Sound Reproduction By Wave Field Synthesis

Group 861 - Spring 2004

Faculty of Engineering and Science
Aalborg University

Institute of electronic systems - department of acoustics

Faculty of Engineering and Science

Aalborg University

Institute of Electronic Systems - Department of Acoustics

TITLE:
Sound Reproduction by Wave Field Synthesis

TIME PERIOD:
8th semester
February 2nd to June 2nd, 2004

GROUP:
861 - 2004

GROUP MEMBERS:
Alfredo Fernández Franco
Sebastian Merchel
Lise Pesqueux
Mathieu Rouaud
Martin Ordell Sørensen

SUPERVISOR:
Rodrigo Ordóñez

COPIES: 8

NO OF PAGES:

- IN MAIN REPORT: 92

- IN APPENDIXES: 38

TOTAL: 130

Abstract:

In this project a Wave Field Synthesis (WFS) system using a linear array of loudspeakers is implemented. It is applied to the reproduction of stereo sound.

The background theory is described and the limitations of practical implementations are introduced. Simulations are carried out in Matlab to visualize the behaviour of a system based on WFS, with special attention to the artefacts: spatial aliasing (SA), truncation effects (TE). Different methods to reduce these artefacts are investigated and compared: increasing the directivity of the array loudspeakers and randomizing the phase of the high frequency content for the SA; applying a tapering window to the array and using a "feeding edge loudspeakers" technique for the TE. Amplitude errors (AE) are also simulated.

The WFS system system is implemented by a linear array of 24 loudspeakers constructed in an anechoic room. The driving signals are calculated by a MATLAB program and played back through 24 channels. Measurements are carried out and the results verify the simulations. They show: the AE, the incidence angle and the variations of the time delay over the listening area.

It is concluded that the system implemented synthesizes the sound field within a limited frequency range of notional sources that can be placed far away.

Det Teknisk-Naturvidenskabelige Fakultet

Aalborg Universitet

Institut for Elektroniske Systemer - Afdelingen for Akustik

TITEL:

Lydgengivelse ved brug af 'Wave Field Synthesis'

TIDSPERIODE:

8. semester

2. februar - 2. juni 2004

GRUPPE:

861 - 2004

GRUPPEMEDLEMMER:

Alfredo Fernández Franco

Sebastian Merchel

Lise Pesqueux

Mathieu Rouaud

Martin Ordell Sørensen

VEJLEDER:

Rodrigo Ordóñez

KOPIER: 8

ANTAL SIDER:

- HOVEDRAPPORT: 92

- APPENDIKS: 38

TOTAL: 130

Synopsis:

Denne rapport omhandler implementationen af et 'Wave Field Synthesis'-system (WFS) benyttet på en række af højttalere. Systemet bliver her benyttet til gengivelse af stereo signaler.

Baggrundsteorien er beskrevet med de omskrivninger der er nødvendige for at gøre det implementérbart og begrænsningerne er introduceret. Simuleringer er udført i Matlab for at visualisere resultatet fra et WFS-system, og der er arbejdet med følgende begrænsninger: spatial aliasering (SA) og trunkeringseffekter (TE). Forskellige metoder til at reducere disse er beskrevet og sammenlignet: Forøgelse af direktiviteten af de enkelte arrayhøjttalere og randomisering af høje frekvenser for SA; vinduesteknikker og 'Feeding Edge Loudspeaker' for TE. Amplitude fejl er også simuleret.

Systemet er opbygget med 24 højttalere på en række i et lyddødt rum og styresignalerne til højttalerne bliver genereret i et MATLAB program. Målingerne er udført og resultaterne underbygger simuleringerne. De viser amplitudefejlene, indgangsvinklen og tidsforsinkelsen over lytteområdet.

Det konkluderes at det implementerede system genererer det ønskede lydfelt i et begrænset frekvensområde fra en nominel kilde, der virtuelt kan placeres langt væk fra lytteren.

Preface

This report is written by group ACO-861 of the 8th semester of the international Master of Science (M.Sc.E.) in acoustics at Aalborg University. It constitutes the documentation of the project work done about Sound Reproduction by Wave Field Synthesis. The report is addressed to anyone with an interest in Wave Field Synthesis, but is primarily intended to students and supervisors of the E-Sector at Aalborg University.

References to literature are in squared brackets with the number that appears in the bibliography, e.g. [3].

Figure, table and equation numbering follows the chapter numbering, e.g. figure 2 of chapter 5 is called Figure 5.2. References to equations are in parenthesis, e.g. (3.11) for equation 11 in chapter 3.

The report is divided into two parts: Main report and Appendix.

Attached is a CD-ROM containing the Matlab source code, the input wave files used for the measurements, the sound files recorded during the measurements, and a copy of the report in PDF/PS format.

Group ACO-861, Aalborg University, 2nd June 2004

<div style="text-align:center">

Alfredo Fernandez Franco   Lise Pesqueux

Sebastian Merchel   Mathieu Rouaud

Martin Ordell Sørensen

</div>

Contents

Introduction

1

Looking back in history, the objective of an audio rendering system has always been to give a good acoustic experience to the listener. These systems are designed to produce sounds as real as possible, so that the listener does not notice that it has been produced by a reproduction system. Therefore, several systems based on different principles have been developed.

All known spatial sound systems are based on three fundamentally different methods or on a mixed form of these methods:

- Loudspeaker stereophony

- Reconstruction of the ear signals (binaural sound)

- Synthesis of the sound field around the listener

Although the loudspeaker stereophony is an industrial standard in today's world, investigations are being made, particularly in both last methods, to design new sound reproduction systems that are able to create more realistic acoustical scenes.

1.1   Wave Field Synthesis

In the late eighties, the concept of wave field synthesis was introduced by Berkhout [1]. He showed that sound control could be based on acoustic holography, which is also called wave field extrapolation. This is done in three steps: the acquisition of the wave field by a microphone array, the
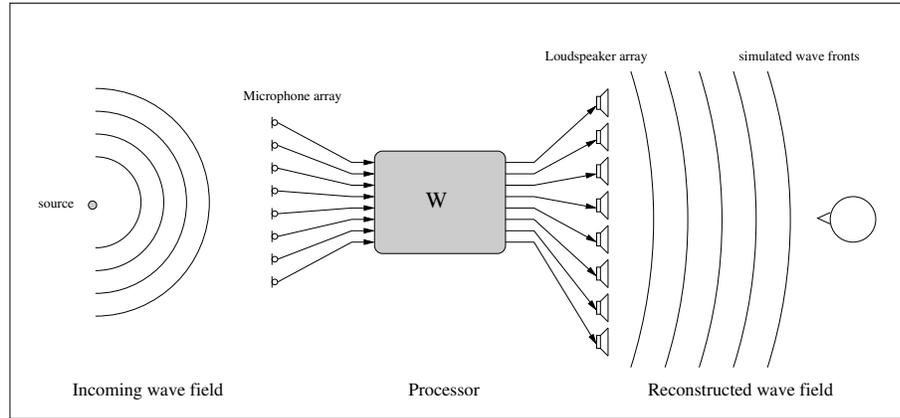
Figure 1.1: *Principle of acoustic holography: the recorded wave field is numerically extrapolated by a processor.*

extrapolation by a processor, and the reconstruction of the wave field by a loudspeaker array (Figure 1.1).

The underlying idea of wave field synthesis is to generate a sound field within a volume bounded by an array of loudspeakers. It is the presence of the transfer function between the microphones and the loudspeakers that enables the synthesis and the control of a sound field. Berkhout uses the so-called Kirchoff-Helmholtz integral (section 2.2.1), which is the mathematical quantization of the Huygens' principle (section 2.1). With such a system, it is possible to synthesize a wave field which is the accurate reproduction of the original wave field within a listening area. This remains valid if there are several sound sources. The general theory can be adapted to a two dimensional representation solution (horizontal plane). The realization of the wave field synthesis reproduction system becomes a linear array of loudspeakers, not a planar array.

1.2 The Stereophonic Principle

In the beginning of the fifties, a lot of research was based on the stereophonic reproduction. The stereophonic sound was vital for the motion-picture and sound-recording industries. In these times, the stereophonic systems used two or more channels. It was found that the effect of adding more than two channels did not produce results good enough to justify the additional

2

technical and economical effort [2]. Most research has been based on the most current stereophonic configuration: the two-channel stereophony.

The sound reproduction method using the stereophonic principle is based on the binaural principle. This enables to localize a sound source, due to the analysis done by the brain of the two different signals received by both ears. When the hears hear the sounds emitted by the two loudspeakers, the brain interprets those signals and determines the position of a phantom source situated between the loudspeakers. The position of the phantom source depends on the time delay, and on the amplitude difference between the two signals.

There are two main problems with the stereophonic reproduction. The first one is that the phantom source can only be placed between the loudspeakers (with a classical two-channel system). The second one, which is the main disadvantage, is the limited size of the listening area, called the "sweet spot". During the seventies, research was carried out to improve the spatial perception with the so called quadraphony using four channels, which adds possible positions for the phantom source. Recently, many new surround standards have been adopted for cinema use and video home theater. The surround standards used on DVDs are Dolby Digital and DTS, and they typically use 5 main channels and one subwoofer channel. These new techniques reduce the disadvantages caused by the restricted area in which the phantom source can be placed, however they do not enlarge the sweet spot.

1.3   Problem Statement

The aim of this project is to implement a working system based on wave field synthesis. Because the available sound signals are stereo signals, wave field synthesis will be applied to a stereophonic system. By using wave field synthesis, it should be possible to enlarge the "sweet spot". Wave fronts can be generated with notional sources. The idea is to simulate two virtual sources far behind the array in order to create plane waves. With plane waves, the propagation direction and the amplitude are equal everywhere in the listening area. Hence, the further the sources are placed from the listener, the less important the position of the listener will be, because there will be a stable stereo image within a listening area. Figure 1.2 shows a basic realization of the wave field synthesis where two notional sources are placed behind the array.

Two virtual sources

Loudspeaker array

30°
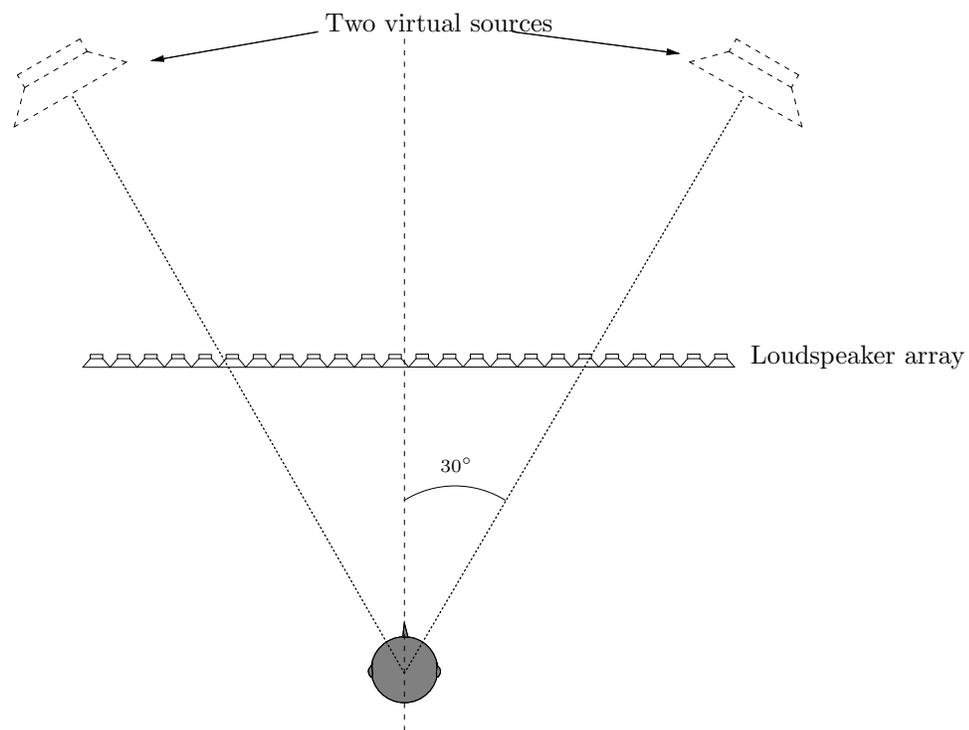
Figure 1.2: *Illustration of the setup for wave field synthesis.*

The system has to be able to reproduce ordinary stereo recordings, by synthesizing two notional sources placed behind the array at stereo position of 30° from the listener at different distances. The distance of the notional sources from the listener is a fundamental parameter. The shape of the wave fronts created by the two virtual sources within the listening area is directly related to this distance.

In order to create the desired wave field, the first task is to investigate the underlying theory of sound propagation and wave field synthesis. To predict the behaviour of the reproduction system, some simulations in Matlab will be carried out. The influence of the different parameters on the behaviour of the system will be studied. The last task is to implement a practical setup in order to validate the Matlab simulations. Measurements are carried out within a listening area to validate the assumptions done about the size of the "sweet spot".

1.4 Implementation of the System

In order to implement the system, some practical conditions have to be presented.

To simplify the model developed, the behaviour of the system has been studied in a free field environment, such as an anechoic room.

Theoretically, a very large number of loudspeakers have to be placed on a surface enclosing a volume. Due to size end equipment limits, the number of reproduction channels have been set to 24. This is done by using a 24-channel sound card to feed the 24 individual loudspeakers in a line array.

The array is neither infinite nor continuous. These practical conditions involve some artefacts which define reproduction limits for the system. the used loudspeakers have a diameter of $154mm$, the minimum distance between two sound sources is $154mm$. This limits the upper frequency which can be synthesized. The line array will have a length of $3.72m$. The lower frequency is limited by the frequency response of the ball loudspeakers. It is situated around $300Hz$.

Theory 2

This chapter describes how, from the Huygens' principle, a system synthesizing wave fields is created, and the consequences of the practical realization of such a system are presented.

The stereophonic principle is also presented with the main limitation of such a system: the size of the "sweet spot". Some considerations about radiation patterns will explain why the sweet spot can be enlarged if plane waves replace spherical waves.

2.1   From the Huygens' principle

In his "Traité de la lumière", in 1690, Huygens explained the diffraction of the waves as follows [3]:

If a wave emitted by a point source P_S with a frequency f is considered, at any time t, all the points on the wave front can be taken as point sources for the production of spherical secondary wavelets of the same frequency. Then, at the next instant, the wave front is the envelope of the secondary wavelets, as shown in Figure 2.1.

The pressure amplitude of the secondary sources is proportional to the pressure due to the original source at these points.

All the secondary wavelets are coherent, which means that they have all the same frequency and phase.
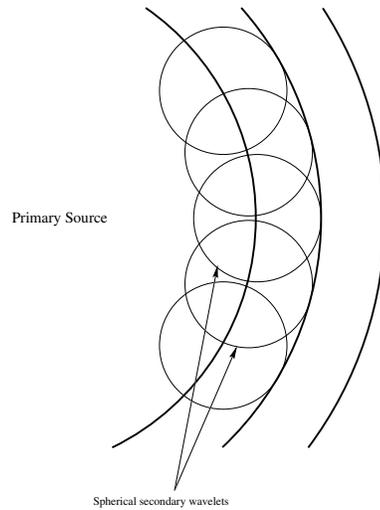
Figure 2.1: *Representation of the Huygens' principle.*

2.2   To the Wave Field Synthesis

Wave Field Synthesis is a direct application of the Huygens' principle. With a loudspeaker array, wavefronts within a volume can be synthesized. Each loudspeaker of the array acts as a secondary source for the production of a wavelet, and the whole system synthesize a sound field which could have been created by one or several sources situated behind the array.

Kirchhoff quantified the Huygens' principle and showed that the sound field inside a source-free volume V can be calculated if the pressure field due to the primary sources on the enclosing surface S is known (Section 2.2.1).

The loudspeakers are not positioned in a wave front of the imaginary source that is synthesized, so the surface used in the Kirchhoff-Helmholtz integral is degenerated to a plane (Section 2.2.2), then to a line. The driving signals of the louspeakers (Section 2.3.1) have to be weightened and delayed.

2.2.1   The Kirchhoff-Helmholtz integral

According to the Huygens' principle, it is possible to calculate the pressure field created by one or several sources, inside a source-free volume V enclosed by a surface S, by considering a distribution of monopole and dipole sources

on the surface for the production of secondary wavelets. The envelope of
these wavelets reproduces the primary sound field. The Fourier transform of
the sound pressure at a listening position L inside V is given by the Kirchhoff-
Helmholtz integral (2.1)[3]:

$$P(\vec{r},\omega) = \frac{1}{4\pi} \oint_S \left[\underbrace{P(\vec{r_S},\omega)\frac{\partial}{\partial n}\left(\frac{e^{-\jmath k|\vec{r}-\vec{r_S}|}}{|\vec{r}-\vec{r_S}|}\right)}_{dipoles} - \underbrace{\frac{\partial P(\vec{r_S},\omega)}{\partial n}\frac{e^{-\jmath k|\vec{r}-\vec{r_S}|}}{|\vec{r}-\vec{r_S}|}}_{monopoles} \right] dS \quad (2.1)$$

The geometry used in this integral is described in Figure 2.2.

k is the wave number defined as ω/c, where ω is the angular frequency of the
wave and c is the sound propagation velocity; r defines the position of the
listening point inside V; and $P(r_S,\omega)$ is the Fourier transform of the pressure
distribution on S.
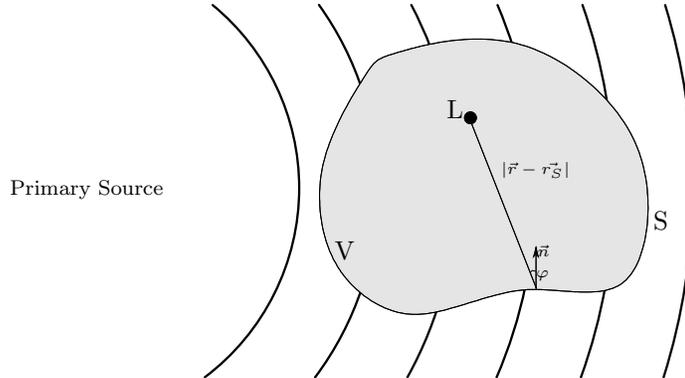
Figure 2.2: *Geometry used for the Kirchhoff-Helmholtz integral.*

In other words, if the sound pressure and its gradient on a surface S enclosing
a source-free volume V are known, then the sound pressure at any point L
inside this volume can be calculated. This integral also means that with an
infinite number of sources, any sound field can be created inside a volume.

Derivation of the integral

The Kirchhoff-Helmoltz integral can be derived by using the wave equation
and the Green's theorem as follows:

8

If u and v are two functions having continuous first and second partial derivatives on a surface S enclosing a volume V, then the Green's integral theorem states that [4]:

$$\int_V (u\nabla^2 v - v\nabla^2 u)dV \quad = \quad \oint_S (u\nabla v - v\nabla u) \cdot \vec{n} \; dS \qquad (2.2)$$

The Fourier transform of the wave equation is:

$$\nabla^2 P + k^2 P \quad = \quad 0 \qquad (2.3)$$

If P_1 and P_2 are the Fourier transforms of two pressure fields, then:

$$P_1\nabla^2 P_2 - P_2\nabla^2 P_1 = P_1(-k^2 P_2) - P_2(-k^2 P_1) = 0$$

which leads to:

$$\oint_S (P_1\nabla P_2 - P_2\nabla P_1) \cdot \vec{n} \; dS \quad = \quad 0 \qquad (2.4)$$

If P_1 is the primary pressure field created by the sources outside V and P_2 is the specific pressure field created by a point source Q inside V, then the surface S is redrawn to exclude Q. Q is now surrounded by a sphere S' of radius ε, as shown in Figure 2.3.
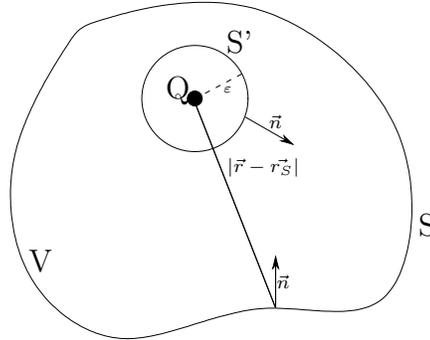
Figure 2.3: *The modified region of integration.*

Note that $\quad P_2 = A\frac{e^{-\jmath k d}}{d}$ with $d = |\vec{r} - \vec{r_S}|$.

Equation (2.4) becomes:

9

$$\oint_{S+S'} \left[\frac{e^{-\jmath kd}}{d} \frac{\partial P_1}{\partial n} - P_1 \frac{\partial}{\partial n} \left(\frac{e^{-\jmath kd}}{d} \right) \right] dS = 0$$

$$\Leftrightarrow \oint_{S'} \left[\frac{e^{-\jmath kd}}{d} \frac{\partial P_1}{\partial n} - P_1 \frac{\partial}{\partial n} \left(\frac{e^{-\jmath kd}}{d} \right) \right] dS' = - \oint_{S} \left[\frac{e^{-\jmath kd}}{d} \frac{\partial P_1}{\partial n} - P_1 \frac{\partial}{\partial n} \left(\frac{e^{-\jmath kd}}{d} \right) \right] dS$$

On S', $d = \varepsilon$, $dS' = \varepsilon^2 d\Omega$ where Ω is the solid angle[1], and $\frac{\partial}{\partial n} = \frac{\partial}{\partial d}$. The equation above becomes:

$$\int_0^{4\pi} \left[\frac{e^{-\jmath k\varepsilon}}{\varepsilon} \frac{\partial P_1}{\partial d} + P_1 \frac{e^{-\jmath k\varepsilon}}{\varepsilon} \left(\jmath k + \frac{1}{\varepsilon} \right) \right] \varepsilon^2 d\Omega = - \oint_{S} \left[\frac{e^{-\jmath kd}}{d} \frac{\partial P_1}{\partial n} - P_1 \frac{\partial}{\partial n} \left(\frac{e^{-\jmath kd}}{d} \right) \right] dS$$

Taking the limit when ε tends to 0 leads to:

$$\int_0^{4\pi} P_1(Q) \, d\Omega = 4\pi P_1(Q) = - \oint_{S} \left[\frac{e^{-\jmath kd}}{d} \frac{\partial P_1}{\partial n} - P_1 \frac{\partial}{\partial n} \left(\frac{e^{-\jmath kd}}{d} \right) \right] dS$$

And the Kirchhoff-Helmholtz integral (2.1) is obtained:

$$P(\vec{r}, \omega) = \frac{1}{4\pi} \oint_{S} \left[P(\vec{r_S}, \omega) \frac{\partial}{\partial n} \left(\frac{e^{-\jmath k|\vec{r}-\vec{r_S}|}}{|\vec{r} - \vec{r_S}|} \right) - \frac{\partial P(\vec{r_S}, \omega)}{\partial n} \frac{e^{-\jmath k|\vec{r}-\vec{r_S}|}}{|\vec{r} - \vec{r_S}|} \right] dS$$

It can also be written as follows:

$$P(\vec{r}, \omega) = \frac{1}{4\pi} \oint_{S} \left(\jmath \omega \rho_0 V_n(\vec{r_S}, \omega) \frac{e^{-\jmath kd}}{d} + P(\vec{r_S}, \omega) \frac{1 + \jmath kd}{d} \cos\varphi \frac{e^{-\jmath kd}}{d} \right) dS \quad (2.5)$$

where ρ_0 is the air density and V_n is the particle velocity in the direction of $\vec{n}$.

---

[1]A solid angle is an angle formed by three or more planes intersecting at a common point.

More specificly, the solid angle Ω subtended by a surface S is defined as the surface area of a unit sphere covered by the surface's projection onto the sphere. Its unit is the steradian. [5]

2.2.2   The Rayleigh's representation theorem

The Kirchhoff-Helmholtz integral shows that by setting the correct pressure distribution $P(\vec{r_S}, \omega)$ and its gradient on a surface S, a sound field in the volume enclosed within this surface can be created.

To engineer a realizable system, the surface S has to be degenerated to a plane $z = z_1$ separating the source area from the listening area, as shown in Figure 2.4. The Kirchhoff-Helmholtz integral (2.1) can be simplified into the Rayleigh I integral for monopoles(2.6) and into the Rayleigh II integral for dipoles (2.7)[6].
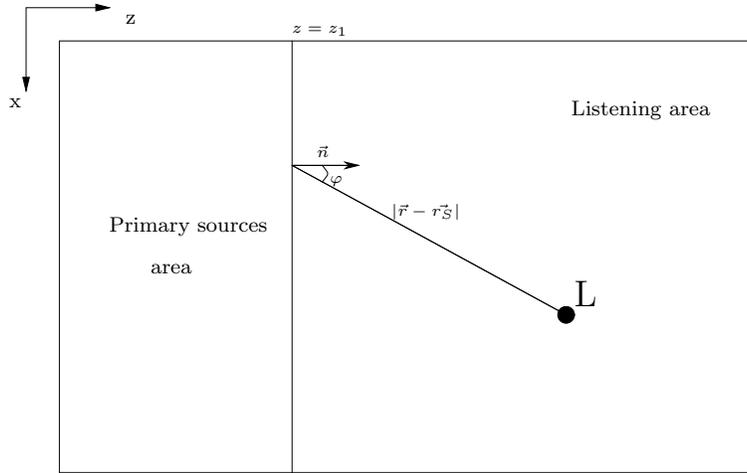
Figure 2.4: *Geometry for the Rayleigh's representation theorem.*

$$P(\vec{r}, \omega) = \rho_0 c \frac{\jmath k}{2\pi} \iint_S \left(V_n(\vec{r_S}, \omega) \frac{e^{-\jmath k|\vec{r}-\vec{r_S}|}}{|\vec{r}-\vec{r_S}|} \right) dS \qquad (2.6)$$

$$P(\vec{r}, \omega) = \frac{\jmath k}{2\pi} \iint_S \left(P(\vec{r_S}, \omega) \frac{1 + \jmath k|\vec{r}-\vec{r_S}|}{\jmath k|\vec{r}-\vec{r_S}|} cos\varphi \frac{e^{-\jmath k|\vec{r}-\vec{r_S}|}}{|\vec{r}-\vec{r_S}|} \right) dS \qquad (2.7)$$

where ρ_0 denotes the air density, c the velocity of sound in air, k the wave number, and V_n the particle velocity in the direction of $\vec{n}$.

11

2.3 Adaptation to practical realization

Discretization

Until now, a continuous distribution of sources on the surface was considered. In reality, the sources in the plane are loudspeakers, so the distribution is discrete.

This leads to the discrete form of the Rayleigh's integrals [6].

For Rayleigh I:

$$P(\vec{r}, \omega) = \frac{\jmath \omega \rho_0}{2\pi} \sum_{n=1}^{\infty} V_n(\vec{r_n}, \omega) \frac{e^{-\jmath k |\vec{r} - \vec{r_n}|}}{|\vec{r} - \vec{r_n}|} \Delta x \Delta y \tag{2.8}$$

And for Rayleigh II:

$$P(\vec{r}, \omega) = \frac{1}{2\pi} \sum_{n=1}^{\infty} P(\vec{r_n}, \omega) \frac{1 + \jmath k |\vec{r} - \vec{r_n}|}{|\vec{r} - \vec{r_n}|} cos\varphi_n \frac{e^{-\jmath k |\vec{r} - \vec{r_n}|}}{|\vec{r} - \vec{r_n}|} \Delta x \Delta y \tag{2.9}$$

From now on, the calculations will be based on the Rayleigh I integral.

Reduction to a line

For practical reasons, the surface is reduced to a line, and the listener is assumed to be in a plane $y = y_1$. Reducing the planar array to a line does not affect the shape of the wave fronts in the xz plane, as shown in Figure 2.5, and only the shape of the wave front in the horizontal ear plane actually affects the perception of sound.

The discrete form of the Rayleigh I integral can be transformed into (2.10):

$$P(\vec{r}, \omega) = \frac{\jmath \rho_0 \omega}{2\pi} \sum_{n=1}^{\infty} \left[V_n(\vec{r_n}, \omega) \frac{e^{-\jmath k |\vec{r} - \vec{r_n}|}}{|\vec{r} - \vec{r_n}|} \right] \Delta x \tag{2.10}$$
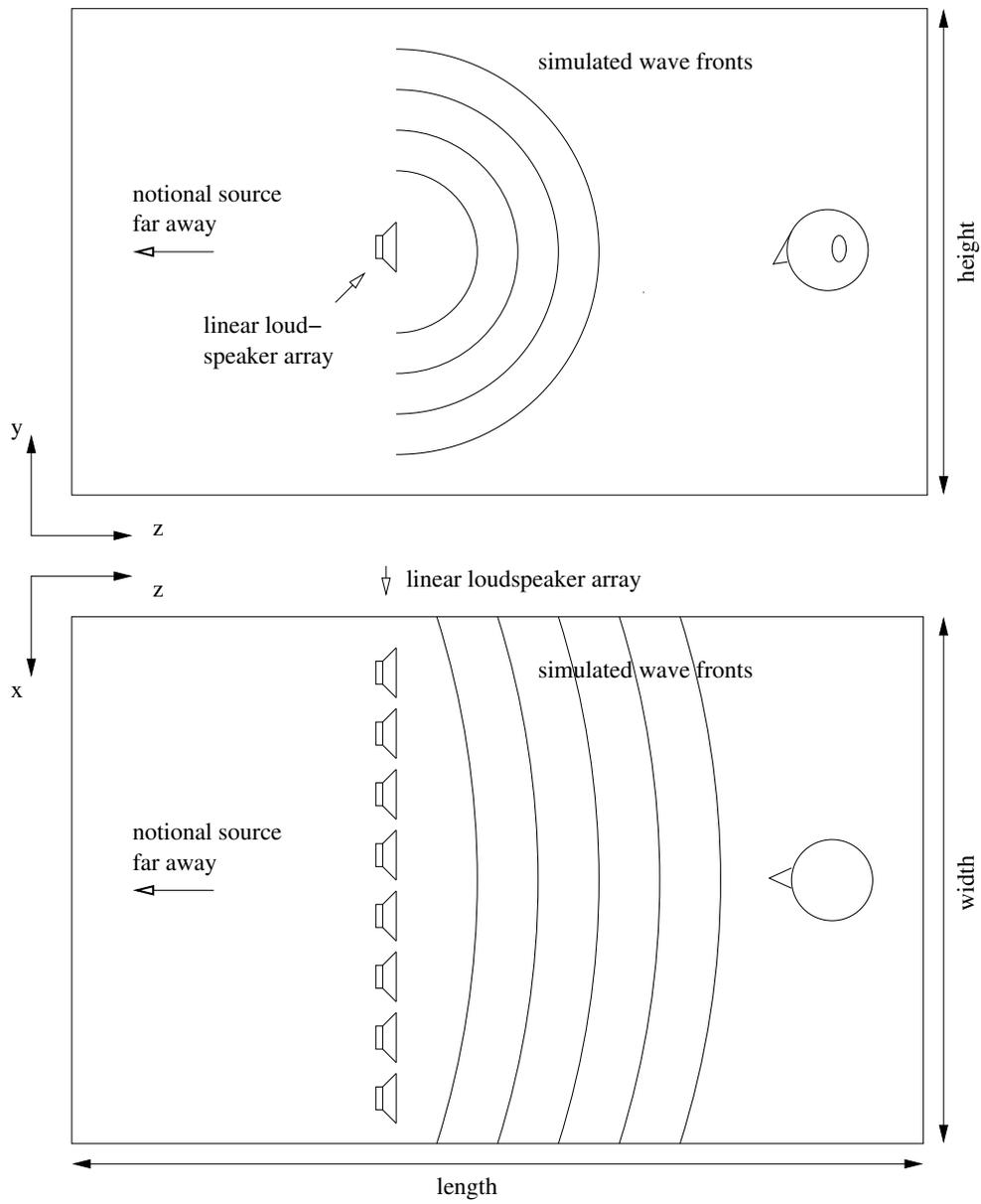
Figure 2.5: *Vertical and horizontal view of the simulated sound field.*

13

2.3.1 Calculation of the driving functions

The sound pressure $P(\vec{r_n}, \omega)$ is linked to the particle velocity $V_n(\vec{r_n}, \omega)$ through the specific acoustic impedance Z [4] as follows:

$$Z(\vec{r}, \omega) = \frac{P(\vec{r}, \omega)}{V(\vec{r}, \omega)} \tag{2.11}$$

For a spherical wave, the specific acoustic impedence is given by [4][p. 128]:

$$Z(\vec{r}, \omega) = \frac{\rho_0 c}{1 + \frac{1}{\jmath k r}} \qquad r \neq 0 \tag{2.12}$$

r being the distance to the point source.

For a pulsating sphere of average radius a and of angular frequency ω, the radial component of the velocity of the fluid in contact with the sphere is calculated using the specific acoustic impedence for a spherical wave evaluated at $r = a$ [4][p. 171].

$$V_n(\vec{r_n}, \omega) = \frac{P(\vec{r_n}, \omega)}{\rho_0 c} \left(1 + \frac{1}{\jmath k a}\right) \tag{2.13}$$

For a discrete distribution of pulsating spheres of average radius a, equation (2.10) becomes:

$$P(\vec{r}, \omega) = \left(\frac{\jmath k}{2\pi} + \frac{1}{2\pi a}\right) \sum_{n=1}^{\infty} \left[P(\vec{r_n}, \omega) \frac{e^{-\jmath k |\vec{r} - \vec{r_n}|}}{|\vec{r} - \vec{r_n}|} \right] \Delta x \tag{2.14}$$

In a practical realization, the sources are loudspeakers with a certain directivity (Appendix D) instead of being ideal pulsating spheres. The pressure has to be weighted by a factor which depends on the directivity $G(\varphi_n, \omega)$. The pressure at each loudspeaker has to be weighted with a factor $A_n(\vec{r_n}, \omega)$

14

to take into account the fact that the sound source is no longer omnidirectional.

Hence, the discrete form of the one-dimension Rayleigh I integral (2.10) can be written as follows:

$$P(\vec{r}, \omega) \;=\; \sum_{n=1}^{\infty} \left[A_n(\vec{r_n}, \omega) P(\vec{r_n}, \omega) G(\varphi_n, \omega) \frac{e^{-jk|\vec{r}-\vec{r_n}|}}{|\vec{r}-\vec{r_n}|} \right] \Delta x \quad (2.15)$$

The geometry used is described in Figure 2.6, adapted from [6].
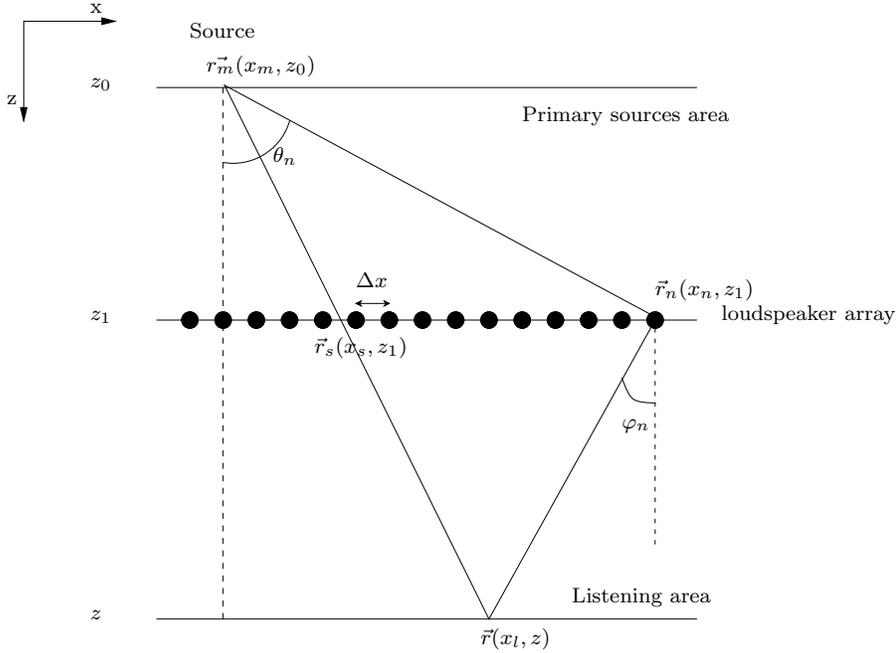
Figure 2.6: *Geometry for the calculation of the driving functions.*

The driving function for the nth loudspeaker is:

$$Q(\vec{r_n}, \omega) \;=\; A_n(\vec{r_n}, \omega) P(\vec{r_n}, \omega) \quad (2.16)$$

The driving functions for the loudspeakers are derived from the synthesized sound field produced by the notional sources.

At a position $\vec{r}$, the pressure field produced by a notional source at the position $\vec{r_m}$ with a spectrum $S(\omega)$ is:

$$P(\vec{r}, \omega) \quad = \quad S(\omega) \frac{e^{-\jmath k |\vec{r} - \vec{r_m}|}}{|\vec{r} - \vec{r_m}|} \tag{2.17}$$

On the array, equation (2.16) becomes:

$$Q(\vec{r_n}, \omega) \quad = \quad A_n(\vec{r_n}, \omega) S(\omega) \frac{e^{-\jmath k |\vec{r_n} - \vec{r_m}|}}{|\vec{r_n} - \vec{r_m}|} \tag{2.18}$$

Given the pressure field of the notional source at a listening position $\vec{r}$, equation (2.15) becomes:

$$S(\omega) \frac{e^{-\jmath k |\vec{r} - \vec{r_m}|}}{|\vec{r} - \vec{r_m}|} \quad = \quad \sum_{n=1}^{N} \left[Q(\vec{r_n}, \omega) G(\varphi_n, \omega) \frac{e^{-\jmath k |\vec{r} - \vec{r_n}|}}{|\vec{r} - \vec{r_n}|} \right] \Delta x$$

or, replacing Q by (2.15) and cancelling out $S(\omega)$:

$$\frac{e^{-\jmath k |\vec{r} - \vec{r_m}|}}{|\vec{r} - \vec{r_m}|} \quad = \quad \sum_{n=1}^{N} \left[A_n(\vec{r_n}, \omega) \frac{e^{-\jmath k |\vec{r_n} - \vec{r_m}|}}{|\vec{r_n} - \vec{r_m}|} G(\varphi_n, \omega) \frac{e^{-\jmath k |\vec{r} - \vec{r_n}|}}{|\vec{r} - \vec{r_n}|} \right] \Delta x \tag{2.19}$$

Using a mathematical method called the stationary-phase approximation[2], the driving function can be calculated. After substantial mathematical manipulations it is found that the driving function can be described by:

---

[2]"Stationary-phase approximation physically means that the wavefront is synthesized by all loudspeakers of the array together, but that a dominant contribution is given by the loudspeaker positioned at the point of stationary phase." [6].

In Figure 2.6, the point of stationary phase is $\vec{r_s}(x_s, z_0)$, and at this specific position, θ_n and φ_n are equal. The wave emitted by this loudspeaker arrives first at the listener position $\vec{r}$. The derivation of the driving function based on (2.19) is totally described in [6] and in Appendix A.

$$Q(\vec{r}_n, \omega) \;=\; S(\omega)\frac{\cos(\theta_n)}{G_n(\theta_n, \omega)}\sqrt{\frac{jk}{2\pi}}\sqrt{\frac{|z - z_1|}{|z - z_0|}}\frac{e^{-jk|\vec{r}_n - \vec{r}_m|}}{\sqrt{|\vec{r}_n - \vec{r}_m|}} \qquad (2.20)$$

The geometry used is described in Figure 2.6.

Interpretation

The driving function $Q(\vec{r}_n, \omega)$ is the sound pressure produced by the notional source at the position of the nth loudspeaker, weighted. This signal with which each loudpeaker is fed is then a filtered version of the original notional source signal.

The driving function's terms can be interpreted as follows:

The factor $e^{-jk|\vec{r}_n - \vec{r}_m|}$ describes the propagation time from the notional source to the nth loudspeaker.

The amplitude factor $\frac{1}{\sqrt{|\vec{r}_n - \vec{r}_m|}}$ is the dispersion factor of a cylindrical wave, hence the notional source can be considered as a line source.

This source obtains a specific directivity, which is inversely proportional to the directivity of one loudspeaker. It is assumed that all the loudspeakers have the same directivity. Since the driving signals are processed separately for each loudspeaker, it should be possible to compensate for their directivity individually.

The factor $\sqrt{\frac{|z - z_1|}{|z - z_0|}}$, weighting the amplitude, depends on the distance between virtual source, loudspeaker array and listener position. It does not change much within the listening area

Equation (2.20) can be written as:

$$Q(\vec{r}_n, \omega) \;=\; S(\omega)C(z, z_0, z_1)G_{line}(\theta_n, \omega)\frac{e^{-jk|\vec{r}_n - \vec{r}_m|}}{\sqrt{|\vec{r}_n - \vec{r}_m|}}$$

17

where

$$C(z, z_0, z_1) \;\; = \;\; \sqrt{\frac{|z - z_1|}{|z - z_0|}} \tag{2.21}$$

and

$$G_{line}(\theta_n, \omega) \;\; = \;\; \frac{\cos(\theta_n)}{G(\theta_n, \omega)} \sqrt{\frac{\jmath k}{2\pi}} \tag{2.22}$$

2.3.2 Artefacts

Spatial Aliasing

In reality it is not possible to use a continuous source to reproduce the ideal wave field. The fact of using a loudspeaker array with discrete sources produces spacial aliasing which occurs above a certain frequency. This frequency is called the spatial aliasing frequency and will be further on called $\mathbf{f_{alias}}$.

As spatial aliasing effects both the recording side and the reproduction side of the wave field synthesis system, it can be calculated for both separately. The frequency which is the lowest gives the overall frequency constraint for the system.

In analogy to sampling in time, the Shannon criterion can be modified for sampling in space[3]. It can be calculated by using the geometry in Figure 2.7 and gives, for the "worst case".

$$\mathbf{f_{alias}} = \frac{c}{2 \, \triangle \, x_{max} \sin(\alpha_{max})}$$

[3]There are destructive interferences when the path difference is equal to $\lambda/2$.

The signal is sampled along the loudspeakerarray where $\triangle x_{max}$ is the distance between two adjacent loudspeakers and the path difference is $\Delta x \sin(\alpha)$. The Shannon criterion is then transposed for spatial sampling.
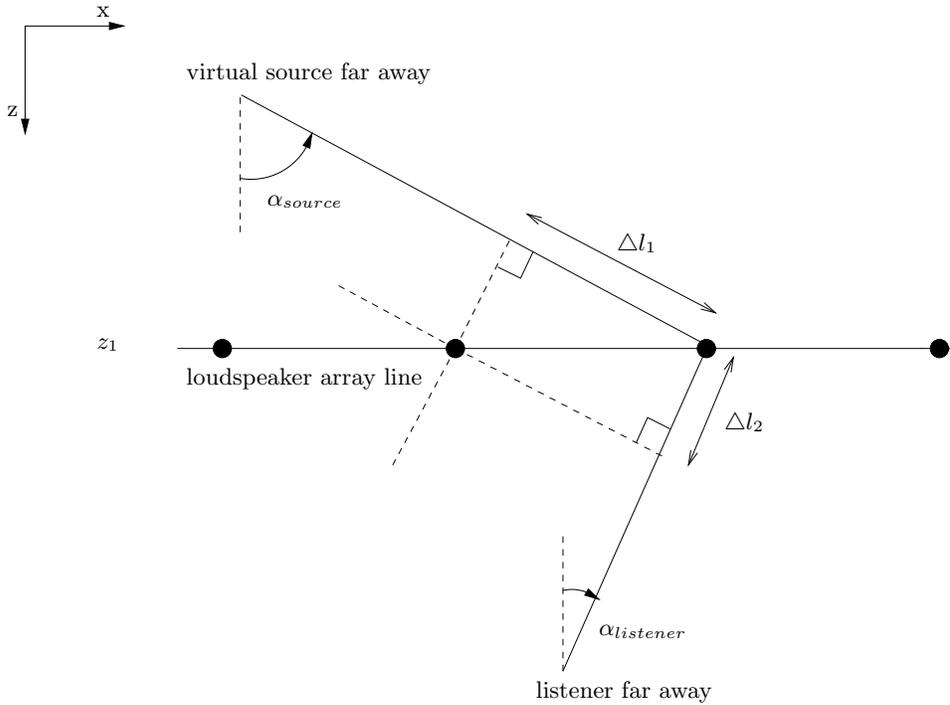
Figure 2.7: *Path difference for the recording side and for the listening side.*

α_{max} is either the maximum angle from a loudspeaker to a notional source $\alpha_{source,max}$ or to any point in the listening area $\alpha_{listener,max}$. For a stereo setup with sources positionned far away, the incidence angle on the recording side is around $30°$ for all the loudspeakers. For the setup described in Figure 4.1, the maximum angle from the listening area to a loudspeaker is $77,6°$. That gives for $\mathbf{f_{alias}}$ $2212,9$ Hz on the recording side and $1132,7$ Hz on the reproduction side. The overall $\mathbf{f_{alias}}$ equals then $1132,7$ Hz. This frequency only changes, if the length of array or size and position of the listening area are changed, or if $\alpha_{source,max}$ becomes bigger than $\alpha_{listener,max}$.

As the perceptual consequence of spatial aliasing is not yet clear, different methods are proposed to minimize this effect. Some simulations can be seen in Section 3.3.

"Spatial Bandwidth Reduction" is proposed in [7] and [8]. It means that the directivity of the notional source is increased, so that interference between the loudspeakers is reduced. Another method is the usage of more directive loudspeakers in the array. In [7], D. de Vries proposes the usage of subarrays that create these directive sources.

19

In [8] the method of "Randomisation of high-frequency content" is proposed. The audible periodicity of the aliasing effect is suppressed by randomizing the time offset for high frequencies between the loudspeakers in the array.

At last method, "OPSI-Phantom source reproduction of high-frequency content", is proposed in [9], meaning that high frequencies will be reproduced with fewer loudspeakers.

There are certain restrictions to the correct reproduction of frequencies due to the physical loudspeaker setup, which cannot be avoided.

Amplitude errors

Reducing the plane containing the distribution of sound sources for the production of the synthesized sound field to a line has two consequences.

The first one is that only virtual sources situated in the horizontal plane can be synthesized.

The second consequence is that amplitude errors occur due to the fact that the waves synthesized are no longer plane or spherical, but have a cylindrical component. The amplitude roll-off increases (3dB/doubling of distance).

Truncation effects

In theory, the line array is infinite, but it is finite in practice, and a truncation effect appears.

The finite array can be compared with a window, through which the virtual source is visible, or invisible to the listener. This analogy with the propagation of light was given in [10]. It is mentioned an area "illuminated" by the virtual source and a corresponding "shadow" area, where the virtual source cannot be perceived. This is illustrated in Figure 2.8. Using this analogy truncation can be visualised as a result of diffraction effects at the edges of the loudspeaker array, where diffraction waves are emitted. These waves occur as echoes close to the original wave front. A simulation is done in chapter 3.
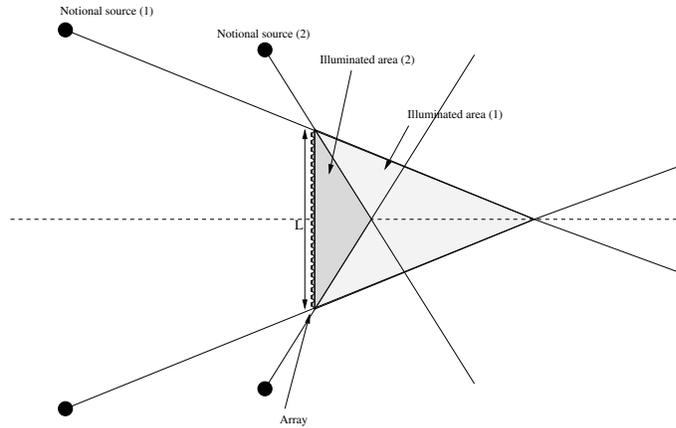
Figure 2.8: *Area where the virtual source can be heard.*

According to A.J. Berkhout, in [11], it can be shown, that "the delay corresponds to the difference in travel time between the direct transmission path and the transmission path via the outer array elements". It is further proposed to use "a weighting function over the array elements" to "suppress this effect effectively". This applied function, which is called tapering-window is applied to the array signals, fades the signal amplitude out near the edges of the array. The disadvantage is, according to [10] a reduction of the listenig area.

According to [12] the diffraction waves coming from the edges of the loudspeakerarray, "can be interpreted as scaled point sources with a specific directivity pattern radiating" in reference to a fixed point in the listening area. Feeding edge loudspeaker is a technique for "the 'missing contribution' of the non-existing loudspeakers outside the array" should be fed "to the edge loudspeakers" according to a proposal by C.P.A. Wapenaar in[13]. Therefore truncation effects can be compensated at the reference point.

2.4 The stereophonic principle

2.4.1 The binaural principle

Having two ears located on either side of the head allows to localize precisely the direction of sound sources. The brain uses the fact that, if the source

is not placed in front of the listener, both ears do not perceive exactly the same sound. A sound which comes from the left reaches the left ear before the right one. It is this small time and amplitude difference which permit to perceive the direction of the sound.

2.4.2 The perception of the stereophony

The stereophonic principle is based on wave interference of two signals: the left and the right signals. With this information, if certain conditions are met, the listener can mentally reconstitute the stereophonic "image". The phantom source is the imaginary source that could have created the sound perceived, and is situated between both loudspeakers. Its position depends on the amplitude and time difference between the loudspeaker signals. If the time difference increases progressively between the two signals, the phantom sources will move towards the loudspeaker which emits the earliest sound. This technique is called the time panning.

The main problem of the stereophonic reproduction is the limited size of the listening area. The listener get a good stereo imaging, if he is ideally placed on a symmetrical axis in regard to the loudspeakers. If he is placed slightly more to the left or the right, there is a time difference, thus compromising the reconstitution of the stereophonic image.

2.4.3 Stereophonic reproduction

In order to reproduce a stereophonic recording which renders not only the sound but also the acoustic spaciousness, it is important to establish specific listening conditions.

The position of the loudspeakers is an important factor. Images cannot be obtained if the angle between the loudspeakers at the listeners position exceeds approximately 60°. Over this angle, stereo systems fail to yield a correct stereo imaging between the two loudspeakers. This phenomenon is called "hole in the middle". It is commonly the result of excessive spacing between the loudspeakers. A large frontal sound scene is needed in order to reproduce a large panning aperture. A compromise is done between the necessity to have a large frontal scene, and the limitation caused by the

"hole in the middle". The point which forms an equilateral triangle with both loudspeakers can be defined as the stereophonic listening point. This is called the "sweet spot". Figure 2.9 shows the two channel stereophonic configuration.
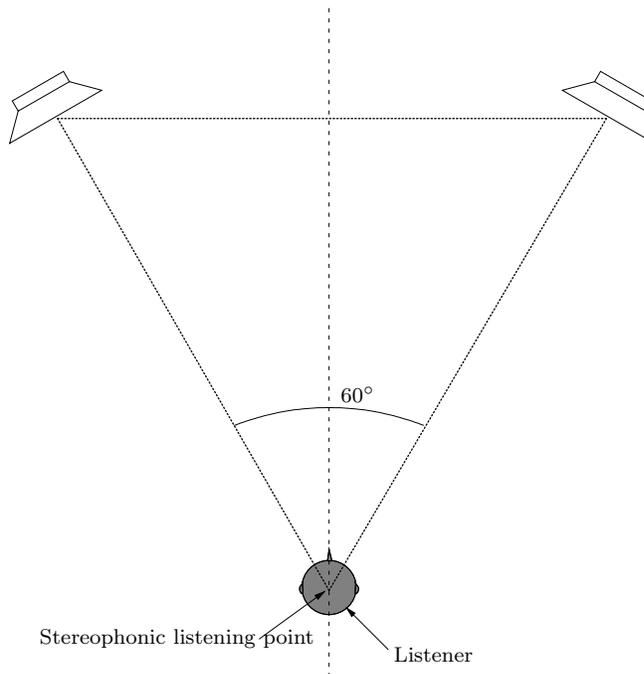
Figure 2.9: *Stereophonic configuration*

2.5 Sound Field consideration

The aim of this project is to enlarge the size of the sweet spot by using wave field synthesis. The idea is to have plane waves coming from the stereo positions, instead of spherical waves. Notional sources far away are synthesized. These sources are considered as two point sources. Some theoretical elements about the radiation of a point source have to be explored in order to understand how plane waves can enlarge the sweet spot.

23

2.5.1 Point source

The sound pressure p at a position $\vec{r}$ created by a point source located at $\vec{r_0}$ is equal to [4]:

$$p(\vec{r}, t) = \frac{A}{|\vec{r} - \vec{r_0}|} e^{\jmath(\omega t - k|\vec{r} - \vec{r_0}|)} \tag{2.23}$$
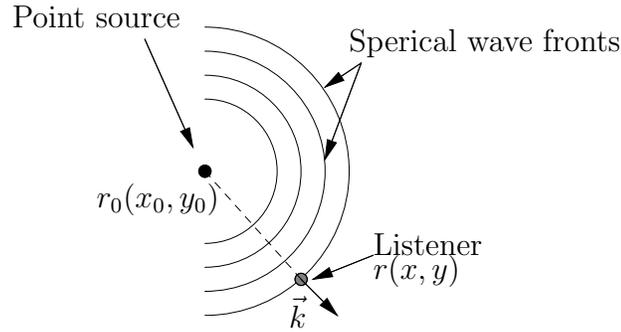
Figure 2.10: *Spherical waves radiated by a point source. The sound pressure is equal on arcs.*

The value of the amplitude is equal to $\frac{A}{|\vec{r}-\vec{r_0}|}$, where A is a constant. The amplitude is inversely proportional to the distance between the point source and the listener. $\vec{k}$ represents the wave vector which is perpendicular to the wave front and defines the wave propagation direction. Its norm is equal to $\frac{\omega}{c}$, where c is the sound propagation velocity.

When the listener is placed far away from the point sources, the wave fronts can be locally considered as plane waves. Theoretically, the far field approximation of Fraunhofer is valid when $k|r - r_0| \gg 1$. With this approximation, the sound pressure at the listener position is given by [4]:

$$p(\vec{r}, t) = A e^{\jmath(\omega t - k|\vec{r} - \vec{r_0}|)} \tag{2.24}$$

Figure 2.11 shows the plane wave fronts. In this case, the amplitude of the pressure does not depend on the distance between the point source and the listener. Furthermore, the direction of the wave vector $\vec{k}$ is constant.
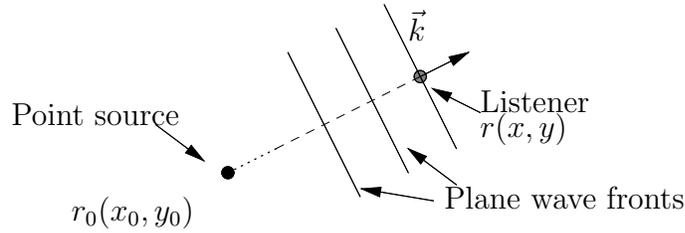
Figure 2.11: *Plane waves radiated by a point source localized far away from the listener.*

2.5.2 Why do plane waves enlarge the sweet spot?

In order to explain the advantages of using plane waves for the stereophonic reproduction, a comparison between the near field and the far field radiations is done. Figure 2.12 shows the two different type of stereo radiations studied. In the case (a), the sources are close to the listener. As it has been explained above, the listener receives spherical waves. In case (b), the sources are placed far away from the listener. Accordingly, the listener receives plane waves .
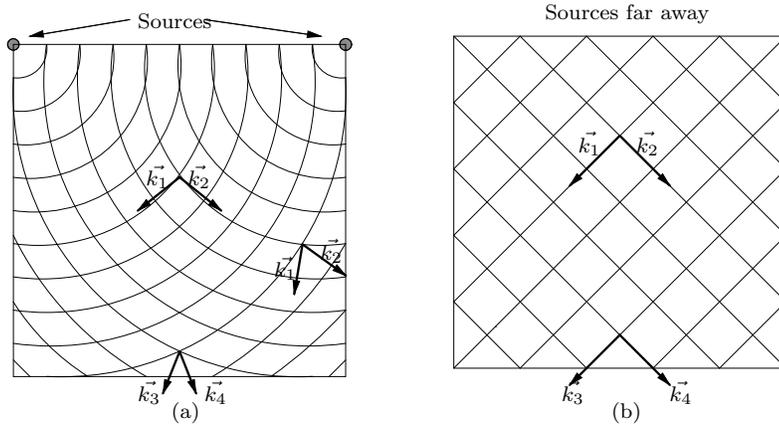
Figure 2.12: *In (a), both sources are placed in both corners of the listening area. In (b), both sources are placed far away.*

Three parameters are investigated in order to understand the advantage of plane waves for having a good stereohonic imaging in the entire listening area:

- The amplitude

- The propagation direction

- The time delay

With spherical waves, the amplitude of the pressure is inversely proportional to the distance between the point source and the listener as shown in Figure 2.13. For example, the pressure difference between 1 m and 2 m distance from a point source is 13.86 dB whereas it is only 0.39 dB between 51 m and 52 m. Far away from the source, the spherical wave front converges to a plane wave. The amplitude can be considered close to constant everywhere in the listening area if the simulated source is that far away. Hence, by using plane waves, the listener can move within the entire listening area without receiving amplitude differences of the two signals. The size of the sweet spot is therefore enlarged.
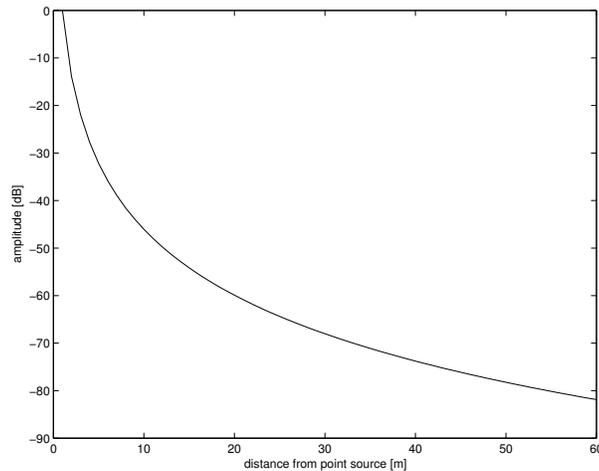
Figure 2.13: *Decaying amplitude with distance for spherical waves.*

The second improvement obtained from plane waves is the constant direction of the wave vector. No matter where the listener is in the listening area, the angle created by the two wave vectors is constant. By using spherical waves, this angle is dependent on the position of the listener. This is illustrated in Figure 2.12.

The time delay Δt is the difference between the time perception of the two signals radiated by both loudspeakers. If the listener is on the symmetry line between both loudspeakers, the distance to each loudspeaker and the corresponding ear is equal. Therefore, both signals, from left and right loudspeaker, arrive at the same time. If the listener's position is now changed

perpendicular to this line, the distance to each speaker changes. That causes a time delay between the signals. It is interesting to note that, if the listener moves on a parallel line to the symetry line between both loudspeakers toward the loudspeakers, the time delay increases in case (a), whereas it is constant in case (b). Figure 2.14 illustrates this characteristic.

(a)   (b)
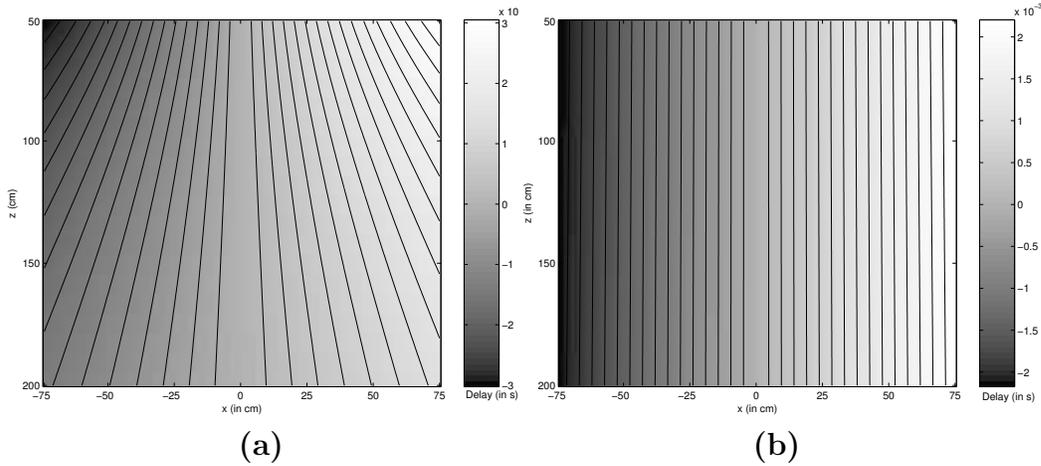
Figure 2.14: *These plots represent lines of equal time delays between the right and the left signal for sources placed* **(a)** *on the array* **(b)** *, and far away.*

An interesting characteristic is the value of the time delay. When the listener moves away from the ideal stereophonic position, for a classical stereo configuration, the time delay increases faster than if the stereo sources are far away. The time delay is smaller with plane waves than with spherical waves. Figure 2.14 shows this characteristic. Indeed, the time delay vary from $-3ms$ to $3ms$ for two close sources and from $-2ms$ to $2ms$ for two sources placed far away.

To conclude, the constant amplitude, the constant propagation angle, and the decrease of the time delay show that the sweet spot is enlarged when using plane waves.

3

Simulation

In this chapter, several simulations concerning wave-field synthesis and its implementation for conventional stereo reproduction, will be presented and discussed. The simulations will focus mainly on the artefacts, due to the practical realization, described in 2.3.2 and different methods to reduce this artefacts.

3.1   Wave field synthesis

The practical setup is described in chapter 4.1. This gives the default parameters for the simulation. Unless stated otherwise, the loudspeaker array has a length of 3.72m and each loudspeaker is drawn as a white circle with a diameter of 148mm. If the loudspeakers are close, these circles, representing the loudspeakers, will overlap. The distance from the array to the middle of the listening area is 1.25m, which is seen as a black and white square. The sides of the listening area each has a length of 1.5m.

With wave-field synthesis, virtual sources can be situated almost anywhere behind the loudspeaker array. The angle to the notional source is called α_{source}. A first example and the general setup for the simulation is shown in Figure 3.1.
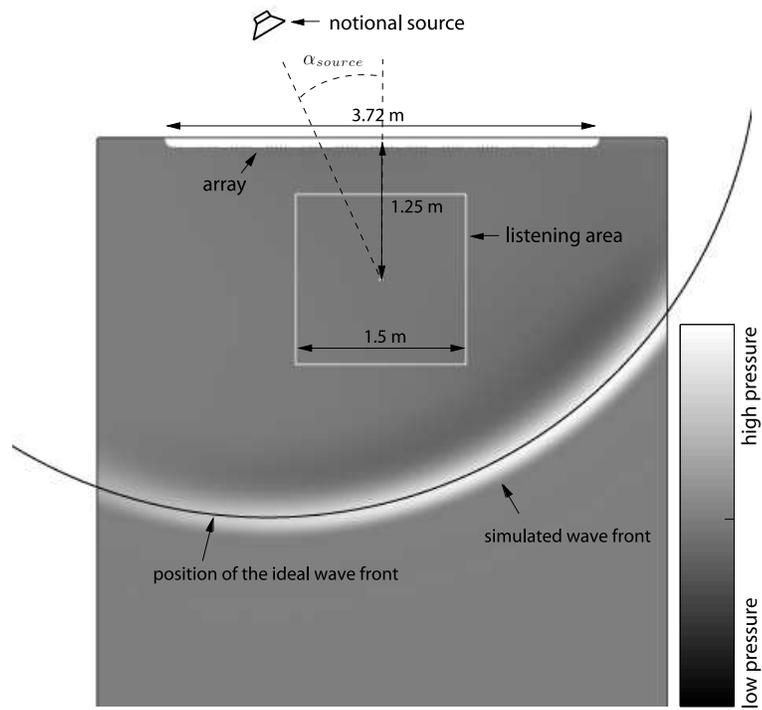
Figure 3.1: *Wave field setup with an array of 96 loudspeakers in a line and a notional source placed 1 m behind the array.*

3.2 Simulation methodology

3.2.1 Overview of the simulation program

To synthesize a sound field in front of the array, the driving functions have to be calculated first. The simulation program uses this driving functions to simulate the behaviour of a wave field synthesis system.

The program is split up into different parts. A general flow chart can be seen in Figure 3.2.
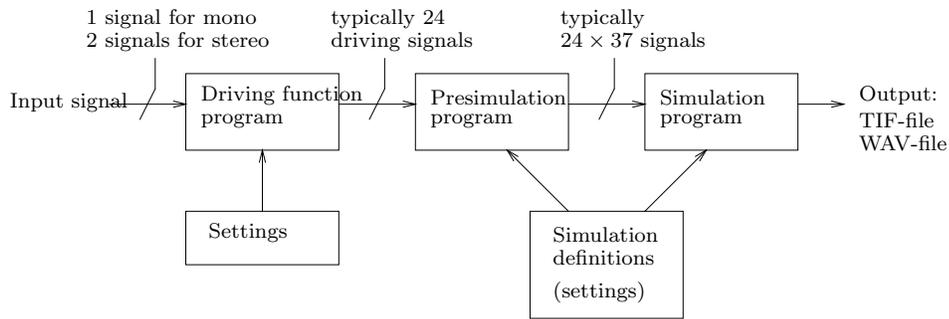
Figure 3.2: *Flow chart showing the overall structure of the signal creation and simulation program.*

A detailed description of the program is included in the Appendixes E and F.

3.2.2 Used signals

Two types of signals were chosen for simulation, pulses low-pass filtered with different cutoff frequencies and sinusoids with different frequencies.

Pulses are used to simulate a wave front of a broadband signal. Sinusoids are used to reveal interference patterns for specific frequencies.

Low-pass filtering is implemented using, a butterworth filter. A second order filter is chosen to minimize oscillation, which would overlay with the effects under observation. Due to the low order filtering, high frequency content remains and could cause aliasing. The pulses used can be seen in Figure 3.3.
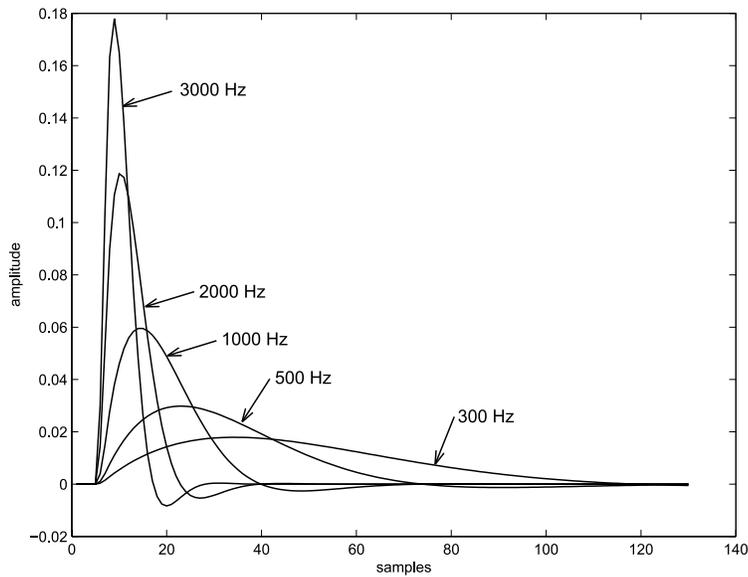
Figure 3.3: *Pulses low-pass filtered with a 2nd order butterworth filter and different cutoff frequencies.*

All sinusoids were low-pass filtered with a butterworth filter of 6th order and a cutoff frequency of their own frequency. This was done to remove the high frequency content at the edges of the signal, due to cutting it off abruptly. A plot of such a filtered sinusoid can be seen in Figure 3.4.

3.3 Aliasing due to spatial sampling

For the practical setup, the distance between the centers of two adjacent loudspeakers, $\triangle x$, is 155 mm. To eliminate truncation effects, a truncation window is applied as explained in section 3.5 on page 44.

In Figures 3.5 and 3.6 on page 33 the notional source is placed 2 m behind the center of the array. The aliasing frequency is 1132.7 Hz as calculated in section 2.3.2. Different input signals are used to show the frequenc dependency of spatial aliasing. Spatial aliasing can be observed as a periodic ringing coming after the desired wave front. The effect decreases as the high frequency content is reduced.

If the number of loudpeakers is halfed and the distance in between is doubled,

(a)

(b)

Figure 3.4: **(a)** *The original, unfiltered sinusoid cutted off at 300 samples.* **(b)** *The filtered sinusoid where the abrupt cutoff is smoothed, containing less high frequencies.*

(a)

(b)
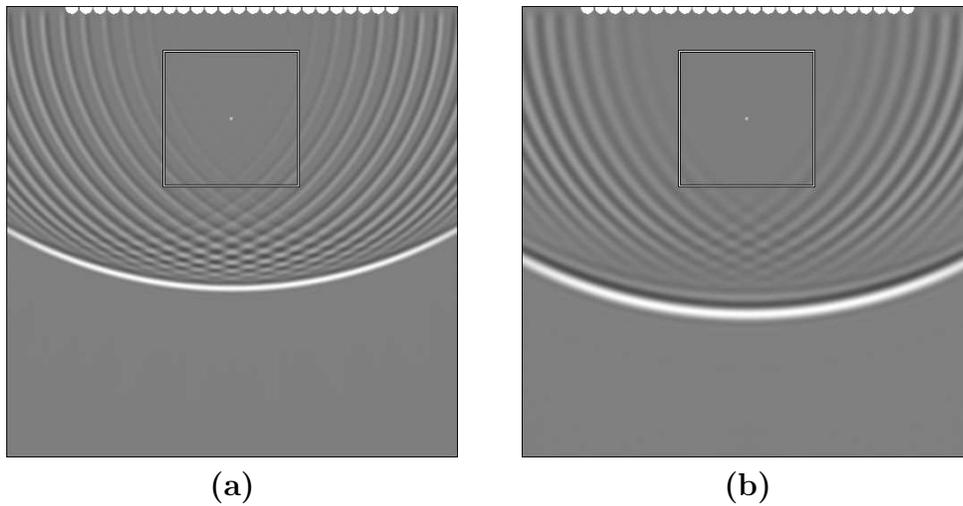
Figure 3.5: *Notional source, placed two meter behind the center of the array, emitting a pulse, low-pass filtered with a cutoff frequency of* **(a)** *3000 Hz,* **(b)** *2000 Hz.*
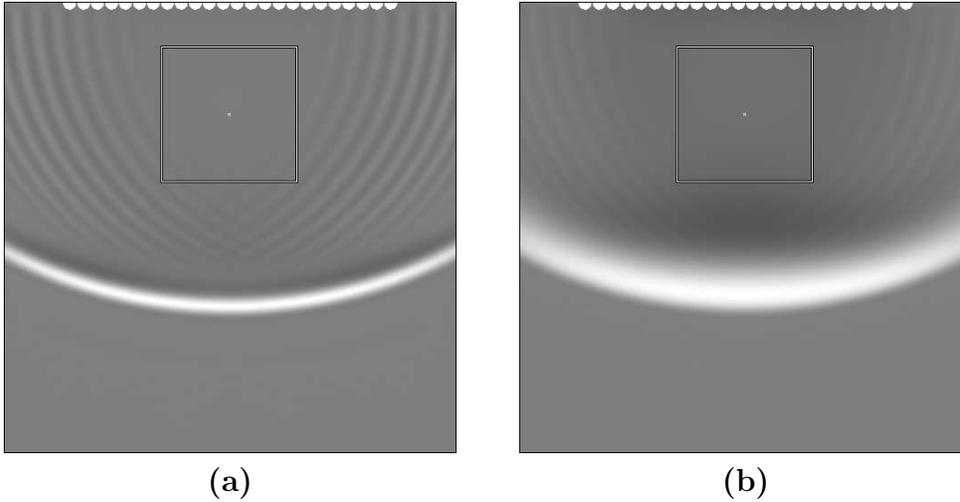
Figure 3.6: *Notional source, placed two meter behind the center of the array, emitting a pulse, low-pass filtered with a cutoff frequency of* **(a)** *1000 Hz,* **(b)** *300 Hz.*

f_{alias} becomes 563.3 Hz. Now spatial aliasing is more visible in Figure 3.7 for a 1000Hz low-pass filtered pulse compared to Figure 3.6a.

As mentioned previously, aliasing is dependent on the angle from the array to the notional source. This angle decreases if the source is virtually moved far away. In Figure 3.8a the source is placed 75 m behind the array, and the aliasing effect is reduced. If the loudspeaker interspacing is doubled, spatial aliasing is increased, as seen in Figure 3.8b.

To study the spatial aliasing effect for single frequencies, a source playing a tone is placed 50 m behind the array with an angle α_{source} of 30° to the left side. In this case the listener position is in the far field of the notional source and receives wave fronts which can be approximated as plane waves. This is shown in Figure 3.9 and 3.10.

In Figure 3.9a a frequency lower than the aliasing frequency is chosen. Spatial aliasing is not visible.

If the frequency increases, the wave field starts to distort. In Figure 3.9b, an unwanted second wave front can be seen, which interferes with the desired one. Spatial information is lost, meaning that waves coming from the direction of the notional source can not longer be correctly reproduced.
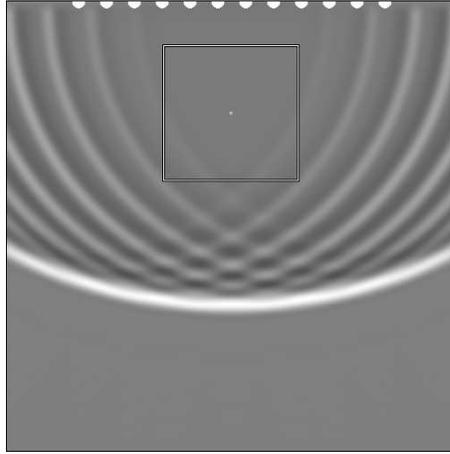
Figure 3.7: *Notional source, placed two meter behind the center of the array, emitting a pulse, low-pass filtered with a cutoff frequency of 1000 Hz reproduced with 12 loudspeakers.*

(a)   (b)
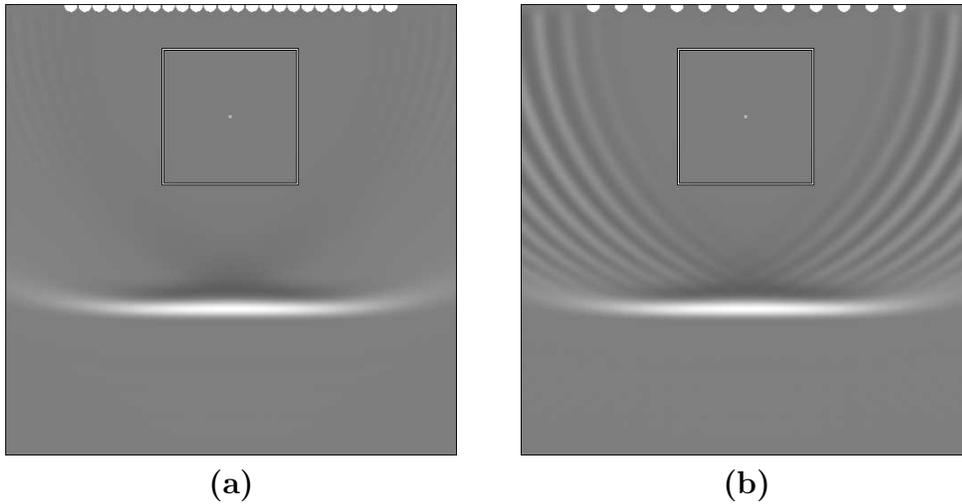
Figure 3.8: *Notional source, placed 75m behind the center of the array, emitting a pulse, low-pass filtered with a cutoff frequency of 1000 Hz reproduced with **(a)** 24 loudspeakers, **(b)** 12 loudspeakers.*
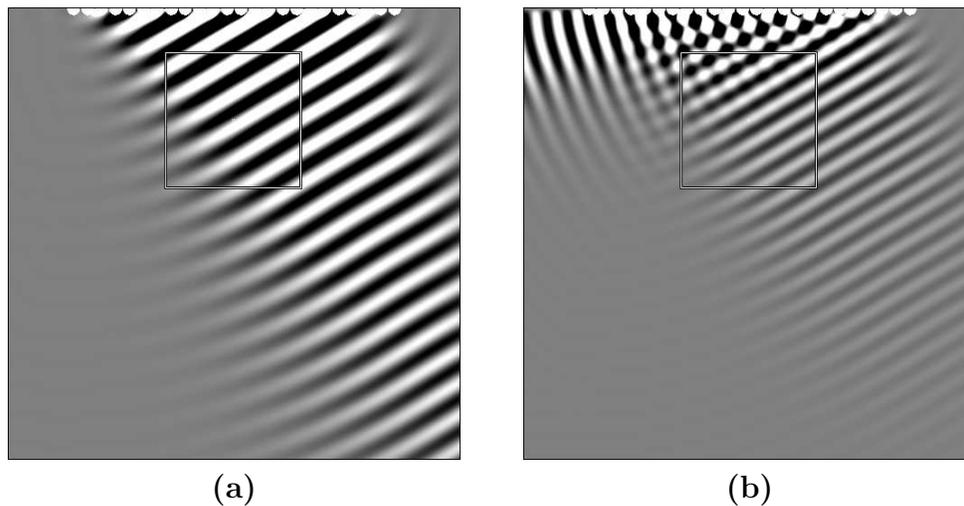
Figure 3.9: *Sinusoidal source placed 50 m behind the array with $\alpha_{source} = 30°$ to the left side* **(a)** *1106.45 Hz,* **(b)** *1500 Hz.*
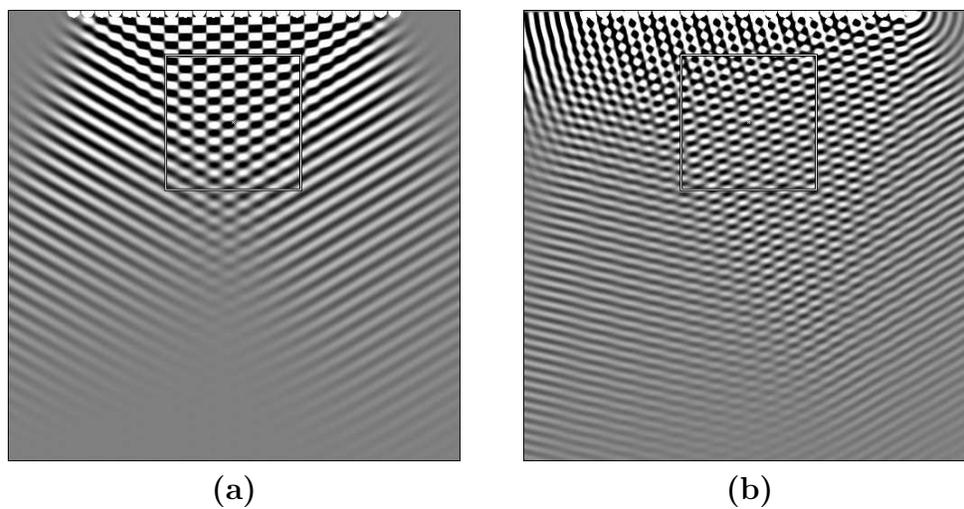
Figure 3.10: *Sinusoidal source placed 50 m behind the array with $\alpha_{source} = 30°$ to the left side* **(a)** *2212.9 Hz,* **(b)** *3000 Hz.*

In Figure 3.10a every second loudspeaker emits a tone of 2212.9 Hz in phase. It is not possible to see, if the notional source is placed on the left or the right side, even though the source direction is still 30° to the left.

If the frequency is increased, more possible directions for the notional source appear. For a frequency of 3000 Hz, a third wave front can be seen in Figure 3.10b.

Figure 3.9b shows that the perceived spatial aliasing by the listener is dependent on the position. For a notional source on the left side, the effect is more noticeable for the listener in the upper left corner of the listening area, because the maximum angle from listener to any loudspeaker is larger than in any other point of the listening area.

If the loudspeaker array is extended, this maximum angle increases for the entire listening area. The spatial aliasing frequency decreases to 1108.6 Hz. This case is shown in Figure 3.11 compared to Figure 3.9b.
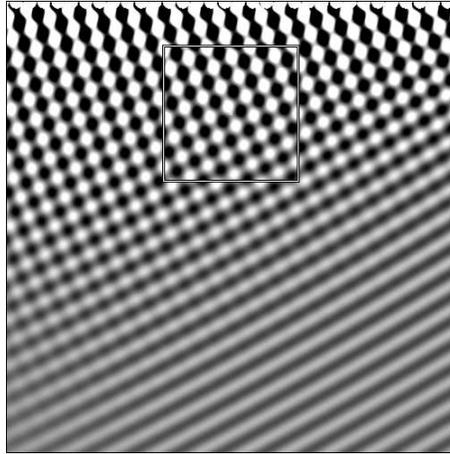
Figure 3.11: *Sinusoidal source 1500 Hz placed 50 m behind the array with $\alpha_{source} = 30°$ to the left side, reproduced with 96 loudspeakers.*

If the angle to the notional source α_{source} is reduced, spatial aliasing decreases. This can be seen in Figure 3.12. The effect disappears for a tone of 2000 Hz in Figure 3.12b.

Wave-field synthesis systems can reproduce notional sources close to the array. The array loudspeakers then receive the waves with a different incidence angle. The effect is shown in Figure 3.13.

36
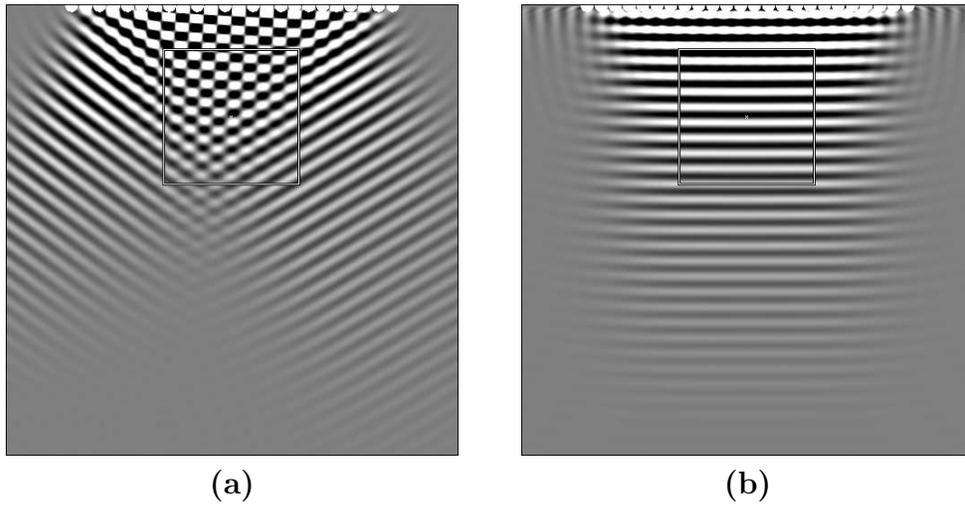
Figure 3.12: *Sinusoidal source 2000 Hz placed 50 m behind the array with α_{source} equals* **(a)** *30°,* **(b)** *0°.*
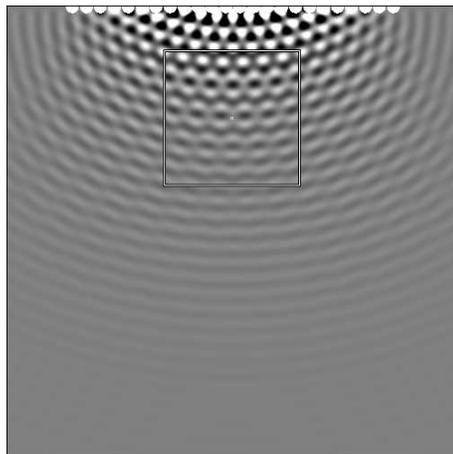
Figure 3.13: *Sinusoidal source 2000 Hz placed 2 m behind the center of the array.*

To give a different view of spatial aliasing, a low-pass filtered pulse was played back by a notional source placed 0.1m behind the center of the array. A sound file was virtually recorded in the center of the listening area and Fourier transformed. In Figure 3.14 the resulting frequency content can be compared with the frequency content of the original input signal. The angle α_{source} equals $86,8°$, which gives an spatial aliasing frequency of 1108.2Hz. Spatial aliasing boosts high frequencies, as it is shown in Figure 3.14b.

(a)   (b)
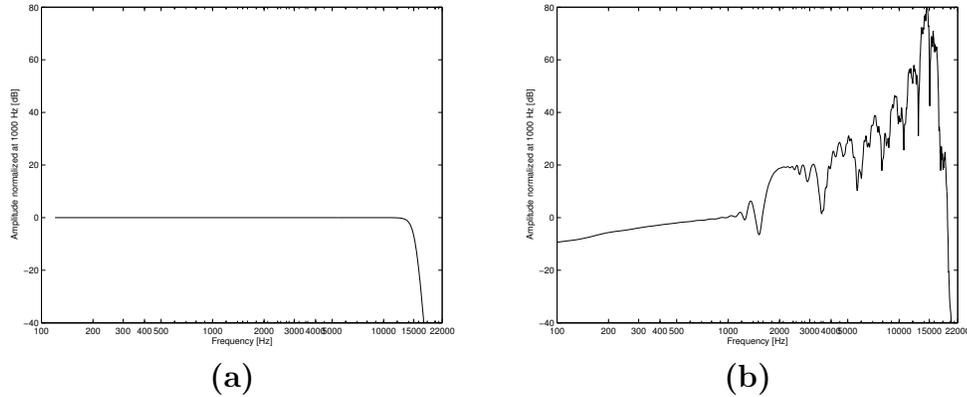
Figure 3.14: *Fourier transform of a* **(a)** *pulse, low-pass filtered with a cutoff frequency of 15000Hz,* **(b)** *the recorded sound in the center of the listening area, if a notional source placed 0.1m behind the center of the array plays back this pulse.*

Different methods to reduce spatial aliasing will be investigated in the following sections.

3.3.1   Spatial bandwidth reduction

It is possible to avoid spatial aliasing if the directivity of the loudspeakers in the array is increased, to reduce the overlapping of their main lobes. Until now, omnidirectional sources have been used for the simulations. Figure 3.16 and 3.17 show what happens if the directivity of the array loudspeakers is increased by applying the directivity patterns shown in Figure 3.15. These created patterns only consists of one mainlobe, neglecting sidelobes that would appear in practice. In practical applications, more directive sources could be generated by using subarrays.

It is shown that with increasing directivity, the spatial aliasing is diminished. One drawback is the decreasing area where the sound field is synthesized as seen in Figure 3.17b, where the edges of the wave front are faded out. The amount of possible positions of the notional source decreases as well, because the propagation angle in the wavefield is now limited due to the loudspeaker directivity. This is shown in Figure 3.18b, where the desired wave front disappears because the array cannot emit sound in that propagation direction.
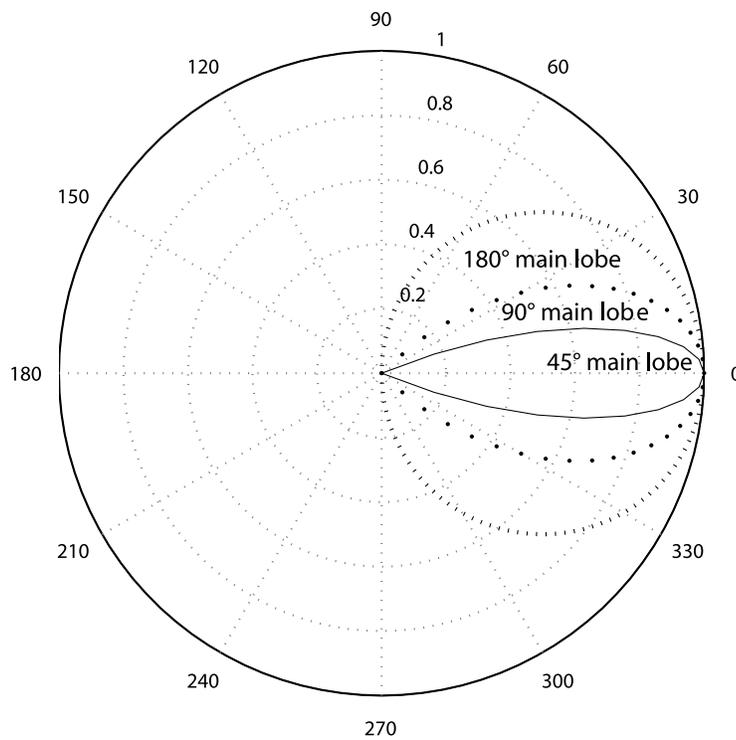
Figure 3.15: *Used ideal directivity patterns for the array loudspeakers with different mainlobe widths.*

A similar effect should be obtained, if the directivity of the notional source is increased, meaning that all wave field components above a specific propagation angle are suppressed.

(a) (b)
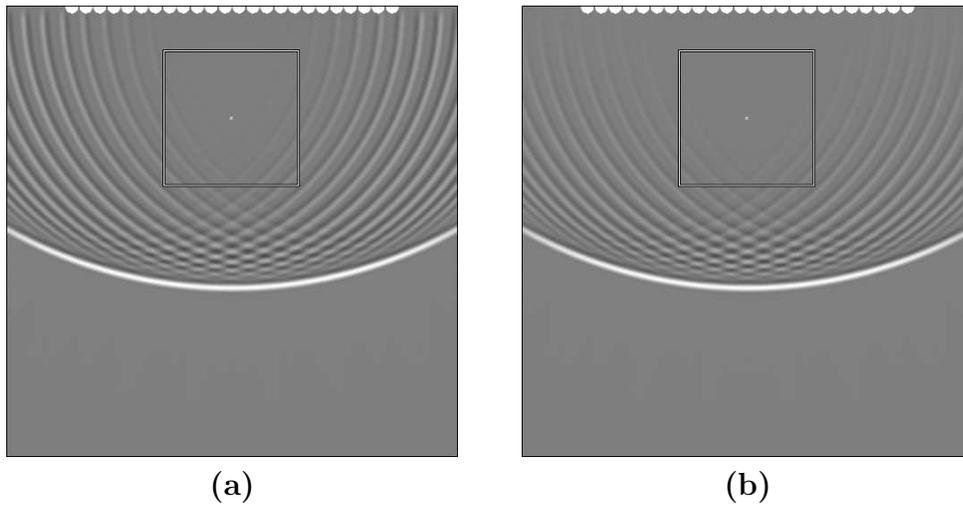
Figure 3.16: *Notional source emitting a pulse, low-pass filtered with a cutoff frequency of 3000 Hz is placed two meters behind the center of the array. The used array loudspeakers* **(a)** *are omnidirectional,* **(b)** *have only one mainlobe with 180°.*

(a) (b)
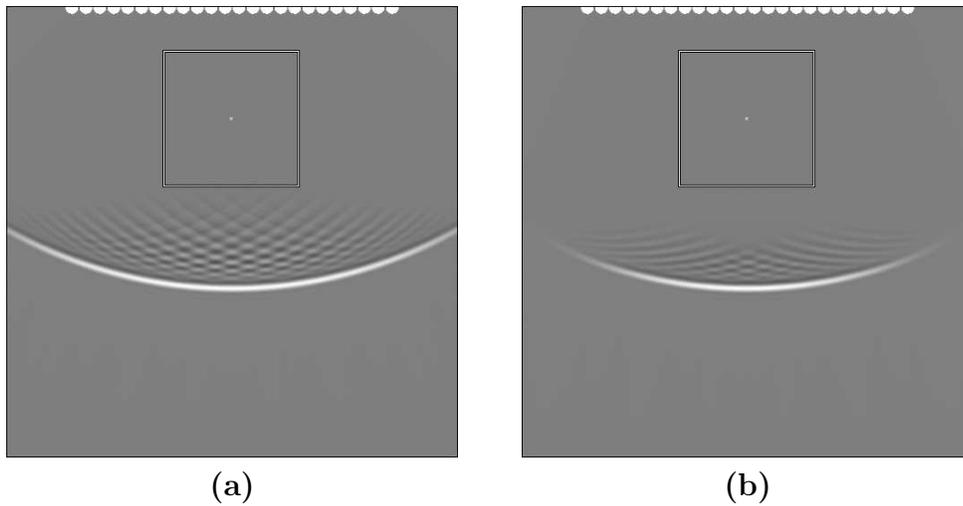
Figure 3.17: *Notional source emitting a pulse, low-pass filtered with a cutoff frequency of 3000 Hz, placed two meters behind the center of the array. The used array loudspeakers have only one mainlobe with* **(a)** *90° angle,* **(b)** *45° angle.*

40

(a)   (b)
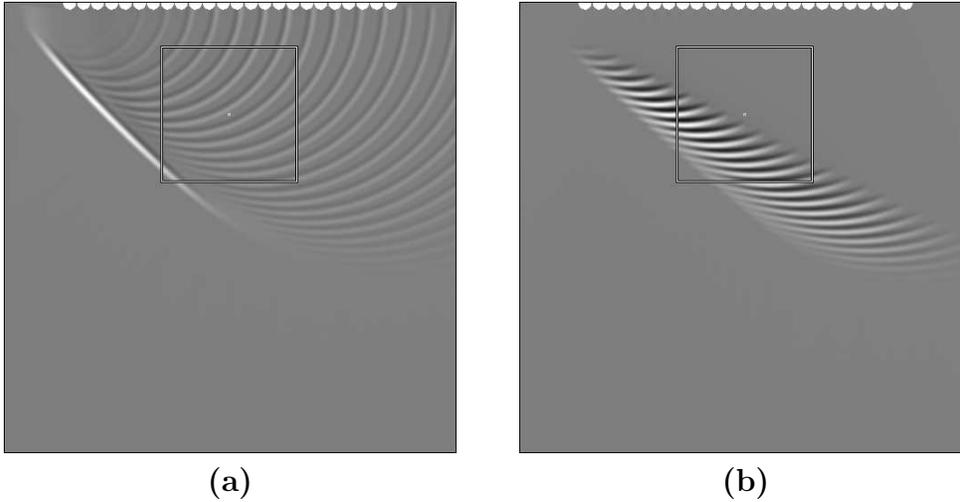
Figure 3.18: *Notional source emitting a pulse, low-pass filtered with a cutoff frequency of 3000 Hz, placed 10 meters behind the array in 45° angle. The used array loudspeakers* **(a)** *are omnidirectional,* **(b)** *have one mainlobe with 45° angle.*

3.3.2 Randomization of high frequency content

An other method to avoid the audible periodicity of spatial aliasing is to randomize the phase of the driving signals for high frequencies [8]. The sound field then looses its spatial properties above a certain frequency. In [8] it is described that this method also has perceptual disadvantages.

Figure 3.19b shows in comparison with 3.19a, that the periodic ringing coming after the wave front is partially replaced with noise, if the phase of each driving signal is manipulated for all frequencies. The used noise which was added to the phase had a range of $\pm\frac{\pi}{2}$. The shape of the wave front itself changes and artefacts appear in front of it.

If the phase is changed only for frequencies above 1kHz, the resulting wave field is shown in Figure 3.20a. As low frequencies play a dominant role in sound localization [14], each driving signal is low-pass filtered with a cutoff frequency of 1kHz after modifying the phase. The resulting sound field is shown in Figure 3.20b. The shape of the resulting wave front is close to the desired shape. Some distortion appears, because of two main reasons. A part of the frequency content above 1kHz is still contained within the

41

low-passeded signal because of the butterworth filter's order. The driving signals are windowed by Hanning windows with a length of 256 samples and an overlap of 128 samples. The phase is then manipulated separately before the windowed sections are added again.

If the phase is shifted with a maximum of $\pm\frac{\pi}{2}$ an overlaying noise is clearly audible.

Variations in the amount of frequency shifting affect the quantity of noise, replacing the periodic aliasing affect.

(a)   (b)
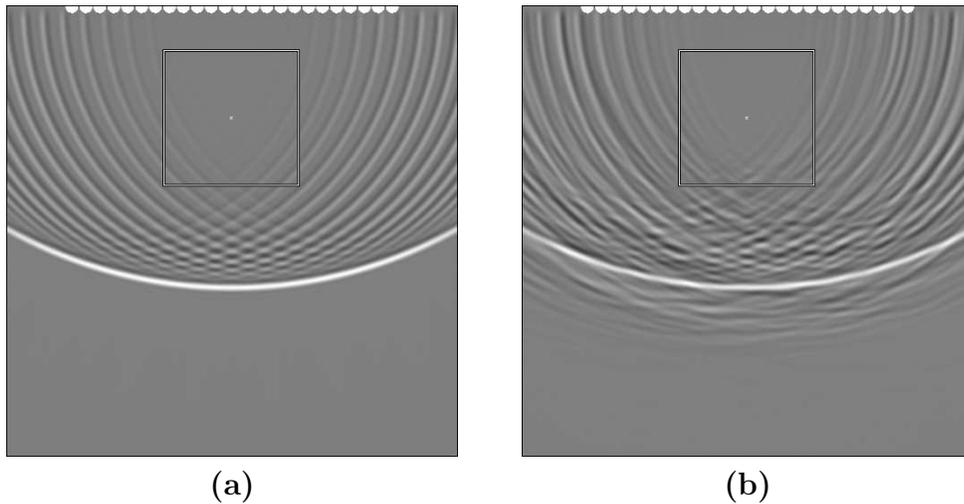
Figure 3.19: *Notional source emitting a pulse, low-pass filtered with a cutoff frequency of 3000 Hz, placed two meters behind the middle of the array. The used array loudspeakers are omnidirectional. (a) Reproduction without randomized phase. (b) All driving signal frequency components are randomly phase shifted with a maximum shifting angle of $\pm\frac{\pi}{2}$.*

3.4 Reflections

If a wave field synthesis setup is implemented in a non anechoic environment, reflections will appear and influence the sound field. That is shown in Figure 3.21 and 3.22, where first order reflections on the side walls are simulated. The reflection coefficient is chosen to 0.5. In practice more reflections of
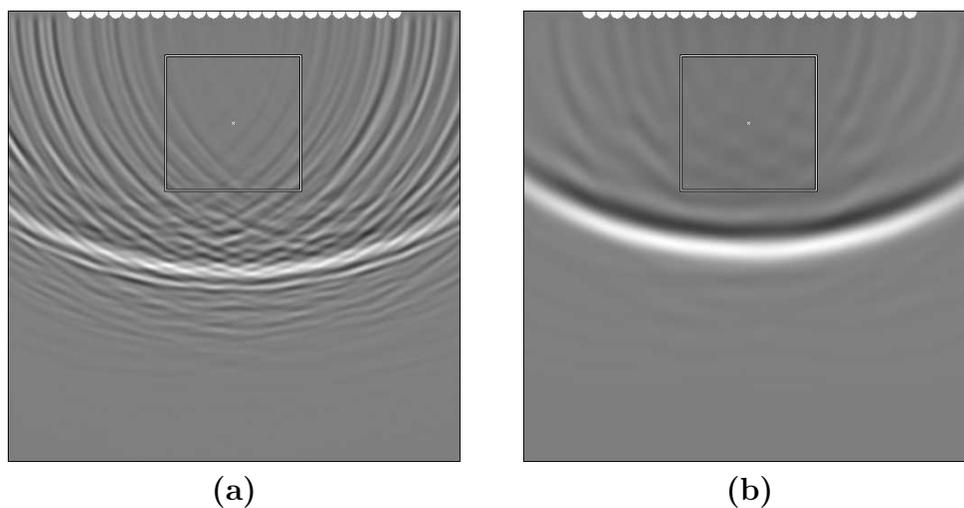
Figure 3.20: *Notional source emitting a pulse, low-pass filtered with a cutoff frequency of 3000 Hz, placed two meters behind the middle of the array. The used array loudspeakers are omnidirectional.* **(a)** *Driving signal components above 1000 Hz are randomly phase shifted with a maximum shifting angle of $\pm\frac{\pi}{2}$.* **(b)** *Same as (a), but signals are again low-pass filtered with 1000 Hz after processing phase.*

43

higher order coming from all surrounding surfaces including the ceiling or floor will overlay.
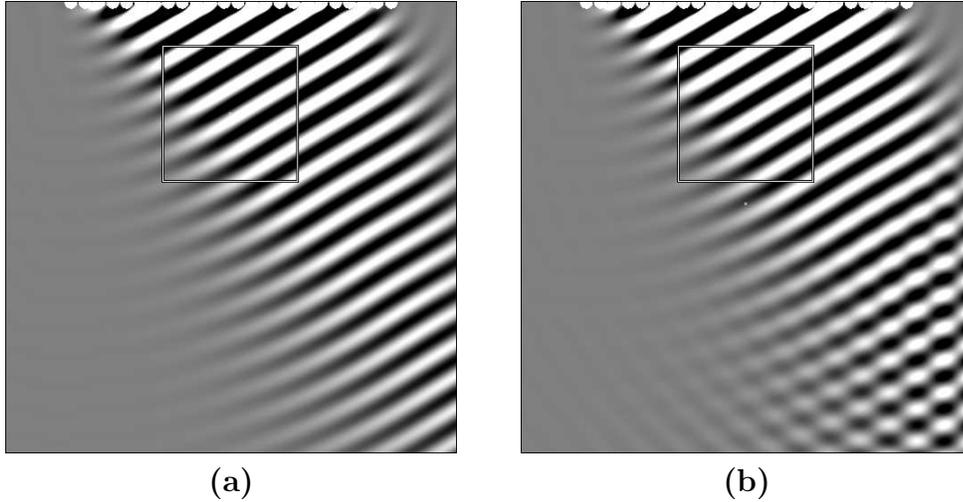
Figure 3.21: *Notional source emitting a sinusoids of 1106.45 Hz is placed 50 m behind the array with 24 loudspeakers in* 30° *angle to the left.* **(a)** *Plot for free field conditions.* **(b)** *Plot with reflection on the right side wall.*

3.5 Truncation Effect

Unless stated otherwise, all simulations are done with an array of 96 loudspeakers and an interspacing of $\frac{155}{4}$mm, so that the array length remains the same as before. The spatial aliasing frequency increases, therefore it is higher than the used frequency range. In this way, the spatial aliasing effects are seperated from the truncation effects.

The notional source is placed 2 m behind the center of the array.

Because of the finite array length, truncation errors occur. This effect is illustrated in Figure 3.24a and 3.24b. It can be seen that the synthesized wave front is followed by two edge events.

(a)   (b)
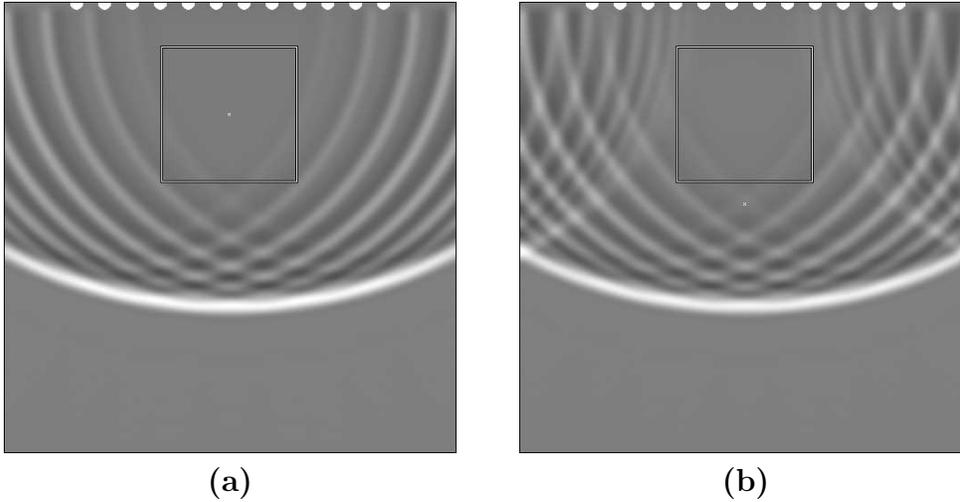
Figure 3.22: *Notional source emitting a pulse low-pass filtered with a cutoff frequency of 1000 Hz is placed 2 m behind the center of the array with 12 loudspeakers. (a) Plot for free field conditions. (b) Plot with reflections on the side walls.*

3.5.1   Tapering window

To suppress the truncation effect, a weighting function is applied over the array elements. A half-sided cosine window is used over one third of the used loudspeakers at each edge of the array. That means that the amplitude of the loudspeakers decreases toward the edges. This was also done to all simulations in the previous sections. An example of a tapering window used is shown in Figure 3.23. The results are shown in Figure 3.25a and 3.25b, where the truncation effects disappear.

In Figure 3.26 the same effect is illustrated by using a sinusoid of 2000 Hz. Small errors in the shape of the wave front can be seen close to the edges. These errors disappear after applying the truncation window, but the effective area of the correct wave field is reduced.

The truncation effect should also diminish if the size of the array is increased. This is shown in Figure 3.27, where the number of loudspeakers is raised to 192 so that the array length doubles. The loudspeaker interspacing remains the same.
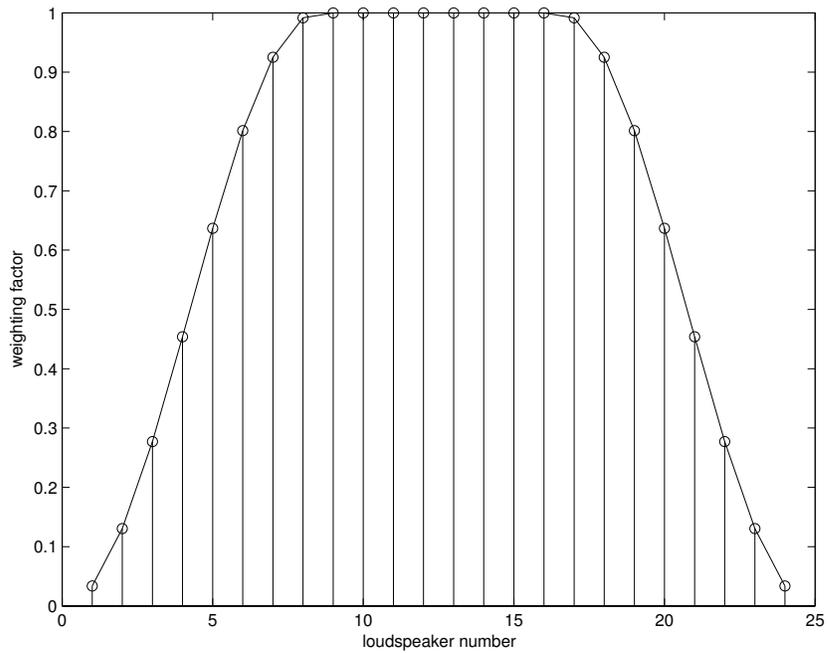
Figure 3.23: *Weighting factors for 24 loudspeakers.*

(a)   (b)
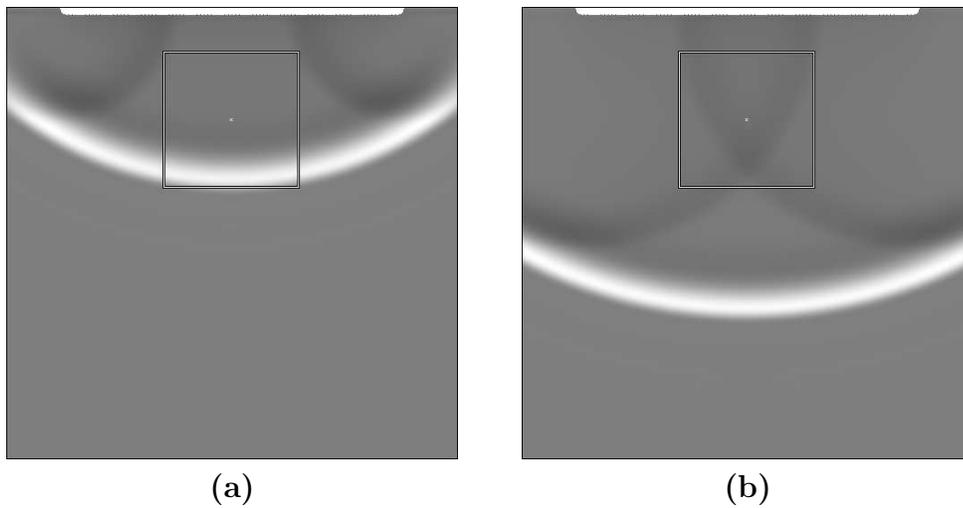
Figure 3.24: *Notional source, placed 2m behind the center of the array emits a pulse, low-pass filtered with a cutoff frequency of 500 Hz* **(a)** *with truncation events,* **(b)** *same as (a) but ≈ 4 ms later.*

(a)   **(b)**
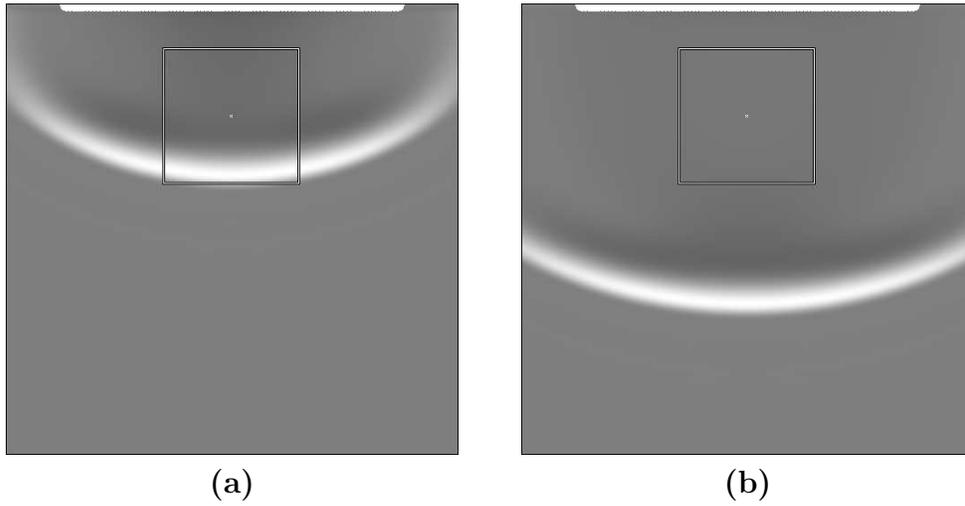
Figure 3.25: *Notional source, placed 2m behind the center of the array emits pulse, low-pass filtered with a cutoff frequency of 500 Hz* **(a)** *with truncation window,* **(b)** *same as (a) ≈ 4 ms later.*

(a)   **(b)**
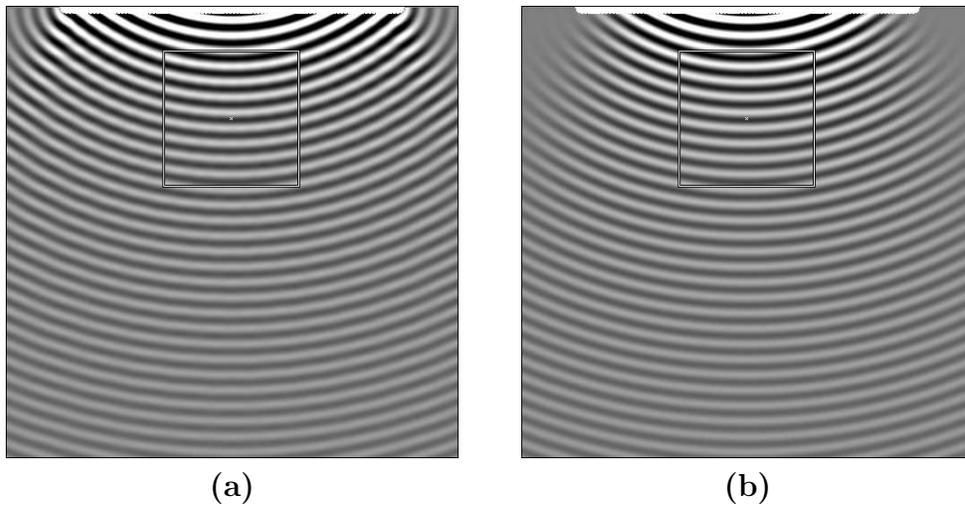
Figure 3.26: *Notional source, placed 2m behind the center of the array emits a sinusoidal of 2000 Hz* **(a)** *without truncation window,* **(b)** *with windowing.*

47

(a)   (b)
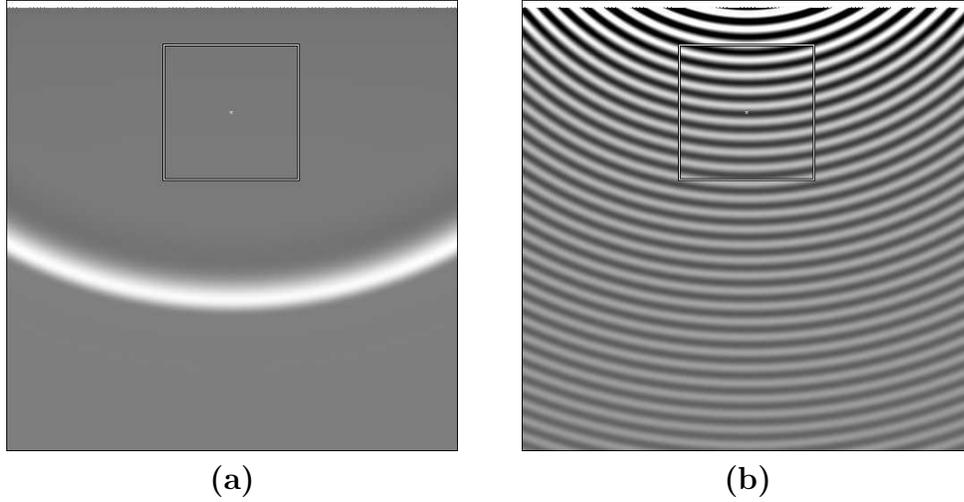
Figure 3.27: *The notional source is placed 2m behind the center of the array of 192 loudspeakers reproducing a* **(a)** *pulse low-pass filtered with a cutoff frequency of 500 Hz,* **(b)** *sinusoidal of 2000 Hz.*

Another method to avoid truncation is the use of side arrays. If the listener is totally surrounded with loudspeakers, the truncation error is avoided.

3.5.2 Feeding Edge Loudspeakers

To overcome the truncation effect, Wapenaar [13] proposed to analytically approximate the contribution of the missing loudspeakers via the "method of integration by parts".

The desired pressure $P_{correct}$ equals the synthesized wave field pressure plus two error terms, introduced by the finite length of the array.

$$P_{correct}(r, \omega) = P(r, \omega) + \epsilon_l(r, \omega) + \epsilon_r(r, \omega) \tag{3.1}$$

ϵ_l represents the error introduced by the missing contribution on the left side of the array. ϵ_r is the error of the right side. They are approximated as follows:

$$\epsilon_l \approx S(\omega) \sqrt{\frac{1}{2\pi jk}} \frac{e^{-jk|r_{N+1/2}|}}{\sqrt{|r_{N+1/2}|^3}} \frac{1}{\sin\theta_{left} - \sin\varphi_0} \frac{e^{-jk|r-r_{N+1/2}|}}{|r - r_{N+1/2}|} \tag{3.2}$$

$$\epsilon_r \approx S(\omega) \sqrt{\frac{1}{2\pi jk}} \frac{e^{-jk|r_{-N-1/2}|}}{\sqrt{|r_{-N-1/2}|^3}} \frac{1}{\sin\theta_{right} - \sin\beta_0} \frac{e^{-jk|r-r_{-N-1/2}|}}{|r - r_{-N-1/2}|} \tag{3.3}$$

The geometry used is shown in Figure 2.6 on page 15.

A fixed listener position $\vec{r}(X_l, Z)$ is chosen so φ_0 is fixed [15]. $\vec{r}$ represents the position of the notional source and $r_{\pm N \pm 1/2}$ represents the position of the virtual loudspeakers which are placed at a distance $\triangle x$ from the edge loudspeakers. The signals resulting from these equations are added to the driving signal of the edge loudspeakers.

In Figure 3.28a a low pass filtered pulse with a cutoff frequency of 500 Hz is shown. The truncation effect is clearly visible as two waves coming from the edges.

In Figure 3.28b the same pulse is shown with the correction factor added to the edge louspeakers. It is possible to see that the waves due to the truncation effect have less amplitude.

Figure 3.29 shows the pulse with the correction factor and with a tapering window applied to the array.

It is shown that the correction achieved by feeding the edge loudspeakers is not as effective as windowing the array.

3.6 Amplitude Errors

A line array emits cylindrical waves, if all loudspeakers play in phase [4]. This causes an amplitude roll off of the reproduced wave. The amplitude on the symmetry axis scales down with $\frac{1}{\sqrt{r}}$, r being the distance to the array.

(a)   (b)
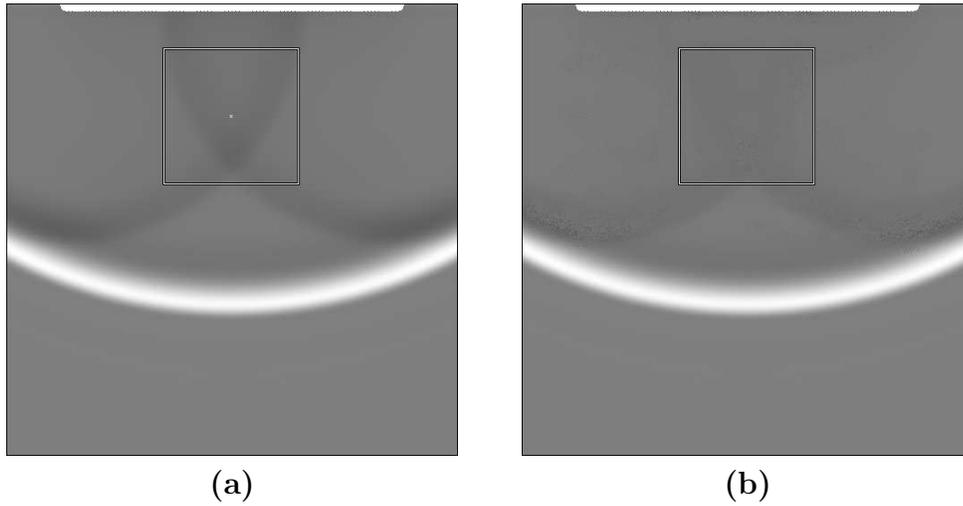
Figure 3.28: *Notional source, placed 2m behind the center of the array emitts a pulse, low-pass filtered with a cutoff frequency of 500 Hz* **(a)** *with truncation events,* **(b)** *same as (a) but with feeding edge loudspeakers technique applied.*
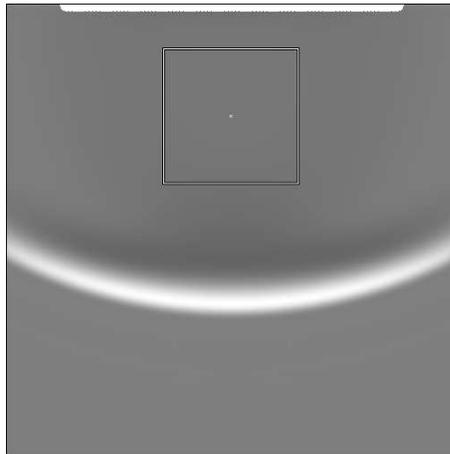
Figure 3.29: *Notional source, placed 2m behind the center of the array emitts pulse, low-pass filtered with a cutoff frequency of 500 Hz with a truncation window applied and the feeding edge loudspeakers technique.*

This is illustrated in Figure 3.30. The simulations are done, using a 600Hz sinusoid.

In Figure 3.30a an overlaying pattern can be seen due to constructive and destructive interference. This fluctuation depends on the used frequency of the sinusoid and can be reduced applying a half sided cosine window, like used to avoid truncation effects. This is done in Figure 3.30b.

(a)   (b)
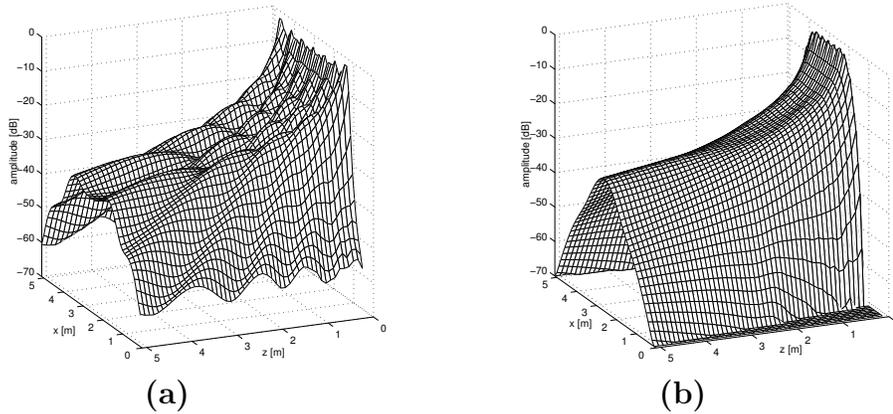
Figure 3.30: *Amplitude roll off for an array of 24 loudspeakers when the notional source is placed 50 m behind it.* **(a)** *The amplitude on the symmetry axis decays with* $\frac{1}{\sqrt{r}}$. **(b)** *Same as (a) but with truncation windowing.*

The roll off on the symmetry axis increases if the notional source is placed closer to the array. This can be seen in Figure 3.31, where the source is placed 2m behind the center of the array. The amplitude roll off increases in the z-direction and decreases in the x-direction, due to the spherical component of the resulting wave front in the horizontal plane.

3.7 Stereo Setup

To simulate a conventional stereo setup, a signal is virtually played back with only two loudspeakers of the array in the traditional 30° position to the middle of the listenig area. In Figure 3.32 the propagation of spherical wavefronts can be seen.
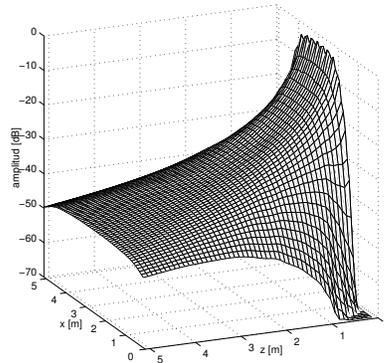
Figure 3.31: *Amplitude roll off for an array of 24 loudspeakers when the notional source is placed 2 m behind the center of the array. The amplitude on the symmetry axis decays with more than $\frac{1}{\sqrt{r}}$. The simulation was made with truncation window.*

(a)   (b)
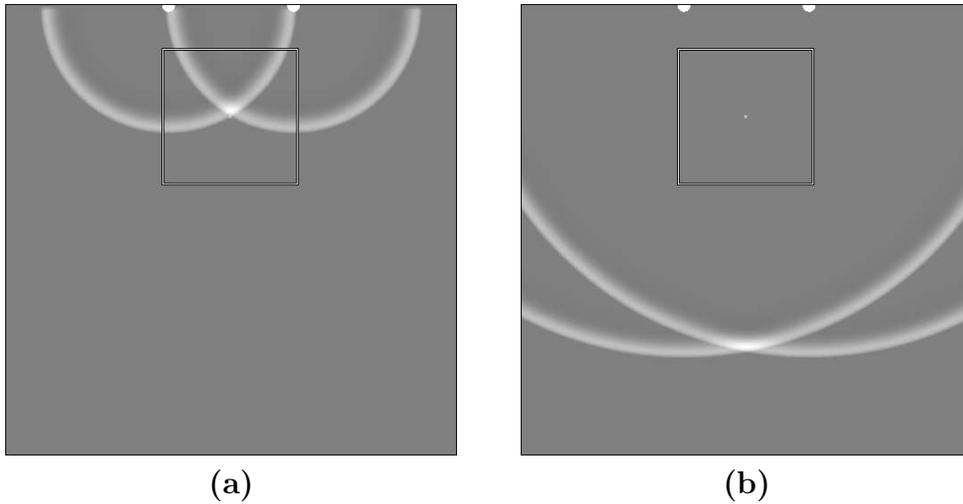
Figure 3.32: **(a)** *Pulse, low-pass filtered with a cutoff frequency of 1000 Hz emitted by 2 loudspeakers at their traditional stereo positions,* **(b)** *same as (a) but $\approx$ 7 ms later.*

52

In Figure 3.33 the interference pattern for a traditional stereo setup is shown.

(a)   (b)
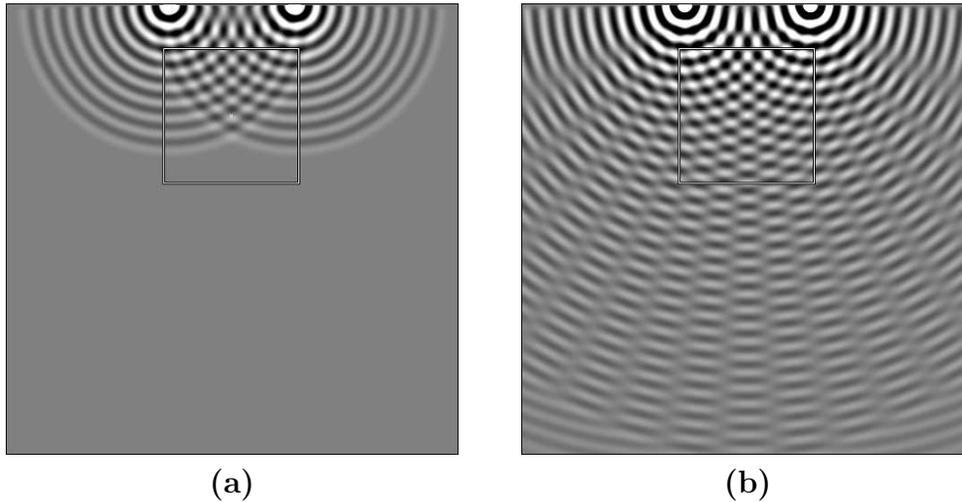
Figure 3.33:   **(a)** *Sinusoid of 2000 Hz emitted by 2 conventional stereo loudspeakers,* **(b)** *same as (a) but ≈ 10 ms later.*

If the loudspeakers are moved far away, the wave fronts converge to plane waves inside the listening area. This can be seen in Figure 3.34.

In practice it is not possible to move the loudspeakers far away, because of physical constraints in the reproduction room. With a wave field synthesis setup, it is possible to simulate the notional sources far away and the wave fronts will be produced like in Figure 3.35. Both simulations were made without truncation windows.
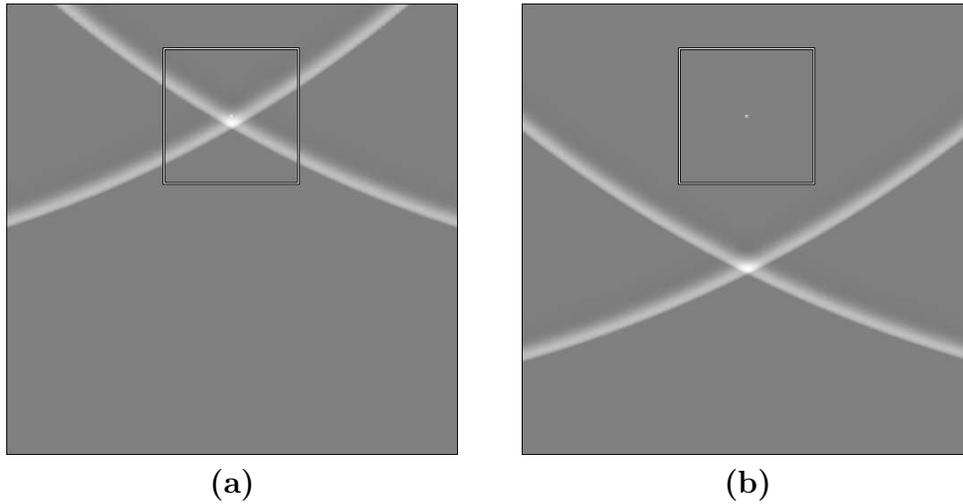
Figure 3.34: **(a)** *Pulse low-pass filtered with a cutoff frequency of 1000 Hz emitted by 2 conventional stereo loudspeakers at 10 m distance,* **(b)** *same as (a) but ≈ 4 ms later.*
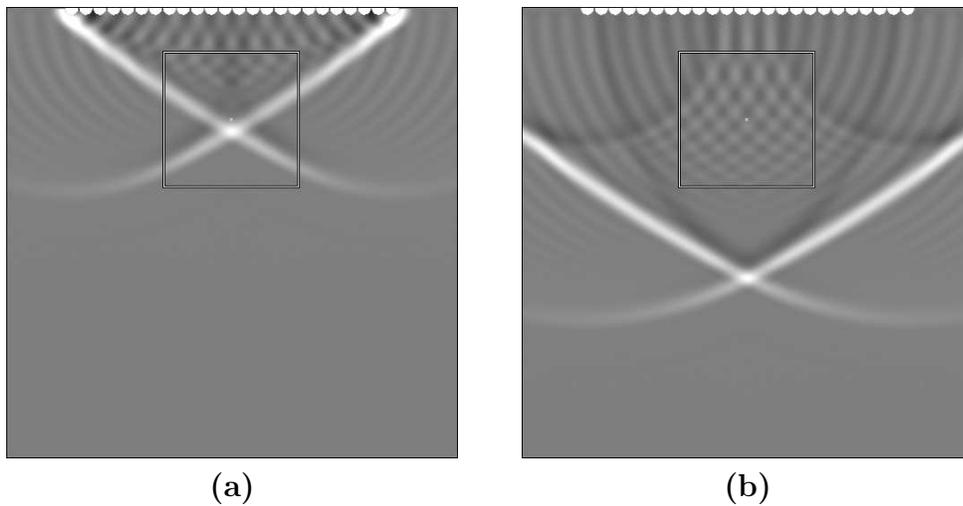
Figure 3.35: *Wave field setup with 24 loudspeakers in an array reproducing a* **(a)** *pulse low-pass filtered with a cutoff frequency of 1000 Hz emitted by 2 virtual sources in 10 m distance* **(b)** *same as (a) but ≈ 4 ms later.*

4

Wave field Synthesis Setup

4.1   Setup configuration

Free field environment allows more accurate measurements on the synthetized wave field. Therefore, an anechoic room is used. Figure 4.1 shows the WFS setup.

The wave field synthesis setup consist of an array of 24 loudspeakers (Figure 4.2). Loudspeakers are fixed on a metal bar with a length of $4m$ by three wires hung on hooks in the ceiling. Due to the position of the hooks, the array is placed from $20cm$ from the wall. According to the properties of the anechoic room, to be so close from the wall is not a problem for the experiments.

Due to the aliasing effect, the high frequency limit is based on the space between the loudspeakers. In order to have a larger frequency range, the space between the loudspeakers has to be minimized. $154mm$ is the diameter of a loudspeaker. In order to minimize the influence of vibration between the loudspeakers, a small piece of acoustilux has been placed between the loudspeakers. Its thickness is 1mm. Hence, the inter loudspeaker distance is $155mm$. This value gives an aliasing frequency of $1132Hz$. As shown in Figure 4.3, the loudspeakers are mounted on the bar by screwing them on T-connectors.

Two parameters are fundamental for defining the listening area: the length of the array L and the maximum distance to the virtual sources D, as shown in Figure 4.4.
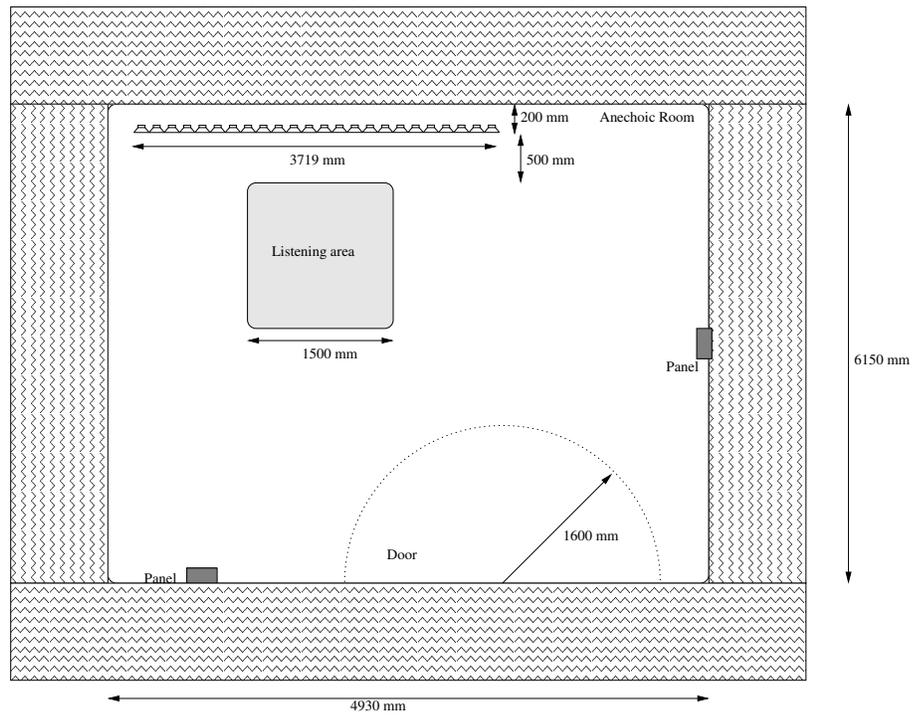
Figure 4.1: *Configuration of the setup in the anechoic room.*
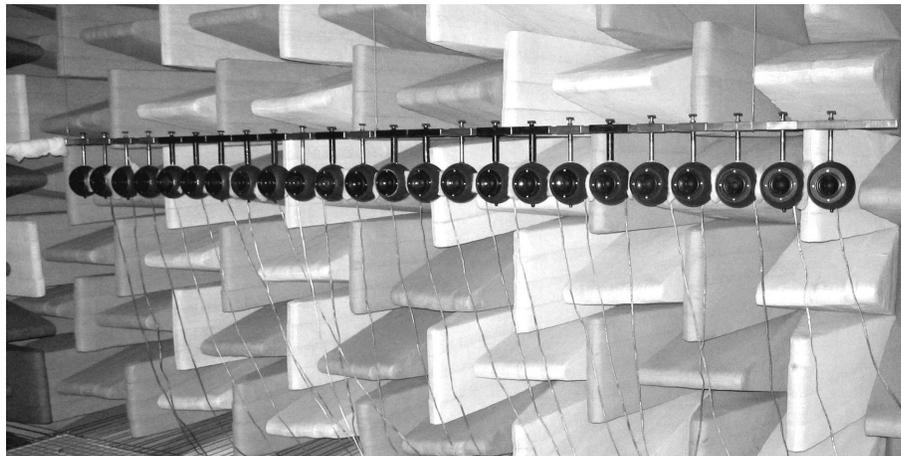
Figure 4.2: *The loudspeaker array.*
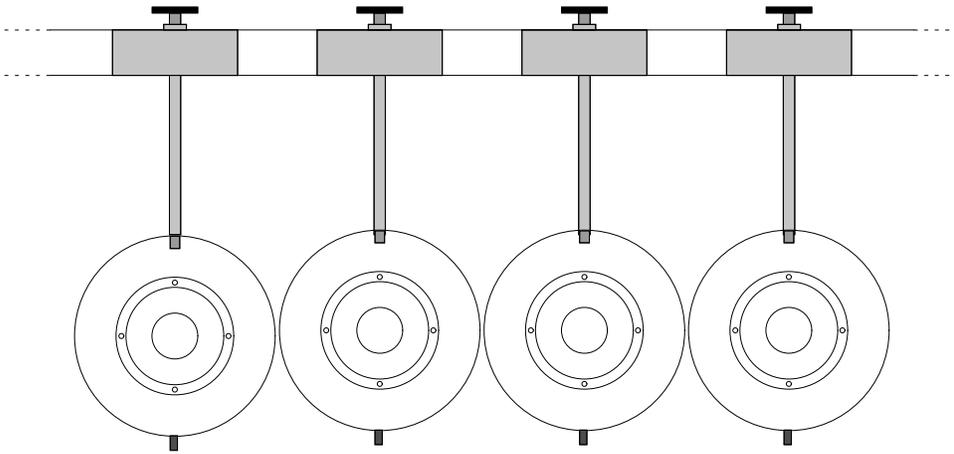
Figure 4.3: *Loudspeakers mounted on a bar with T-connectors.*
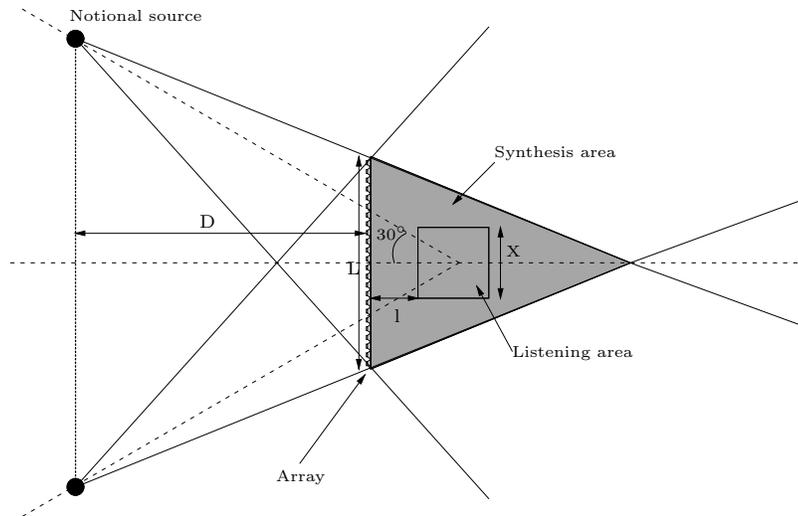
Figure 4.4: *The synthesis area depends of the place of the notional sources and the size of the array.*

The bigger the distance between the notional source and the array is, the smaller is the synthesis area. The larger the array is, the wider is the synthesis area. When the distance between the notional sources and the array tends to infinity, the size of the synthesis area becomes limited. The equation (4.1) gives the side length X of the listening area when the notional source is far away.

$$\lim_{D \to \infty} X = \frac{L - 2 \times l \times \tan 30}{1 + 2 \times \tan 30} \qquad (4.1)$$

D is the distance between the array and the plane where the notional sources are. L is the width of the array. l is the distance between the array and the listening area.

The listening area chosen is a square of $1.5m \times 1.5m$. Its center is situated $1.25m$ from the array.

4.2 General equipment

4.2.1 Presentation of the equipment used

A general presentation of the equipment used can be seen in Figure 4.5. The optical to XLR and XLR to optical converters are needed for transmiting the signals from the control room to the anechoic room. A documentation of these converters is situated in Appendix B on page 86.

4.2.2 Loudspeakers

Knowing the features of each transducer is important in order to calculate the driving signals and to synthesize accurately the sound field. Loudspeaker measurements have been carried out. These measurements are described in Appendix D. The MLSSA measurement system, which is used, is described in Appendix C. This system obtains the frequency response of the loudspeaker

Adobe Audition 1.0
using 24 channels playback

No level adjustment with software.

Sound card RME DIGI9652
3x ADAT optical out

converter optical to XLR

3x optical cables (Toslink)

connection to anechoic chamber

3x 2m XLR/XLR signal cables

using control panel 56 to panel 57
Signal B, sockets 1-3

converter XLR to optical

3x 10m XLR/XLR signal cables

3x DA converters

3x optical cables

BEHRINGER ultragain pro 8 digital

ADA 8000   channel 1-8: 56543-00
             channel 9-16: 56545-00
             channel 17-24: 56551-00

4x amplifiers

24x XLR/RCA cables

ROTEL 6-channel poweramp

RB-976 MK II   channel 1-6: 33980-00
               channel 7-12: 33982-00
               channel 13-18: 33975-00
RB-976   channel 19-24: 33973-00

All volume controls were set to the lowest horizontal value.

24 XLR/XLR speaker cables

loudspeaker array

array with 24 loudspeakers

Figure 4.5: *Reproduction chain.*

(Figure D.2 on page 102). This plot shows that the frequency response of this type of loudspeaker is limited. The frequency response is not flat in a wide range of frequencies. However, it is acceptable for the frequency range needed. This plot shows the lowest usable frequency (around $300Hz$).
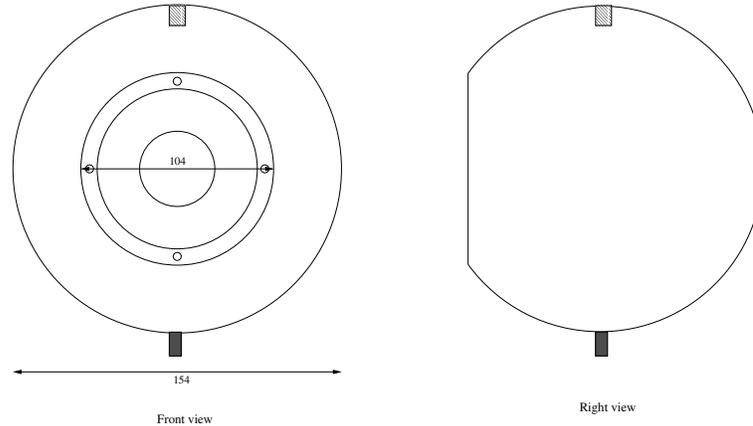
Figure 4.6: *The Loudspeaker.*

The directivity of one loudspeaker is measured with a step-size of 2.5°. As shown in Figure D.4 on page 104, this transducers is relatively omnidirectional for low frequencies (around $500Hz$). The directivity has to be used for calculating the driving signals.

4.3 Calibration of the setup

To get a reproduceable result out of the measurements, it is necessary to calibrate all 24 channels including loudspeakers, amplifiers and play-back equipment to the same output level. The sound pressure level measured at $1m$ distance in front of each loudspeaker, one at a time, using a $10s$ time average. Band limited white noise that is within the frequency range of loudspeaker is used. 24 equal driving signals are reproduced in Adobe Audition by creating white noise and using a 6th order Butterworth bandpass filter from $400Hz$ until $10kHz$. The signal is chosen to measure the overall gain of each channel. A wider frequency range would cause higher losses inside the loudspeaker and could give rise to distortion.The measurement is made with the B&K 2238 Mediator SPL meter (No. 33948-00). The measures sound pressure levels are listed in Table 4.1 together with the calculated correction factors

$A_{correction}(n)$. The AAU number in the table gives the last two digits of (02017-xx).

loudspeaker n	1	2	3	4	5	6	7	8
AAU number	32	25	24	01	09	08	27	30
$L_{Leq}(n)$ / dB	71.3	70.0	72.3	70.9	70.0	68.4	69.7	69.3
$A_{correction}(n)$	0.63	0.73	0.56	0.66	0.73	0.88	0.76	0.79
loudspeaker n	9	10	11	12	13	14	15	16
AAU number	29	18	17	00	11	06	03	36
$L_{Leq}(n)$ / dB	71.7	70.7	71.1	71.2	70.8	71.1	70.2	70.4
$A_{correction}(n)$	0.60	0.68	0.65	0.64	0.67	0.65	0.72	0.70
loudspeaker n	17	18	18	20	21	22	23	24
AAU number	27	31	12	16	10	13	23	26
$L_{Leq}(n)$ / dB	71.3	69.8	67.9	67.3	68.0	69.0	68.0	68.5
$A_{correction}(n)$	0.63	0.75	0.93	1.00	0.92	0.82	0.92	0.87

Table 4.1: Measured sound pressure levels at 1m distance in front of each loudspeaker and correction factors.

L_{Leq} is equivalent sound pressure level averaged over $10s$ without any frequency weighting. $A_{correction}$ is the amplitude correction factor with which each driving signal will be multiplied. The reference is given by channel 20 with the lowest output value $L_{Leq}(20) = L_{Leqref} = 67.3dB$. $A_{correction}$ is defined as:

$$\frac{1}{A_{correction}(n)} = 10^{\frac{L_{Leq}(n) - L_{Leqref}}{20}}$$

Experiments 5

Experiments were carried out to verify some characteristics of the synthesized wave field, and the assumption made concerning the enlargement of the stereo "sweet spot".

Four things were tested :

- The evolution of the pressure amplitude over the listening area.

- The time delay between left and right signals.

- The incidence angle of the wave fronts.

- The stereophonic image through an informal listening test

5.1   Measurements

5.1.1   General informations

The appropriate signals were played back by the wave field synthesis system, and recorded with a microphone to be analyzed.

For the theory (Chapter 2), as for the simulations (Chapter 3), it was assumed that the wave field synthesis system is situated in a free field environment, avoiding the effects introduced by reflexions.

The signals used are:

- $600Hz$ tone played by a source placed $10cm$ behind the array.

- $600Hz$ tone played by a source placed $50m$ behind the array.

- $2kHz$ tone played by a source placed $10cm$ behind the array.

- $2kHz$ tone played by a source placed $50m$ behind the array.

- Speech signal [16] at left stereo position, placed $10m$ behind the array.

- Speech signal [16] at right stereo position, placed $10m$ behind the array.

The position of the six virtual sources synthesized are shown in Figure 5.1.
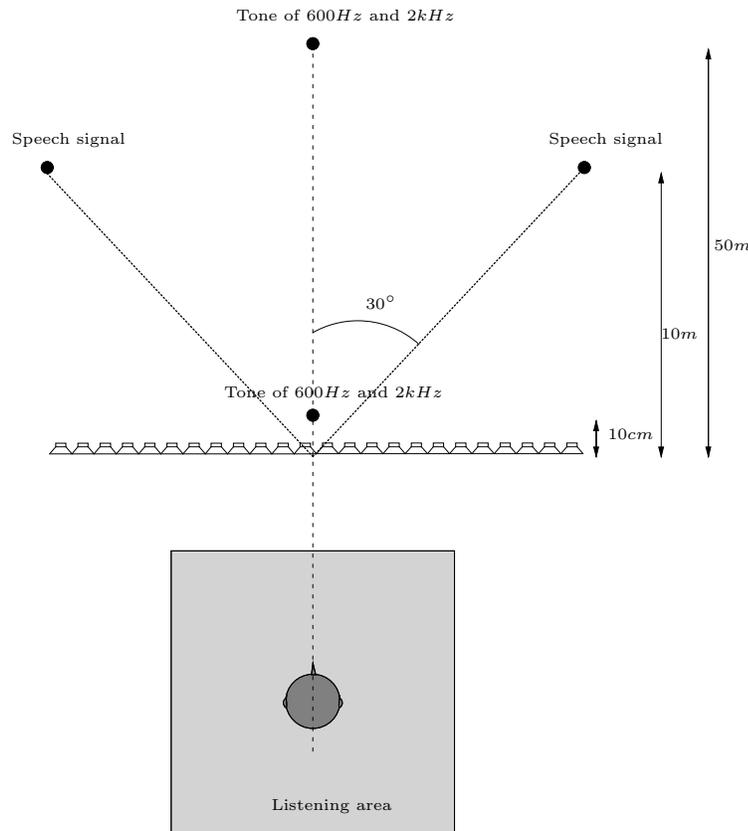
Figure 5.1: *Four different positions of the virtual sources.*

Note that all these signals were low-pass filtered, and the truncation window was not used when the driving signals were created.

The cut off frequency for a tone is equal to its frequency, as explained in Section 3.2.2 on page 30.

The speech signals are low-pass filtered with a cut-off frequency of $4kHz$.

From the simulations,the wave fields of the signals can be predicted, and then, a comparison can be made between the simulations and the results. The expected wave fields for the tones,are shown in Figures 5.2 and 5.3.

(a)   (b)
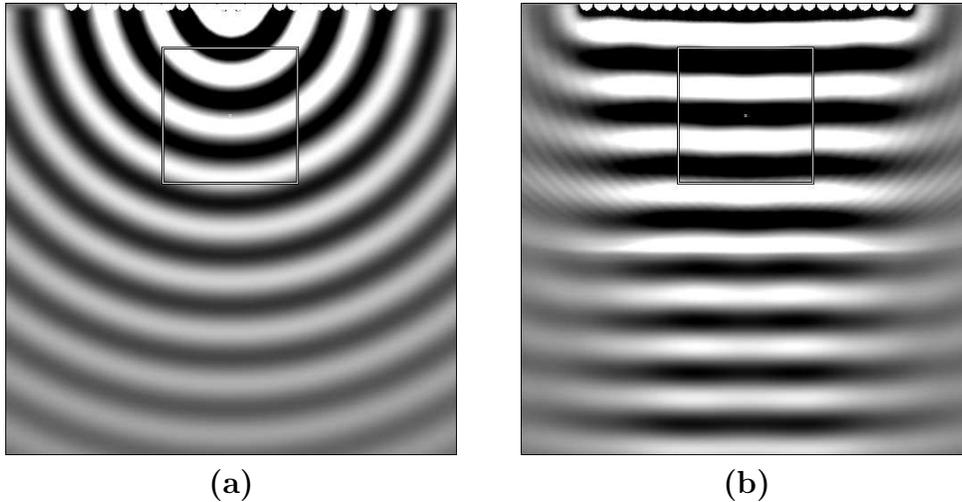
Figure 5.2: *Expected wave field of a source situated at* **(a)** $10cm$ *and* **(b)** $50m$ *behind the array, playing a* $600Hz$ *tone.*

Recording the tones at discrete points gave the necessary data to study the evolution of the pressure amplitude over the listening area. With such signals, a comparison with the simulations was possible. Two kind of tones were played back by the system, one of them, with a frequency high enough to create artefacts such as spatial aliasing. Unfortunately, the spatial sampling of the listening area was too low to actually measure these artefacts.

The analysis of the reproduced speech signals showed the actual wave front emitted by the array, and also showed the propagation time of the wave within the listening area. Thus, the time delay between two waves emitted by sources at stereo position could calculated. Speech signals allows the calculation of the incidence angle of the wave fronts.

An informal listening test was realized. First, the listener hears the speech
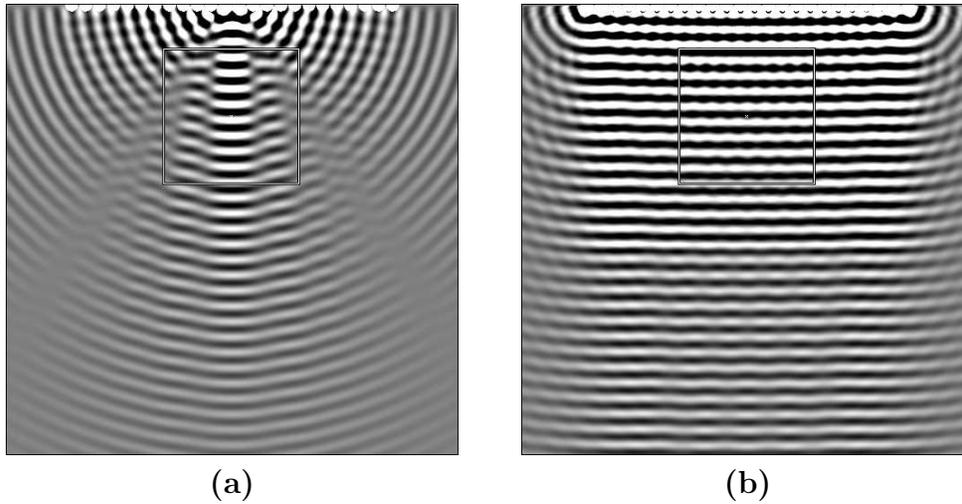
Figure 5.3: *Expected wave field of a source situated at* **(a)** *10cm and* **(b)** *50m behind the array, playing a 2kHz tone.*

signals played by the notional sources situated at stereo position, $10m$ behind the array. Then, a moving source, playing a tone was synthesized.

5.1.2   Measurement protocol

The measurement were performed in the anechoic chamber, room B4-111 at AAU. Most of the equipment for the wave field synthesis system (i.e the power amplifiers and the D/A converters) were situated inside the anechoic chamber due to the limited number of connections to the control room. The remaining equipment (i.e the computer, the measuring amplifier and the A/D D/A converter) was situated in the adjacent control room. See Figure 4.1 on page 56 to have an overview of the setup.

The signals sent to the system were composed by 12 stereo signals which were generated by the driving function program (see Appendix E). The stereo signals were separated and routed to 24 loudspeakers. The six different kind of signals were played one after the other, with a short pause in between.

They were then recorded, using the recording chain illustrated in Figure 5.4 and composed by:
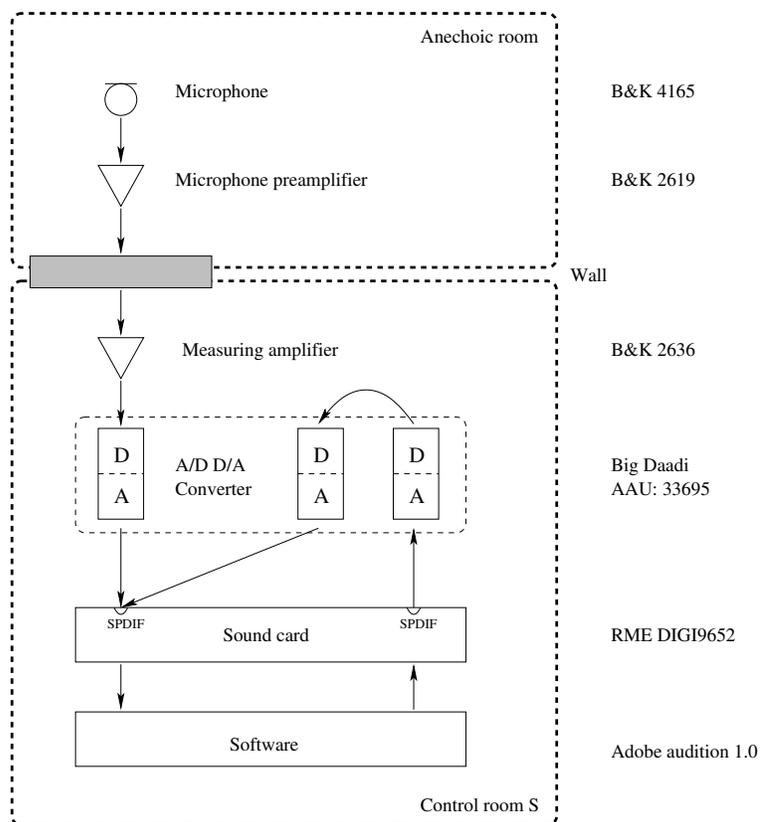
Figure 5.4: *The recording chain*

The used equipment for recording consisted of:

- A free field microphone B&K type 4165 (and pointing towards the ceiling)

- The corresponding microphone preamplifier B&K type 2619

- A measuring amplifier B&K type 2636

- A A/D D/A converter Trace Technologies type BIG DAADI

- A computer with the sound card RME DIGI 9652 (containing 3 optical I/O of 8 channels each and one SPDIF I/O of 2 channels each)

- The software Adobe Audition 1.0

A noise signal is played back simultaneously and then recorded by the computer, to be used as a time reference. Then, the analysis of the recorded noise signal is performed to detect possible delay differences introduced by the recording chain.

The signal synthesized by the system was recorded at several listening points within the listening area. A grid of $1 \times 1m$ was used to set the different listening positions. As shown in Figure 5.5, the microphone was moved by intervals of $20cm$, and $10cm$ for the last row and the last column, so that the entire area defined in Section 4.1 is covered. The signal has been recorded at 81 different points.

The microphone was placed above the grid with the help of a laser pointing downwards to the grid, small errors were expected due to the positioning of the grid with respect to the position of the array and the positioning of the microphone from the position of the grid. This error was considered to be $\pm 1cm$. The microphone itself has a width of 1/2", which also introduces a small error.
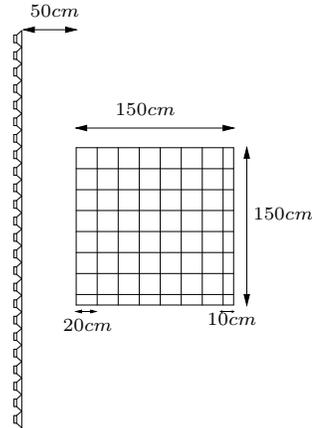
Figure 5.5: *The grid shows the different places where the microphone was positioned.*

5.2 Results

5.2.1 Amplitude

An analysis of the amplitude for the different signals was made in order to evaluate the synthesized sound field in terms of the amplitude variations over the listening area, and to visualize the artefacts introduced by the system.

The amplitude of the tones is calculated at each measuring position of the listening area, and the results are plotted in the Figures 5.6, 5.7, 5.8 and 5.9. Similar plots are made in the simulations, but with doubled resolution, and are shown here for comparison with the results.

In Figure 5.6, a spherical wave field is created with a frequency below the aliasing frequency, emitted by a notional source close to the center of the line array. A level roll-off can be observed in the x and z direction. A smooth amplitude decay is expected, but ripples appear in the measurements. These ripples are a result of constructive and destructive interference between the loudspeakers in each listening position. For the simulation the loudspeakers are assumed identical. In practice their position and behaviour deviates from the ideal.

If the source is moved further away, the amplitude fluctuation is expected to increase (see Figure 5.8a). This is verified by the measurements, as can be
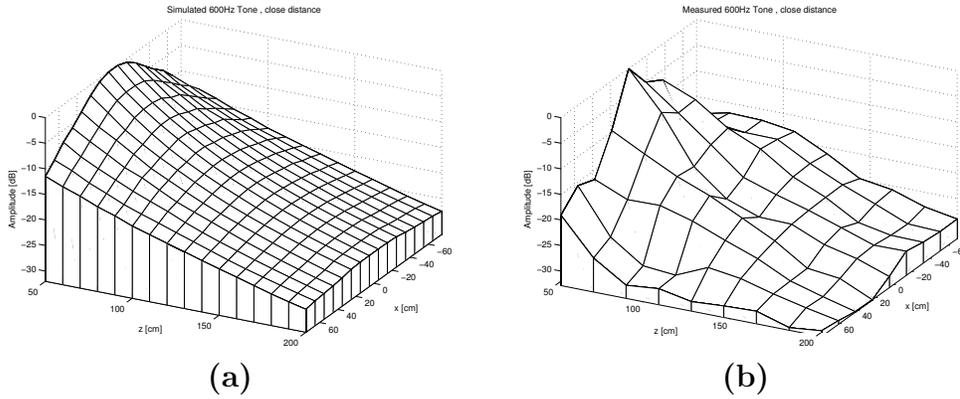
Figure 5.6: *Variations of the amplitude of a 600 Hz tone played by a source close to the array. Plot* **(a)** *was obtained from simulation, and plot* **(b)** *was generated with the measurements.*

seen in Figure 5.8b.
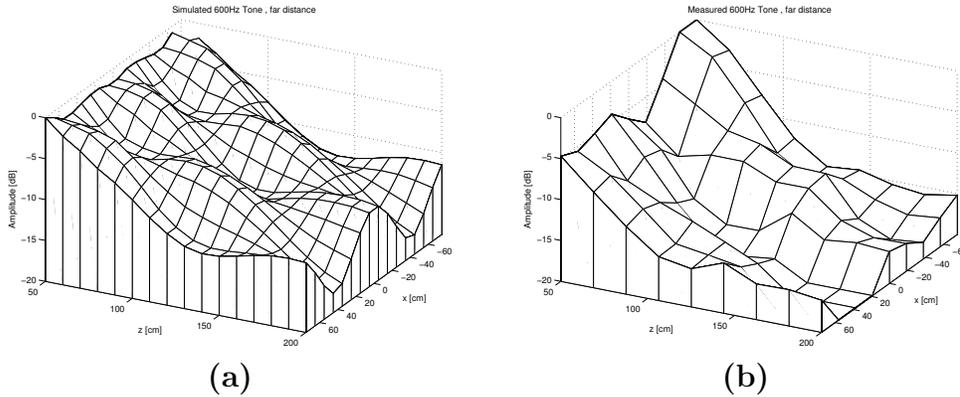
Figure 5.7: *Amplitude variations of a 600 Hz tone played by a far away source. Plot* **(a)** *was obtained by the simulation, and plot* **(b)** *was generated with the measurements.*

In Figure 5.7, an overall level roll-off appears only in the z direction, because the spherical component in the horizontal plane decreases as the source is positioned further away.

In Figure 5.8, the level roll-off in the x direction seems bigger than in Figure 5.6, but in fact, in both cases, spherical waves of identical shape are supposed to be synthesized, so that the level roll-off should remain the same.

69
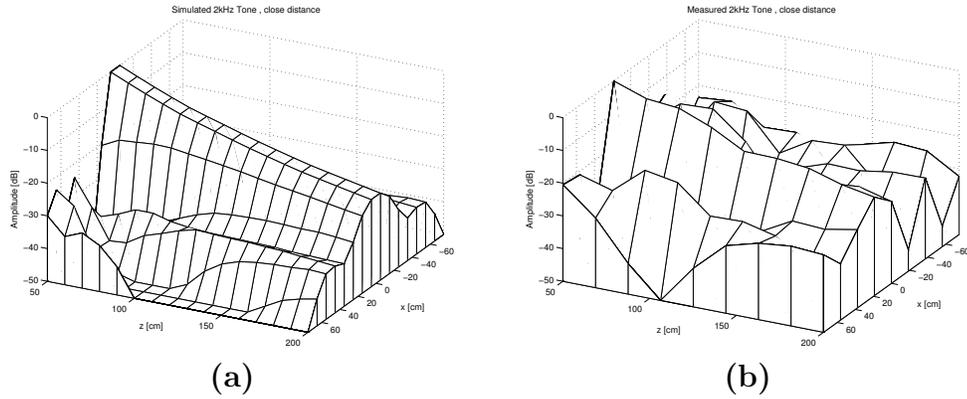
Figure 5.8: *Amplitude variations of a 2 kHz tone played by a source close to the array. Plot* **(a)** *was obtained by the simulation, and plot* **(b)** *was generated with the measurements.*

The interference increase is due to the high frequency content, so that an interference pattern overlays and the amplitude variations increase.
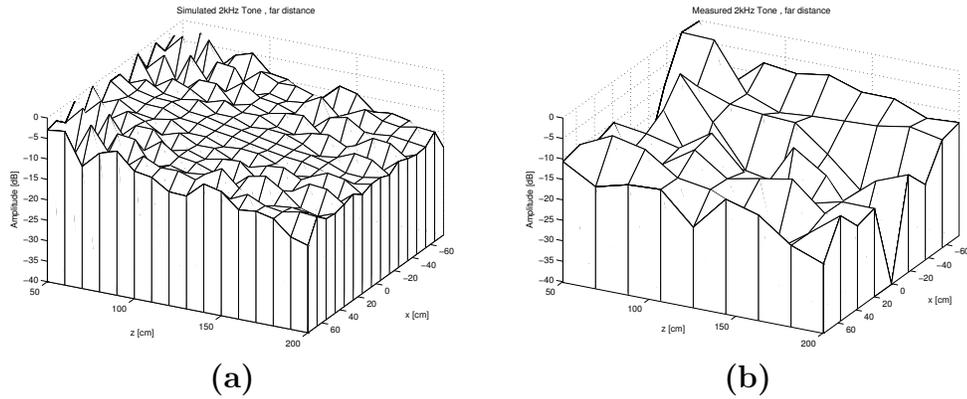
Figure 5.9: *Amplitude variations of a 2 kHz tone played by a far away source. Plot* **(a)** *was obtained by the simulation, and plot* **(b)** *was generated with the measurements.*

In Figure 5.9, the source is moved far away. The ripples are even more numerous than in the previous figure. The resolution in both pictures is too low to visualize the resulting interference patterns.

Note that these interference patterns are not necessarily the same as spatial aliasing or truncation, but are influenced by these effects.

5.2.2 Time delay Evaluation

In a classical two-channel stereo system, the time delay Δt is defined as the difference between the time propagation of the signals coming from the left and the right loudspeakers.

It was assumed in Subsection 2.5.2 that one of the advantages of having plane waves instead of spherical waves in a stereo system, was that they reduce the variation of the time delay between the right and left signals when the listener moves within the listening area.

Two matrices of time propagation data were calculated with the speech signals. Then, the time delays were obtained by calculating the difference of those two matrices.

In order to determine both matrices of time propagation data, the time delay on each recording of the same speech signal at two different position is calculated. All speech signals are time referenced to the noise signal. By fixing a reference point, such as one of the corners of the listening area, a matrix of the time differences between the reference point and the other recording points was obtained. To obtain this result, the cross correlation of both signals was calculated. Figure 5.10 shows the value of the time differences between the recorded signal in the top left corner and the other 80 recorded signals. In this case, the notional source was placed at the right stereo position. Figure 5.10 illustrates the propagation of the signal within the listening area.
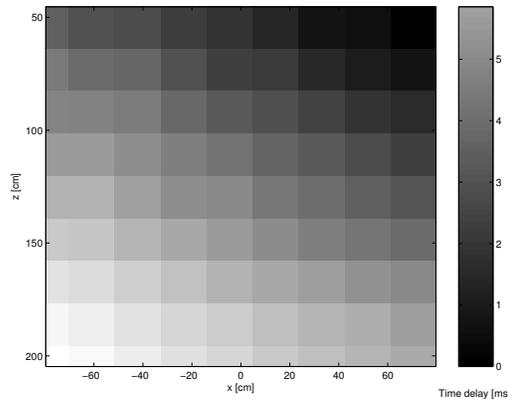
Figure 5.10: *Plot representing the propagation of the right speech signal.*

For both notional sources, a matrix of the time propagation was calculated . The top left corner and the top right corner recording points were respectively

chosen as the two reference points for the right and the left speech signals. Considering the symmetry of the setup, the distances between each notional source and the corresponding reference position are equal. Thus, the time delay between the two stereo signals was obtained by subtracting the time differences at each position.

The delay of the wave coming from the right notional source with respect to the wave coming from the left notional source is represented in Figure 5.11, and in a 3D plot in Figure 5.12.
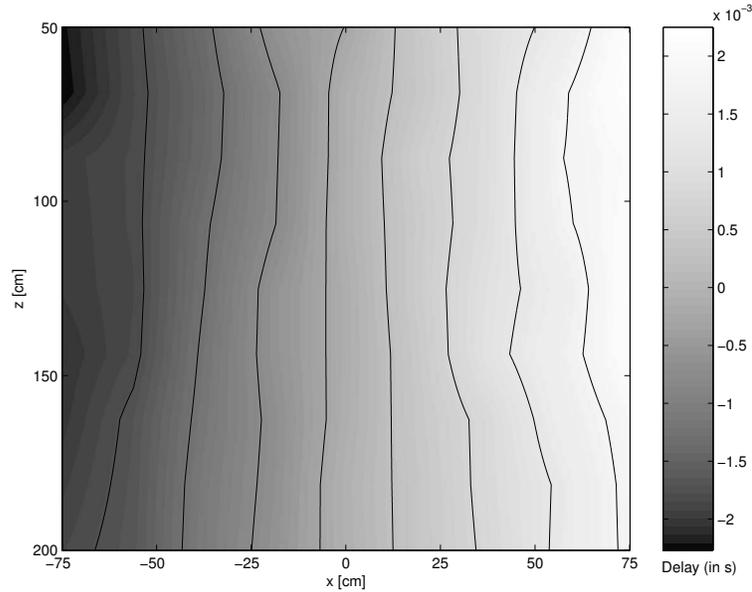
Figure 5.11: *Evolution of the delay of the right signal compared to the left signal over the listening area.*

Because the measurements were carried out in a free field environment (in an anechoic room), no errors due to reflexions are introduced while the time delay is calculated. When the listener is situated in the symmetry axis, there is no delay between the two signals. When the listener moves in the perpendicular direction of that axis, the absolute value of the delay increases and reaches a maximum of around $2ms$ when the listener is located at one side of the listening area.

The lines of equal delay are parallel to the symmetry axis of the listening area. Then the assumption that the waves emitted by the notional sources are, in the listening area, plane waves, is validated.

(a)   (b)
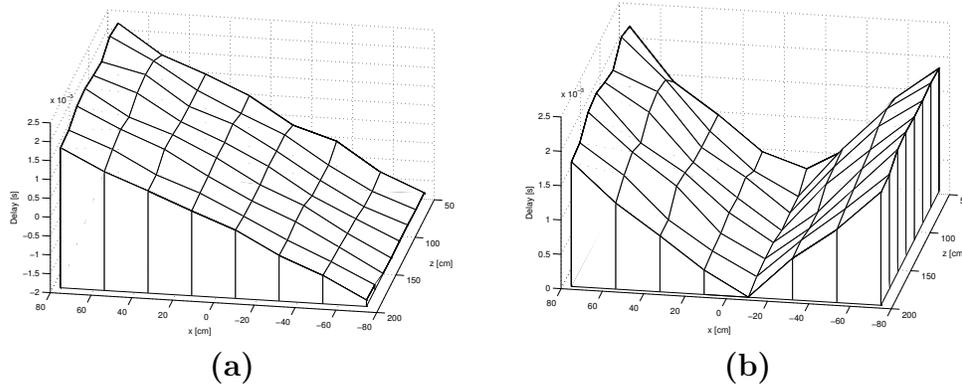
Figure 5.12: *The delay of the right wave compared to the left one is plotted in* **(a)** *relative value* **(b)** *and in absolute value.*

5.2.3   Incidence Angle of the wave fronts

The mean incidence angle of the wave front is calculated from the time delay matrix. The method used, is to calculate the mean slope of all columns and the mean slope of all rows. The angle between these two slopes will then give the approximate incidence angle[1].

As it is illustrated in the Figure 5.1 on page 63, the incidence angle of both speech signals at the center of the array is equal to 30°. Theoretically, the extreme values of the propagation angle are equal to 22.7°, and 31.9° due to not completely plane waves. The results were compared with the mean of the incidence angle at all the recorded points. The mean is equal to 26.9°.

Figure 5.13 illustrates the propagation of the two speech signals. Both matrices shows the time differences and the incidence angle of propagation for both speech signals.

The mean values measured were respectively equal to 27.8° and 29.5° for

---

[1]To calculate the slope of a row or column, a line $y = a + bx$ is fitted to n points with coordinates $(x_1, y_1), \cdots, (x_n, y_n)$ using the following normal equations: $an + b \sum x_j = \sum y_j$; $a \sum x_j + b \sum x_j^2 = \sum x_j y_j$. These equations will calculate the least squares best fit of the line $y = a + bx$ to the data points.

The method is repeated for all rows and colums to obtain a mean slope in both the x and z-direction. The angle between these slopes gives the approximate incidence angle, as is calculated in the following manner: $\theta_{measured} = \arctan\left(\frac{b_z}{b_y}\right)$.
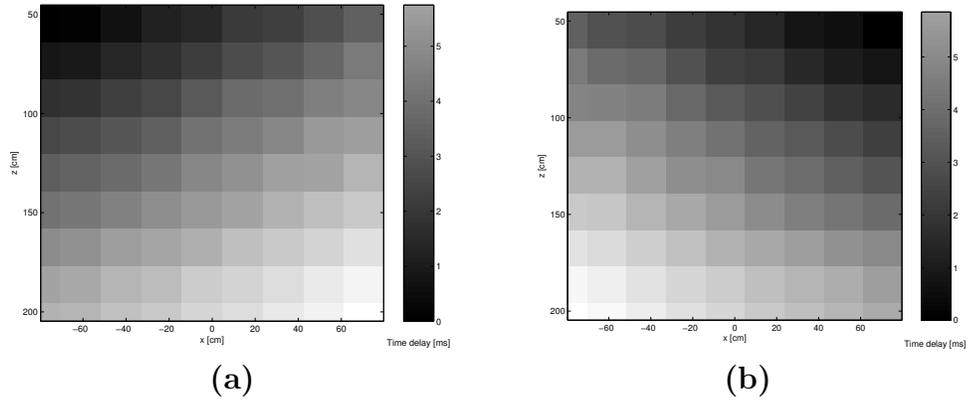
Figure 5.13: *These plots represent the time propagation of each speech signal* **(a)** *notional source left,* **(b)** *notional source right.*

the left and the right speech signals. The incidence angles were close to the expected values of 26.9°. The mean measured values lie within the theoretical range.

5.2.4 Informal listening test

The last part of the experiments consisted of an informal listening test in order to subjectively evaluate the performance of the wave field synthesis system.

Both speech signals were played separately. When moving to the left and right within the listening area, the source is perceived as a virtual source placed behind the array at the stereo position. Thus, the impression is consistent with the theoretically expected result.

A moving source was played. The signal used was a tone. In the beginning, the source was situated $10m$ behind the array on the left and moved progressively approaching the right side of the array. At different listening points, the impression given to the listener was that the source moved as it was expected.

In both tests, the acoustic experience was consistent with the results expected. The image of the created virtual source was stable and in a specific location.

Conclusion and Discussion

6

This report presents an approach to the realization of a sound reproduction system based on the concept of wave field synthesis (WFS).

6.1 Conclusion

The report begins with a description of the underlying theory in Chapter 2, where it is shown that it is possible to control the sound field within a volume with a distribution of sound sources on a surface enclosing the volume. Several adaptations are performed to simplify the theory and make it realizable with a line array of loudspeakers, but some artefacts and limitations appear. The system becomes limited by truncation effects, amplitude errors and spatial aliasing. Methods exist to reduce these artefacts, and are studied in Chapter 3.

A simulation program has been created to visualize the behaviour of a WFS system composed by a line array of loudspeakers, playing independent signals processed by the driving function program. Through simulations, the artefacts and different methods to reduce them have been studied.

It is shown that spatial aliasing effects can be reduced by increasing the directivity of the array loudspeakers. This method omits signal contributions which could cause errors, but reduces the possible number of positions for the notional source. Another method is to randomize the phase of the high frequencies. The periodicity of the aliasing effect is removed, but random noise is introduced to the system.

To avoid truncation errors, it is shown in simulations that it is efficient to apply a tapering window to the array, even though it reduces the synthesis area. Another method which consists of feeding the edge loudspeakers with the inverse of the truncation error at a reference point is only suitable for that specific point. As seen in the simulations, it is not as effective as tapering window.

The simulations show that the amplitude decreases with distance due to the property of a line array. The roll-off is less than that of a spherical source.

A wave field synthesis system has been implemented in an anechoic room, and it can synthesize notional sources at different positions. Measurements have been carried out to verify the simulations. The results confirm that the incidence angles of the wave fronts correspond with the expected value. Within the limitation of the measurements it is also shown that the amplitude roll-offs are as expected.

One goal of the project was to improve stereo imaging by using WFS, and more specifically to enlarge the "sweet spot". With sources far away, waves received can be approximated by plane waves in the listening area. It is shown that the amplitude, the incidence angle and the time delay between the two stereo signals, vary less over the entire listening area, compared to a traditional stereo setup. With WFS, it is thus possible to enlarge the "sweet spot" by synthesizing sources far away. However, the reproducable frequency range of a WFS system is limited because of spatial aliasing.

6.2 Discussion

A Wave Field Synthesis system can improve a stereo system, but is limited by the artefacts that appear in practice.

In order to reduce the truncation effects, one possible way of enhancement is to add side arrays or to create a circular array of loudspeakers. Due to the fact that a circular array has no edges, the truncation effects can be totally avoided.

The main limitation of a WFS system remains the spatial aliasing. To reduce it, it is necessary to get closer to a continuous distribution of sound sources.

A solution using several arrays for the reproduction of different frequency bands can be a possible way to enhance the system, but such a solution is quite expensive. Another solution could be the use of Discrete Mode Loudspeakers (DML), as suggested in [17]. They consist of small exciters attached to a panel. DML arrays are suitable to WFS, but the broad directivity characteristic increases spatial aliasing compared to a loudspeaker array solution. According to [17], it seems that no distracting effect due to spatial aliasing could be perceived.

In this project, WFS was applied to stereo reproduction, but it has been shown that sound field from any notional source behind the array can be reproduced. The main application of WFS is then the reproduction of entire sound scenes. Individual instruments can be placed accurately in the desired positions, so the rendering of an orchestral concert becomes more realistic.

Bibliography

[1] A. J. Berkhout. *A holographic approach to acoustic control.* J. Audio Eng. Society, Vol.36, 1988, pp977-995, 1988.

[2] Marinus M. Boone. *Acoustic rendering with wave field synthesis.* Lab. of acoustical imaging and sound control, Delft, The Netherlands, 2001.

[3] S. Spors, H. Teusch, and R. Rabenstein. *High-Quality Acoustic Rendering with Wave Field Synthesis.*

[4] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders. *Fundamentals of Acoustics Fourth Edition.* John Wiley & Sons, 2000.

[5] Eric W. Weisstein. *MathWorld – A Wolfram Web Resource, Solid Angle.* http://mathworld.wolfram.com/SolidAngle.html verified 1/6-2004.

[6] D. de Vries. *Sound Reinforcement by Wavefield Synthesis: Adaption of the Synthesis Operator to the Loudspeaker Directivity Characteristics.* Laboratory of Seismics and Acoustics, Delft University of Technology, Delft, The Netherlands, 1996.

[7] D. de Vries et. al. *The wavefield synthesis concept applied to sound reinforcement: restrictions and solutions.* Preprint No.3812, 96th AES Convention, Amsterdam, 1994.

[8] E. Start. *Direct Sound Enhancement by Wave Field Synthesis.* Dissertation, Delft University of Technology, Delft, The Netherlands, 1997.

[9] H. Wittek. *OPSI - A Hybrid WFS / Phantom Source Solution to Avoid Spatial Aliasing.* Institut für Rundfunktechnik, Munich, Germany, 2002. http://www.irt.de/wittek/hauptmikrofon verified 1/6-2004.

[10] J. J. Sonke. *Variable acoustics by wave field synthesis.* Thela Thesis, Amsterdam, The Netherlands, ISBN 90-9014138-3, 2000.

[11] A. J. Berkhout, D. de Vries, and P. Vogel. *Acoustic control by wave field synthesis.* Laboratory of Seismics and Acoustics, Delft University of Technology, Delft, The Netherlands, 1993.

[12] D. de Vries, E. Start, and Valster. *The wave field synthesis concept applied to sound reinforcement: restrictions and solutions.* Preprint No.5370, 110th AES Convention, Amsterdam, 1994.

[13] C. P. A. Wapenaar. *On aperture artefacts.* Amoco Production Company, 1991.

[14] F. L. Wightman and D. J. Kistler. *The dominant role of low-frequency interaural time differences in sound localization.* Journal of the Acoustical Society of America, Vol.91(3), 1991.

[15] De Vries, W. Start, and G. Valstar. *The wave field synthesis concept applied to sound reinforcement: Restrictions and Solutions.* Laboratory of Seismics and Acoustics, Delft, The Netherlands.

[16] Bang & Olufsen. *Music for Archimedes - Track 5.* Bang & Olufsen, 1992.

[17] Marinus M. Boone and Werner P.J. de Bruijn. *On the applicability of distributed mode loudspeaker panels for wave field synthesis based sound reproduction.* Lab. of acoustical imaging and sound control, Delft, The Netherlands, 2000.

[18] H. Wittek. *Perception of Spatially Synthesized Sound Fields.* Institut of Sound Recording, University of Surrey, Guildford, Surrey, UK; Institut für Rundfunktechnik, Munich, Germany, 2003. www.hauptmikrofon.de/wittek.html verified 1/6-2004.

[19] M. Tuckerman. *The stationary phase approximation.* 1999. http://www.nyu.edu/classes/tuckerman/stat.mech/lectures/lecture_16/node2.html verified 1/6-2004.

Glossary

Virtual source: Imaginary source synthesized by the array.

Notional source: : similar to virtual source.

Phantom source: The imaginary source position as perceived in stereophony by the interpretation of the two loudspeaker signals.

Sweet spot: The area where a correct stereo image is perceived.

Synthesis area: Area where the sound field is correctly reproduced.

Listening area: Defined area inside the synthesis area used for the simulations and the measurements.

Stability [18]: "The degree to which the perceived location of a source changes with time".

Driving Function

Mathematical manipulations of the derivation of the driving function are described in this appendix. The method of the stationary phase approximation is presented. Then, by using this mathematical method, the calculation of the driving function is described

A.1   Stationary phase approximation

The stationary phase approximation is a method for evaluating the following integral I [19]:

$$I = \int_{-\infty}^{-\infty} e^{-jf(x)} dx$$

First, assume that $f(x)$ has a minimum at $x = x_s$, such that $f'(x_s) = 0$. This minimum is well separated from the other minima of $f(x)$ and its value is significantly lower than the value of the other minima. Thus, the dominant contribution of the above integral will come from the integration region around x_s, which is called the stationary phase point.

$f(x)$ can be approximated about this point by using a Taylor series expansion:

$$f(x) \approx f(x_s) + f'(x_s)(x - x_s) + \frac{1}{2}f''(x_s)(x - x_s)^2$$

Since $f'(x_s) = 0$, $f(x)$ becomes:

$$f(x) \approx f(x_s) + \frac{1}{2}f''(x_s)(x - x_s)^2$$

Inserting the expansion into the expression for I gives:

$$I = e^{-\jmath f(x_s)} \int_{-\infty}^{-\infty} e^{-\frac{\jmath}{2}f''(x_s)(x-x_s))^2} dx$$

Which leads to:

$$I = \sqrt{\frac{2\pi}{\jmath f''(x_s)}} e^{-\jmath f(x_s)}$$

A.2   Calculation of the driving function

The driving function is defined by:

$$Q(\vec{r_n}, \omega) \;\; = \;\; A_n(\vec{r_n}, \omega)S(\omega)\frac{e^{-\jmath k|\vec{r_n}-\vec{r_m}|}}{|\vec{r_n} - \vec{r_m}|} \tag{A.1}$$

The geometry is described in Figure 2.6 on page 15.

In order to find the weighting function $A_n(\vec{r_n}, \omega)$, the following equation (equivalent to equation (2.19 on page 16)) has to be solved.

$$\frac{e^{-\jmath k|\vec{r}-\vec{r_m}|}}{|\vec{r} - \vec{r_m}|} \;\; = \;\; \sum_{n=1}^{N} \left[\frac{A_n(\vec{r_n}, \omega)G(\varphi_n, \omega)}{|\vec{r_n} - \vec{r_m}||\vec{r} - \vec{r_n}|} e^{-\jmath k(|\vec{r_n}-\vec{r_m}|+|\vec{r}-\vec{r_n}|)} \right] \Delta x \tag{A.2}$$

Assuming:

$$I = \frac{e^{-\jmath k|\vec{r}-\vec{r_m}|}}{|\vec{r} - \vec{r_m}|}$$

82

$$\alpha(x_n) = k(|\vec{r_n} - \vec{r_m}| + |\vec{r} - \vec{r_n}|)$$

$$f(x_n) = \frac{A_n(\vec{r_n}, \omega)G(\varphi_n, \omega)}{|\vec{r_n} - \vec{r_m}||\vec{r} - \vec{r_n}|}$$

Equation (A.2) becomes :

$$I = \sum_{n=1}^{N} f(x_n)e^{-\jmath\alpha(x_n)}\Delta x \qquad (A.3)$$

$\alpha(x_n)$ is a rapidly varying function which have a minimum. This minimum is significantly lower than the other value of $\alpha(x_n)$. Hence, the stationary phase approximation can be used. Only the area of the minimum contributes significantly to the integral. x_s corresponds to the point of stationary phase, such that $\alpha'(x_s) = 0$. $f(x_n)$ is a slowly varying function. Therefore, it is assumed that $f(x_n) = f(x_s)$.

Hence, equation (A.3) becomes :

$$I = f(x_s)\sqrt{\frac{2\pi}{\jmath\alpha_s''}}e^{-\jmath\alpha(x_s)} \qquad (A.4)$$

Physically, the stationary phase approximation means that most of the energy radiated by the array to a specific listening position is produced by the loudspeaker of the array which is in the direct path from the source to the receiver.

Mathematical manipulations give the value of the stationary point:

$$x_s = x_m + \frac{|x_l - x_m||z_1 - z_0|}{|z - z_0|} \qquad (A.5)$$

Then, α_s'' is calculated. It is equal to:

83

$$\alpha''(x_s) = k \left[\left(\frac{|x_l + x_m|}{|z - z_0|} \right)^2 + 1 \right]^{-\frac{3}{2}} \frac{|z - z_0|}{|z - z_1||z_1 - z_0|} \qquad (A.6)$$

After some manipulations, $|\vec{r_s} - \vec{r_m}|$ and $|\vec{r} - \vec{r_s}|$ are derivated and are equal to :

$$|\vec{r} - \vec{r_s}| = |z - z_1| \sqrt{\left(\frac{|x_l - x_m|}{|z - z_0|} \right)^2 + 1}$$

$$|\vec{r_s} - \vec{r_m}| = |z_1 - z_0| \sqrt{\left(\frac{|x_l - x_m|}{|z - z_0|} \right)^2 + 1}$$

Thus, $f(x_s)$ can be calculated:

$$f(x_s) = \frac{A_s(\vec{r_s}, \omega) G(\varphi_s, \omega)}{|\vec{r_s} - \vec{r_m}||\vec{r} - \vec{r_s}|} = \frac{A_s(\vec{r_s}, \omega) G(\varphi_s, \omega)}{|z - z_1||z_1 - z_0|[(\frac{|x_l - x_m|}{|z - z_0|})^2 + 1]} \qquad (A.7)$$

At the point of stationary phase, φ_s is equivalent to θ_s and $(|\vec{r_s} - \vec{r_m}| + |\vec{r} - \vec{r_s}|)$ is equal to $|\vec{r} - \vec{r_m}|$.

Using these last results and the equations (A.6) and (A.7), equation (A.4) becomes:

$$I = \frac{A_s(\vec{r_s}, \omega) G(\theta_s, \omega) e^{-jk|\vec{r} - \vec{r_m}|}}{|z - z_1||z_1 - z_0|[(\frac{|x_l - x_m|}{|z - z_0|})^2 + 1]} \sqrt{\frac{2\pi}{j}} \left[\left(\frac{x_l + x_m}{z - z_0} \right)^2 + 1 \right]^{\frac{3}{4}} \sqrt{\frac{|z - z_1||z_1 - z_0|}{k|z - z_0|}}$$

Thanks to the definition of I, the last equation can be simplified into:

$$\frac{1}{|r - r_m|} = \frac{A_s(\vec{r_s}, \omega) G(\theta_s, \omega)}{\sqrt{|z - z_1||z_1 - z_0||z - z_0|}} \sqrt{\frac{2\pi}{jk}} \left[\left(\frac{|x_l + x_m|}{|z - z_0|} \right)^2 + 1 \right]^{-\frac{1}{4}}$$

After some manipulations, the weighting function is obtained:

$$A_s(\vec{r_s}, \omega) = \sqrt{\frac{\jmath k}{2\pi}} \frac{1}{G(\theta_s, \omega)} \sqrt{\frac{|z - z_1|}{|z - z_0|}} \frac{|z_1 - z_0|}{\sqrt{|r_s - r_m|}}$$

Knowing that every loudspeaker can be a stationary point by varying the listening position, the last equation can be generalized for each loudspeaker:

$$A_n(\vec{r_n}, \omega) = \sqrt{\frac{\jmath k}{2\pi}} \frac{1}{G(\theta_n, \omega)} \sqrt{\frac{|z - z_1|}{|z - z_0|}} \frac{|z_1 - z_0|}{\sqrt{|r_n - r_m|}} \tag{A.8}$$

Hence, by substituting (A.8) into (A.1), the driving function is obtained:

$$Q(\vec{r_n}, \omega) = S(\omega) \frac{cos(\theta_n)}{G(\theta_n, \omega)} \sqrt{\frac{\jmath k}{2\pi}} \sqrt{\frac{|z - z_1|}{|z - z_0|}} \frac{e^{-\jmath k|\vec{r_n} - \vec{r_m}|}}{\sqrt{|\vec{r_n} - \vec{r_m}|}} \tag{A.9}$$

B
Documentation
XLR⇔optical
converters

The computer's soundcard has 3 optical ADAT outputs, each one containing 8 channels. There is no optical connectors between the control room, where the computer is located, and the anechoic room, so the signals must be converted into electrical signals to connect the two rooms.

It is not possible to D/A-convert the signals in the control room and transmit 24 analog channels seperately, as it exceeds the number of signal-connectors between the two rooms. To solve this problem, two converters have been made, which converts the optical ADAT signal into an electrical signal which can then be converted back into an optical signal.

With this solution, the signals will have to be D/A-converted and amplified in the anechoic room.

The electrical signals will have to travel through at least 15-20 m of cable from the control room to the anechoic room, so a balanced transmission is preferred to maximize the S/N-ratio and avoid bit-errors.

The two boards have been designed based on the UA9638C and UA9637AC balanced line driver / receiver. The optical receiver / transmitter is TORX173 and TOTX173 from Toshiba, and they interface directly to UA9638C and UA9637AC with TTL level digital signals.

To provide a stable +5V supply voltage to the components, the universal LM7805 voltage regulator is used as a power supply in both converters. This

dictates that the lower limit of the input voltage is around 7V for the LM7805
to work properly, and the upper limit is determined by the power dissipated
in the voltage regulator or 35V as the absolute maximum in accordance
with the data sheet. The LM7805 acts as an over-voltage protection for
the other components. It features internal short-circuit protection and will
automatically shut down if the regulator over-heats due to too high input
voltage.

As each balanced driver/receiver consists of two channels, it was decided to
build four channels in total, which give the possibility to convert 4 optical
signals (32 audio channels when using ADAT).

B.1   Optical to XLR converter

The voltage regulator circuit for the optical⇒XLR converter can be seen in
Figure B.1.
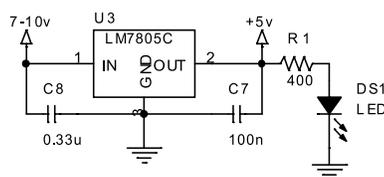
Figure B.1: *The voltage regulator circuit and Power-On LED for the
optical⇒XLR converter.*

C7 and C8 serve as decoupling capacitors and R1 adjusts the current for the
power-LED, DS1.

The rest of the circuit for the optical⇒XLR converter can be seen in Fig-
ure B.2 on the next page.

All capacitors are decoupling/filter capacitors, and L1 to L4 are used for
additional filtering of the supply current as recommended in the data sheet.
R2 to R5 are termination resistors, which are soldered directly to the back
of the XLR connectors.

The entire circuit has been implemented on a 2-layer PCB. The layout can
be seen in figures B.3 (top layer), B.4 (bottom layer) and B.5 (component
placing).

Figure B.2: *The circuit for the optical⇒XLR converter as it is implemented.*
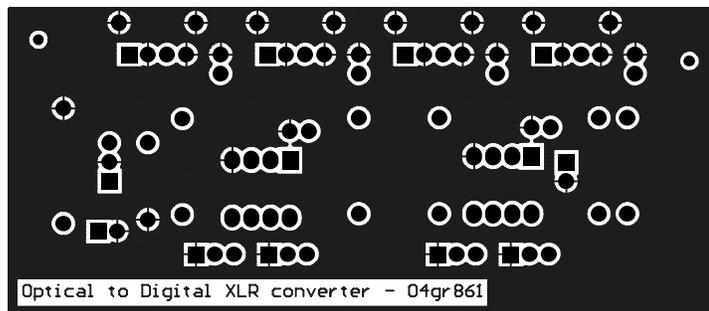
Figure B.3: *Optical⇒XLR converter PCB top layer.*
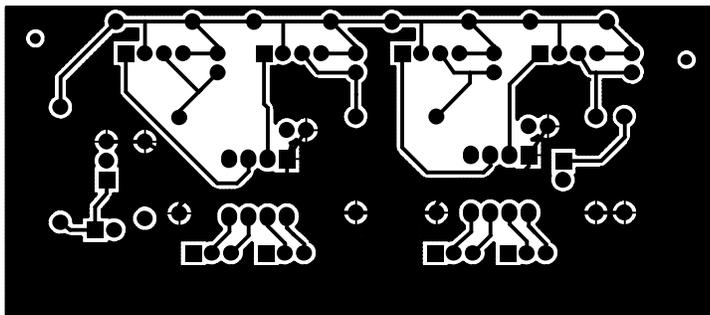
Figure B.4: *Optical⇒XLR converter PCB bottom layer.*
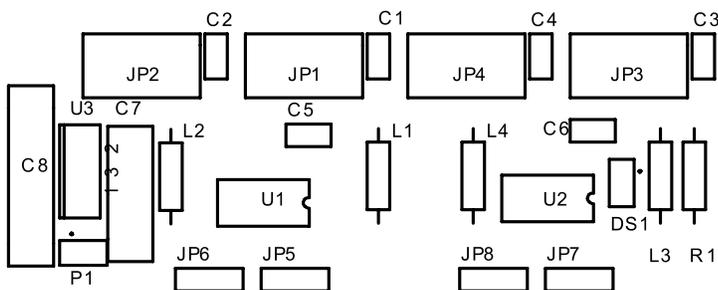
Figure B.5: *Optical⇒XLR converter PCB component placing.*

B.2   XLR to optical converter

The voltage regulator circuit for the XLR⇒optical converter is shown in
Figure B.6.
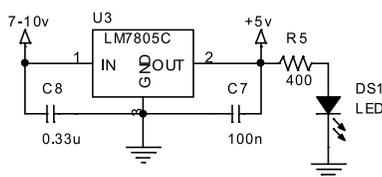
Figure B.6:   *The voltage regulator circuit and Power-On LED for the
XLR⇒optical converter.*

C7 and C8 serve as decoupling capacitors and R5 adjusts the current for the
power-LED, DS1.

The remaining circuit for the XLR⇒optical converter can be seen in Figure
B.7.
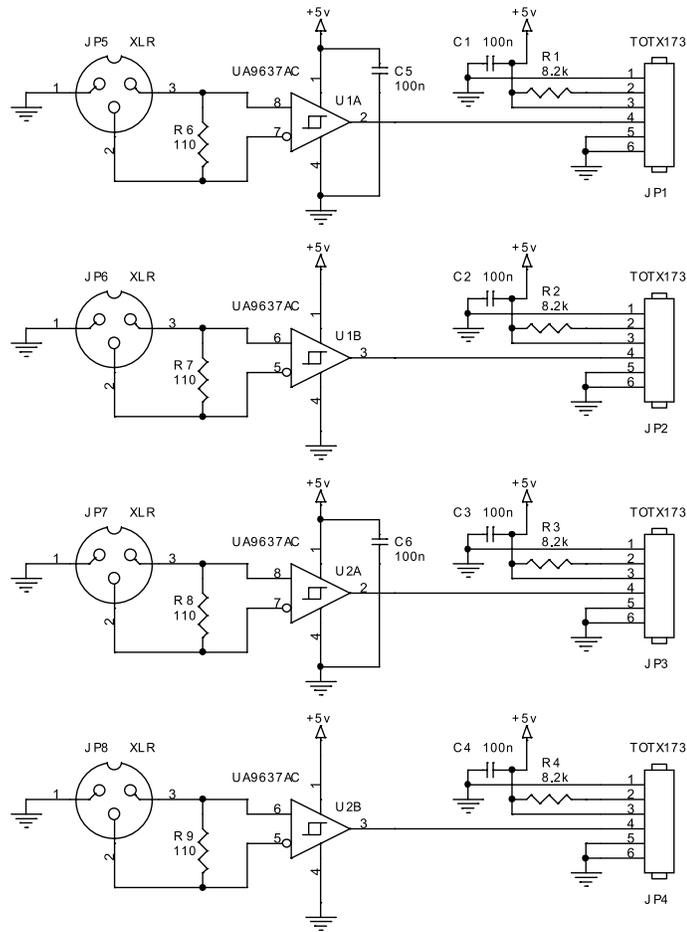
Figure B.7: *The circuit for the XLR⇒optical converter as it is implemented.*

All capacitors are decoupling/filter capacitors, and R1 to R4 are the current-limiting resistor for the transmitter LED.

As in the other converter, the XLR connectors are also terminated with 110Ω resistors, R6 to R9.

The PCB layout of the XLR⇒optical converter can be seen in figures B.8 (top layer), B.9 (bottom layer) and B.10 (component placing).
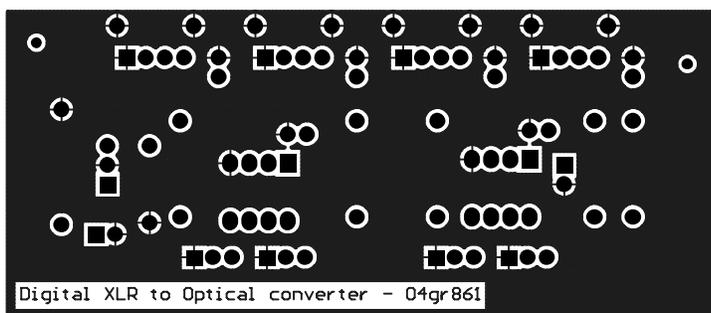
Figure B.8: *XLR⇒optical converter PCB top layer.*
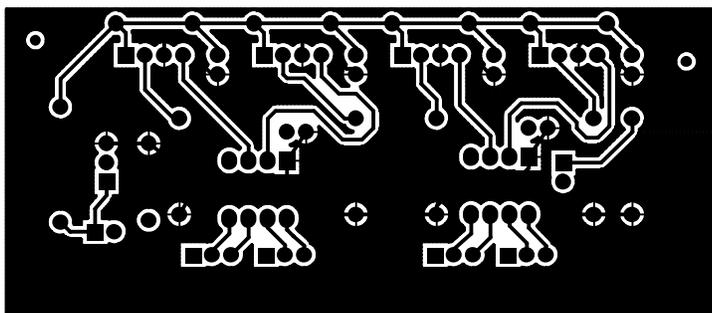
Figure B.9: *XLR⇒optical converter PCB bottom layer.*

B.3   Testing

The converters have been tested together with good results with a 25m XLR signal cable (the longest available single cable). A flow chart of the test-setup is shown in Figure B.11 and a list of the equipment in Table B.3.

When connected, the receiving D/A-converter immediately locks to the optical signal it receives, and the sound from the loudspeaker is as expected.
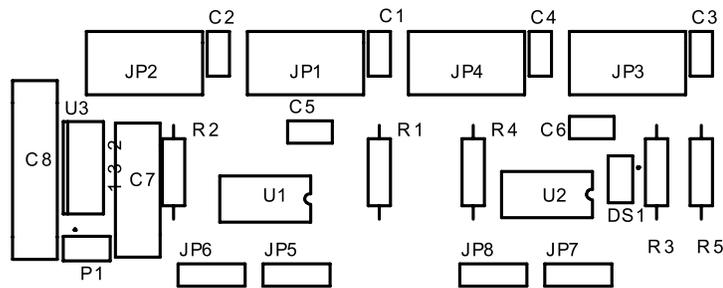
Figure B.10: *XLR⇒optical converter PCB component placing.*
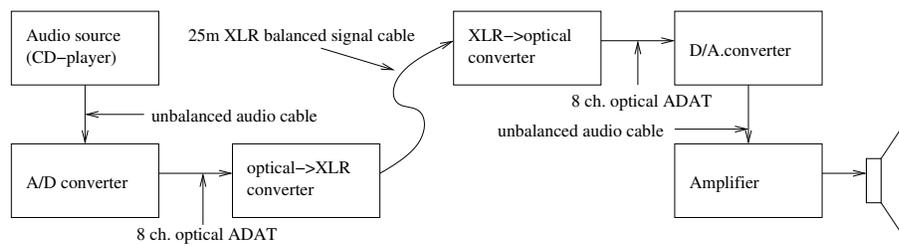
Figure B.11: *A flowchart of the equipment used to test the two converters.*

Description	AAU no.	Comment
CD/DVD player	33995	Technics DVD-A10
A/D-D/A-converter	56543	Behringer ADA8000
A/D-D/A-converter	56545	Behringer ADA8000
Power Amplifier	08340	Pioneer A-616
Loudspeaker	02144-03	B&W 601s3
25m XLR signal cable	-	For connecting the converters
Loudspeaker and signal cables	-	-

Table B.1: *Equipment used to test the optical⇒XLR and XLR⇒optical converters together.*

If bit errors did occur in the signal path, there would probably not be any
output from the loudspeaker because of the 8 channel encoding in the ADAT
format.

All 4 converters in both PCB's were tested to make sure they all worked
correctly. No problems was found, so it is concluded that they all work as
expected.

An issue that has not been adressed, is jitter in the digital stream. The D/A-
converter has internal clock recovery circuits which will remove most of the
jitter before converting, so this is considered to be of little or no significance
to the system. The difference in channel delay in the individual channels
according to the data sheets of the components, is also very small (much less
than a period in the ADAT clock which is $\approx 163 \mu s$), so this should not be
an issue either.

C

The MLSSA measuring method

The MLSSA[1]-system uses a special signal, an MLS-signal, to measure the frequency response of a system.

The recorded signal from the measuring microphone can be considered as the MLS-signal convolved with the impulse response:

$$MS[n] = MLS[n] * IR[n] \qquad \text{(C.1)}$$

where $MS[n]$ is the measured signal, $MLS[n]$ is the generated MLS signal, and $IR[n]$ is the system's impulse response.

The MLS-signal is constructed as a series of 0's and 1's, but the 0's are changed into -1's to create a signal with no DC-component. It has some special properties, which makes it suitable for loudspeaker measurements, if it is used together with a high-quality measuring microphone. Some of these properties are:

1. High energy content

2. Flat frequency spectrum except for DC

3. Easy calculation of the impulse-response

These properties will be briefly explained in the following sections.

---

[1]Maximum Length Sequency System Analyzer

C.1 Property 1: High Energy Content

To ensure a high S/N-ratio it is important to use a test-signal with a high energy content relative to its peak amplitude.

A value used to define this energy ratio is the crest factor:

$$crest\,factor = \frac{peakvalue}{RMSvalue} \qquad (C.2)$$

The best S/N-ratio is achieved with a crest factor of 1. Only a square-wave has a crest factor of 1, and the MLS-signal is a square wave. It therefore utilizes the entire dynamic range. In comparison, a sine wave has a crest factor of $\sqrt{2}$, and a normally distributed random signal roughly has a crest factor of 8. If an impulse $\delta[n]$ is sent to test the loudspeaker, it would give a low S/N-ratio because of the very high crest factor, which again is due to the very low RMS value of such an impulse.

C.2 Property 2: Flat Frequency Spectrum

An MLS-signal is mathematically constructed to have a completely flat frequency spectrum, where a gaussian white noise signal only averages out to a flat spectrum. The MLS-signal is therefore only pseudo-random, periodic and can be recreated unlike true white noise. It also has a higher energy content than white noise.

C.3 Property 3: Easy Calculation of the Impulse Response

The impulse response can be calculated in different ways. The typically used method is to calculate the cross-correlation between the measured signal and the original MLS signal. As seen in equation (C.1), the impulse-response can be calculated by deconvoluting the original MLS signal from the measured signal. The cross-correlation gives the same result, but is much faster.

Deconvolution or cross-correlation correspond to division in the frequency
domain, indicating that the frequency response of the system should be seen
as the top envelope of the Fourier transform of the recorded signal.

A simple test is performed to calculate the impulse-response from a loud-
speaker measurement (in the example, a response from one of the loudspeak-
ers is used).

The original measured impulse response is shown in Figure C.1.
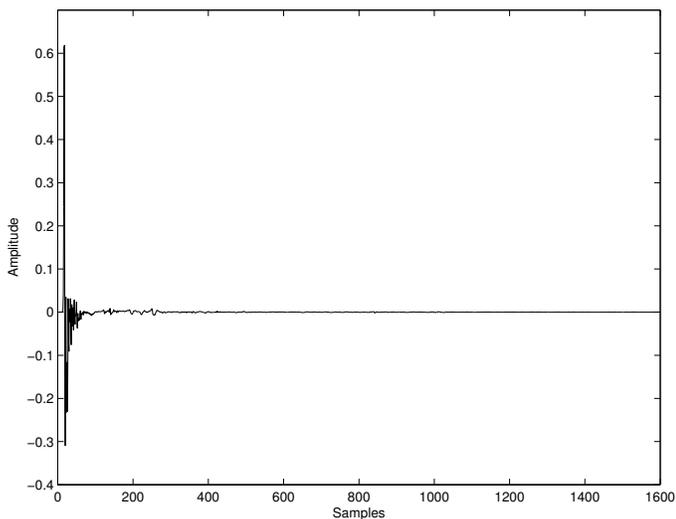
Figure C.1: *Original measured impulse response of the loudspeaker.*

As only the impulse response, and not the recorded signal, could be aquired
from MLSSA, a signal has been constructed by convolution of the measured
impulse-response with an MLS sequence (order 15). The constructed signal
can be seen in Figure C.2 on the next page.

To calculate the impulse response, the cross-correlation between the mea-
sured signal and the original signal is calculated using Matlab. This results
in the impulse response shown in Figure C.3 on the facing page. The response
has been cut to 1600 samples for an easier view of the details in the result.

The Fourier transform of the original and the calculated impulse response
are plotted together in Figure C.4 on page 98 for comparison. As it can
be seen, the two responses are almost identical, but small differences exist,
which derives from the cutoff of the last part of the measured signal. The

96
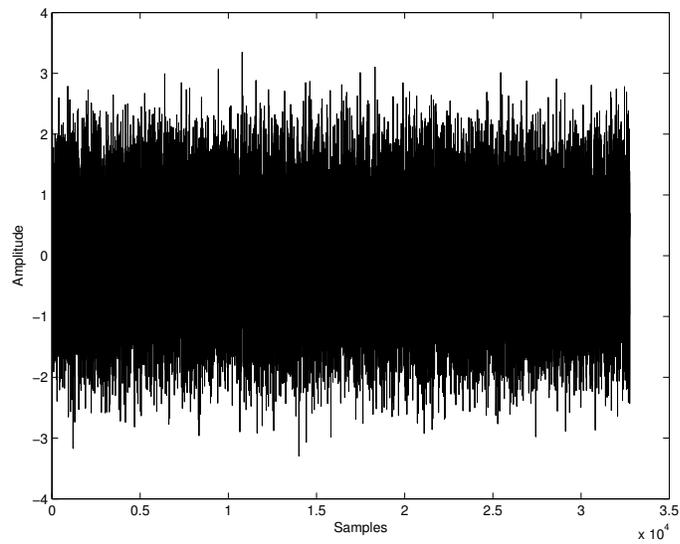
Figure C.2: *Measured response of the MLS signal.*
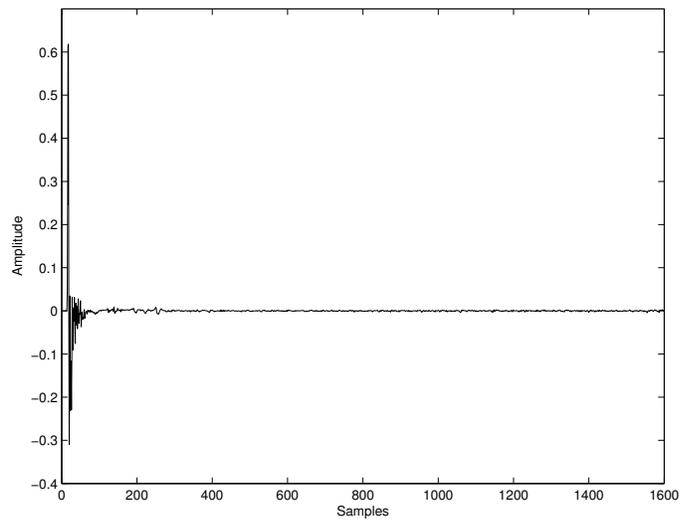
Figure C.3: *The calculated impulse-response.*

97

microphone signal is only sampled as the MLS-signal is played back, resulting in a small loss of information.
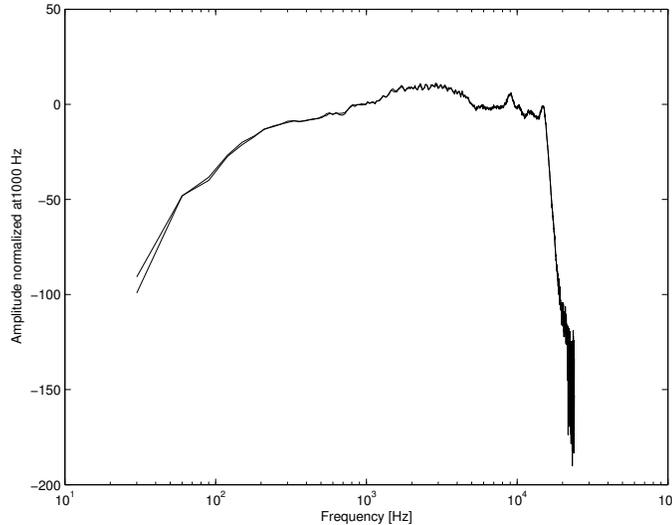
Figure C.4: *Fourier transform of the two impulse-responses from Figure C.1 and C.3.*

This method of manually calculating the MLS-response and back again to the impulse response through the cross-correlation is only to illustrate how MLSSA works internally. When using MLSSA, it simply outputs an impulse response which can then be loaded into Matlab for further analysis.

C.4   Non-Linearities

If non-linearities are present in the system, the MLS-measurement could result in a wrong impulse response. In such cases, small spikes in the calculated impulse response can be seen, which looks like many small reflections, but at seemingly random places.

If this behaviour becomes a problem, the speaker test should be remade at a lower volume to keep the speaker within its linear work limits.

If the system being tested is non-linear, a traditional sine sweep test may be better to determine the frequency response, as the MLS method only applies

to linear systems. In any case, if the system is nonlinear, the system cannot
be defined correctly by a frequency response.

D

Loudspeaker measurements

The array is composed of 24 identical[1] loudspeakers.

To create proper driving functions for the loudspeaker array, it is necessary to have a good model for the used loudspeakers. It is assumed that the radiation pattern of the loudspeaker is symmetrical around the on-axis. The measurements needed are thereby limited to measurements between 0° (front) and 180° (back).

All measurements were performed in the anechoic room in AAU's acoustics laboratory and controlled from the adjacent Control Room S.

The used equipment is listed in Table D.

The MLSSA system was set to default settings, except the output voltage (stimulus voltage), which was $\pm 0.6V$. The power amplifier was set to 0 dB amplification and only served as current supply for the loudspeaker. An external clock generator was used to create a 48kHz clock frequency for the MLSSA system.

To reduce the noise floor, each measurement was made with an average over 5 impulse responses, which lowers the noise floor by approximately 6 dB.

The stepsize was chosen to 2.5°, therefore 73 impulse response measurements was performed.

A sketch of the measurement setup can be seen in Figure D.1.

---

[1]They have been checked for apparent defects, and otherwise supposed to be identical.

Description	AAU no.	Comment
Microphone	08712-00	B&K 4165
Microphone preamp	06560	B&K 2619
Measuring Amplifier	08717	B&K 2636
Power Amplifier	08340	Pioneer A-616
MLSSA	26827	Toshiba T-3200SX
Ext. clock generator	-	48 kHz
Loudspeaker	02017-27	Homemade from trawlnet ball
Microphone cable	02123-10	-
Misc. speaker and signal cables	-	-

Table D.1: *Equipment used to measure the frequency response of the loud-speaker.*
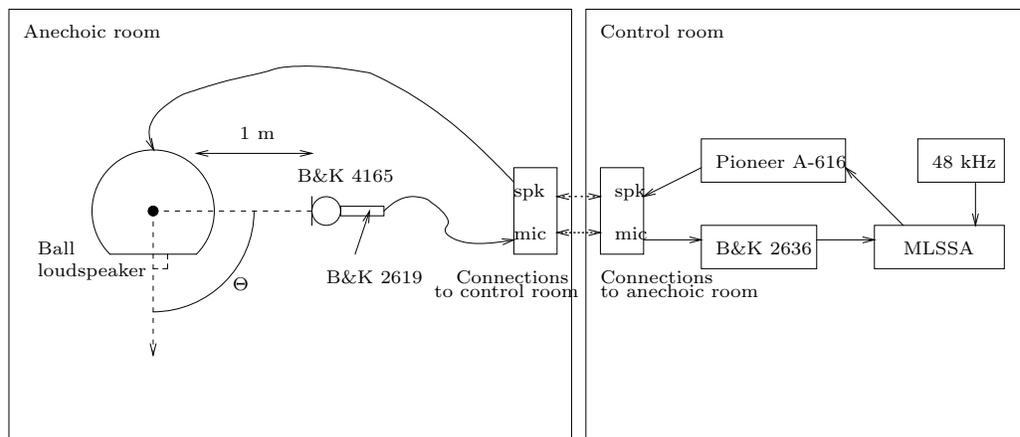
Figure D.1: *The setup of the equipment used to measure the loudspeaker impulse response. The measurement was conducted for $\theta = [0°; 180°]$ in 2.5° steps.*

101

The impulse responses were loaded into Matlab and the first ≈140 samples were removed, as there should not be any signal before the impulse reaches the microphone. At a distance of $1m$ and a sample frequency of $48kHz$, this delay is approximately 140 samples. This was confirmed in a plot of the impulse response to ensure that no samples of the actual impulse were removed.

The remaining impulse response was then cut down to 1600 samples to reduce the noise. The low-frequency reproduction of the loudspeaker is very poor, so a longer impulse response would only constitute more noise and not add usable information.

The Fourier transformation of the impulse response reveals the loudspeaker's frequency response. A selection of these frequency responses can be seen in Figure D.2 for different angles.
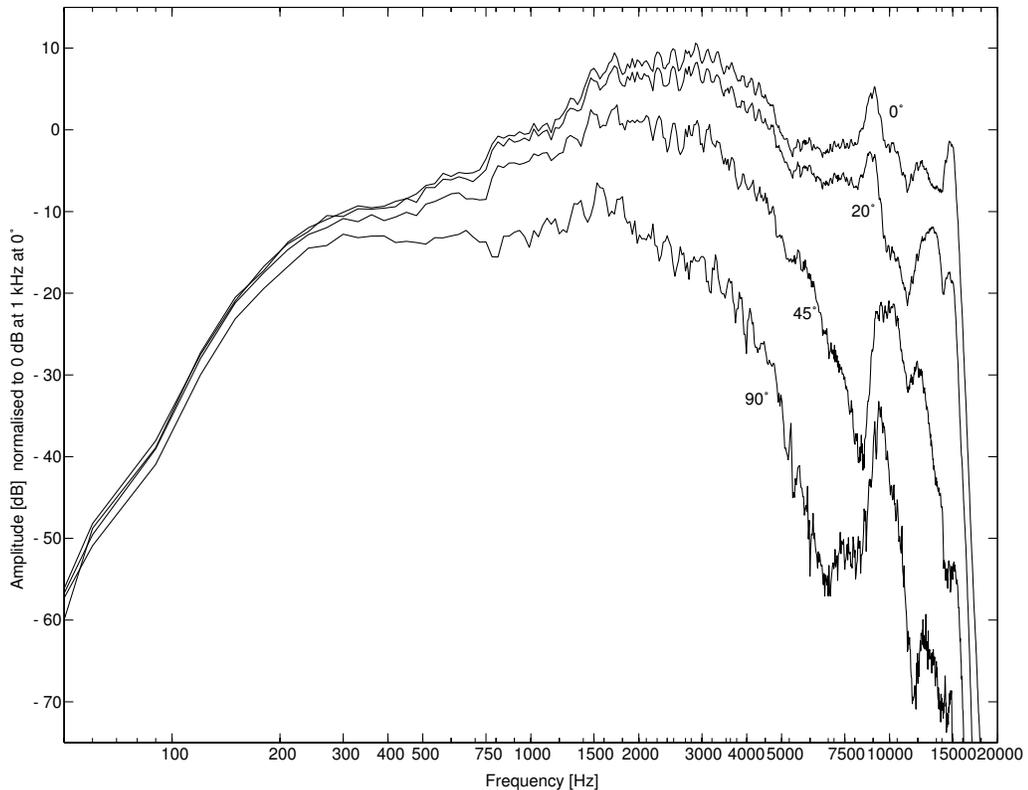
Figure D.2: *Frequency response plots for loudspeaker with AAU no. 02017-27. Plots are shown for 0°, 20°, 45° and 90°.*

All the measured frequency responses is shown together in Figure D.3 on the next page.
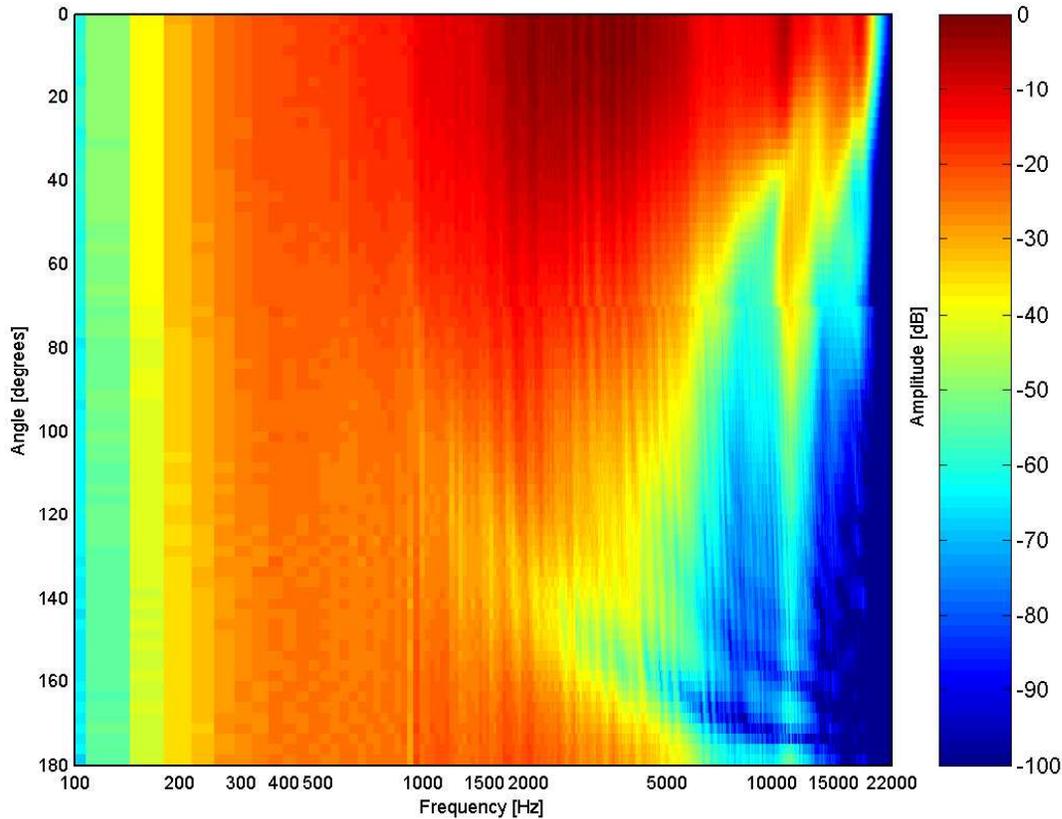
Figure D.3: *A plot showing the frequency responses of the loudspeaker in 2.5° steps from 0° to 180°.*

To visualize the behaviour of the loudspeaker, directivity plots are made for specific frequencies. Figure D.4 shows the directivity for $500Hz$, Figure D.5 for $2kHz$, Figure D.6 for $8kHz$, and Figure D.7 for $15kHz$. The radius is normalized to 1 as the maximum amplitude, and the directivity has been mirrored around the on-axis, as the loudspeaker is considered symmetrical.

All these measurements give the foundation for creating the directivity function $G(\varphi, \omega)$ used to calculate the driving signals as explained in section 2.3.1 on page 14.
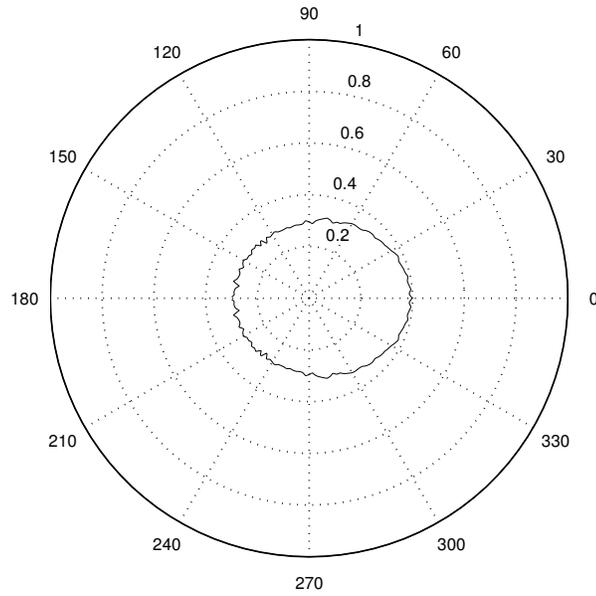
Figure D.4: *Directivity of the loudspeaker at 500 Hz.*
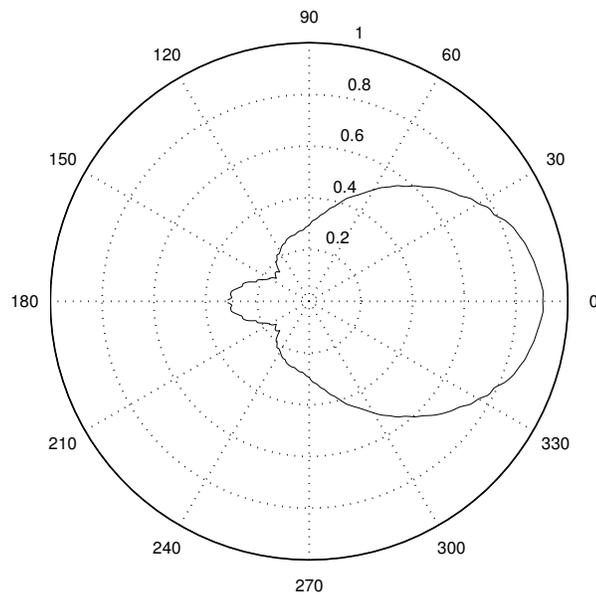
Figure D.5: *Directivity of the loudspeaker at 2000 Hz.*
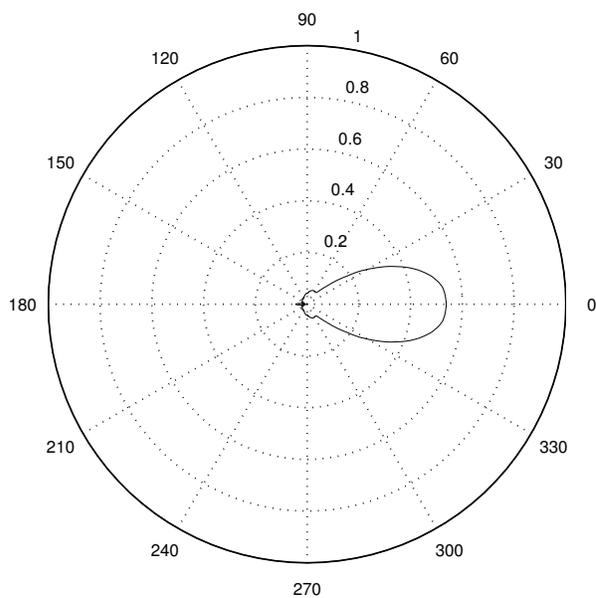
Figure D.6: *Directivity of the loudspeaker at 8000 Hz.*
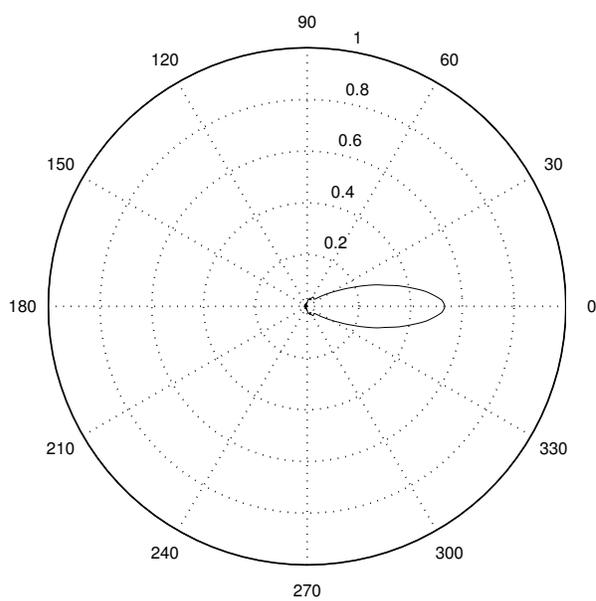
Figure D.7: *Directivity of the loudspeaker at 15000 Hz.*

DSP for Driving Signals generation

The calculation of the driving signals for the loudspeakers is based on equation (E.1) derived in Chapter (2 on page 6).

$$Q(\vec{r}_n, \omega) \;\;=\;\; S(\omega)\frac{cos(\theta_n)}{G(\theta_n, \omega)}\sqrt{\frac{jk}{2\pi}}\sqrt{\frac{|z - z_1|}{|z - z_0|}}\frac{e^{-jk|\vec{r}_n - \vec{r}_m|}}{\sqrt{|\vec{r}_n - \vec{r}_m|}} \qquad (\text{E.1})$$

The code for generating the driving functions requires 5 Matlab generated functions (See Figure E.1):

- dfprogram

- drivingfunction

- specgram

- g_function

- invspecgram

dfprogram has a user interface, which searches for *.wav files on the same directory as the m-file is located, then "asks" the user to input the data required for processing the signals. dfprogram automatically detects if the file has stereo content. If it does, dfprogram calculates the x_m position of both notional sources on the z_0 plane in order to position them 30° from the listener position line, perpendicular to the z_1 array plane (See Figure E.2).
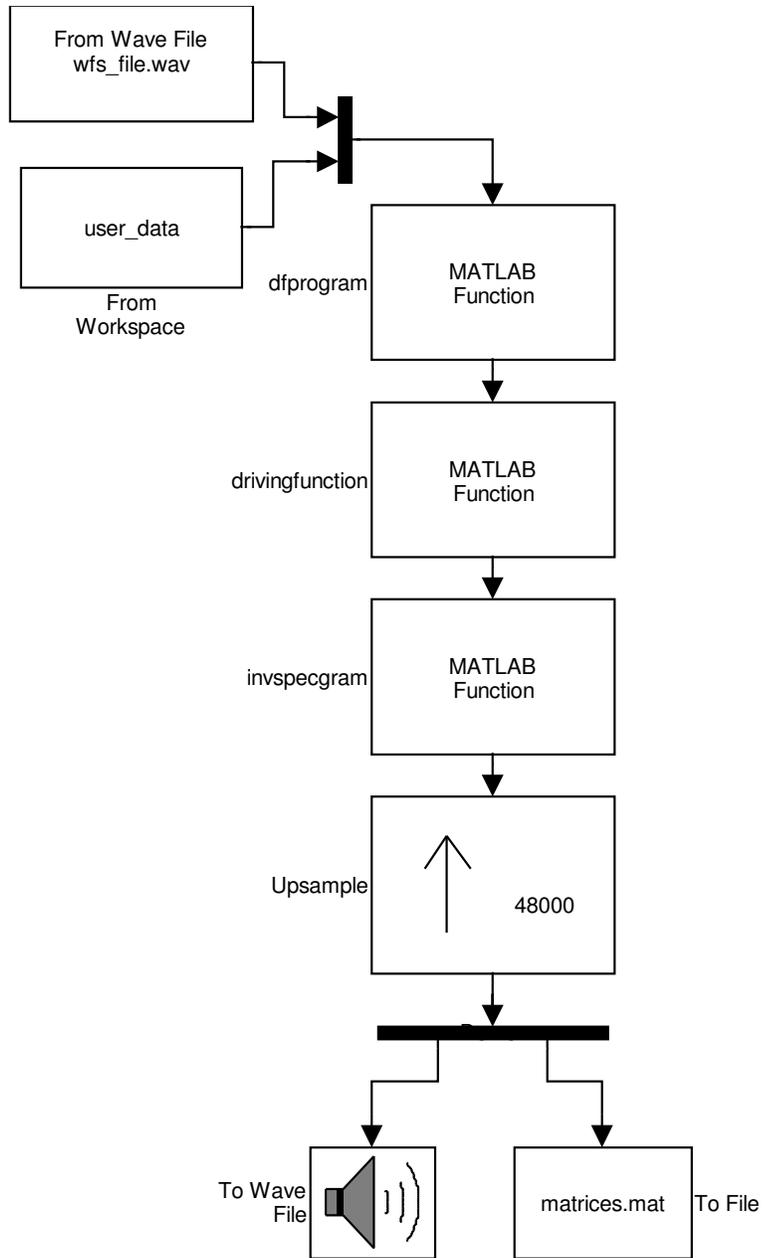
Figure E.1: *Data flow for driving signals generation.*
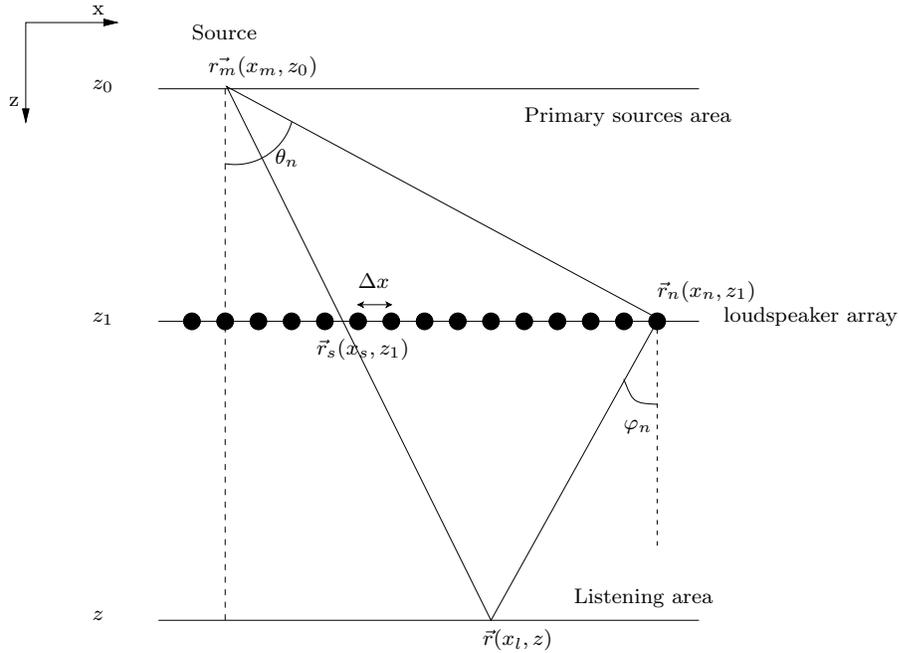
Figure E.2: *Geometry used to calculate the driving signals.*

The data loaded from the `*.wav` file consists of the audio information, two channels in the stereo case, and the sampling frequency.

The data input by the user consists of the distance from the notional source plane to the loudspeaker array $(z_1 - z_0)$, the distance from the loudspeaker array to the listener position plane $(z - z_0)$, the number and distance between loudspeakers, and in the case of a single notional source the position on the x axis, with the center of the array placed at the origo.

`dfprogram` calculates the time in samples that takes the sound from the notional source to arrive to the loudspeaker array. If the audio data vector is not long enough, and the distance is considerable, audio is lost because of the information shifting caused by the delay, `dfprogram` zero-pads the audio content at the end of the file to avoid this.

Then `dfprogram` checks if the audio input file is stereo. If it is, it processes each source at a time, generating 24 vectors of audio signals for each source. Then the results are summed and scaled.

The delay of the sound traveling from the notional source to the array is removed by eliminating the same amount of samples that were added at the

end of the input signals, from the beginning of the output audio vectors.

The resized audio vectors are then upsampled to $48kHz$, because the impulse response of the loudspeakers has a sample frequency of $48kHz$, and the convolution of both signals is required.

The audio data for the loudspeaker array is saved as stereo files and into a matrix with size: number of loudspeakers $\times$ length of the output audio vectors.

`dfprogram` requires the utilization of `drivingfunction` this function calculates the driving signals with the data introduced by the user. `drivingfunction` calculates the spectrogram of the audio signal with Matlab's `specgram` function.

`specgram` splits the signal into overlapping segments (overlap of 50%), windows each segment with a Hanning window of 256 samples length and forms the columns of a matrix B with the N-point discrete Fourier transforms of these segments. N depends on the sampling frequency of the `*.wav` file to be processed.

`drivingfunction` enters the data into a loop that is repeated for as many loudspeakers as there are on the array. The first step into this loop is the calculation of θ see Figure E.2. Then calculates the distance between the loudspeaker and the source.

In order to use the directivity of the loudspeakers in the calculations, it is necessary to use the `g_function`, which outputs the directivity of the loudspeaker according to the frequency and the angle of emission. Considerations were taken to limit the frequencies feeded into `g_function` to its range of operation.

K number of time-frequency domain matrices are obtained, where K stands for the number of loudspeakers in the array.

These matrices are transformed into time domain with `invspecgram`. This function calculates the inverse FFT of the columns of this matrices and then rearranges the columns by transposing them, taking one at a time and placing them into an output vector, shifting each rows by the number of samples overlapped in the spectrogram and summing them, therefore overlapping them to generate a time domain vector.

These output vectors are returned to `dfprogram` to be saved as audio files.

The simulation program

The simulation program is intended to simulate any number of loudspeakers, playing independent signals. The simulation can use the impulse responses measured in Appendix D on page 100 to simulate an array of these loudspeakers. It can also simulate ideal loudspeakers, which are characterized by an impulse response $\delta[n]$ in all directions.

A sketch of the simulation area can be seen in Figure F.1. The angle and distance are shown for one specific simulation point and loudspeaker in a 14-loudspeaker array setup.

To calculate the sound pressure from a single loudspeaker in any point in the simulation area, a transfer function $h[n]$ from the loudspeaker position to the simulation point is needed. The transfer functions from the loudspeakers to the simulation points are linear, which makes it possible to superpose the sound pressures generated by the individual loudspeakers.

The impulse responses measured in Appendix D on page 100 describe the transfer function from the electrical input of the loudspeaker to the sound pressure $1m$ from the loudspeaker. This transfer function is extrapolated for other distances by shifting the measured impulse response in time and multiplying it with a dispersion term depending on r, the radius or distance to the loudspeaker. For spherical waves, this dispersion term is r^{-1}.

The base impulse responses are the measured far-field impulse responses shifted 135 samples to the left, to set the initial distance close to $0m$. This compensates for the traveling time between the loudspeaker and the microphone in the measurement of the loudspeaker. These base impulse responses
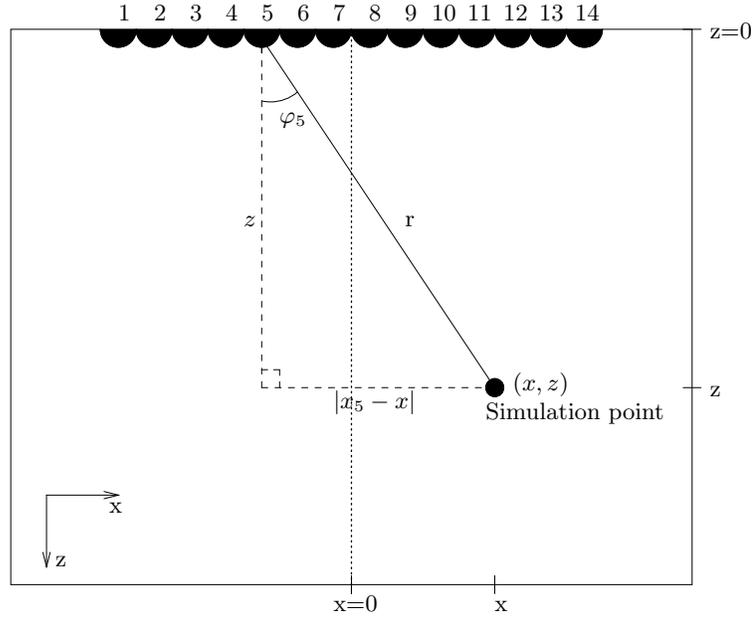
Figure F.1: *A sketch of the simulation area with the angle and distance as it is used in the simulation.*

are called $h_{0_\varphi}[n]$ where φ is the angle to the loudspeaker in degrees. In the loudspeaker measurements, φ is discretisized in 2.5 degree steps, and these are all saved in the file 'ir27.mat' and used in the simulation program.

The transfer function $h[n]$ from a single loudspeaker to a given point in the simulation area is calculated in the following manner:

$$h[n] = \frac{h_{0_\varphi}[n]z^{-r\frac{F_s}{c}}}{r} \tag{F.1}$$

where r is the distance, F_s is the sample frequency ($48000Hz$) and c is the speed of sound ($343\frac{m}{s}$). $z^{-r\frac{F_s}{c}}$ is a delay-term that delays h_{0_φ} by $r\frac{F_s}{c}$ samples.

To calculate the loudspeaker response to any given input signal $x[n]$, the input signal to the loudspeaker is convolved with the transfer function $h[n]$. This is written as:

$$P[n] = h[n] * x[n] \tag{F.2}$$

112

where $P[n]$ is the output signal from the loudspeaker.

When (F.1) is combined with (F.2), it results in equation (F.3):

$$P_{\varphi,r}[n] = \frac{(h_\varphi[n] * x[n]) \, z^{-r\frac{F_s}{c}}}{r} \tag{F.3}$$

Equation (F.3) calculates the resulting sound pressure for a specific angle and distance to a loudspeaker with input signal $x[n]$.

Because the loudspeaker is considered to be a point source, the sound pressure is not defined in equation (F.1) as r goes toward zero. In the simulation, this is avoided by not calculating the sound pressure of a point if r is less than $77mm$ (the loudspeaker radius). Instead the pixel color is set to white in those points, which results in white circles where the loudspeakers are positioned.

To derive the angle φ and distance r from one loudspeaker m in coordinates[1] $(x_m, 0)$ to a specific point in the simulation area with coordinates (x, z), the following equations are used:

$$
\begin{aligned}
r_m &= \sqrt{z^2 + (x_m - x)^2} \quad [m] & \text{(F.4)}\\
\varphi_m &= \frac{180}{\pi} \arcsin\left(\frac{|x_m - x|}{r_m}\right) \quad [°] & \text{(F.5)}
\end{aligned}
$$

The equations (F.4) and (F.5) are only valid for points in front of the array, so it is not possible to simulate the sound pressure behind the array, unless these equations are expanded to give the correct values in exchange for longer calculation times.

Eq. (F.3) is now expanded to all 24 loudspeakers in the array by adding the contributions of all loudspeakers (superpositioning) in the same point:

$$P[n] = \sum_{m=1}^{M} \frac{(h_{\varphi_m}[n] * x_m[n]) \, z^{-r_m\frac{F_s}{c}}}{r_m} \tag{F.6}$$

---

[1]The origin of the coordinate system lies in the center of the array, so all loudspeakers have a z-coordinate of 0.

where 'm' is the loudspeaker number and 'M' is the number of loudspeakers.

An overall flow chart of the simulation program can be seen in Figure F.2. Figures F.3 and F.4 shows the snapshot and sound file parts respectively.
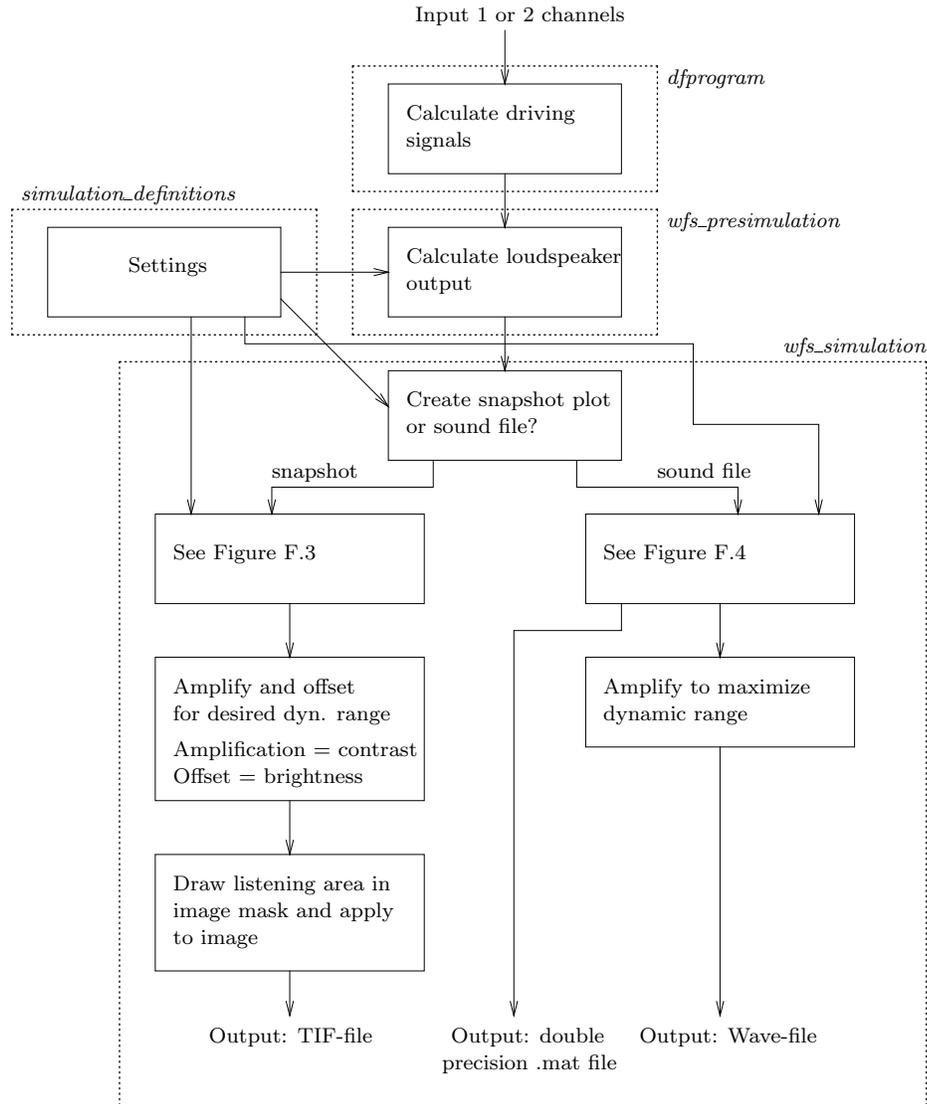
Figure F.2: *Overall flow chart of the simulation program, from notional source to output, which can be either a picture or sound file.*

The settings file *simulation_definitions* contains all the settings for the simulation program. A list of the most important variables is shown in Table F.

From figure F.2

Iterate over all pixels in
the simulation area

Calculate distance to all
loudspeakers from point

Is distance < minimum
distance (loudsp. radius)?

yes

Set white pixel in image
mask matrix to mask
loudspeaker

no

Calculate angles to all
loudspeakers from point

Set sound pressure
in current point to 0

Calculate sample value
from all loudspeakers based
on angle and distance, add
together and save to output
matrix

Finished with all pixels?

no

yes

Back to figure F.2

Figure F.3: *Flow chart of the part of the simulation program that creates a TIF-file to illustrate the sound pressure in front of the array at some specific time instant.*

Name	Unit	Default value	Description
simulation_type	-	1	'1' snapshot or '2' sound file
no_speakers_array	-	24	Number of loudspeakers in array
interpolation	-	1	'0' nearest neighbor (fastest) or '1' bilinear interpolation
snapshot_time	samples	1500	Time of snapshot image
snapshot_resolution	mm/pixel	10	Resolution in output image
delta_x	mm	155	Loudspeaker spacing
dist_area_array	mm	0	Distance from loudspeaker array to simulation area
size_area	mm	5000	Simulation area side length
sampling_point_x	mm	0	X-coordinate of sampling point
sampling_point_z	mm	0	Z-coordinate of sampling point
sampling_time_start	samples	0	Start time of sampling
sampling_time_end	samples	4000	End time of sampling
order_reflections	-	0	Order of reflections, either 0 (fastest), 1 or 2
reflection_koefficient	-	0.5	Common reflection koefficient used for all side walls

Table F.1: *A list of the most important simulation settings as set in 'simulation_definitions'.*
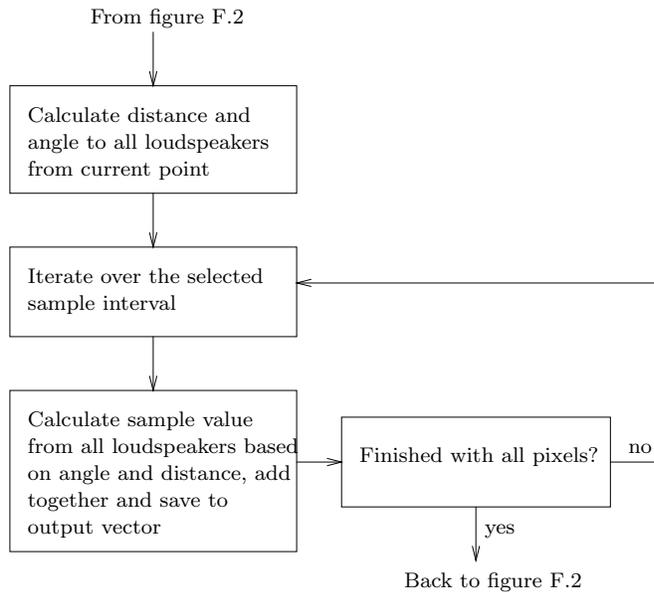
Figure F.4: *Flow chart of the part of the simulation program that creates a sound file virtually sampled in a specific point in the simulation area.*

To make a plot of the simulation area, which could be a square $5m \times 5m$ in front of the array, equation (F.4), (F.5) and (F.6) are calculated for $x = [-2.5; 2.5]$ and $z = [0; 5]$ in steps of i.e. $0.01m$, which will create a plot of 501×501 pixel.

When all pressures for the simulation area have been calculated, this matrix is amplified to have a peak absolute value of 0.5, and then offset with +0.5 to utilize the dynamic range $[0; 1]$ in the output picture. The center value will then be 0.5, which will be seen as gray on the picture and indicate average room pressure. Many plots have had an additional amplification to increase the contrast in the picture. This causes clipping of the information in the matrix when it is written to an image file, but is in many cases preferred to increase the contrast of the image.

F.1   Optimizations

To decrease the program execution time, optimizations have been performed. The simulation program has been split into two parts, one that is called the 'presimulation' and the actual simulation.

The presimulation calculates the individual loudspeaker outputs in 2.5° steps, using equation (F.2) for all $h_\varphi[n]$. This generates a matrix of size $no_{speakers} \times no_{angles} \times length_{signal}$ which is saved to the file 'speaker_output.mat'. This file is loaded into the simulation.

The advantage of splitting the simulation program into two parts, is that the simulation can be repeated with different settings, e.g. the size of the simulation area, the resolution, the snapshot[2] time etc., without having to recalculate the loudspeaker output signals. As long as the driving signals and the array size remains unchanged, the presimulation does not need to be run again. If the input signal is long, these convolutions can take a lot of time to calculate, so it is preferred to not recalculate the loudspeaker output each time the simulation is run, unless it is necessary.

The simulation images used in the report are typically 501×501 pixel and takes around $30s$ to be calculated on a modern PC. This also makes it possible to calculate animations, i.e. to visualize how a wavefront propagates through the listening area.

F.2   Virtual recording

As seen in Figure F.2 on page 115, the simulation program can be configured to produce a sound-file that contains the sampled sound in any selected point in the simulation area. With this feature, it is possible for example to calculate the frequency response in any given point by playing back a pulse on the virtual array and then make a Fourier-transformation of the output signal that was virtually recorded in that point.

The sound files for several points can also be analyzed to create amplitude plots that shows how the amplitude rolls off with distance. This has been used in section 5.2.1 on page 68.

---

[2]The time of the plot, where 0 is the start of the driving signals.