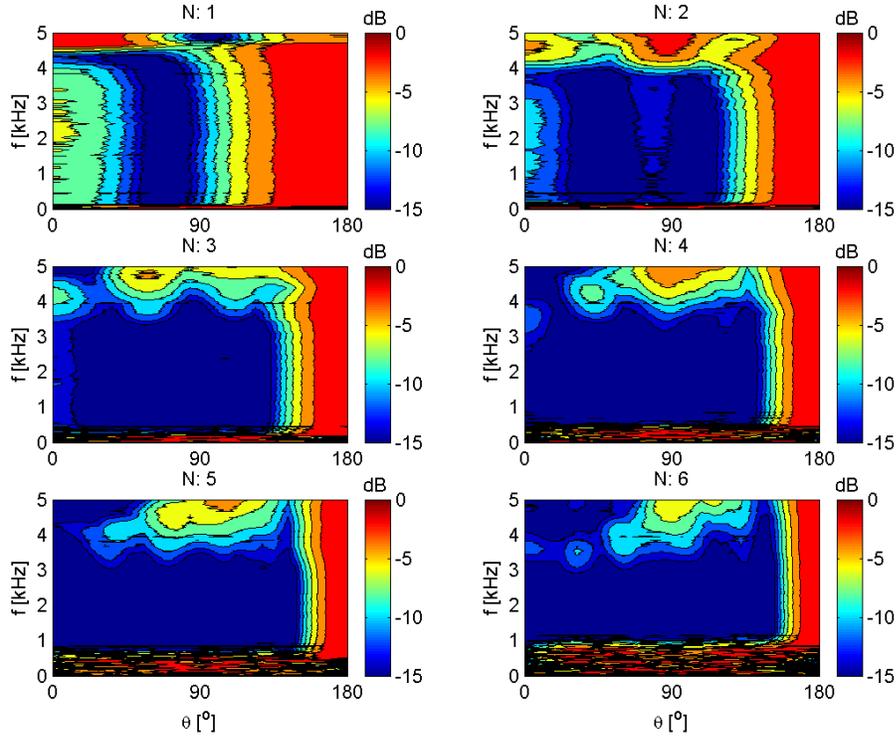


Figure A.12: Relative beamformer response level, as a function of the angle and the frequency, using development of spherical harmonics to a degree of max $N = 1$ to $N = 6$ respectively. The angular position of the sound source is $\theta_0 = 180^\circ$. In the dark blue areas the beamformer response level is lower than -15 dB relatively to the highest value.

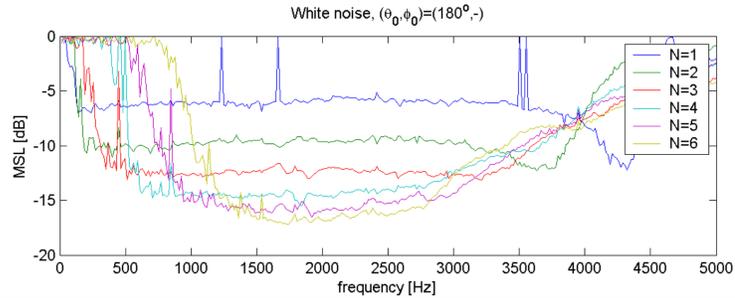


Figure A.13: Maximum Sidelobe Level of beamformer response, from white noise sound source with angular position $\theta_0 = 180^\circ$, as a function of the frequency, and at different degree $N = 1, \dots, 6$.

A.9 Test G - White noise and sinus source

Purpose: To investigate having both a sinus sound source and a white noise source present at the same time.

Setup: The same setup as figure A.1, but having two sets of signal generator and loudspeaker. The two loudspeakers are placed at the distance $r_0 = 1.00m$, and at the elevation $\theta_0 = 90^\circ$. Azimuth of source one is $\phi = 210^\circ$ and source two is $\phi = 158^\circ$. (see figure A.14)

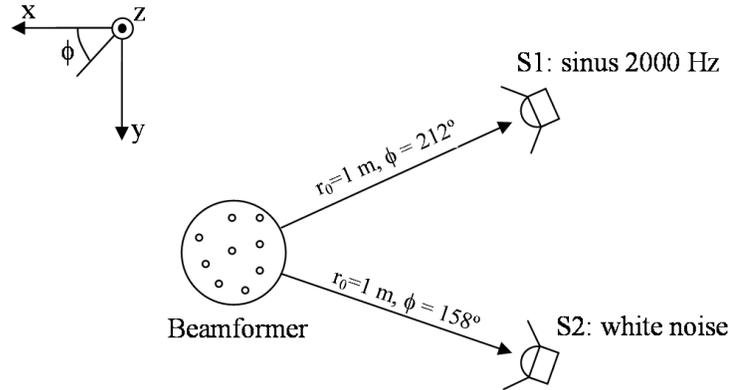


Figure A.14: Setup for test with both sinus source and white noise source seen from above.

Test conditions: The measurements are carried out for 1 second, at the samplings frequency $f_s = 16384$. The results are mixed with a picture of the surroundings. The sinus signal at source one is at frequency $f = 2032Hz$.

Results: Figure A.15 shows the average of the 64 resulting FFT's of the signals from the 64 channels. It is easy to see that the signal from source one stands out at 2032 Hz. Figure A.16 illustrates the result mixed with the photo of the surroundings. In the top picture the results of the FFT at 2032 Hz are used, and in the bottom picture the results at 2736 Hz are used.

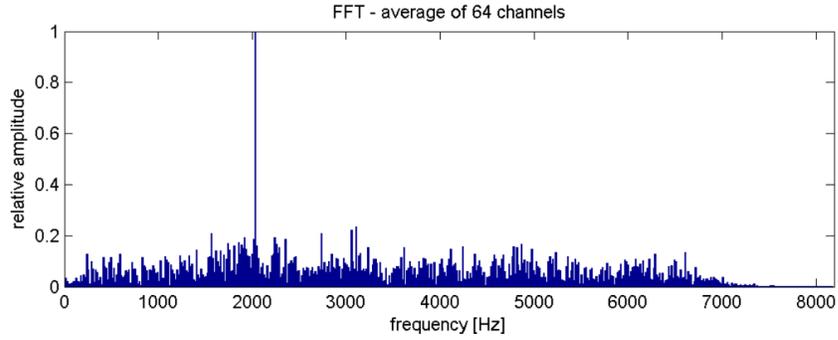


Figure A.15: Average of the 64 channel FFT, scaled relative to the maximum value at 2032 Hz.

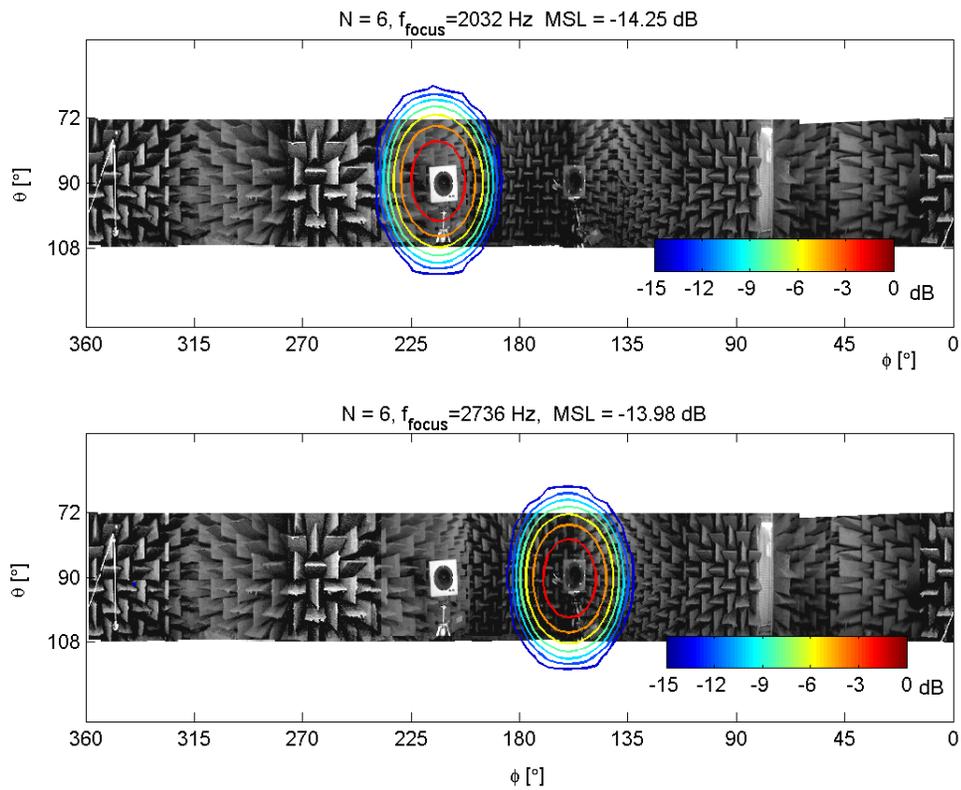


Figure A.16: Combination of contour plot and picture of surroundings. Top: Contour-plot based on FFT result at 2032 Hz. Bottom: based on FFT results at 2736 Hz.

A.10 Test H - Reflecting plate

Purpose: To investigate how a wave reflected on a hard plate affects the beamformer response.

Setup: As in figure A.1, but with an hard reflecting wooden plate placed as in figure A.18. The sound source is placed in the distance $r_0 = 1.00m$ and at the angular position $(\theta_0, \phi_0) = (78^\circ, 180^\circ)$. The reflecting plate is placed at around $(\theta_0, \phi_0) = (100^\circ, 50^\circ)$ at a distance of around $r_0 = 0.75m$. The plate is tilted around 5 degrees in order to reflect the sound wave originating from the sound source in such a way, so that it travels towards the center of the beamformer after the reflection.

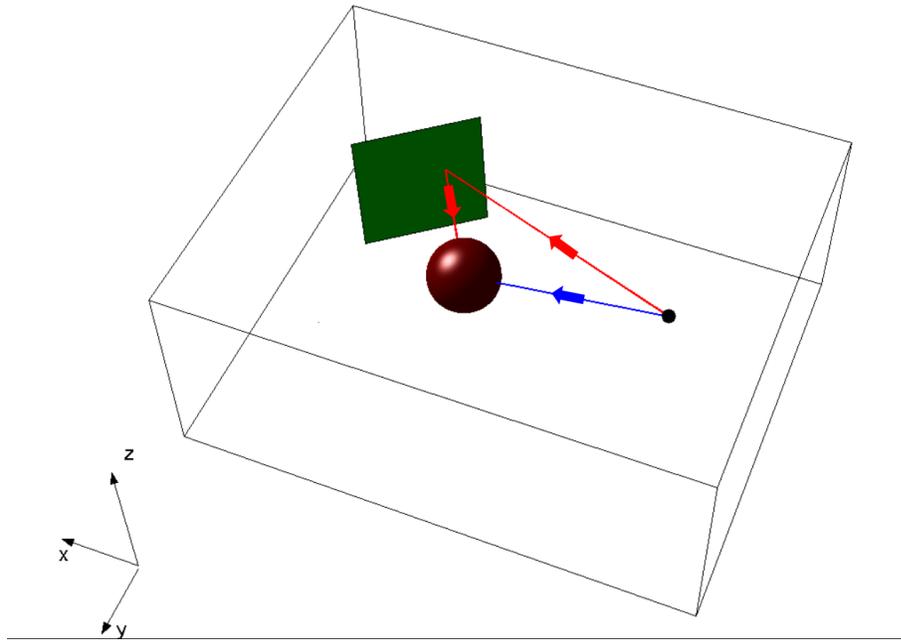


Figure A.17: View of room having the spherical beamformer at the center represented by the red sphere. The black dot represents the position of the sound source, and the green plate represents the reflecting wooden plate. The blue line represents the path of the wave travelling directly, from the sound source to the beamformer. The red line represents the path of the reflected wave.

Test conditions: The sound source signal is a 2192 Hz sinusoidal signal, sampled at $f_s = 16384Hz$ for 1 second. Eight pictures of the surroundings in the region from $\theta = 72^\circ$ to 108° and $\phi = 0^\circ$ to 360° , are taken in order to see how the beamformer response shows the expected peaks in the direction of the sound source and the reflecting plate.

Results: Figure A.18 shows the result of the spherical beamforming made from the measured data. The absolute value of the beamformer response shows a maximum in the direction of the sound source and a local peak in the direction of the reflecting plate, around 0.5 the value in the direction of the sound source.

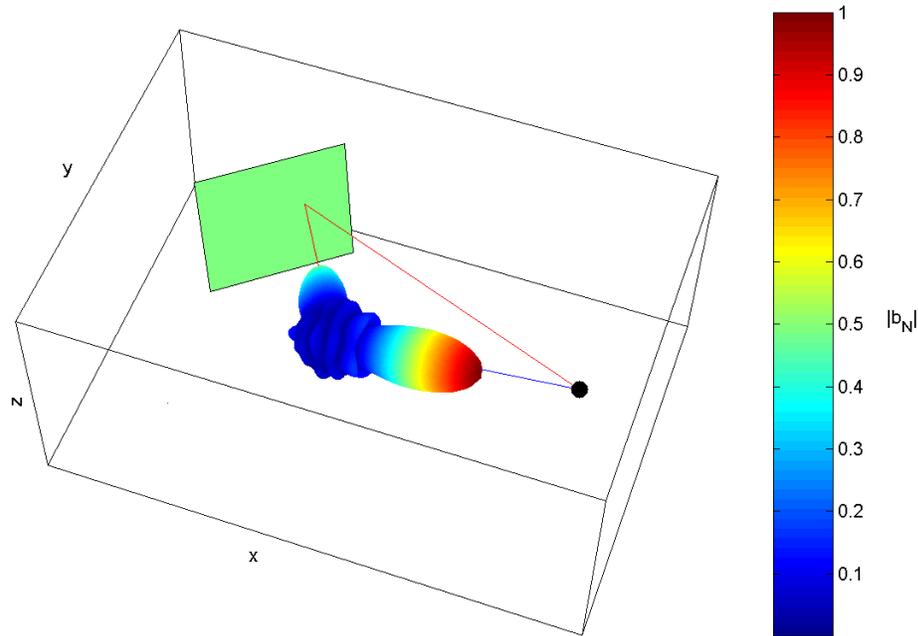


Figure A.18: Absolute value of the beamformer response, having max in the direction of the sound source (big black dot), and another big local max in the direction of the reflecting plate (green square).

In figure A.19 the beamformer response level, found as $20\log\left(\frac{|b(\theta,\phi)|}{|b(\theta,\phi)|_{max}}\right)$, is shown. $|b(\theta,\phi)|_{max}$ is the maximum value of the beamformer response and should be the

same as the direction of the main sound source. The beamformer response level is represented as a contour plot, where the contour lines represent the values defined on the colourbar ranging from -15 to 0 dB. The beamformer response level outside the dark blue contour line, is lower than -15 dB, relative to the maximum of the mainlobe. The contour plot is mixed with the picture of the surroundings, and shows to have maximum in the direction of the sound source. The reflected wave from the plate, is around -6 dB relative to the wave signal directly from the sound source. (see figure A.19).

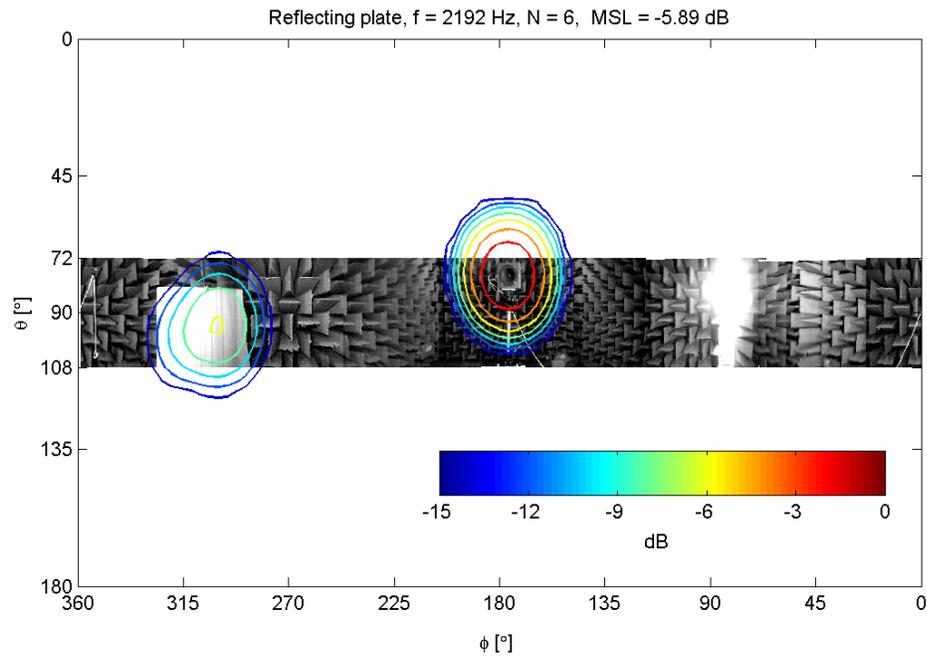


Figure A.19: Combination of contour plot of the beamformer response level, $20\log\left(\frac{|b(\theta,\phi)|}{|b(\theta,\phi)|_{max}}\right)$, and a picture of the surrounding environment around the equator.

Appendix B

Symbols, units and abbreviations

B.1 Symbols

a	Radius of sphere, [m]
c	Speed of sound, [m/s]
f	Frequency, [Hz]
f_s	Sampling frequency, [Hz]
i	Complex unit, $i^2 \equiv -1$
k	Wave number, [1/m]
\vec{k}	Wave number vector, [1/m]
n	Degree of spherical harmonics and order of Bessel functions
m	Order of spherical harmonics
q	Microphone number
r	Radius in spherical coordinates [m]
r_0	Radius to where sound source originates [m]
r_{foc}	Focusing radius [m]
t	time, [s]
\vec{u}	Particle velocity, [m/s]
w	Cubature weights

(x, y, z)	Cartesian coordinates, [m]
$z(\vec{\kappa}, t)$	DSB beamformer response, time domain
$\overline{\overline{G}}$	Reproducing kernel matrix
L	Number of samples
N	Maximum degree of spherical harmonics
Q	Number of microphones/points on sphere
$Z(\vec{\kappa}, \omega)$	DSB beamformer response, frequency domain
<hr/>	
h_n	Spherical Hankel function of order n
j_n	Spherical Bessel function of order n
n_n	Spherical Neumann function of order n
A_n^m	Decomposition constant
H_n	Cylindrical Hankel function of order n
P_n^m	Associated Legendre polynomials of degree n and order m
J_n	Cylindrical Bessel function of order n
N_n	Cylindrical Neumann function of order n
Y_n^m	Spherical harmonics of degree n and order m
<hr/>	
α_n^m	Decomposition constant
$\delta_{n,m}$	Kronecker delta
$\delta(x - x_0)$	Delta function
ϵ	Mean error of the beamformer response
γ	Random number
$\vec{\kappa}$	Steering direction unit vector
λ	Wavelength
$\nu(ka, N)$	Error function for decomposition constants
∇	Gradient
θ	Elevation in spherical coordinates, [°]
θ_0	Elevation of incoming wave field, [°]
ϕ	Azimuth in spherical coordinates, [°]
ϕ_0	Azimuth of incoming wave field, [°]
ρ_0	Density, [kg/m ³]

σ	Standard deviation
ω	Angular frequency, [rad]
$\Delta_q(\vec{\kappa})$	Time delay, [s]
Φ	Velocity potential
∞	Infinity

B.2 Units

dB	Decibel
s	Seconds
Hz	Hertz
kHz	Kilohertz
MHz	Megahertz
kg	Kilogram
cm	Centimeter
m	Meter
rad	Radians
°	Degree

B.3 Abbreviations

B&K	Brüel & Kjær
DFT	Discrete Fourier Transformation
DG	Directional Gain
DSB	Delay and Sum Beamforming
FFT	Fast Fourier Transformation
FWHM	Full Width Half Max
MSE	Mean Square Error
MSL	Maximum Sidelobe Level
SHB	Spherical Harmonics Beamforming

Bibliography

- [1] J. J. Bowman, T. B. A. Senior, and P. L. E. Uslenghi, *Electromagnetic and acoustic scattering by simple shapes*, Hemisphere Publishing Corporation, 1987.
- [2] James G. Ryan Bradford N. Govern and Michael R. Stinson, *Microphone array measurement system for analysis of directional and spatial variations of sound fields.*, J. Acoust. Soc. Am. **112**, no. 5.
- [3] M. Brandstein and D. Ward, *Microphone arrays*, Springer, 2001.
- [4] H. S. M. Coxeter, *Generators and relations for discrete groups*, Berlin, Springer, 1972.
- [5] P. Thushara D. Abhayapala, *Modal analysis and synthesis of broadband nearfield beamforming arrays*, Ph.D. thesis, The Australian National University, Telecommunication Engineering Group, Research School of Information Science and Engineering.
- [6] Diemer De Vries Edo Hulsebos and Emmanuelle Bourdillat, *Improved microphone array configurations for auralization of sound fields by wave-field synthesis*, J. Audio Eng. Soc. **50**, no. 10.
- [7] R.H. Hardin and N.J.A. Sloane, *McLaren's improved snub cube and other new spherical designs in three dimensions*, Discrete Computational Geometry **15** (1996), 429–441.

-
- [8] I.H. Sloan I.G. Graham, *Fully discrete spectral boundary integral methods on smooth closed surfaces in r^3* , Numer. Math. **92**, no. 1.
- [9] J. J. Christensen J. Hald, *A class of optimal broadband phased array geometries designed for easy construction*, Proc. Inter-Noise (2002).
- [10] Don H. Johnson and Dan E. Dudgeon, *Array signal processing*, Prentice Hall, 1993.
- [11] P. M. Juhl, *The boundary element method for sound field calculations*, Ph.D. thesis, The Acoustics Laboratory, Technical University of Denmark, 1993.
- [12] Sophocles J. Orfanidis, *Introduction to signal processing*, Prentice-Hall, 1996.
- [13] G. R. Baldock and T. Bridgeman, *The mathematical theory of wave motion*, Ellis Horwood Ltd., 1987.
- [14] Eugen Skudrzyk, *The foundation of acoustics*, Springer - Verlag, 1971.
- [15] Ian H. Sloan and Robert S. Womersley, *Extremal systems of points and numerical integration on the sphere*, Journal of Computational and Applied Mathematics **21**, no. 1.

Index

- Amplitude error, 48
- Angular frequency, 5, 8, 58, 63
- Array signal processing, 1
- Azimuth, 4
- Backlobe, 24
- Beamformer, 1
- Beamformer response, 19
- Beamforming, 1
- Beampattern, 19
- Bessel functions, 40
 - cylindrical, 8
 - spherical, 5, 7, 8
- Boundary condition, 10
- Cartesian coordinate system, 4
- Cubature rule, 30
 - quasi-Monte Carlo, 33
- Cubature weight, 30
- Decomposition constants, 6
- Delaunay triangulation, 56
- Delay and Sum Beamforming, 61
- Delta function, 20, 24
- Directional gain, 20
- Discrete Fourier Transformation, 58
- Elevation, 4
- Fast Fourier Transformation, 58
- Focusing radius, 22, 25
- Frequency resolution, 58
- Full Width Half Max, 20
- Hard sphere, 7
- Implementation, 57
- Inner product, 31
 - discrete, 32
- Kronecker delta, 6
- Legendre polynomials
 - associated, 6
- Mainlobe, 20
- Maximum Sidelobe Level, 20
- Neumann functions
 - cylindrical, 8
 - spherical, 8
- Omnidirectional, 23

- Optimization, 49
- Origo, 4
- Particle velocity, 5
- Phase error, 45
 - random normally distributed, 45
- Planner arrays, 1
- Reproducing kernel, 30
 - basis, 32
 - matrix, 30
- Resolution, 20
- Rotational symmetry axis, 35
- Sampling frequency, 58
- Sidelobes, 20
- Sound pressure, 5
- Spatial filtering, 1, 18
- Spherical coordinate system, 4
- Spherical coordinates, 4
- Spherical Hankel function, 9
- Spherical harmonics, 6
- Spherical Harmonics Beamforming, 18
- Spherical polynomials, 33, 34
- Spherical t-design, 33, 34
- Stability, 45
- Standard deviation, 45
- Steered response, 19
- Surface of sphere, 43
- Symmetry operators, 35
- Tetrahedra
 - structure, 35
- Time dependency, 5, 7
- Transparent sphere, 7
- Velocity potential, 5, 7
- Wave equation, 4