An Ultrasonic Tracking System for a Personal Sound Zone Simulator

An implementation using Direction of Arrival and Uniform Circular Arrays.

Master's Thesis

NANNA GODIKSEN DAHL
SMCA171030

Aalborg University
Sound and Music Computing

AALBORG UNIVERSITY
STUDENT REPORT

Department for Architecture, Design, and Media Technology
Rendsburggade 14
DK-9000 Aalborg
Aalborg University
http://www.create.aau.dk

Title:
An Ultrasonic Tracking System for a Personal Sound Zone Simulator.

Theme:
Master's thesis

Project Period:
September 2016 - June 2017

Project Group:
SMCA171030

Participant(s):
Nanna Godiksen Dahl

Supervisor(s):
Jesper Kjær Nielsen

Copies: 4

Page Numbers: 56

Date of Completion:
May 21, 2017

Abstract:

This project examines an ultrasonic tracking system for a personal sound zone simulator using Direction of Arrival and Uniform Circular Arrays. The project investigates Personal Sound Zone systems, what has been done, and what can be done in the future. For this project an amplifier for six ultrasonic sensors, a 3D printed mount for the sensors, and an algorithm were created. The results of the simulations of the algorithm showed that it should be possible to create a system this way with ultrasonic sensors. However, it has not yet been possible to create a system that works. The results of the test were unsatisfying due to a problem with the amplifier of the sensors. This needs to be fixed before concluding if it is possible to create such a system in reality.

Contents

Preface

This report was created as a part of a 50 ECTS-point Master's thesis in Sound and Music Computing. According to the curriculum of Sound and Music Computing, the thesis must use scientific theories and methods. The thesis should critically analyze existing work and synthesize new knowledge (Study Board of Media Technology, 2014).

This project examines an ultrasonic tracking system for a personal sound zone simulator using Direction of Arrival and Uniform Circular Arrays. Together with this report, an amplifier for six ultrasonic sensors, a 3D printed mount for the sensors, and an algorithm were created. The time span for this project was from September 2016 to June 2017, and it was carried out at the facilities of Sound and Music Computing at Aalborg University in Aalborg.

Submitted together with the report is an AV production, which gives a visual representation of the project.

Figures, tables, and equations are referred to by the number of the chapter that they are occurring in and then followed by the number of the figure, table or equation in that specific chapter. The bibliography and all the sources referred to uses APA-style, which looks like this: ([Author's last name], [year published]). It is expected that the reader has a knowledge equivalent to or greater than a 10*th* semester Sound and Music Computing student.

Acknowledgements
A special thanks go to my supervisor Jesper Kjær Nielsen who have been very helpful throughout the project period. Our weekly meetings helped me keep going even though everything seemed impossible at the time. Special thanks also go to Thomas Kristensen, who have been saving me with his immense knowledge in the electronics laboratory several times. Thank you!

Aalborg University, May 21, 2017

Nanna Godiksen Dahl
<ndahl12@student.aau.dk>

Chapter 1

Introduction

According to Betlehem, Zhang, Poletti, and Abhayapala, 2015, sound reproduction has increasingly become a universal part of our daily lives. One area that has especially increased is spatial sound reproduction which is about reproducing sound over larger regions of space using an array of loudspeakers. This progress in the field of spatial sound has opened up for the idea of personal sound zones.

Personal sound zones is a concept of sound being controlled within certain regions of space for multiple listeners without causing too much interference for listeners within other regions (Betlehem et al., 2015). There are several reasons for the growing interest in personal sound zones. For one, the desire to be able to deliver various audio sources to multiple people in one room without the physical separation of headphones. For this especially, there is a big range of possibilities of audio applications. An example where personal sound zones could be desired in relation to this can be seen in a living room where some are watching television while others desire to listen to music (Druyvesteyn & Garas, 1997). Personal sound zones can also be used in exhibition centers and museums, where audio is a part of the exhibition. This can allow the visitors to listen to the audio attached to one exhibition and still communicate with each other without the restraint of wearing headphones and without the sound from one exhibition is interfering with the other exhibitions (Betlehem et al., 2015; Druyvesteyn & Garas, 1997).

According to Berg and Serpanos, 2011, the use of personal listening devices (MP3's, iPod's, phones) have in 2008 increased four times the percentage that was reported in 2001 in adolescent girls. This increase in the use of listening devices has been found to be related to an increase in tinnitus and high-frequency hearing loss in adolescent girls. They reported that it was the high sound pressure level (SPL), and many hours a day spent on using personal listening devices that caused this. Personal sound zones could be a solution to this problem. Situations, where personal listening devices are used for several hours at a time such as public transport, can be a good place to implement personal sound zones. Also, passenger cars could benefit from this solution (Betlehem et al., 2015).

Open and shared offices have become very common today (Pierrette, Parizet, Chevret, & Chatillon, 2014). A common problem in these offices is noise. This noise has been noted to cause discomfort for the employees and it affected the concentration of the employees. One of the main factors of noise that the employees declared that they often hear was comprehensible conversations, and many were inconvenienced by this. Pierrette et al., 2014 state as a conclusion to their investigation, that open workspaces could among others be improved by installing noise-control systems. In this situation, personal sound zones could be used as a noise-control system to create quiet zones. These quiet zones could then keep the noise from the rest of the office from interfering without having to wear headphones with loud masking music which also can be a source of inconvenience when working (Betlehem et al., 2015; Druyvesteyn & Garas, 1997; Pierrette et al., 2014).

As of today, there have been several attempts at creating personal sound zones, and according to Olik et al., 2013, there are several methods for creating these sound zones. These methods can be categorized into three major groups.

Sound Focusing	Sound Canceling	Sound Field Synthesis
Delay and Sum Beamforming (Van Veen & Buckley, 1988)	Acoustic Contrast Control (Choi & Kim, 2002)	Analytical Sound Field Synthesis (Wu & Abhayapala, 2010)
Brightness Control (Choi & Kim, 2002)	Acoustic Energy Difference Maximisation (Shin et al., 2010)	Pressure Matching (Poletti, 2008)

Table 1.1: The three major groups of sound zone methods.

Table 1.1 shows the three major groups of sound zone methods, with two noteworthy methods for each group. The first group consists of methods that attempt to focus the sound energy at the desired target. This can be done using beam forming as described in Van Veen and Buckley, 1988 or it can be done by maximizing the potential energy under the given input power as described in Choi and Kim, 2002. The methods in the second group are focusing on controlling sound. As described in Choi and Kim (2002) the first method Acoustic Contrast Control (ACC) controls the sound by maximizing the contrast between the bright zone and the dark zone. According to Shin et al., 2010 the second method Acoustic Energy Difference Maximization build upon the ACC method but it differs in the way that it focuses on the difference between the mean square pressures between the bright and the dark zone whereas the ACC focuses on the ratio of the mean squared pressure between the bright and the dark zones. The last group is focusing on creating multiple spatial regions using one loudspeaker array. The first method Analytical Sound Field Synthesis (Wu & Abhayapala, 2010) assumes that we already know the sound field beforehand. This knowledge is used to analyze the sound field. When this analysis is done they use an already existing single-zone sound field reproduction technique. The other method Pressure Matching (PM) aims to control sound pressure at multiple matching points within one zone while attenuating the sound field in other zones which also contains multiple matching points (Poletti, 2008).

In the study of Olik et al., 2013 they compared the Delay and Sum, ACC and PM methods in a real room. This was done in order to find the performance difference in a real room because most of the methods had typically been tested under anechoic conditions. They found that the ACC method created the highest contrast between the zones while PM generated a uniform bright zone. They performed a listening test where they found that the distraction scores of the ACC method were the lowest of the three methods. They also found that when speech was interfering it would cause a high distraction. As a conclusion, they found that there should be more development of procedures to evaluate sound zone systems.

This study shows that there are several methods for creating sound zones and that each individual method has both advantages and disadvantages for different needs. This suggests that a design of the sound zone system should be created before choosing the methods to create it. It also suggests that multiple methods can be combined in order to create the best solution for the given system. Here a simulation of a sound zone system could be useful in order to find the best fitting methods for the given design of the sound zone system.

The next section will take a look at the history of personal sound zones and give an overview of what has already been done.

1.1 State of the art

According to Betlehem et al., 2015, personal sound zones were first proposed by Druyvesteyn and Garas, 1997 where a separation of sound sources was required in order to create personal sound. Druyvesteyn and Garas, 1997 considered a person A positioned in listening area A listening to program a at a certain loudness level, and a person B in the same room but positioned in listening area B, who did not want to listen to program a. In order to keep the sound pressure level (SPL) of program a as low as possible at listening area B, they tried different methods. They found that active noise control worked well for frequencies below $1000Hz$ when they used a multi-input multi-output feed-forward control system. For the high-frequency region ($4000Hz$ and above) they found that a directional behavior of each loudspeaker was very effective while a loudspeaker array worked well in the mid-frequency region ($1000Hz$ - $4000Hz$). Even though Druyvesteyn and Garas, 1997 may have been the first to propose a personal sound system, a lot of research had already been done in the field of active noise and sound control which may have been the first stepping stones to the research of personal sound systems. Before Druyvesteyn and Garas, 1997 proposed personal sound zones, Elliott and Nelson, 1988 were researching the active control of sound in aircraft cabins and vehicle cabins. Here the object was to reduce the noise from the engine and other sources of noise coming from the aircraft or vehicle inside the cabin. This research have led both Nelson and Elliott in the direction of personal sound zones in their later research, where they among others have been researching the generation of personal sound zones using mobile devices and the generation of personal sound zones in a vehicle cabin (Nelson, 2014; Elliott, 2009).

When talking about personal sound zones, the term bright zone and dark zone often comes up. According to Shin et al., 2010 these terms were first introduced by Choi and Kim, 2002 where they defined the bright zone to be where the acoustic energy was high and the dark zone to be where the acoustic energy was low. Here they described two methods for obtaining these two zones. The first method focused on maximizing the energy in the bright zone using the input source power also described in Chapter 1 and the second method was about maximizing the contrast between the two zones also described in Chapter 1.

So far there have been a lot of research in the field where the sound zones have been fixed, which of course makes the problem much more simple, than if the sound zones were to move around in the room or follow a person around. A simple solution for creating fixed sound zones could, for example, be by placing loudspeakers in the ceiling above the sound zone. This would mean that the distance between the zones would be masking the sound from another zone. This method would, however, not make it possible to change the direction of the sound arrival, but it would be possible to create binaural sound using head tracking. Another simple method could be to create independent loudspeaker arrays for each sound zone. This could be done using separate monopole and dipole speakers because they can create a sound field inside the zone without generating sound outside the speaker array. This, however, makes it difficult to alter the position of the zones and the number of zones (Poletti, 2008). With multi-zone systems, it is easier to alter the position of the zones. In these multi-zone systems either one big or multiple smaller loudspeaker arrays can be used. However, the multiple smaller loudspeaker arrays will as the speakers in the ceiling put some constraints on the system (Poletti, 2008).

Most research so far has in mind that the systems zones should be adaptable and able to move around. Still, a lot of the research have still just focused on two or more fixated zones. So what is needed in order to take the next step and make the zones move around with the people in the room? The next section will take a look at what can be done in this direction and how this project can be different compared to what already has been done.

1.2 How can this project be different?

According to Section 1.1 most of the current research within personal sound zones have focused on creating sound zones that are fixed. Most of them have two or more zones that either function as a bright zone or a dark zone. However, in an ideal sound zone system situation, the user would not want to be constrained to a fixed position all the time. In fact, in an ideal sound zone system, the user should be able to move around and have the sound zone follow.

In order to create a sound zone system that would be able to move the zones around in the room and preferably follow the user around in the room, a tracking system would be needed. The tracking system would need to be able to track multiple people at once and be able to know where people are situated in the room at all times. There are multiple ways such a tracking system can be created, which all have pros and cons. If a visual tracking system would have to be implemented, for instance, cameras would have to be put up in the room. This may be undesirable in some rooms, such as living rooms, where some people may feel uncomfortable with being filmed all the time. Another drawback to visual tracking is occlusion, which of course can be solved by implementing multiple cameras that shoot from different positions in the room and by that from different angles. A pro to visual tracking is that the user not necessarily need to wear any device in order to be tracked depending on the visual tracking method.

Another way to implement a tracking system is to use acoustic tracking. There are also multiple ways of doing acoustic tracking, and some methods are better suitable for one thing and not other things, which means that acoustic tracking also has some pros and cons. One method is to use audible frequencies. This form of tracking can be used to determine the position or direction of a sound source. This information can be used in hearing aids, as well as in surveillance systems (Jensen, Nielsen, Heusdens, & Christensen, 2016). Another method is to use inaudible sound such as ultrasound, which is above 20kHz. When doing a quick search on ultrasonic tracking it is mostly used for either medical purposes, like ultrasonic images of fetuses in the womb, or used for underwater tracking. However, according to Deak, Curran, and Condell, 2012, ultrasound can also be used for indoor localization. They describe a system called Active Bats (Hodges & Hopper, 2001) where several sensor receivers are placed in the ceiling one square meter apart. The user wears a transmitter and from that, the sensors in the ceiling are calculating the time of flight in order to determine the position of the user. This, however, is very costly and the fixed sensors in the ceiling require to be positioned very precisely (Deak et al., 2012).

One advantage of using ultrasound compared to audible sound is that even though there will be microphones recording in the room it is possible to only record in the ultrasonic spectrum which means that the user itself is not being recorded. A downside to ultrasonic tracking as could be seen with the active bats is that the user will have to wear a device that emits ultrasound.

As mentioned in Chapter 1 there are several methods for creating a sound zone system and that the methods that should be used are depending on the design of the sound system. It was also mentioned that a simulation of a sound zone system could help testing which methods would fit the design best. Therefore, this project could differ from the other projects by creating a tracking system for a simulation of a personal sound zone system, where the sound zones should be able to move around.

1.3 Problem Statement

As mentioned in Section 1.2 there are pros and cons to all tracking methods, which means that the specific project should be in mind when considering which tracking system should be used. The nature of this masters thesis is Sound and Music Computing, which means that a natural choice is to look into acoustic tracking and since personal sound zones is a fairly new concept there have not been much research in how people interact with such a system. Building a personal sound zone system is difficult and so far there has not been build a system where it is possible to move the sound zones around. Therefore it is very interesting to create a simulation of such a system to see how people would react to it and interact with it before building it. Also, to be able to determine which methods should be used to creating it. In a simulation of such a system, the user will have to wear headphones in order to simulate the zones. Since they would have to be wearing headphones, the ultrasonic device could easily be attached to them and the user will not notice the ultrasonic device. Furthermore, a sufficient acoustic tracking system has not been found. The aim of this project should, therefore, be looking into how a cost-efficient acoustic tracking system can be created for a personal sound system where the user should be able to move around in an indoor environment. This leads to a problem statement stating:

"How can a tracking system for a simulation of a personal sound system be made using acoustic tracking?"

1.4 Requirement specifications

In order to solve this problem and be able to conclude on this project some requirements will have to be specified.

Because of the inaudible advantage that ultrasound has, the tracking system should be using ultrasound. This will not disturb the user and it makes it possible to only record in that spectrum and not record the speech of the user.

The indoor ultrasonic tracking system described in Section 1.2 was very costly and not very adaptive because of the many sensors that needed to be put into the ceiling. Therefore, this system should be designed so that it is adaptable and cost-efficient. It should be possible to move the system from room to room and easily expand it if a bigger room is used.

With these requirement specifications, the design of the system can start. The design of the system is described in the following chapter.

Chapter 2

Analysis & Design

Chapter 1 gave an introduction to the problem that a tracking system for personal sound zones needs to be made in order to make the sound zones able to move around with the user. It was decided that this tracking system should be made with ultrasound, and it should be cost-efficient because according to Deak et al., 2012 they are very costly. This chapter will analyze some of the existing solutions and from that design a system that would fit the requirements of this project.

According to Hodges and Hopper, 2001, there are several indoor sensor tracking systems. These include radio systems, infrared systems and ultrasonic systems such as Cricket (Priyantha, Chakraborty, & Balakrishnan, 2000). However, what is similar for all of them is the fact that they are not very accurate in terms of 3D location. Cricket is a location-support system, which helps devices learn their location. It uses beacons mounted on the ceiling or the wall. Each beacon proclaims a geographic space where they distribute information about this place. The information in the beacon can be obtained by a listener, which each device has attached to it. The Cricket system is a low-cost system, as the beacons are inexpensive. The Cricket system does not explicitly track the user location, but it helps the different devices learn where they are, and it can be decided who the information in the beacon should be conveyed to (Priyantha et al., 2000). The Active Bats System described in Hodges and Hopper, 2001, is another system that is able to track people using 3D locating data, which has an accuracy of less than one meter. Actually, 95% of the readings are accurate to within $3cm$ (Hodges & Hopper, 2001) of the position. This system, as mentioned before, uses sensor receivers mounted in the ceiling where they are all connected to a wired network. The user of the system will have to wear a wireless transmitter that emits ultrasonic information to the receivers in the ceiling. The system then uses time of flight (TOF) to estimate the location of the transmitter. Because of power-saving, the transmitters are not emitting ultrasound all the time. A wireless network is used to trigger the transmitters whenever the transmitter is in the proximity of a relevant area. This wireless network is also used to update the transmitters. Even though this system is like the Cricket system, the Active Bats system performs better in terms of the accuracy of locating the transmitters. However, this better performance comes with a cost. The Active Bats system is costly compared to the other systems and the scalability of the system is limited because of the mounted sensors in the ceiling which are demanding precisely positioning (Deak et al., 2012; Hodges & Hopper, 2001).

As mentioned in Section 1.2, a way of tracking audible sound sources was to find the direction of arrival (DOA) of the source. In Jensen et al., 2016 they propose a model to determine the DOA of an audible signal with reverberation using a microphone array. In this paper, they assume a Uniform Linear Array (ULA) but mentions that the methods they developed could just as well be used with any other array structure. They concluded that their model outperformed existing DOA models that did not consider the early reverberations.

This way of determining DOA of an ultrasonic sound source for indoor localization have not been researched extensively, which makes it interesting to look into this. The following section gives an introduction to different microphone array geometries.

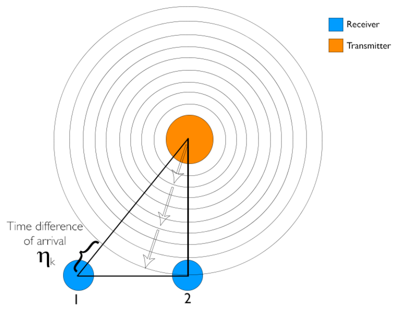
2.1 Microphone Arrays

Compared to a single sensor, an array of sensors have an enhanced performance when it comes to location performance (Stoica & Moses, 2005). These arrays of sensors can come in all kinds of geometries which all perform differently. According to Kawitkar, 2009, the most investigation has been on the Uniform Linear Array (ULA) and he aims to compare this array geometry to less investigated geometries such as Uniform Circular Arrays (UCA) and Uniform Rectangular Arrays (URA). He found that the ULA performed best at the center of the array, while it performed worse near the end fire, this was due to a loss of spatial resolution at each end of the array (Kawitkar, 2009). This problem was also noted in a paper by Baysal and Moses, 2003. The symmetry of the UCA and URA resulted in a better uniform performance in this matter. He also found that the UCA had some advantages towards the ULA and URA, but because the structure of the UCA does not possess the Vandermonde form as the other two geometries, it is not possible to use fast mathematical algorithms such as fast Fourier transform on this kind of geometry (Kawitkar, 2009; Lau, Leung, Liu, & Teo, 2006). However, according to Lau et al., 2006, it is possible to transform the steering vector of the UCA into a Vandermonde form. There are two main approaches to this transform, but this, however, is outside the scope of this report and can be read about in the paper by Lau et al., 2006.

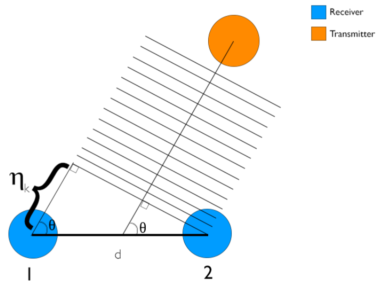
In this project, a UCA will be used using the DOA estimation model described in the following section.

2.2 Direction of Arrival Estimation

Direction of Arrival (DOA) estimation has been used for array signal processing for some time now. It has been used in several applications such as acoustical tracking, radar detection, sonar localization and in mobile communication systems (Lee, Hudson, & Yao, 2014).



(a) Wave propagation from the sound source.



(b) The far field assumption model.

Figure 2.1: The wave propagation and far field assumption (Kunin, Jia, Turqueti, Saniie, & Oruklu, 2011).

Normally sound waves propagate spherically from a source. But if we only consider signal sources in the far field we can assume that by the time the spherical sound waves reach the microphone array they can be approximated by plane waves (Kunin, Jia, Turqueti, Saniie, & Oruklu, 2011). Figure 2.1a demonstrates how the spherical sound waves are propagating from the sound source towards the microphones. It can be seen that the sound will reach microphone 2 before microphone 1. The delay with which the sound is arriving at the two microphones is used to calculate the DOA. If it is assumed that the source is in the far field, Figure 2.1b shows how the spherical waves are approximated by plane waves, which creates a triangle between the plane waves and the two microphones.

With this assumption the signal model can be written

$$x_k(n) = \beta s(n - \eta_k) + e_k(n) \tag{2.1}$$

for $n = 0, 1, \cdots, N-1$, where β is the attenuation of the signal from the source to the microphone, s is the source signal, η is the delay from the source to the microphone k in samples and $e_k(n)$ is added noise (Stoica & Moses, 2005). This noise can be internal noise from the microphones, but it can also be background noise.

As mentioned in 2.1, a uniform circular array (UCA) will be used for the DOA estimation. This has an influence on the next equation because this equation is different whether it is a linear array, circular array or something third. As seen on Figure 2.1b this equation calculates the delay η_k between the plane sound wave reaching one microphone and then the other microphone (Kunin et al., 2011). The delay η_k for $k = 1, \ldots, K$ microphones can be calculated by

$$\eta_k = \frac{f_s}{c}(d - r\cos(\theta - 2\pi\frac{k-1}{K})) \tag{2.2}$$

where f_s is the sampling frequency [Hz], c is the propagation speed [m/s], d is the distance to the source [m], r is the radius of the UCA [m] and θ is the DOA [rad] (Nielsen, Jensen, Jensen, Christensen, & Jensen, 2016).

We assume that the source signal is narrow band because we only need the transmitter to emit a continuous sinus wave of some frequency above $20kHz$ and not a broadband signal such as speech. The signal should have a frequency above $20kHz$ because the signal should be ultrasound as stated in Section 1.4. The fact that the signal is narrow band makes the model simpler as we do not need to take the multiple components of a broadband signal into consideration. This means, that the signal $s(n)$ in Equation 2.1, which is the signal we are looking for in the far field, can be written as $s(n) = \alpha e^{j\omega n}$ where $\alpha = Ae^{j\phi}$ with A being the amplitude and ϕ being the phase of the source signal. If we replace the above and using that $e^{a+b} = e^a e^b$ Equation 2.1 can be rewritten as

$$x_k(n) = A\beta e^{j\phi}e^{j\omega n}e^{-j(\epsilon-\rho\cos(\theta-2\pi\frac{k-1}{K}))} + e_k(n) \tag{2.3}$$

where $\epsilon = \omega\frac{f_s}{c}d$ [rad], and $\rho = \omega\frac{f_s}{c}r$ [rad]. Using $e^{a+b} = e^a e^b$ again, $x_k(n)$ can be narrowed further down to

$$\begin{aligned} x_k(n) &= A\beta e^{j\phi}e^{j\omega n}e^{-j\epsilon}e^{j\rho\cos(\theta-2\pi\frac{k-1}{K})} + e_k(n) \\ &= A\beta e^{j(\phi-\epsilon)}e^{j\omega n}e^{j\rho\cos(\theta-2\pi\frac{k-1}{K})} + e_k(n) \\ &= e^{j\omega n}e^{j\rho\cos(\theta-2\pi\frac{k-1}{K})}z + e_k(n) \end{aligned} \tag{2.4}$$

where $z = A\beta e^{j(\phi-\epsilon)}$. θ and z are unknown in this equation. In order to find these two unknowns we first stack the signal for all microphones in a vector

$$\underline{x} = \begin{bmatrix} \underline{x}_1 \\ \vdots \\ \underline{x}_K \end{bmatrix}. \tag{2.5}$$

For each microphone, we have a vector containing the signal for $n = 0, 1, \cdots, N-1$ using this we can write Equation 2.4 as

$$x = \begin{bmatrix} x_1(0) \\ x_1(1) \\ \vdots \\ x_K(N-1) \end{bmatrix} = \begin{bmatrix} e^{j\omega 0}e^{j\rho cos(\theta-2\pi\frac{1-1}{K})} \\ e^{j\omega 1}e^{j\rho cos(\theta-2\pi\frac{1-1}{K})} \\ \vdots \\ e^{j\omega N-1}e^{j\rho cos(\theta-2\pi\frac{K-1}{K})} \end{bmatrix} z + \begin{bmatrix} e_1(0) \\ e_1(1) \\ \vdots \\ e_K(N-1) \end{bmatrix} \tag{2.6}$$

Simplifying this we can write

$$\underline{x} = \underline{h}(\theta)z + \underline{e}. \tag{2.7}$$

To be able to find θ and z we first need to minimize Equation 2.7 and find z

$$\begin{aligned} \underline{e}^H\underline{e} &= (\underline{x} - \underline{h}(\theta)z)^H(\underline{x} - \underline{h}(\theta)z) \\ &= \underline{x}^H\underline{x} - z^*\underline{h}^H(\theta)x - \underline{x}^H\underline{h}(\theta)z + |z|^2\underline{h}^H(\theta)\underline{h}(\theta) \end{aligned} \tag{2.8}$$

First we isolate z using the following equation where we find the complex derivative.

$$\begin{aligned} \frac{\partial f(z)}{\partial z^*} &= \frac{\partial}{\partial z^*}\Big(a|z|^2 + b^*z + bz^* + c\Big) \\ &= \frac{\partial}{\partial z^*}\Big(azz^* + b^*z + bz^* + c\Big) = az + b = 0 \end{aligned} \tag{2.9}$$

Equation 2.9 can then be solved for z which gives

$$z = \frac{-b}{a} \tag{2.10}$$

The values of a and b can be found in Equation 2.8 and be put into Equation 2.10, which gives us

$$z = \frac{\underline{h}^H(\theta)\underline{x}}{\underline{h}^H(\theta)\underline{h}(\theta)} \tag{2.11}$$

Now that we know z we can put it back into Equation 2.8 and then we get

$$\underline{x}^H\underline{x} - \underline{x}^H\underline{h}(\theta)[\underline{h}^H(\theta)\underline{h}(\theta)]^{-1}\underline{h}^H(\theta)\underline{x} \tag{2.12}$$

where $\underline{x}^H\underline{x}$ is just a scalar and $[\underline{h}^H(\theta)\underline{h}(\theta)]^{-1}$ is equal to KN. This means that we can just disregard them and the rest we can write as the cost function $|\underline{h}^H(\theta)\underline{x}|^2$. We can then find θ by taking the maximum of the cost function

$$\theta = \underset{\theta\in\left[0,\,\frac{2}{\pi}\right]}{\arg\max}\, |\underline{h}^H(\theta)\underline{x}|^2. \tag{2.13}$$

This model is going to be used to estimate the direction the sound source is coming from in the tracking system. But before implementing it into the system a simulation can tell if the model is working for a scenario where the sound source is ultrasound. Furthermore, the simulation of the model can also be used to determine some of the parameters of the microphone array. The following section will look at the simulation of the signal model, described in this section.

2.3 Simulation

This section gives an in-depth description of the DOA simulation created in order to design the microphone array that we are going to use. When designing the microphone array different parameters such as how many microphones and the radius is important. These parameters change dependently of each other.

Simulation of some mathematical model is often used to test the normal behavior of a system. Often in a mathematical model, a random component is included (Voss, 2014). The random component of the signal model described in Section 2.2 is the noise that we add to the model. In order to be able to test the behavior of the system a large number of samples from the model needs to be generated. Questions such as how the performance of the system differs at different sensor array radiuses can then be answered by analyzing this large set of samples from the model. The method for this is called the Monte Carlo Method (Voss, 2014). Simulating using the Monte Carlo method means that we can learn about the behavior of the model by studying the samples instead of the model itself (Voss, 2014).

As mentioned earlier the objective of doing this simulation is to test the performance of the signal model at different sensor array radiuses in order to design the sensor array so that it performs best possible. The radius of the array is important because if the radius is not at the right size some undesired spatial aliasing can occur when the signal is sampled. For a uniform linear array the spacing between the microphones is connected to the frequency of the signal and the spacing should obey $d \leq \frac{\lambda}{2} = \frac{c}{2f}$, where λ is the wavelength of the signal and c is the propagation speed, in order to avoid spatial aliasing (Benesty, Chen, & Huang, 2008). When spatial aliasing occurs we will see a peak in the cost function at multiple different degrees. The equation shows that if the frequency gets higher, the spacing of the microphones should get smaller. We assume that this is the same case for a uniform circular array.

The different radiuses will also be tested with different signal to noise ratios (SNR), in order to take different noise scenarios into account when testing the performance. The Monte Carlo method calls for a large set of samples for analyzing, so the simulation of the signal model will be iterated 500 times.

The signal model described in Section 2.2 was implemented in MATLAB, where it was tested for 20 different radius values going from $0.005m$ to $0.1m$ with an increment of $0.005m$ and with 16 SNR values going from $-10dB$ to $20dB$ with an increment of $2dB$. The frequency of the signal we were looking for was $40kHz$. Two different simulations were created, one with 3 microphones in the sensor array and one with 6 microphones in the sensor array. The 500 samples of theta values gathered from the signal model for each SNR in each radius were gathered in a $500x16x20$ array. The theta that the signal model needed to compute was where the position of a source was known. This means that the true theta is also known and that we can compare the true theta with the calculated theta for each SNR in each radius. The code for the simulation can be found in Appendix A.

2.3.1 Results

A root-mean-squared-error (RMSE) of the 500 iterations of each radius and SNR were made in order to analyze the behavior of the signal model. The 500 iterations were matched against the true theta value for the given position of the source. The values for each SNR were then plotted with the radius on the x-axis and the RMSE on the y-axis to give a visual representation of the data.

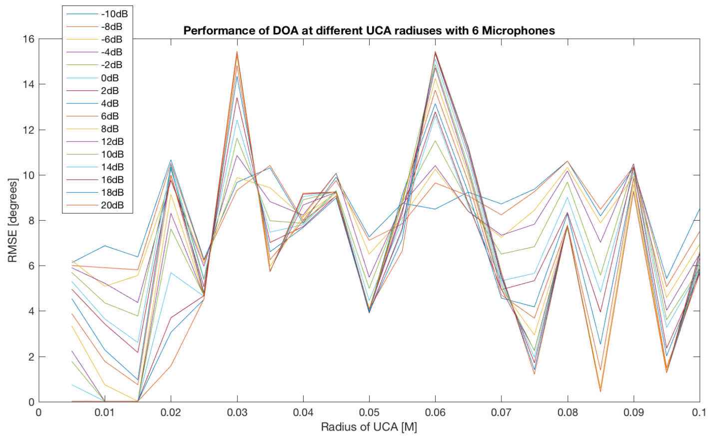


Figure 2.2: Simulation with 3 microphones.

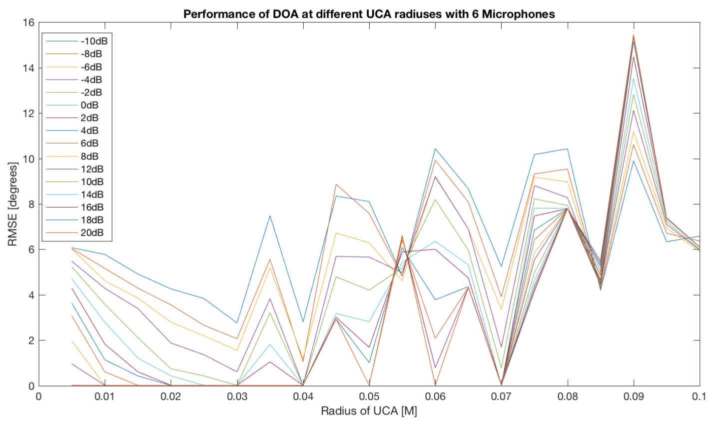


Figure 2.3: Simulation with 6 microphones.

Figure 2.2 shows the RMSE for the simulation with 3 microphones. The better the performance of the array the closer to zero it gets. It is visible in Figure 2.2 that most of the SNR's except for -10, -8, and $-6dB$ are all going towards zero until the radius is $0.015m$. After the radius of $0.015m$ all values are going up and down interchangeably for the rest of the radiuses. It is seemingly here that the aliasing mentioned in Section 2.3 occurs and we can not rely on the results of the model being correct when the radius of the sensor array is above $0.015m$ for 3 microphones.

In Figure 2.3, however, it seems that the radius of the sensor array can be significantly bigger. Here all the SNR's are going towards zero up until $0.03m$ radius. The $20dB$ however, are zero up until $0.04m$ radius. Hereafter the plot also interchangeably goes up and down for the rest of the radiuses.

2.3.2 Conclusion

It was visible from the results of the simulation that if an UCA was created with a radius of $0.015m$ it would be possible to use it with both three microphones and 6 microphones. Furthermore, it was also visible from the two plots that the radius of the UCA is dependent on the number of microphones in the sensor array.
In future simulations, it could be a good idea to add another random element to the simulation. In this simulation, the source position was fixed. In future simulations, it could, therefore, be interesting to look at how the signal model and UCA performs if the source position is random for each iteration of the Monte Carlo simulation.

2.4 Designing the Microphone Array

The Monte Carlo simulation described in Section 2.3 gave some parameters that have to be taken into consideration when designing the initial microphone array. These parameters were chosen after the simulation showed that the DOA estimation was performing best with these parameters. The simulations showed that if the radius of the simulation was $0.015m$ it would be able to perform well with both 3 and 6 microphones. The microphone array should, therefore, have a radius of $0.015m$ because this would allow for making tests with both three microphones and six microphones, without having to create two different microphone arrays.

Figure 2.4 shows the first design of the microphone array. The array should be able to sit in a regular microphone stand. Measurements of the microphone stands showed that the handle of the microphone array should be approximately $0.02m$ in diameter. The handle should also be round so that it would fit best into the microphone stand. The holes for the transducers should, of course, be the size of the transducers so that they fit. The radius $0.015m$ is measured from the center of the array to the center of the transducer.

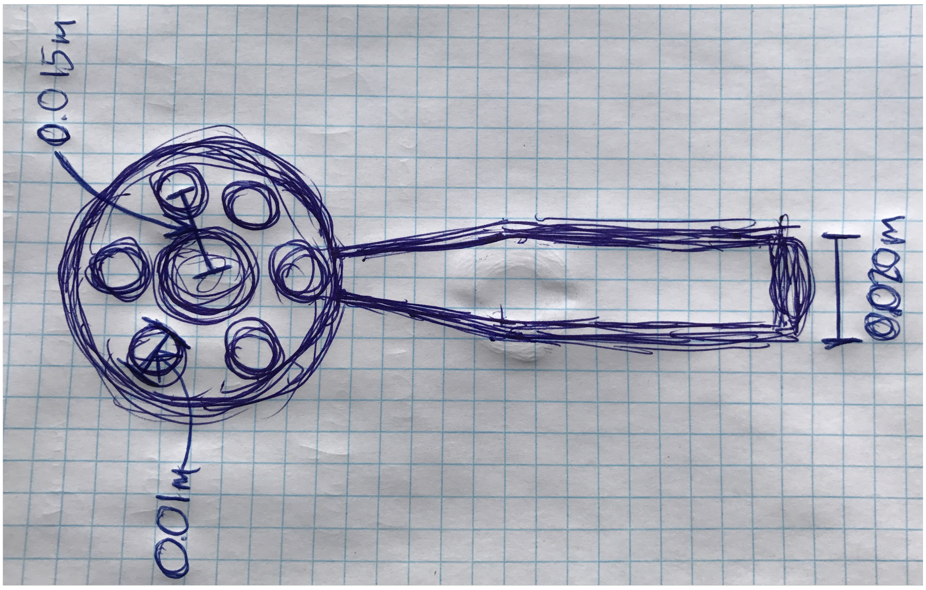


Figure 2.4: Sketch of the first design of the microphone array.

2.5 Designing the Transmitter

Because a tracking system using ultrasound requires the user to wear an ultrasound transmitter, the transmitter should be as small as possible and as discrete as possible. The components for this should, therefore, be picked with this in mind. Since this tracking system is being built for a simulation of a sound zone system, it is possible to attach the transmitter to the headphones that the user would have to wear. Therefore the first iteration of the design of the transmitter should just be a small box that can fit the transmitter and the components for it.

If this tracking system were to be used in a real sound zone system the ideal would be to find a way to implement it in devices that the user already carries around, such as a watch or a mobile phone. But for now, the focus is on getting the system to work for a simulation of a sound zone system.

This chapter has now covered some analysis of the problem and designed a system that now needs to be implemented. What should be kept in mind when starting to implement the system is that the system should be built at low cost.

Chapter 3

Implementation

In Chapter 2 the problem of creating a tracking system for a simulation of a personal sound zone system was analyzed and the system was then designed hereafter. The next step is to implement this design. Since the signal coming into the microphones will be a weak signal due to the distance an amplifier for each sensor in the microphone array needs to be built. The amplified signal should then be loaded into the computer using an audio input device. The audio input device converts the analog signal into a digital signal that we can use in the algorithm explained in Section 2.2, which is being implemented in Matlab. A mount for the microphones was also needed in order to hold them. This chapter will look at how the microphone array and the amplifiers were built and how the algorithm was implemented in Matlab. There will first be a section about the building of the microphone array. This will go into depth with realizing the design described in Section 2.4 and then give a description of the amplifiers built for the microphone array. In the end of this chapter, there will be a description of the implementation of the signal model into Matlab and what extensions there might be to the algorithm.

3.1 Building the Microphone Array

In Section 2.4 the microphone array was designed. The requirements was that the array should be circular, fit into a microphone stand, and have a radius of 0.015m. The ultrasonic receivers gotten for this project was Murata's MA40S4R. They have a center frequency of $40kHz$ which means that this is the frequency that they work best around. The diameter of the sensor is $0.0099m$, which means that the holes in the microphone array should be slightly bigger so that they can fit inside (Murata, 2017). It took some trial prints before getting the size for the holes completely right. The radius of each hole should be $0.0052m$, in order for the sensors to be able to fit without falling out.

Because of one of the requirements in Section 1.4 being that the system should be low cost, it was decided to use a 3D printer to create the microphone array. A modeling tool called SketchUp was used to create the 3D model of the array. Figure 3.1 and 3.2 shows the 3D model of the designed microphone array with room for up to 6 microphones. The 3D printer used to print this model was a Ultimaker 2 Extended+. Figure 3.3 shows the printed microphone array with 6 sensors fitted in.

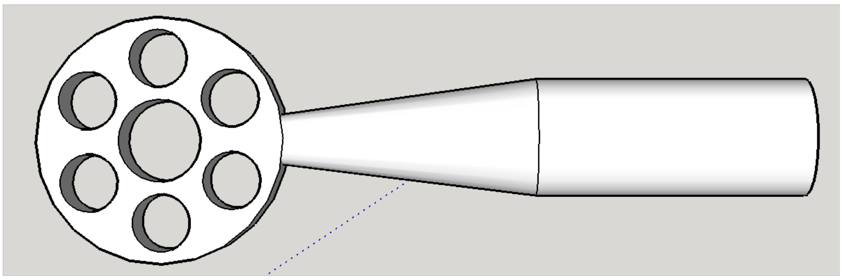


Figure 3.1: Top view of the 3D model in SketchUp.

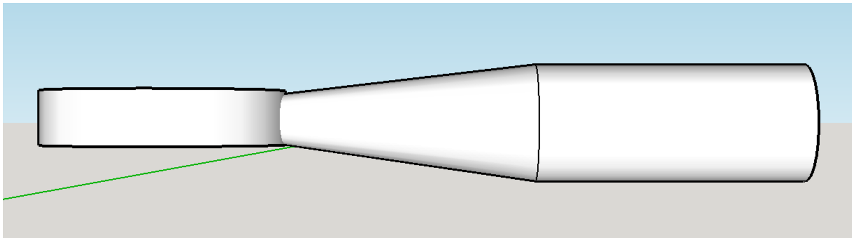


Figure 3.2: Front view of the 3D model in SketchUp.

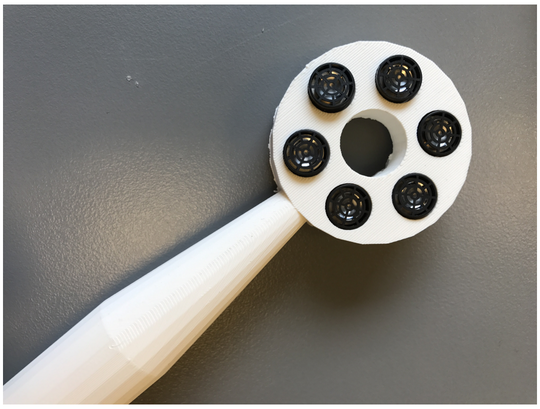


Figure 3.3: The 3D printed version of the first design of the array.

After printing the first design of the array, it was realized that since the array would have to be vertical as seen in Figure 3.2 the handle of the array should be placed differently. The handle in the first design would be interfering with the signals coming from that direction because it is higher than the surface of the sensors. Therefore a second design was developed. Some initial tests of the system also showed that the radius of the array could be smaller in order to avoid spatial aliasing. It was therefore decided to decrease the radius as much as possible. However, the size of the sensors only allowed for making the radius $0.0038m$ smaller, which means that the new radius was $0.0112m$. In the new design, it was decided to place the handle underneath the sensors. The second design can be seen in Figure 3.4, 3.5 and 3.6.

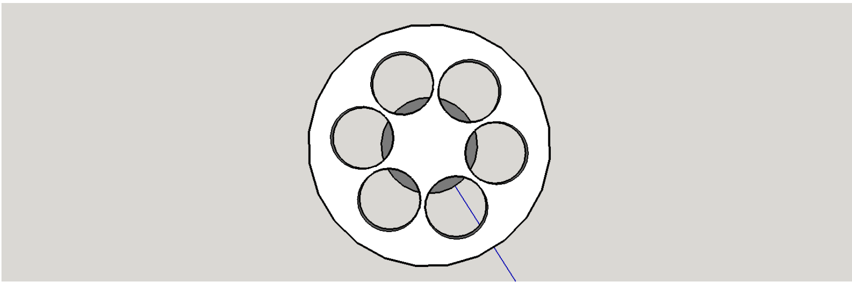


Figure 3.4: Top view of the second design of the 3D model in SketchUp.

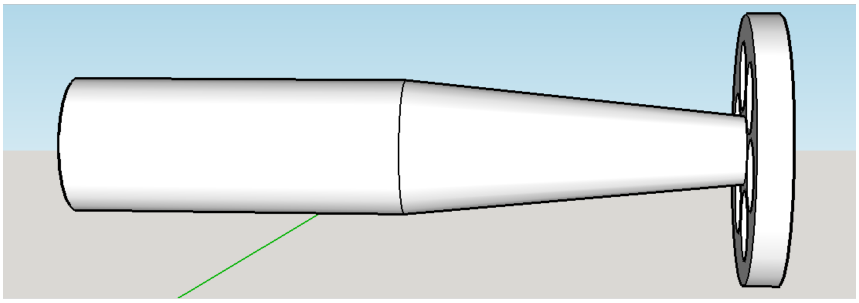
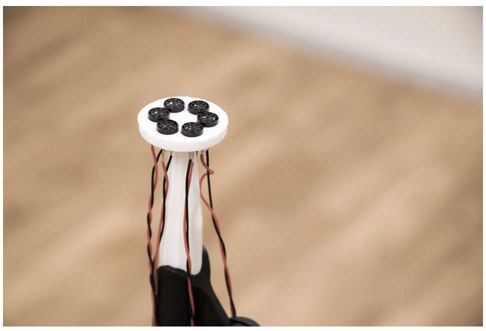


Figure 3.5: Front view of the second design of the 3D model in SketchUp.



(a) Pespective view of the second design printed.

(b) Front view of the second design printed.

Figure 3.6: The second design after being printed.

The signal that is received through the sensors needs to be amplified before it is sampled and sent to the computer in order to get a decent reading. This will be described in the next section.

3.2 Amplifying the signal

Since the signal that is being received by the microphones is very weak, it needs to be amplified. For this, we have built an amplifier for each microphone. Before building the amplifiers, the circuit was drawn in a simulation program called LT-spice. With this program, you can see the signals behavior throughout the circuit. This is very useful when you have to debug your actual circuit because you can compare the signal at different points in the circuit.

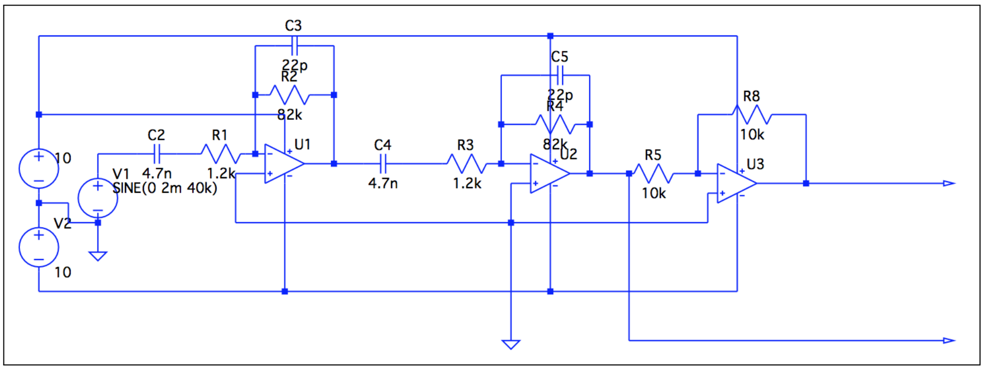
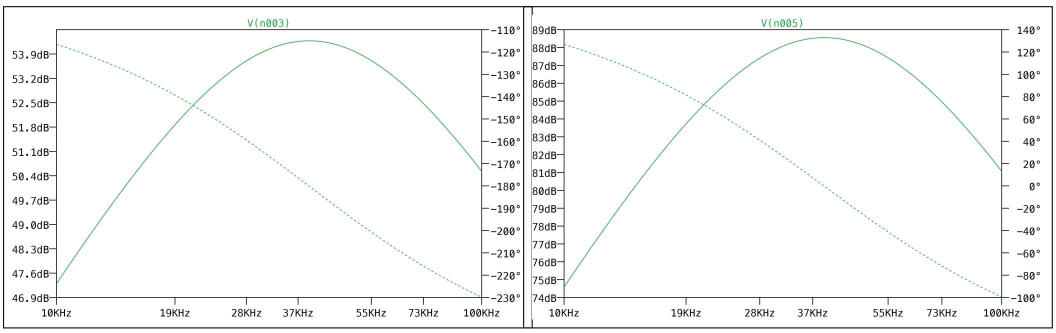


Figure 3.7: The schematics of the amplifying circuit.

Figure 3.7 shows the schematics of the amplifying circuit. In this circuit, we have used an oscillating voltage source to simulate the sinus wave that will be received through the microphones. The circuit uses three operational amplifiers. The first two functions as a two stage inverted bandpass filter (Wong, 2011). Each stage amplifies with approximately $20log(82k/1.2k) = 36.69dB$, which means that the signal is being amplified with a total gain of $73.38dB$.



(a) Bandpass filter after stage 1. **(b)** Bandpass filter after stage 2.

Figure 3.8: The frequency response after the first and second stage showing the bandpass filter.

In LTspice it is possible to do an AC sweep, which makes it possible to plot the frequency response of the output. This can be used to see the bandpass filter at each stage. Figure 3.8 shows the frequency response of the output at each of the first two stages of the amplifier. Both have a center frequency around $40kHz$, however, the frequency response in Figure 3.8b is amplified.

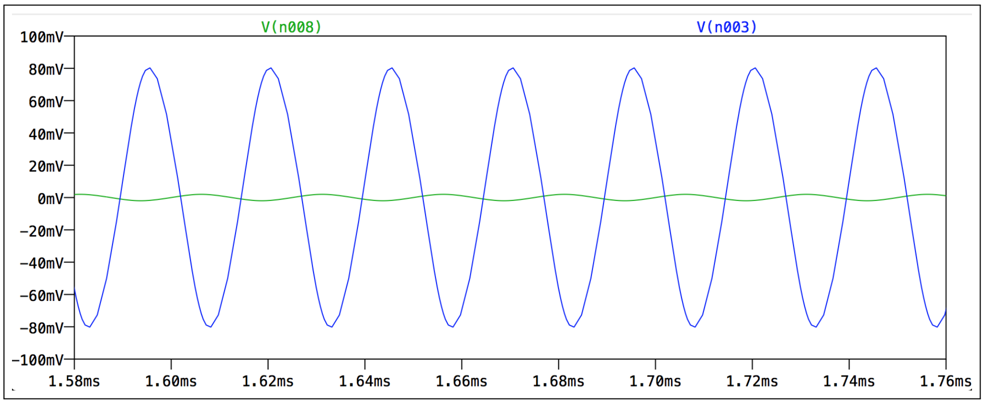


Figure 3.9: The signal before and after the first stage.

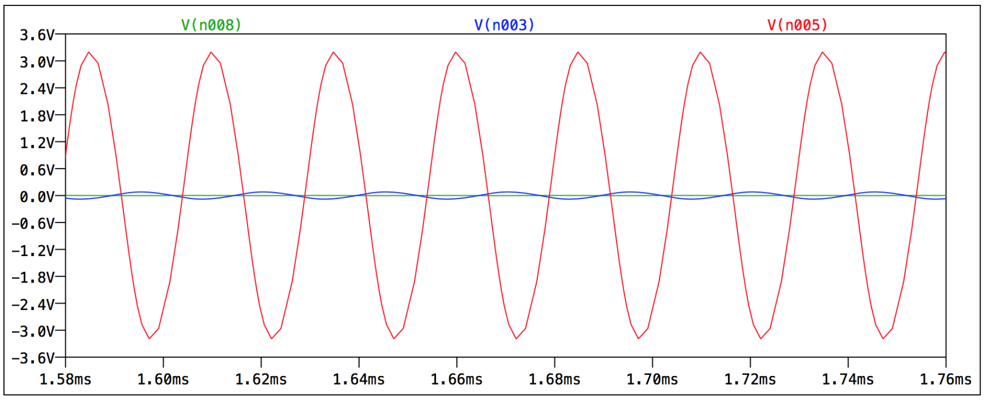


Figure 3.10: The signal before in between and after the second stage.

Figure 3.9, 3.10 and 3.11 is from a simulation in the program LTspice. Figure 3.9 shows the signal before being amplified (green line) and then after the first stage where it has been amplified with the gain of $36.69dB$ (blue line). In Figure 3.10 the signal has been amplified additionally $36.69dB$ (red line). The last stage is used to create a differential signal which creates a stable signal that is more robust to noise. This means that this stage does not have a gain or bandpass filter on it, it just inverts the phase of the signal so that it looks like the Figure 3.11.

After testing the circuit in LTspice, the circuit was built on a Veroboard. For the operational amplifier, a TL074 Quad Low Noise Op-amp was used. This op-amp has four op-amps build-in. This makes it less space consumable. This op-amp was chosen because it produces low noise, which is very desirable.

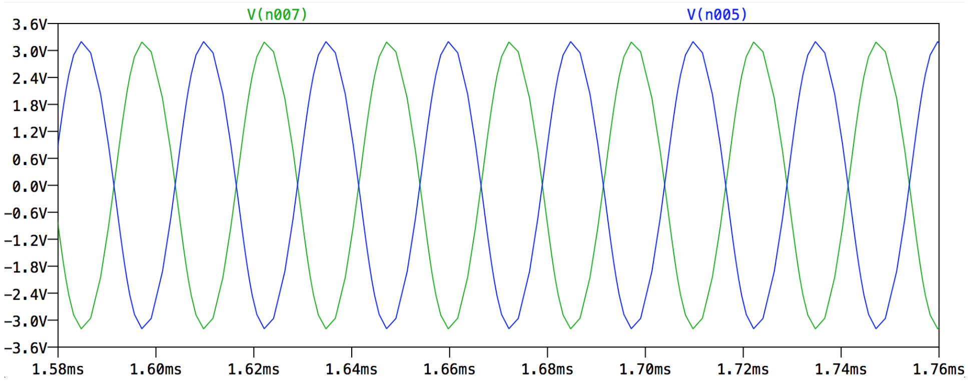


Figure 3.11: The amplified signal together with the inverted signal.

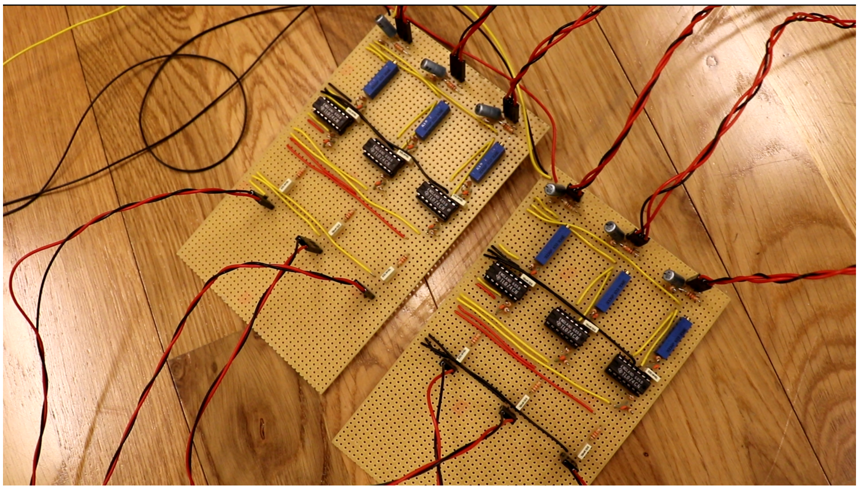


Figure 3.12: The circuit build on a veroboard.

Figure 3.12 shows the circuit after it had been built on a Veroboard. Each microphone is amplified by its own op-amp. The wires coming in, in the bottom of Figure 3.12 is from the microphones, and the wires going out are XLR cables that go into an audio input device, which is converting the signals from analog to digital so that we can get the signals into the computer.

3.3 Building the Transmitter

As of now the implementation of the transmitter is only consisting of a waveform generator which is attached to an MA40S4S sensor from Murata, 2017. Figure 3.13 shows the Murata transmitter which is attached to the generator in Figure 3.14.

As mentioned in Section 2.5 the transmitter should be wireless and should be able to be worn by the user on a pair of headphones. Making it wireless can be

(a) The transmitter from the front. **(b)** A closer look at the transmitter.

Figure 3.13: The transmitter.

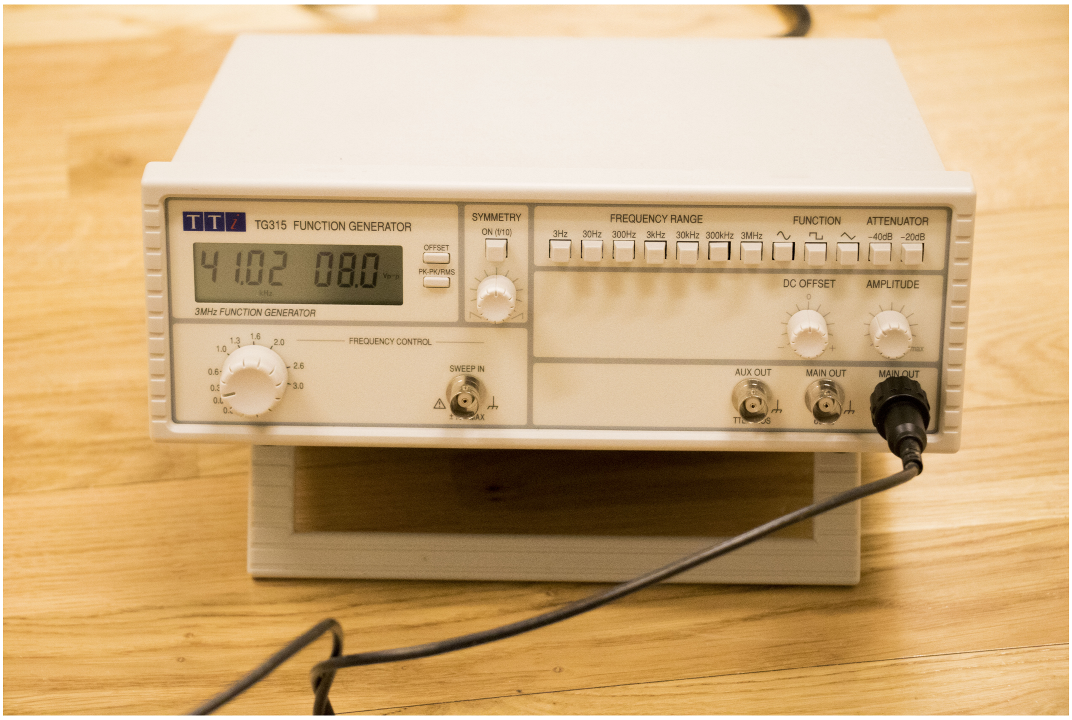


Figure 3.14: The waveform generator.

done by using a microcontroller such as an Arduino Nano. The Arduino Nano has a good size for this project, being only $18x45mm$ and weighing no more than $7g$. It is possible to power the micro-controller with a $9V$ battery, which means that the power supply should be inside the wearable box as well. The manufacturing cost of the transmitter would be low as well since the Arduinos are easy to get and cheap. The box that should contain the Arduino, the transmitter, and the battery, can be 3D printed just like the microphone array. The material in the 3D printers are cheap and is very light, so this will not add much weight to the device.

3.4 Algorithm

The signal model described in Section 2.2 is the model that is going to be used for determining the angle that the transmitter has towards the microphone array.

An algorithm was made in Matlab for the simulation of the signal model. In the simulation algorithm, the signals we were working on were generated on the computer and then some noise were put onto the signal. With just some smaller alterations this algorithm could be used with real life signals as well. With the 2017a edition of Matlab, it is possible to record audio from an audio input device. The control of the command prompt and the calling function is returned immediately after calling the record function. This means that other commands can be run while still recording (Mathworks, 2013).

```
fs = 96000;                                   %sampling freq
n = 96000;                                    %number of samples
freq = 41110;                                 %frequency that we are searching for in Hz
omega = (2*pi)*freq/fs;                       %frequency that we are searching for in omega range
micNum = 6;                                   %number of microphones
c = 343;                                      %propagation speed
thetaGrid = (1: 1: 360);                      % Theta range
radius = 0.0112;
rho = omega*(fs/c)*radius;


%% Setup audio recorder
deviceReader = audioDeviceReader(fs, n);
setup(deviceReader);
release(deviceReader);
deviceReader.ChannelMappingSource = 'Property';
release(deviceReader);
deviceReader.ChannelMapping = [1,2,3,4,5,6];


costFunc = zeros(1,numel(thetaGrid));         % preallocate costFunc

phase = pi;
N = (1:n)';
f = exp(1j*omega*N);                          % signal we are looking for
X = deviceReader();                           % the signal from channel 1-6 in the soundcard

 %% filter Signal
fc = [37000,45000];                           % cut-off frequencies
[b,a] = butter(2, fc/(fs/2), 'bandpass');     % butterworth bandpass filter

dataOut = filter(b,a,X);

X = hilbert(dataOut);                         % hilbert transform of the signal
X = X(:);

%% DOA
for thetaIdx = 1:length(thetaGrid)
    h = hFuncTrack(f, rho, (thetaGrid(thetaIdx)*pi/180),micNum);
    costFunc(thetaIdx) = abs((h'*X))^2;
end
[M,I] = max(costFunc);
index = thetaGrid(I);
```

Listing 3.1: TrackingSystem.m

```
function h = hFuncTrack(f, rho, theta, micNum)
N= length(f);
h = zeros(N*micNum, 1);
index = 1:N;
for k = 1:micNum
    h(index) = f*exp(1j*rho*cos(theta-2*pi*((k-1)/micNum)));     % calculates h in the signal function x=h(theta)*z+e
    index = index+N;
end
end
```

Listing 3.2: hFuncTrack.m

In line $13-18$ in Listing 3.1 the audio recorder is being set up. The audio recorder is told that it should be recording channel $1-6$ on the audio input device. Line 26 is the one that records and assigns the data to X. The recorder is set to record the same amount of samples as the sampling frequency, which means that every time this line is called, the recorder will record 1 second of sound. The signal is put through a bandpass filter in order to filter unnecessary noise out. This is what happens in line $29-32$ in Listing 3.1, where the cutoff frequency is set to 37000 and 45000. The Hilbert transformation in line 34 transforms the signal from real numbers to complex numbers. The for loop in line $38-41$ calculates the cost function between the real signal and the h vector described in Equation 2.6 in Section 2.2. In line 39 the function `hFuncTrack` is called. This function can be seen in Listing 3.2 and contains a for loop that calculates the h vector in the signal model in Equation 2.7 in Section 2.2. After the for loop in line $38-41$ in Listing 3.1 the maximum of the cost function is found just as in Equation 2.13 in Section 2.2. With this maximum, we can find the index in the theta grid which is going from $0-360\,\text{deg}$.

3.4.1 Extensions to the algorithm

This algorithm only calculates the direction of the sound. In order to be able to track the precise position of the user, two microphone arrays would need to be used. The second microphone array would have to be placed in another position in the room. When the direction of arrival is calculated for both arrays, it is possible to compute where these two angles are intersecting with each other.

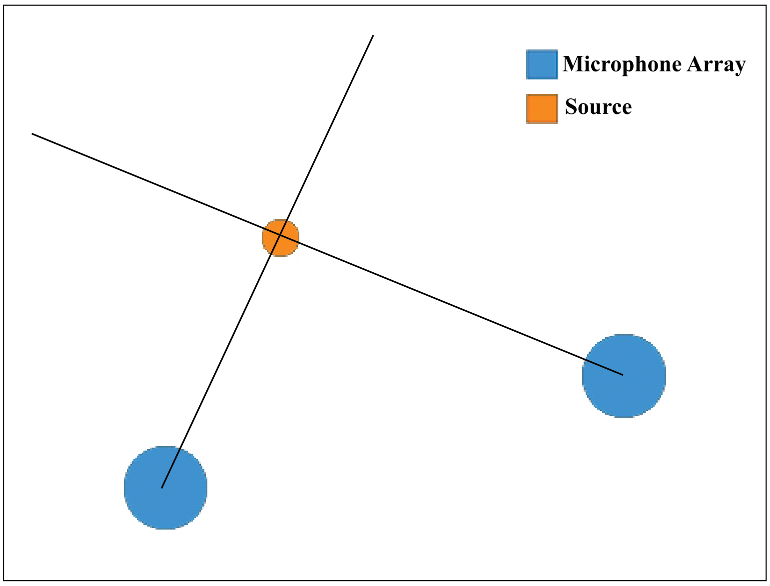


Figure 3.15: Illustration of the cross section of two microphone array's DOA estimation.

Figure 3.15 illustrates how the two DOAs will intersect exactly where the sound source is placed. The test in Chapter 4 will be performed on the algorithm in Section 3.4 in order to find out if the DOA estimation works. If this is the case, the next step will be to develop the algorithm where the precise position is being triangulated. This algorithm will need to know the size of the room, and where the microphone arrays are placed in the room. The DOA algorithm will then be turned into a function that this new algorithm can use to determine the DOA for each array.

Chapter 4

Testing

4.1 Initial Tests

During the building of the amplifier, several informal tests were made in order to see if the amplifier worked and the correct data was loaded into the algorithm. However, even though the data looked correct, the results from the DOA were not correct. The angles that were measured were consistent when doing several readings in the same place, but they did not correlate with the angle they were supposed to.

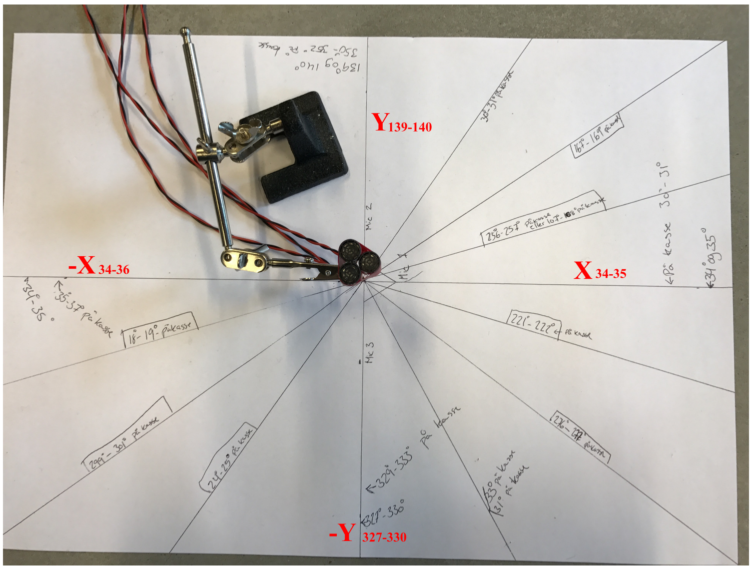


Figure 4.1: Informal test of the amplifiers with the DOA algorithm.

Figure 4.1 shows the measured angles from one of the informal tests. Here it is obvious, that there is no consistency in what angle is measured where. The setup for the informal test was in an electronics laboratory on the floor. This means that there were a lot of surfaces where the sound could reflect. Therefore, it was decided to test it in a big room with sound isolation on the walls, in order to be sure that the misreadings were not because of the reflections in the room. In the bigger room, there again was a lot of misreadings. Therefore it was decided to do some tests with the audible sound and some measurement condenser microphones. It was also in these informal tests that it was found out that the radius of the array could be even smaller than the initial $0.015m$.

4.2 Audible Test

In the beginning, the same audio input device was used for the audible test, as the one used for the initial tests with ultrasound. However, it quickly became obvious that it somehow was the audio input device that could be the source of the misreadings. The test was therefore done on another audio input device.

4.2.1 Setup

The test was performed in the sound lab at the Aalborg University facility at Rendsburggade 14. Six measurement condenser microphones from Behringer (Music Group, 2017) were used in the microphone array. They were placed in a 3D printed mount with the radius of $0.066m$, the 3D printed mount can be seen in Figure 4.2b. The signal that was looked for was a $1000Hz$ sinus wave transmitted by an iPhone 7 plus at highest gain using an app called Signal Generator by Media Punk Studios (iTunes Apple, 2017). Figure 4.2a shows the setup of the transmitter. In order to connect the microphones to the computer a Focusrite Scarlett 18i20 USB Audio Interface (Focusrite, 2017) was used. The audio interface was set to sample with a sample rate of $48kHz$ and provide the microphones with $48V$ phantom power. The angles that were measured was approximately 0 deg, 90 deg, 180 deg, and 270 deg.

(a) The transmitter. **(b)** The microphone array.

Figure 4.2: The transmitter and the microphone array in the audible test.

Figure 4.3a shows the setup of the audible test. The red tape on the floor marked where the approximate angle was, and where the transmitter should be placed. The distance between the front foot of the microphone stand, see Figure 4.3b, the transmitter was mounted on, to the middle of the big red cross in Figure 4.3a was $1.87m$. The microphone array was placed above the big red cross in Figure

4.3a. The distance from the floor to the top of the microphones was $1.56m$ and the distance from the floor to the transmitter was $1.61m$.

(a) Test setup for the audible test.

(b) Placement of the transmitter.

Figure 4.3: The setup of the audible test, with placement of the transmitter.

4.2.2 Method

The transmitter was placed at one of the four angles seen in Figure 4.3a. The algorithm which was made in Matlab was situated on a Macbook Pro. It calculated the DOA twenty times for each of the four angles. The 20 measurements were collected in a vector in order to compute the variance and the mean of the measurements. The cost function and data from the last of the 20 readings were also saved in order to look at the plots and analyze them.

4.2.3 Results

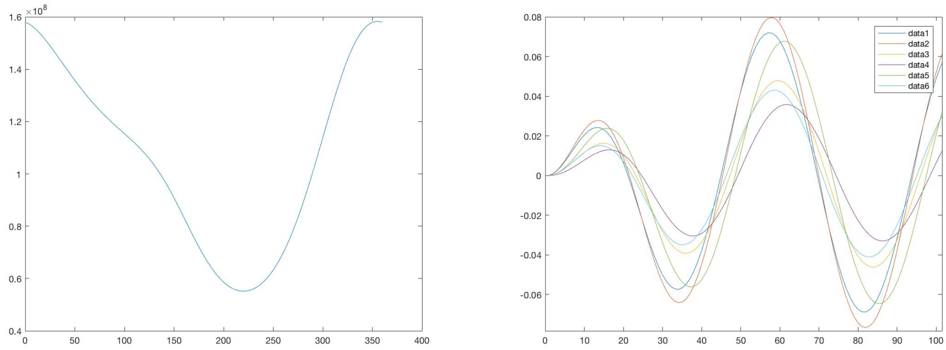
Table 4.1 shows the mean and the variance of the 20 measurements of each angle. The variance is a measure of how much the measurements are spread out compared to the mean.

	0/360 degrees	**90 degrees**	**180 degrees**	**270 degrees**
Mean	355.30	103	168.30	290.80
Variance	0.22	0.32	0.22	0.17

Table 4.1: The mean and variance of the 20 measurements for each angle.

Figure 4.4 shows the cost function and the data from all 6 microphones for the last measurement at 0 deg. In Figure 4.4a it can be seen that there is a peak around 350 deg and one at 0 deg. This fits well with the mean value seen in Table 4.1. When the transmitter is placed at zero degrees the signal are expected to hit microphone 1 first, then hit microphone 2 and 6 simultaneously, hereafter microphone 3 and 5 simultaneously, and lastly microphone 4. Looking at Figure 4.4b microphone 1 is slightly before but still almost simultaneous with microphone 2. Then the signal

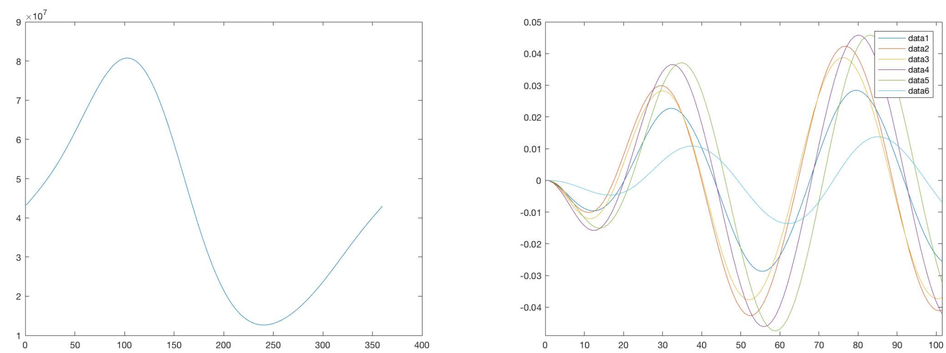
hits microphone 6 almost simultaneous with microphone 3 and 5, and lastly, the signal hits microphone 4.
In the cost function of the last of the 20 measurements at 90 deg there can be seen



(a) The cost function of 0 deg. **(b)** Data from all 6 microphones at 0 deg.

Figure 4.4: The plot of the cost function and the signals from all 6 microphones at 0 deg.

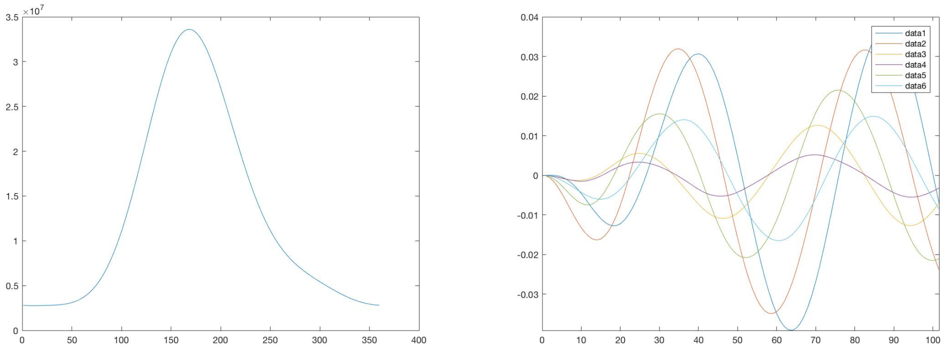
a peak around 100 deg if looking at Figure 4.5a. This fits well with the mean of the 20 measurements shown in Table 4.1. At 90 deg the signal from the transmitter should hit microphone 2 and 3 first simultaneously, then 1 and 4 simultaneously and lastly 5 and 6 simultaneously. Looking at Figure 4.5b this is almost the case.



(a) The cost function of 90 deg. **(b)** Data from all 6 microphones at 90 deg.

Figure 4.5: The plot of the cost function and the signals from all 6 microphones at 90 deg.

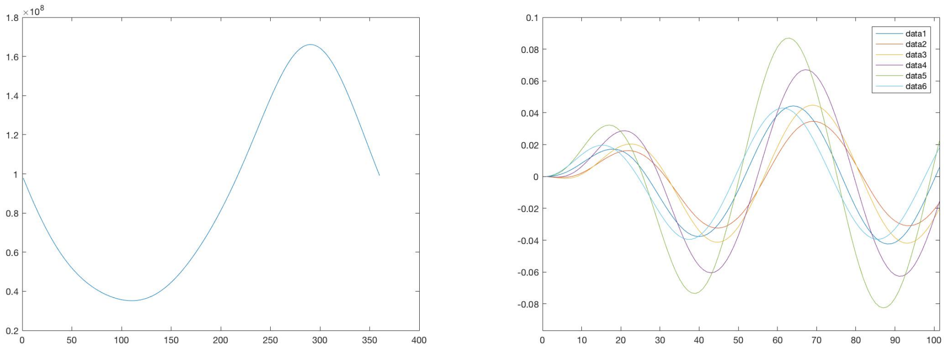
Figure 4.6a shows the cost function of the last measurement at the 180 deg angle. The peak of this cost function is lying around 170 deg. The mean for 180 deg in Table 4.1 is 168.30 deg, which means that the cost function agrees with the mean. When looking at how the signal would hit the different microphones, this would be the opposite of the transmitter being placed at 0 deg. The signal should first hit microphone 4, then microphone 3 and 5 simultaneously, then microphone 2 and 6 simultaneously and last microphone 1. Looking at Figure 4.6b the signal hits microphone 4 simultaneously with microphone 3. Furthermore, the amplitude of the signal at microphone 4 is very low compared to the other microphones. The signal then hits microphone 5, 2, 6 and then 1 in that order.



(a) The cost function of 180 deg. **(b)** Data from all 6 microphones at 180 deg.

Figure 4.6: The plot of the cost function and the signals from all 6 microphones at 180 deg.

In the cost function of 270 deg in Figure 4.7a there is a peak at approximately 290 deg, which is almost the same as the mean in Table 4.1. When the transmitter is placed in 270 deg the signal should hit microphone 5 and 6 first. Then it should hit microphone 1 and 4 and lastly microphone 2 and 3. Just the opposite of when the transmitter is placed at 90 deg. According to Figure 4.7b the signal first hits microphone 6 and 5, and then 1 and 4 and lastly 3 and 2.



(a) The cost function of 270 deg. **(b)** Data from all 6 microphones at 270 deg.

Figure 4.7: The plot of the cost function and the signals from all 6 microphones at 270 deg.

4.2.4 Discussion

The results of this test showed that the algorithm was working for audible sound sources. The reason why the results were not entirely 0, 90, 180 or 270 deg was because it was difficult to determine if the transmitter was placed at the exact angle. However, looking at the mean in Table 4.1 the measured results are close to the angle that they were supposed to. Because the transmitter is placed in the far-field it is difficult to measure if it actually is placed at the exact angle that it is supposed to, so it might as well have been placed a bit to either side of the supposed angle. With these results, we can assume that the algorithm works relatively well. What is more interesting to look at is the plots of the data we get from each microphone. The plots of the data from the six microphones showed that the data behaved as we expected it to. We could see that the signal was hitting the microphones in the right order.

In table 4.1 we can see the mean and the variance of the twenty measurements at each angle. The variance of the measurements was significantly small which means that there is a significantly small change in the measurements relatively to the mean. This shows that the readings were consistent and not random readings. Furthermore, the cost function plots showed no sign of spatial aliasing which means that the radius of the microphone array was correct.

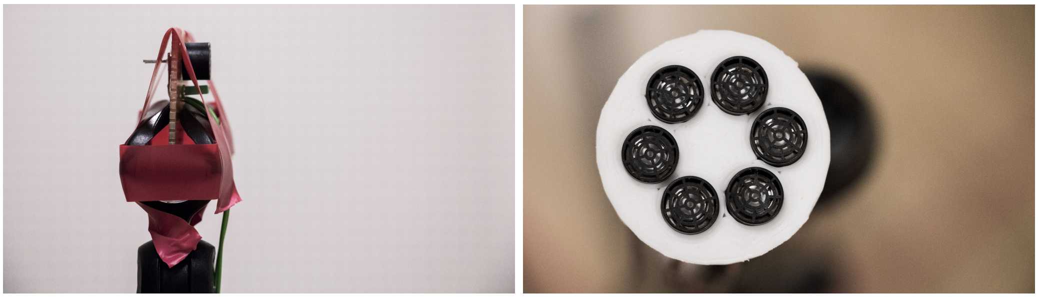
All these results confirmed the simulations made earlier, that the algorithm in fact works, and that it is not the algorithm that might have been a factor in the misreadings in the initial tests.

4.3 Ultrasound Test

The audible test showed that the algorithm was working, which means that it should not be the reason behind the misreadings in the initial tests. However, it was found out that the audio input device used for the initial tests might have been a factor to why it was not working in the initial tests. One theory was as well that the sample rate of $96kHz$ on the first audio input device was not high enough. So, when getting a new audio input device, it was desired to get one with a higher sampling rate. The new audio input device Focusrite Scarlett 18i20, with a sampling rate up to $192kHz$, was therefore very useful.

4.3.1 Setup

The ultrasound test was performed at the same facility and room as in Section 4.2. This time the microphone array designed and built in Chapter 2 and 3 was used instead of the Behringer microphones in Section 4.2. The radius of the microphone mount was $0.0112m$ and the microphone mount can be seen in Figure 4.8b. The signal that we were looking for was a $40420Hz$ sinus wave generated by a waveform generator, which can be seen in Figure 3.14, and transmitted by a Murata Sensor (Murata, 2017) seen in Figure 4.8a. The Focusrite Scarlett 18i20 USB Audio Interface (Focusrite, 2017) was used as an audio input device to acquire the data into the computer. The angles that were measured was again approximately



(a) The transmitter. **(b)** The microphone array.

Figure 4.8: The transmitter and the microphone array in the audible test.

(a) Test setup for the ultrasound test. **(b)** Placement of the transmitter.

Figure 4.9: The setup of the audible test, with placement of the transmitter.

0 deg, 90 deg, 180 deg and 270 deg. Figure 4.9 shows the setup of the ultrasound test. The red tape on the floor marked the approximate angle and where the transmitter should be placed. The transmitter was placed $0.70m$ from the center of the big red cross. The distance was measured as in Section 4.2, see Figure 4.3b. The distance from the floor to the top of the sensors in the microphone array was $1.52m$ and the distance from the floor to the transmitter was $1.57m$.

4.3.2 Method

The transmitter was placed at one angle and then that angle was measured with a sampling rate of $96kHz$, $176.4kHz$ and $192kHz$ with the algorithm made in Matlab. When the DOA had been measured at each sampling rate the transmitter was placed at the next angle where the procedure was repeated. As in Section 4.2 each angle and sampling rate was measured 20 times. These measurements were collected in a vector, where the mean and variance were computed. The cost function and data from the 6 microphones from the last measurement at each angle and sampling rate were also saved. These two were plotted in order to see how the data behaved.

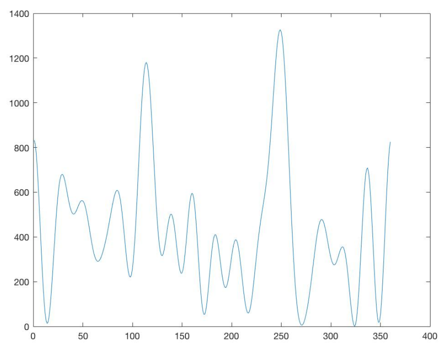
4.3.3 Results

As in Section 4.2, Table 4.2 shows the mean and variance of the 20 measurements for each sample rate and angle. Here, the variance is also a measure of how much the measurements are spread out compared to the mean.

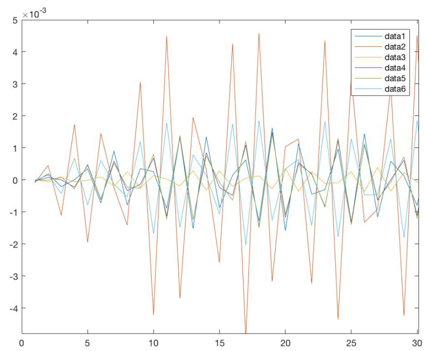
	0/360 degrees		**90 degrees**		**180 degrees**		**270 degrees**	
	Mean	**Variance**	**Mean**	**Variance**	**Mean**	**Variance**	**Mean**	**Variance**
96kHz	249.15	0.13	321	4756.70	249	419.58	149.85	1363.60
176.4kHz	249.25	0.20	335.95	0.05	252.95	0.05	21.15	934.34
192kHz	249	0.42	320.55	4774.30	253.15	0.66	28.20	1792.80

Table 4.2: The mean and variance of the 20 measurements for each sample rate and angle.

In the following figures, the plots for each sample rate at 0 deg will be displayed. The rest of the figures can be seen in the Appendix B. Since the following figures are the plots for the 0 deg angle, the signal coming from the transmitter should first hit microphone 1, then it should hit microphone 2 and 6 simultaneously, then microphone 3 and 5 simultaneously, and then hit microphone 4 as the last one assuming the data is read into the computer correctly.



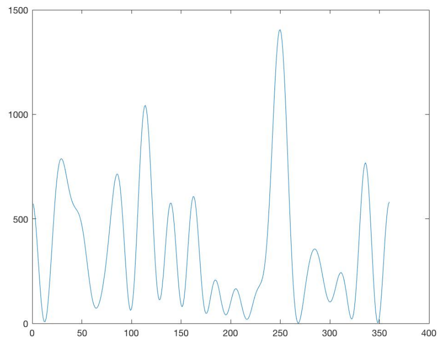
(a) The cost function of 0 deg at sample rate 96kHz.



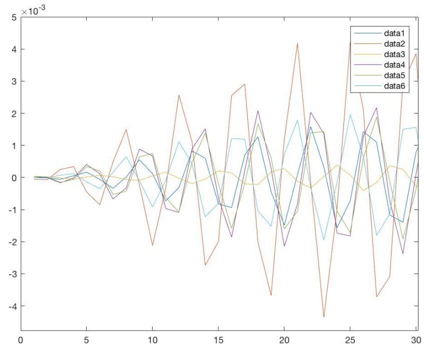
(b) Data from all 6 microphones at 0 deg with a sample rate of 96kHz.

Figure 4.10: The plot of the cost function and the signals from all 6 microphones at 0 deg and sample rate 96kHz.

Figure 4.10a shows the cost function of 0 deg at the sample rate of 96kHz. This cost function is containing a lot of peaks, with two of them being twice as high as the others. The highest peak is around 250 deg and the second highest is around 120 deg. The 250 deg peak corresponds well with the mean in Table 4.2. The data from all the microphones in Figure 4.10b shows that the amplitude of the signals are very different, which makes it difficult to see which microphone that the signal hits first. It is also obvious that the plots are a lot more pointy than the plots of the data in Section 4.2. The signal seems to be hitting microphone 1, 4, and 5 simultaneously, and microphone 2 and 6 simultaneously. Microphone 3 is very low in amplitude and is hard to read.



(a) The cost function of 0 deg at sample rate 176.4kHz.



(b) Data from all 6 microphones at 0 deg with a sample rate of 176.4kHz.

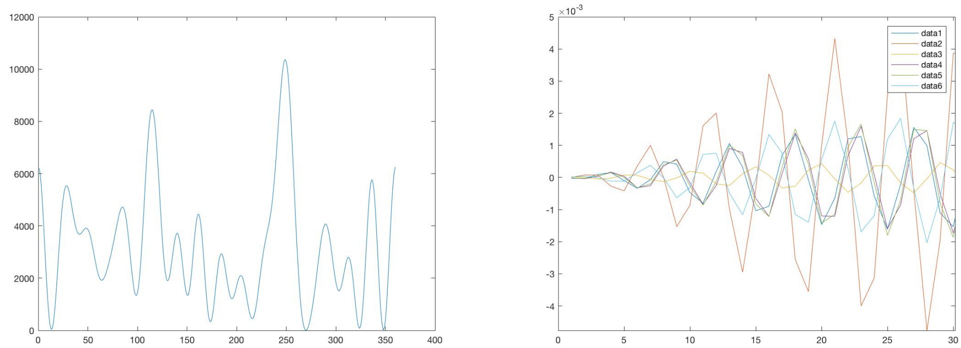
Figure 4.11: The plot of the cost function and the signals from all 6 microphones at 0 deg and sample rate 176.4kHz.

In Figure 4.11a the cost function for the 176.4kHz sample rate has a lot of peaks just like Figure 4.10a. Again there is a peak around 250 deg and the second highest peak is around 120 deg. Also here the 250 deg peak fits well with the mean. In this cost function, the smaller peaks are almost as high as the second peak. In Figure 4.11b the amplitude of the data is again very different, however, the curves of the data points are not as pointy as in Figure 4.10b. It seems in Figure 4.11b that the signal is hitting microphone 2 and 6 first simultaneously. The signal then hits microphone 1, 4, and 5 simultaneously afterward. Again microphone three had a very low amplitude, and it is not easy to read anything from that signal.

The plots in Figure 4.12 are very similar to the plots in Figure 4.11. The peaks in the cost function in Figure 4.12a are around 250 deg and 120 deg as they are at the other two sampling rates. Again, here it also fits with the mean in Table 4.2. The curves of the plots in Figure 4.12b are a bit more smooth than the curves of the plots in Figure 4.11b, but otherwise the same.

4.3.4 Discussion

The results of the ultrasound test are not that good. First of all, the mean of the 20 measurements is not even close to the angle that they are supposed to. These can be seen in Table 4.2. In this table, it is also obvious that some of the variances are very high. When looking at the raw data, it can be seen that there are some outliers in the data where there have been measured a very different angle one or two times out of the 20 measurements. There does not seem to be any noticeable



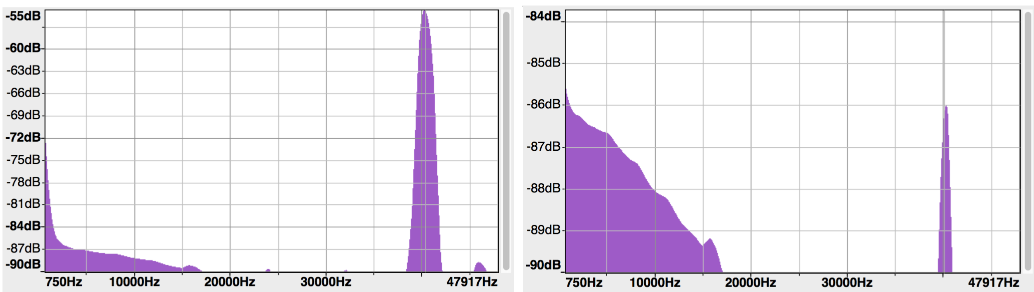
(a) The cost function of 0 deg at sample rate 192kHz.

(b) Data from all 6 microphones at 0 deg with a sample rate of 192kHz.

Figure 4.12: The plot of the cost function and the signals from all 6 microphones at 0 deg and sample rate 192kHz.

difference in the performance when looking at the sample rate either. We assumed that the performance would be better as the sampling rate was increased. Looking at all the plots of the cost function, there were a lot of peaks. This can indicate that there is spatial aliasing going on and that the microphone array would need to be even smaller. This is however not possible with the sensors used for the microphone array. Some other factors that could also cause these peaks are noise or model inaccuracies.

The plots of the data from each microphone showed that the amplitude of the data was very different for each microphone. Microphone three was even amplified so little that it was difficult to see the behavior of it in the plots. In the initial tests, it was checked several times that the amplitude was approximately the same. However, a change to the amplifiers in the last second before the test might have caused this irregularity. A recording was made in Audacity in order to plot the spectrum of the signals.



(a) Spectrum plot of the signal at microphone 1.

(b) Spectrum plot of the signal at microphone 3.

Figure 4.13: The plot of the spectrum of microphone 1 and 3 in Audacity.

In Figure 4.13 it is obvious that microphone 3 is not amplified as much as microphone 1. Even though microphone 1 is amplified more it is still not amplified much. It can be seen in Figure 4.13a that the signal peaks at $-55dB$ which is not much. This suggests that something happened to the amplifier when some components were changed just before the test. This also explains why the transmitter would have to stand no further than $0.70m$ away from the microphone array to be able to get a decent signal into the computer.

All in all the test was not satisfying, and a new test needs to be conducted when the amplifier has been fixed in order to get reliable results.

Chapter 5

Discussion

In this chapter, we will discuss the process of the entire project. The project was a long master's thesis running for two semesters. The initial idea was to work with Personal Sound Zones and actually creating a working Sound Zone system that would be able to move around with the user. The initial idea was to divide the project into three larger milestones. The first milestone was to create a tracking system, the second milestone was to create the simulation of the sound zone system, and the last was to create the actual sound zone system. During the research phase of the project, it became obvious that this was a little too ambitious, having the time constraints and the project being a solo project in mind. Therefore the problem was narrowed down to how we could create a tracking system for a simulation of a personal sound zone system using acoustic tracking. For this, a microphone array and an amplifier for ultrasonic sensors were made. The making of this amplifier took a lot of hours because a new problem always arrived when one problem had been solved. This called for a lot of problem-solving throughout the building of the amplifier. At last, we had an amplifier that seemed to work, but the results still were not good when doing small tests in the electronics laboratory. Therefore it was decided to test it in a sound laboratory where there would be fewer reflections. When there still were problems with the results, in the small informal tests, it was decided to try out with audible sound and microphones for that. When this also failed to give good results, it was decided to try another audio input device. The only other audio input device was one with a max sampling rate of $48kHz$, but this one worked with the audible sound. In order to be able to do tests with the ultrasound as well, a new audio input device was ordered. This device had a sampling rate of up to $192kHz$. While waiting on the new audio input device, some tests of the simulation of the circuit in LTspice showed that the bandpass filter that we had put on the amplifier actually were centered around $26kHz$ where it should have been centered around $40kHz$. This was fixed just before the new audio input device came, and the simulation showed that there should not be a difference in the amplification of the signal even though the bandpass filter had been moved. However, when testing with the new audio input device, the signals were not amplified enough, which might have influenced the results that we got in the ultrasound test. The audible sound was also tested with the new audio interface in order to verify that this indeed was working as it was supposed to. The results of that test were positive and we assumed that the new device was working and loading the data correctly.

In the ultrasound test results, it was clear that especially microphone 3 were not being amplified enough and the overall amplitude of the different microphones was varying a lot. The fact that the transmitter had to be placed so close to the microphone array did not set off any alarms at first. This was because, in the initial tests, the transmitter had not been placed further than $0.70m$ away from the microphone array. But looking at the results and then afterward looking at the recordings in Audacity, showed that something was completely wrong. In order to find out why the amplifier did not work properly, the amplifier has to go back to the electronics laboratory. Here we will use an oscilloscope to look at the signal going through the circuit at different points in the circuit. We will then compare signals in the circuit to the signals in the simulation program LTspice. Maybe there is a short circuit somewhere or maybe the op-amps are broken.

In the results from the ultrasound test, it was also obvious that there was a lot of spatial aliasing going on. This happened even though we had done the simulations to determine the size of the array. It even seemed like that there were more aliasing going on with the smaller array than the bigger array first built. This was not formally tested but is just an observation from the author. Some thoughts about this could be that the hole in the middle of the first array could have avoided some reflections or that we made the second array too small for 6 microphones. The last thought is based on the simulation made in Section 2.3, which showed that the radius for 6 microphones could be up to $0.03m$. There would, of course, need to be some testing in this area when the amplifiers are being fixed.

Another area that needs testing is the performance of the microphone array with only three microphones. If the performance of three microphones is as good or almost as good as 6 microphones, this would be a benefit, as this would lower the costs of the system.

The future of this project would be to get the amplifier fixed so that a decent test can be performed. If the new tests prove that it is indeed possible to do DOA with ultrasound, the extensions described in Section 3.4 would need to be implemented. When this is implemented it again needs some testing. The tests would need to test the performance of the tracking system and how accurate it is able to track people. It would also need to be tested in the ability to track multiple people at once. When the tracking system is built, it can start to be used in sound zone system simulations and help create future sound zone systems where it is possible to move around in the room and have the sound zone following the person.

Chapter 6

Conclusion

As a conclusion to this report, an attempt was made to solve the problem of creating an acoustic tracking system for a simulation of a personal sound system. This was done by using Direction of Arrival (DOA) and a Uniform Circular Array. It was not possible to make the system work with ultrasound within the time limit of the project. However, it could be seen in the simulations of the algorithm, that it should be possible to create a working system using this algorithm.

One of the requirements were that the system should be low cost, which was kept in mind during the design and development of the system. The only part that is a bit costly is the audio input device. The design of the system has also taken the adaptable aspect into account when developing the system.

In conclusion, it should be possible to create an acoustic tracking system for a simulation of a personal sound system. But in reality, it has not been possible yet due to a problem with the amplifiers of the sensors. These need to be fixed before concluding if it is possible to create such a system in reality.

Bibliography

Baysal, U. & Moses, R. [R.L.]. (2003). On the geometry of isotropic arrays. *IEEE Transactions on Signal Processing*, *51*(6), 1469–1478. doi:10.1109/TSP.2003.811227

Benesty, J., Chen, J., & Huang, Y. (2008). *Microphone Array Signal Processing*. doi:10.1007/978-3-540-78612-2

Berg, A. L. & Serpanos, Y. C. (2011). High frequency hearing sensitivity in adolescent females of a lower socioeconomic status over a period of 24 years (1985-2008). *Journal of Adolescent Health*, *48*(2), 203–208. doi:10.1016/j.jadohealth.2010.06.014

Betlehem, T., Zhang, W., Poletti, M. A., & Abhayapala, T. D. (2015). Personal sound zones: Delivering interface-free audio to multiple listeners. *IEEE Signal Processing Magazine*, *32*(2), 81–91. doi:10.1109/MSP.2014.2360707

Choi, J.-W. & Kim, Y.-H. (2002). Generation of an acoustically bright zone with an illuminated region using multiple sources. *The Journal of the Acoustical Society of America*, *111*(4), 1695–1700. doi:10.1121/1.1456926

Deak, G., Curran, K., & Condell, J. (2012). A survey of active and passive indoor localisation systems. *Computer Communications*, *35*(16), 1939–1954. doi:10.1016/j.comcom.2012.06.004

Druyvesteyn, W. F. & Garas, J. (1997). Personal Sound. *Journal of the Audio Engineering Society*, *45*(9), 685–701. Retrieved from http://www.aes.org/e-lib/inst/browse.cfm?elib=7843

Elliott, S. (2009). Professor Stephen Elliott BSc(Hons), PhD, FREng.

Elliott, S. & Nelson, P. (1988). Multichannel active sound control using adaptive filtering. *ICASSP-88., International Conference on Acoustics, Speech, and Signal Processing*, 2590–2593. doi:10.1109/ICASSP.1988.197176

Focusrite. (2017). Scarlett 18i20. Retrieved from https://us.focusrite.com/usb-audio-interfaces/scarlett-18i20

Hodges, S. & Hopper, A. (2001). a Sentient Computing System.

iTunes Apple. (2017). Signal Generator. Retrieved from https://itunes.apple.com/us/app/signal-generator-audio-test-tone-utility/id543661843?mt=8

Jensen, J. R., Nielsen, J. K., Heusdens, R., & Christensen, M. G. (2016). DOA estimation of audio sources in reverberant environments. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, *2016-May*, 176–180. doi:10.1109/ICASSP.2016.7471660

Kawitkar, R. (2009). Performance of Different Types of Array Structures Based on Multiple Signal Classification (MUSIC) Algorithm. *2009 Fifth International Conference on MEMS NANO, and Smart Systems*, 159–161. doi:10.1109/ICMENS.2009.39

Kunin, V., Jia, W., Turqueti, M., Saniie, J., & Oruklu, E. (2011). 3D Direction of Arrival Estimation and Localization Using Ultrasonic Sensors in an Anechoic Chamber. *Proceedings of the IEEE International Ultrasonics Symposium*, 756–759. doi:10.4236/jst.2011.13010

Lau, B. K., Leung, Y. H., Liu, Y., & Teo, K. L. (2006). Transformations for nonideal uniform circular arrays operating in correlated signal environments. *IEEE Transactions on Signal Processing*, *54*(1), 34–48. doi:10.1109/TSP.2005.861104

Lee, J. Y., Hudson, R. E., & Yao, K. (2014). Acoustic DOA estimation: An approximate maximum likelihood approach. *IEEE Systems Journal*, *8*(1), 131–141. doi:10.1109/JSYST.2013.2260630

Mathworks. (2013). Record and Play Audio. Retrieved from http://www.mathworks.com/help/matlab/import_export/record-and-play-audio.html

Murata. (2017). Ultrasonic Sensors. Retrieved from http://www.farnell.com/datasheets/484560.pdf

Music Group. (2017). Condenser Microphones ECM8000.

Nelson, P. (2014). Professor Philip Nelson FREng.

Nielsen, J. K., Jensen, T. L., Jensen, J. R., Christensen, M. G., & Jensen, S. H. (2016). Grid Size Selection For Nonlinear Least-Squares Optimisation in Spectral Estimation and Array Processing, 1653–1657.

Olik, M., Francombe, J., Coleman, P., Jackson, P. J. B., Olsen, M., Møller, M., … Bech, S. (2013). A Comparative Performance Study of Sound Zoning Methods in a Reflective Environment. *Audio Engineering Society Conference: 52nd International Conference: Sound Field Control - Engineering and Perception*, (September), 1–10. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=16914

Pierrette, M., Parizet, E., Chevret, P., & Chatillon, J. (2014). Noise effect on comfort in open-space offices: development of an assessment questionnaire. *Ergonomics*, *0139*(January), 1–11. doi:10.1080/00140139.2014.961972

Poletti, M. A. (2008). An Investigation of 2D Multizone Surround Sound Systems. *125th Audio Engineering Society Convention*, 1–9. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=14703

Priyantha, N. B., Chakraborty, A., & Balakrishnan, H. (2000). The Cricket location-support system. *Proceedings of the 6th annual international conference on Mobile computing and networking - MobiCom '00*, *2000*(August), 32–43. doi:10.1145/345910.345917

Shin, M., Lee, S. Q., Fazi, F. M., Nelson, P. a., Kim, D., Wang, S., … Seo, J. (2010). Maximization of acoustic energy difference between two spaces. *The Journal of the Acoustical Society of America*, *128*(1), 121–31. doi:10.1121/1.3438479

Stoica, P. & Moses, R. [Randolph]. (2005). *Spectral Analysis of Signals*. Prentice Hall, Inc. Retrieved from http://user.it.uu.se/~ps/SAS-new.pdf

Study Board of Media Technology. (2014). Curriculum for Master of Science in Sound and Music Computing. (September).

Van Veen, B. D. & Buckley, K. M. (1988). Beamforming: a versatile approach to spatial filtering. doi:10.1109/53.665

Voss, J. (2014). *An Introduction to Statistical Computing* (1st ed.). Wiley.

Wong, K. D. (2011). Kerry D. Wong. Retrieved from http://www.kerrywong.com/2011/01/22/a-sensitive-diy-ultrasonic-range-sensor/

Wu, Y. & Abhayapala, T. (2010). Simultaneous soundfield reproduction at multiple spatial regions. *Audio Engineering Society Convention 128*, 1–7. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=15264

Appendix A

Simulation code

This appendix chapter is featuring the code for the simulation done in Section 2.3.

```
clear all;
close all;
clc

tic
%% Known Parameters
fs = 96000;             % sampling rate 176400
n = 1:100;              % timerange
freq = 40000;            % frequency that we are searching for in Hz
omega = (2*pi)*freq/fs; % frequency that we are searching for in omega range
phase = pi;              % in rad
micNum = 3;              % amount of microphones
f = exp(1j*omega*n);     % signal
c = 343;

%% Changing parameters
radius = [0.005: 0.005: 0.1];   % radius of the microphone array
SNR = [-10: 2: 20];             % Signal to noise ratio
thetaGrid = [1: 0.1: 360];               % Theta range

for i = 1:numel(radius)
    rho(i) = omega*(fs/c)*radius(i);    % creates an array with rho for all radiuses
end



%% CostFunc
simIteration = 500;                                              % how many simulation iterations
indexArray = zeros(simIteration,numel(SNR),numel(radius));       % preallocate indexArray
trueThetaArray = zeros(simIteration,numel(SNR),numel(radius));   % preallocate trueThetaArray
costFunc = zeros(1,numel(thetaGrid));                            % preallocate costFunc
for i = 1:numel(radius)
    for j= 1:numel(SNR)
        [i, j]
        for k = 1:simIteration
            [X trueTheta] = datageneration(f, micNum, radius, omega, phase, rho, fs, c); %datageneration
            Xnoise = awgn(X,SNR(j),'measured');                 % adds noise to the data signal
            for thetaIdx = 1:length(thetaGrid) %theta should be in rad which is why it is written (theta*pi/180) below
                h=hFunc(f, rho(i), (thetaGrid(thetaIdx)*pi/180), micNum);
                costFunc(thetaIdx) = abs((h'*Xnoise(:,i)))^2;        % cost function
            end
            [M,I] = max(costFunc);            % finds the max of the cost function
            indexArray(k,j,i) = thetaGrid(I);           % saves the index in the 3D matrix
            trueThetaArray(k,j,i) = trueTheta*(180/pi);              % creates a 3D matrix with the true theta value
        end
    end
end

%% RMSE - Root mean square estimat



rmse = sqrt(sum((indexArray(:,:,:)-trueThetaArray(:,:,:)).^2)/numel(indexArray)); % calculates the root mean square estimat

```

```
toc
%% Save

save('DOAsimulation500w6mic20000.mat')
```

Listing A.1: DOASimulation.m

Listing A.1 is the program that was used do the Monte Carlo simulation. It uses two functions that can be seen in Listing A.2 and A.3. Listing A.2 computes the simulated data that are the data which is gonna be transmitted by the ultrasound transmitter in real life. Listing A.3 computes the h vector in the signal function in Equation 2.7.

```
function [X trueTheta] = datageneration(f, micNum, radius, omega, phase, rho, fs, c)

%% initialize parameters
trueTheta = [3,2];
%trueTheta = 2*pi*rand(1);                               % creates random theta
SourcePosition = [2*cos(trueTheta), 2*sin(trueTheta)];  % true simulation position of the source
SourcePosition = [0.7,0];
d = norm(SourcePosition);                               % distance to the source in meter
epsilon = omega*(fs/c)*d;                               % in rad
amplitude = 1;                                          % amplitude: unknown
alpha = amplitude*exp(1j*phase);                        % unknown because of amplitude
beta = 1;                                               % beta: unknown
z = amplitude*beta*exp(1j^(phase-epsilon));             % unknown because of amplitude and beta


X=zeros(micNum*length(f),length(radius));                   % preallocates X

for i = 1:numel(radius)
    h = hFunc(f, rho(i), trueTheta, micNum);            % Calculates X for every radius where trueTheta is fixed
    X(:,i) = h*z;
end


end
```

Listing A.2: datageneration.m

```
function h = hFunc(f, rho, theta, micNum)
N= length(f);
h = zeros(N*micNum, 1);
index = 1:N;
for k = 1:micNum
    h(index) = f*exp(1j*rho*cos(theta-2*pi*((k-1)/micNum)));     % calculates h in the signal function x=h(theta)*z+e
    index = index+N;
end


end
```

Listing A.3: hFunc.m

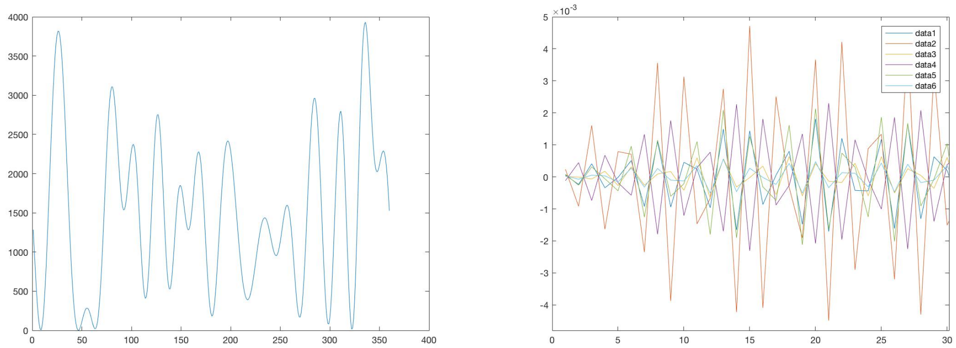
Appendix B

Ultrasound Test

This appendix chapter features the plots of the remaining angles and sample rates from the ultrasound test.

B.1 Plots at 90 deg angle

When the transmitter is placed at a 90 deg angle the signal from it is expected to hit microphone 2 and 3 simultaneously first. Thereafter it should hit microphone 1 and 4 simultaneously, and then microphone 5 and 6 simultaneously.



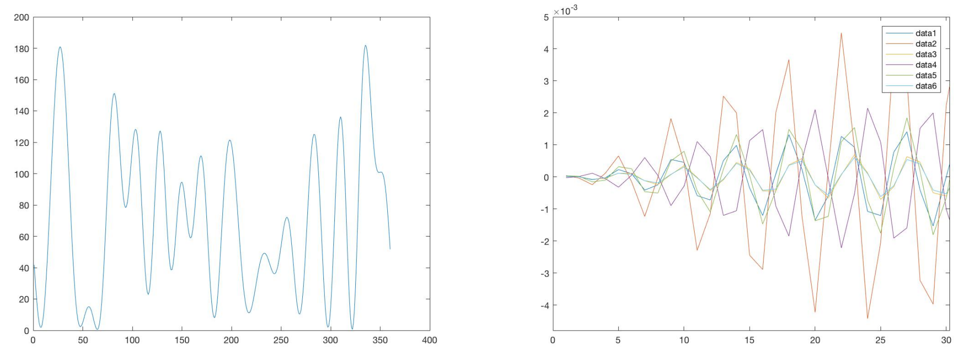
(a) The cost function of 90 deg at sample rate 96kHz.

(b) Data from all 6 microphones at 90 deg with a sample rate of 96kHz.

Figure B.1: The plot of the cost function and the signals from all 6 microphones at 90 deg and sample rate 96kHz.

Figure B.1a shows the cost function at a sample rate of 96kHz. There are a lot of peaks and the two highest peaks are very close to each other in terms of height. The highest peak is around 340 deg and the second highest is around 30 deg. The mean of the 20 measurements is 321 and the variance is 4756.7, which means that there is a big spread from the mean in the data.

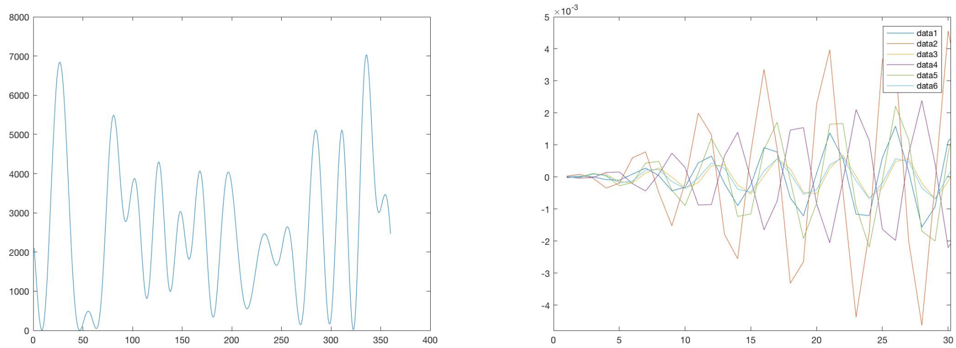
Figure B.1b shows that the signal hits microphone 2, 1, 5, 6, and 3 simultaneously and microphone 4 is out of phase. The signal at microphone 3 and 6 are very low in amplitude, which makes them difficult to read. The very pointy structure of the curve makes it difficult to see what microphone is hit first.



(a) The cost function of 90 deg at sample rate $176.4kHz$. **(b)** Data from all 6 microphones at 90 deg with a sample rate of $176.4kHz$.

Figure B.2: The plot of the cost function and the signals from all 6 microphones at 90 deg and sample rate $176.4kHz$.

The cost function in Figure B.2a looks like the one in Figure B.1a. The mean for the 20 measurements is 335.95 and the variance is 0.05. The variance is low which means that the spread is not far from the mean. The curve of Figure B.2b is a little less pointy and gives a better opportunity to see which microphone is hit first. Microphone 4 is the microphone that is hit first, then microphone 2 is hit a little before it hits the rest of the microphones simultaneously. Again the signal is small in amplitude at microphone 3 and 6.



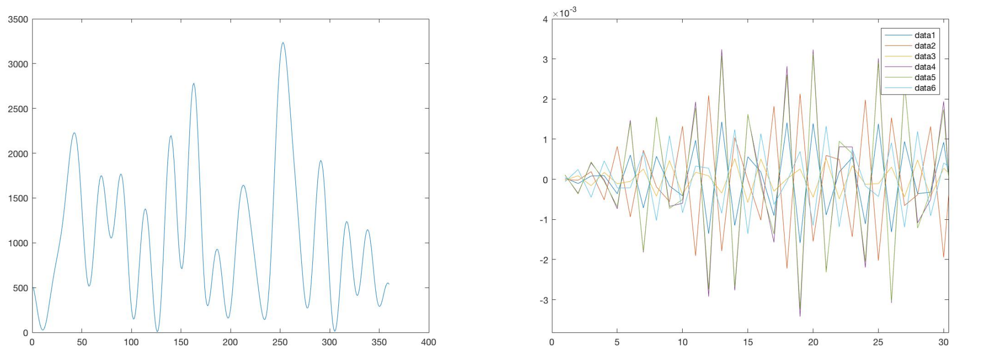
(a) The cost function of 90 deg at sample rate $192kHz$. **(b)** Data from all 6 microphones at 90 deg with a sample rate of $192kHz$.

Figure B.3: The plot of the cost function and the signals from all 6 microphones at 90 deg and sample rate $192kHz$.

The cost function in Figure B.3a is almost identical to the ones in Figure B.1a and B.2a. The mean of the 20 measurements for the sample rate of 192kHz is 320.55 and the variance is 4774.30. The high variance again means that there is a large spread from the mean. In Figure B.3b it is actually possible to see a difference between all the microphones where the signals hit simultaneously at the other sample rates. The signal first hits microphone 4 and then in $2, 1, 5, 6, 3$ in that order. Also here the signal is low in amplitude at microphone 3 and 6.

B.2 Plots at 180 deg angle

When the transmitter is placed at a 180 deg angle the signal from it is expected to hit microphone 4 first. Thereafter it should hit microphone 3 and 5 simultaneously, and then microphone 2 and 6 simultaneously and lastly microphone 1.



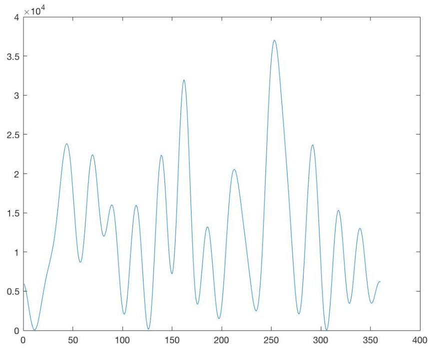
(a) The cost function of 180 deg at sample rate 96kHz.

(b) Data from all 6 microphones at 180 deg with a sample rate of 96kHz.

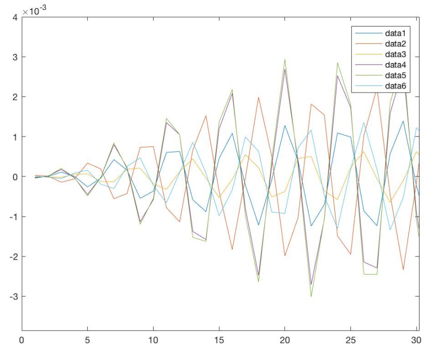
Figure B.4: The plot of the cost function and the signals from all 6 microphones at 180 deg and sample rate 96kHz.

Figure B.4a shows the cost function at a sample rate of 96kHz, which shows a lot of peaks. The highest peak is around 255 deg and the second highest is around 160 deg. The mean of the 20 measurements is 249 and the variance is 419.58, which means that there is a big spread from the mean in the data but not as big as in two of the plots at the 90 deg angle. The pointy curve in Figure B.4b makes it difficult to see which microphone the signal hits first. However, it appears that the signal hits microphone 6 and 3 first, then it hits microphone 2 and at last, it hits microphone $4, 5$, and 1 simultaneously. The signal at microphone 3 is again very low in amplitude.

The cost function in Figure B.5a is very similar to the cost function in Figure B.4a. The peaks are at the same places. The mean of the 20 measurements at sample rate 176.4kHz is 252.95 and the variance is 0.05. The variance is small which means that there is a small spread from the mean. The curve in Figure B.5b is more smooth than the one in Figure B.4b. The signal first hits microphone $4, 5$,



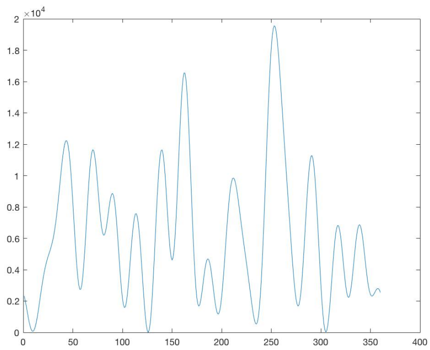
(a) The cost function of 180 deg at sample rate 176.4kHz.



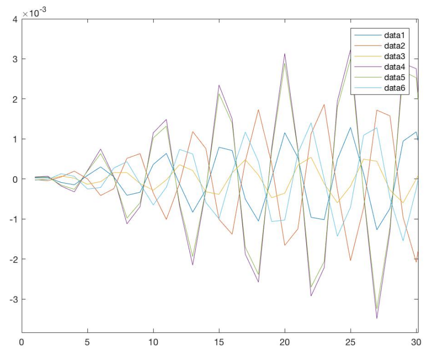
(b) Data from all 6 microphones at 0 deg with a sample rate of 176.4kHz.

Figure B.5: The plot of the cost function and the signals from all 6 microphones at 180 deg and sample rate 176.4kHz.

and 1 simultaneously. Then it hits microphone 3, then 6 and lastly microphone 2. Here the signal at microphone 3 is also very low in amplitude.



(a) The cost function of 180 deg at sample rate 192kHz.



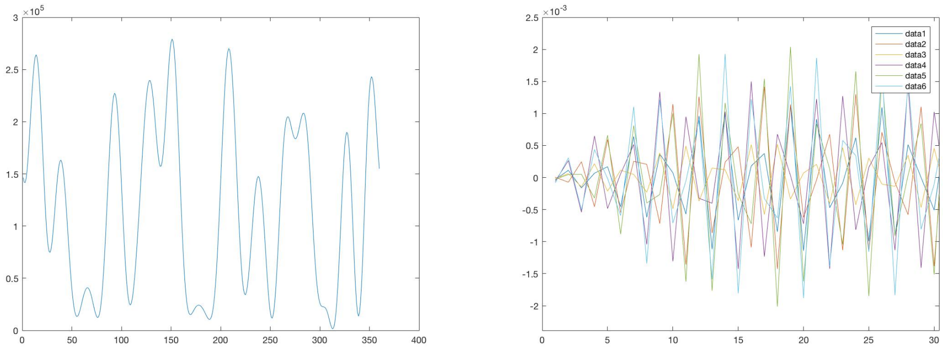
(b) Data from all 6 microphones at 180 deg with a sample rate of 192kHz.

Figure B.6: The plot of the cost function and the signals from all 6 microphones at 180 deg and sample rate 192kHz.

The cost function in Figure B.6a is similar to the cost functions in Figure B.4a and B.5a. The mean of the 20 measurements is 253.15 and the variance is 0.66. The variance is small, which means that the data does not spread much from the mean. In Figure B.6b the signal first hits 6 and 3 simultaneously, the it hits microphone 2, and then microphone 1, 5, and 4 simultaneously. Here the signal at microphone 3 is also low in amplitude.

B.3 Plots at 270 deg angle

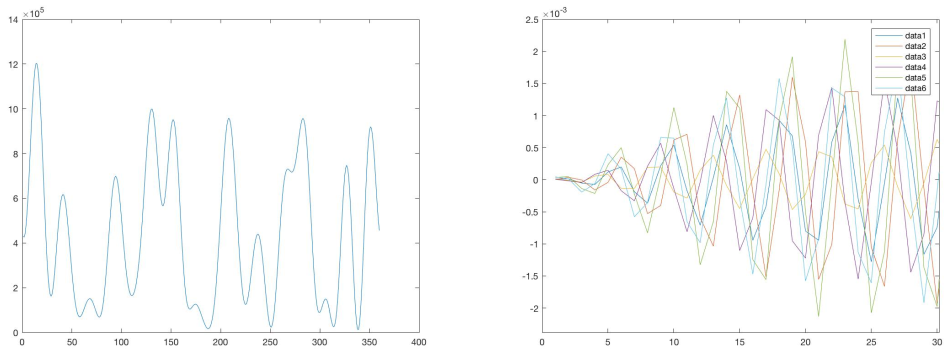
When the transmitter is placed at a 270 deg angle the signal from it is expected to hit microphone 5 and 6 simultaneously first. Thereafter it should hit microphone 1 and 4 simultaneously, and then microphone 2 and 3 simultaneously.



(a) The cost function of 270 deg at sample rate 96kHz.

(b) Data from all 6 microphones at 270 deg with a sample rate of 96kHz.

Figure B.7: The plot of the cost function and the signals from all 6 microphones at 270 deg and sample rate 96kHz.



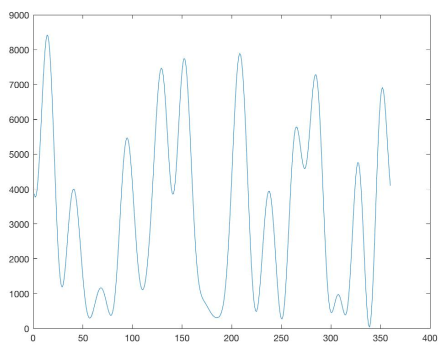
(a) The cost function of 270 deg at sample rate 176.4kHz.

(b) Data from all 6 microphones at 270 deg with a sample rate of 176.4kHz.

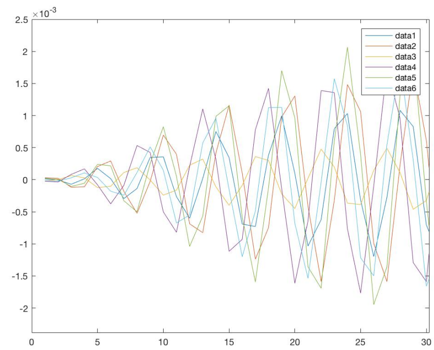
Figure B.8: The plot of the cost function and the signals from all 6 microphones at 270 deg and sample rate 176.4kHz.

In Figure B.7a there is a lot of peaks in the cost function. The peaks are so close to each other in terms of height that it is almost impossible to determine which one is the highest and the second highest. However, looking closely it seems that the highest peak is around 150 deg, which correlates well with the mean being 149.85. The variance, however, is very high 1363.60, which means that there is a big spread from the mean. Again the pointy structure of the curve in Figure B.7a makes it difficult to see what microphone the signal is hitting first, but it seems like it hits microphone $6, 4, 1, 3$, and 5 almost simultaneously and then it hits microphone 2.

The cost function in Figure B.8a looks almost like the cost function in Figure B.7a. However, the highest peak is now around 20 deg. The mean of the 20 measurements is 21.15 and the variance is 934.34. The variance is still high, but not as high as with a sample rate at $96kHz$. The curve in Figure B.8b is less pointy, which makes it easier to read. The signal first hits microphone 3 and 4 simultaneously. Then it hits microphone 6, then 1, then 5, and lastly microphone 2.



(a) The cost function of 270 deg at sample rate $192kHz$.



(b) Data from all 6 microphones at 270 deg with a sample rate of $192kHz$.

Figure B.9: The plot of the cost function and the signals from all 6 microphones at 270 deg and sample rate $192kHz$.

The cost function in Figure B.9a is similar to the one in Figure B.8a. However, the peaks are closer to each other in terms of height. The mean for the 20 measurements is 28.20 and the variance is 1792.80. The variance is very high, which means that there is a big spread from the mean in the data. Looking at the data in Figure B.9b the signal seems to hit microphone $4, 6$, and 3 simultaneously, and then microphone $1, 5$, and 2 simultaneously. Here the signal at microphone 3 is also low in amplitude.