Bachelor Thesis

The Mathematics of Multi-Input Sound Field Recording

Lydia Krott

May 8, 2013
Blacksburg, Virginia

Advisor: Prof. Robert C. Rogers
Department of Mathematics

Advisor: Prof. Dr. Lutz Weis
Fakultät für Mathematik

Karlsruhe Institute of Technology
I would like to gratefully thank my advisor Prof. Robert C. Rogers from Virginia Tech for his continuous guidance and support for this research. He encouraged me to participate in poster symposiums and an undergraduate research conference, which helped me better understand the subject and improved my presentation skills.

Also, I want to express my special thanks to Prof. Dr. Lutz Weis from Karlsruhe Institute of Technology (KIT) for being the second advisor for this thesis and ensuring that the KIT guidelines are met.

I would like to thank the Department of Mathematics of Virginia Tech for supporting me financially and allowing me to participate in graduate level events.

I want to acknowledge the Institute of Mathematics of KIT for offering me this rewarding exchange program and making this experience possible by providing financial support through a scholarship. I would particularly like to thank Prof. Dr. Wolfang Reichel for his support before and during my two semesters abroad.
Contents

List of Figures iv

1 Introduction 1

2 Horizontal-Only Recording 5
  2.1 Exact Pickup Function 6
  2.2 Simplified Pickup Function 9
    2.2.1 Full Rank Encoding with N Microphones 11
    2.2.2 Analysis of Common Stereo Recording Techniques 12
  2.3 Modified Pickup Function 14

3 Full-Sphere Ambisonic Recording 20
  3.1 Spherical Harmonics 20
  3.2 Simplified Pickup Function 22
    3.2.1 Voltage Function \textit{v}\textsubscript{omni} 23
    3.2.2 Voltage Function \textit{v}\textsubscript{bi} 24
    3.2.3 Voltage Function \textit{v}\textsubscript{card} 27
  3.3 Analysis of the First Order B-Format 28

4 Conclusions 31
  4.1 Overview of Results 31
  4.2 Open Problems 31

Bibliography 33
List of Figures

1.1 Polar pattern of the Shure SM58 microphone ........................................ 2
2.1 Pickup patterns of different types of microphones ................................. 6
2.2 Coincident arrays with ideal number of microphones ............................ 12
2.3 Common stereo recording technique: Mid/Side ...................................... 13
2.4 Common stereo recording technique: XY .............................................. 14
2.5 Simplified and modified bidirectional recordings by comparison .............. 15
2.6 Simplified and modified cardioid recordings by comparison .................. 15
3.1 Spherical coordinate system .............................................................. 20
3.2 B-Format Microphone Types ............................................................. 28
3.3 B-Format Microphone Array .............................................................. 29
1 Introduction

This work deals with the encoding process of audio recording and reproduction in surround sound technology. That is, we examine ways in which various recording techniques encode the information from sound fields in two and three dimensions into discrete signals. The aim of this thesis is to study recording with coincident microphone arrays in which we employ various types of microphones. This encoding process is the first stage in both coincident microphone stereo recording and Ambisonic sound reproduction. Our goal is to clarify the connection between the Fourier expansions of the directional sound fields and the discrete channel recordings made with different microphones in different configurations.

In traditional multichannel audio recording, the channels (or tracks), each generated by a microphone, directly correspond to a designated loudspeaker arrangement from which the sound is played back. In contrast, the encoding signals in Ambisonics are processed without regard to a particular speaker layout. The encoded information can suit any layout which makes it possible to use one recording for multiple speaker configurations. One reason is the ability of Ambisonic reproduction systems to separately reproduce the pressure and velocity components of recorded audio signals [BLH06]. So Ambisonic is a two part solution to audio reproduction. Here, the term solution is highly appropriate since we have the problem of extracting as much information from the sound field as possible and then feeding it to loudspeakers in a way that fools the human ear into hearing the original sound. Psychoacoustics deals with the relationship between the sound field that is presented to a listener and his/her perception, as well as with the numerous mechanisms by which the human ears localize sounds. As experience, expectations, and the degree of attention also play roles in sound localization, providing high quality sound reproduction is a difficult task [Ger74]. We are in need of mathematical descriptions of the various processes to obtain or approximate the best possible results. This thesis describes a mathematical model of the encoding process, which is the first step of recording the sound field with the help of microphones. Here, the advantage of the Ambisonic theory is that there is no need to consider the actual details of the reproduction system when dealing with the recording.

Within the mathematical model it will be necessary to include simplifying assumptions such as the terminology of coincident microphones. It is clear that true coincidence cannot be achieved due to the finite size of any practical microphone. The basic idea in overcoming this issue is to compensate for the spatial mismatch by means of signal processing [Bat09], which will not be part of this thesis. Furthermore, we will assume a very general representation of the sound field, which does not particularly deal with the interactions of
pressure and velocity.

Every microphone is described by a polar pattern that specifies how sensitive the microphone is to incident sound coming from a particular direction. For example, an omnidirectional microphone registers the full range in lower frequencies and reproduces them equally at any point on its surface. So its response is a perfect sphere and we can categorize it as a nondirectional microphone. In contrast, the class of directional microphones includes such well-known types as the bidirectional (or figure-eight) and cardioid microphone. Note that the polar pattern of a microphone is frequency dependent. Figure 1.1 gives an example of a polar pattern and illustrates the frequency dependence of Shure’s SM58 microphone [Inc09]. In our mathematical model we will neglect the frequency dependence in the first step and then offer an approach that takes it into account by modifying the model. However, this will not provide an exact reflection of frequency dependence.

![Figure 1.1: Polar pattern of the Shure SM58 microphone](image)

We need the following functions to describe the audio recording process mathematically.

1. A sound field function $f$ which represents the sound information as a function of direction and time.$^1$

2. A pickup function $m$ that describes the pickup pattern of the recording microphone.

3. A voltage function $v$ that outlines the interaction of the two previous functions and provides discrete signals to a surround sound system.

As we contemplate on coincident microphone arrays, it is practical to do the calculations within the unit circle or unit sphere, respectively. The sound field function is dependent on time $t \geq 0$ and direction, given by an angle $\Omega$. We write $f = f(t, \Omega)$. The pickup function representing the directional response of the microphone is time independent, but its direction is given by a reference angle $\bar{\Omega}$ which provides the information about where

---

$^1$De facto, the sound field consists of the physical components velocity and pressure. But in this thesis we assume that the sound field can be described as a function of direction and time.
the microphone is pointed, for instance to the front, sideways or any other direction. We write \( m = m(\Omega; \bar{\Omega}) \). Eventually, we use the following ansatz for the voltage function

\[
v(t, \Omega) = \int_{\Omega \in S} f(t, \Omega)m(\Omega; \bar{\Omega}) \, dS
\]

where \( dS \) denotes the line integral over the unit circle in two dimensions and the surface integral over the unit sphere in three dimensions. For every time \( t \geq 0 \) and a microphone with pickup pattern \( m(\Omega; \bar{\Omega}) \) pointed in direction \( \bar{\Omega} \), the electrical voltage is given as the integral over the unit circle or sphere of the product of sound field and pickup function. This approach adequately takes into account that a sound wave cannot be reproduced without either measuring it over an area at one moment or measuring it over a time interval at one point.

In the literature, the emphasis is on simplified models of picking up sound field information with a dirac delta characteristic. Michael A. Gerzon, a pioneer of audio reproduction, conveniently assumed the function of the gain for the recording system to be a directional characteristic or a so called dirac delta microphone pickup. However, he mentions that this might not be accurate:

“\text{The gain with which a sound from the direction } (x,y,z) \text{ is applied to a channel of the periphonic recording will be a complex function of the direction, i.e., a function of the unit sphere. It is convenient to think of this function as the directional characteristic (i.e., the amplitude gain in various directions) of a directional microphone, although the recording may in fact be made by “pan-potting” sounds into the channels to stimulate pickup by directional microphones.”} [Ger73]

We provide a new approach by thinking of said function as an integral over the circle or sphere rather than the isolated amplitude gain at one point.

In chapter 2 we deal with horizontal-only audio recording. In particular, we model the recording of coincident multiple-microphone systems in two dimensions, that is multiple microphones of the same or different types arranged in a plane and all centered at one point (the origin). The general theory is provided by classical Fourier series. A simplified model of microphone pickup is applied and the emphasis is put on the optimal use of multiple microphones. Encoding matrices are obtained from simplified, modified, and generalized forms of the pickup function, which contribute to the main result: to obtain necessary conditions on the design of the microphones for full rank encoding of the sound field information. We also utilize this information for the analysis of commonly used and commercially available microphone arrays and make statements about their efficiency with regard to Ambisonic encoding.

In chapter 3 we need some analog to the theory of Fourier analysis, which will be provided by spherical harmonics. We extend the theory of horizontal-only recording in chapter 2 to full-sphere (or periphonic) recording in order to analyse the Ambisonic B-Format. The
B-Format microphone was invented by Peter G. Craven and Michael A. Gerzon and was patented in 1977 [CG77]. The goal of the B-Format was to invent a microphone that provides a full three-dimensional listening experience. In the context of describing the sound field in terms of a spherical harmonics decomposition, also known as Higher Order Ambisonics (HOA), the B-Format holds a first order Ambisonic signal [Bat09]. The equivalent of the B-Format in two dimensions is the so called native B-Format, that is the double mid-side arrangement, which is analysed in chapter 2.

An overview of the conclusions made in the previous chapters is given in chapter 4. Also, we give an outline for further reading and recommendations for future work based on the results of this thesis.
2 Horizontal-Only Recording

In this chapter we apply a simplified model to mathematically describe the encoding process of a coincident microphone array in a plane. We model the sound field as being dependent on time \( t \geq 0 \) and direction \( \Omega = \theta \in [-\pi, \pi) \). Since dependence on direction is periodic, we can use a classical Fourier series representation for the sound field. So we write the sound field function \( f(t, \theta) \) as its Fourier series

\[
f(t, \theta) = a_0(t) + \sum_{n=1}^{\infty} a_n(t) \cos(n\theta) + b_n(t) \sin(n\theta).
\]

A microphone will pick up some or all of the information that is contained in the Fourier coefficients. This depends on its pickup pattern. The bidirectional microphone, for example, picks up front and back information equally, but with opposite polarity. A cardioid or hypercardioid microphone primarily picks up sound information from the front, thus making it an unidirectional microphone, whereas the omnidirectional microphone picks up all incident sound equally. Their polar patterns are displayed in Figure 2.1. We define the displayed position as default. That is, with the microphone pointed in the positive direction of the horizontal \( x \)-axis which we will refer to as frontwards. On the basis of their polar patterns we classify three types of microphones.

**Omnidirectional microphones**

These microphones pick up the sound field within \( 360^\circ \) weighting all information equally. Thus the pickup does not depend on direction and

\[
m(\theta; \bar{\theta}) = \text{constant}. \tag{2.1}
\]

**Unidirectional microphones**

The defining property of this class is a symmetry to the horizontal axis, that is

\[
m(\theta + \bar{\theta}; \bar{\theta}) = m(-\theta + \bar{\theta}; \bar{\theta}). \tag{2.2}
\]

**Bidirectional microphones**

We define this class to be the microphones with a symmetry to the horizontal and the vertical axis. Thus we have

\[
m(\theta + \bar{\theta}; \bar{\theta}) = m(-\theta + \bar{\theta}; \bar{\theta}), \tag{2.3}
m(\theta + \bar{\theta}; \bar{\theta}) = -m(\theta + \bar{\theta} + \pi; \bar{\theta}). \tag{2.4}
\]
2.1 Exact Pickup Function

We are in search of the $2\pi$-periodic function $m(\theta; \bar{\theta})$ that describes the pickup pattern of a microphone. We can express this function as a classical Fourier series

$$m(\theta; \bar{\theta}) = A_0 + \sum_{n=1}^{\infty} A_n \cos(n(\theta - \bar{\theta})) + B_n \sin(n(\theta - \bar{\theta})).$$

We define the voltage function, which provides each of the signals to the surround sound system, by

$$v(t, \bar{\theta}) = \int_{-\pi}^{\pi} f(t, \theta)m(\theta; \bar{\theta})d\theta.$$
Now we can deduce properties for the three classes of microphones. We prove the following lemma in order to do so. Throughout this thesis, \( \delta_{n,l} \) denotes the Kronecker delta.

**Lemma 2.1** Let \( n, l \in \mathbb{N}, \bar{\theta} \in (-\pi, \pi] \) and \( y \in \mathbb{R} \) be independent of \( \theta \). Then

(a) \( \int_{-\pi}^{\pi} \cos(n\theta + y) d\theta = 0 \)
(b) \( \int_{-\pi}^{\pi} \sin(n\theta + y) d\theta = 0 \)
(c) \( \int_{-\pi}^{\pi} \cos(n\theta) \cos(l(\theta - \bar{\theta})) d\theta = \pi \cos(l\bar{\theta}) \delta_{n,l} \)
(d) \( \int_{-\pi}^{\pi} \sin(n\theta) \cos(l(\theta - \bar{\theta})) d\theta = \pi \sin(l\bar{\theta}) \delta_{n,l} \)

**Proof**

We prove (a) by utilizing basic trigonometric identities.

\[
\int_{-\pi}^{\pi} \cos(n\theta + y) d\theta = \frac{\sin(n\theta + y)}{n} \bigg|_{\theta = \pi}^{\theta = -\pi} = \frac{1}{n} (\sin(n\pi + y) - \sin(-n\pi + y)) = \frac{1}{n} (\sin(n\pi + y) + \sin(n\pi - y)) = \frac{2}{n} \sin(n\pi) \cos(y) = 0.
\]

We apply this identity to prove (c).

\[
\int_{-\pi}^{\pi} \cos(n\theta) \cos(l(\theta - \bar{\theta})) d\theta = \frac{1}{2} \int_{-\pi}^{\pi} \cos((n-l)\theta + l\bar{\theta}) + \cos((n+l)\theta - l\bar{\theta}) d\theta = \frac{1}{2} \int_{-\pi}^{\pi} \cos((n-l)\theta + l\bar{\theta}) d\theta = \frac{1}{2} \int_{-\pi}^{\pi} \cos(l\bar{\theta}) d\theta = \frac{1}{2} \delta_{n,l} \cos(l\bar{\theta}) 2\pi = \pi \cos(l\bar{\theta}) \delta_{n,l}.
\]

The proofs of (b) and (d) are similar.

**Omnidirectional microphones**

Property (2.1) yields

\[ m(\theta; \bar{\theta}) = A_0. \]

Then for the voltage function we have

\[
v_{\text{omni}}(t) = \int_{-\pi}^{\pi} f(t, \theta) A_0 d\theta = A_0 \left( a_0(t) \int_{-\pi}^{\pi} d\theta + \sum_{n=1}^{\infty} a_n(t) \int_{-\pi}^{\pi} \cos(n\theta) d\theta + b_n(t) \int_{-\pi}^{\pi} \sin(n\theta) d\theta \right) = a_0(t) 2\pi A_0.
\]
2.1 Exact Pickup Function

Unidirectional microphones

Property (2.2) translates into

\[ A_0 + \sum_{n=1}^{\infty} A_n \cos(n\theta) + B_n \sin(n\theta) \overset{\dagger}{=} A_0 + \sum_{n=1}^{\infty} A_n \cos(-n\theta) + B_n \sin(-n\theta) \]

\[ = A_0 + \sum_{n=1}^{\infty} A_n \cos(n\theta) + (-B_n) \sin(n\theta). \]

We conclude \( B_n = 0 \) for all \( n \geq 1 \), which means that the pickup function of a unidirectional microphone has the general form

\[ m(\theta; \bar{\theta}) = A_0 + \sum_{n=1}^{\infty} A_n \cos(n(\theta - \bar{\theta})). \]

and therefore\(^1\)

\[ v_{\text{uni}}(t, \bar{\theta}) = a_0(t)2\pi A_0 + \sum_{n=1}^{\infty} A_n \int_{-\pi}^{\pi} f(t, \theta) \cos(n(\theta - \bar{\theta})) d\theta \]

\[ = a_0(t)2\pi A_0 + \sum_{n=1}^{\infty} A_n \left\{ a_0(t) \int_{-\pi}^{\pi} \cos(n(\theta - \bar{\theta})) d\theta \right. \]

\[ + \sum_{m=1}^{\infty} a_m(t) \int_{-\pi}^{\pi} \cos(m\theta) \cos(n(\theta - \bar{\theta})) d\theta \]

\[ + \sum_{l=1}^{\infty} b_l(t) \int_{-\pi}^{\pi} \sin(l\theta) \cos(n(\theta - \bar{\theta})) d\theta \left. \right\} \]

\[ = a_0(t)2\pi A_0 + \sum_{n=1}^{\infty} a_n(t)\pi \cos(n\bar{\theta}) A_n + b_n(t)\pi \sin(n\bar{\theta}) A_n. \]

Bidirectional microphones

The two properties (2.3) and (2.4) result into \( B_n = 0 \) for all \( n \geq 1 \) as above, as well as

\[ A_0 + \sum_{n=1}^{\infty} A_n \cos(n\theta) \overset{\dagger}{=} -A_0 - \sum_{n=1}^{\infty} A_n \cos(n(\theta + \pi)). \]  \hfill (2.5)

Applying trigonometric identities we have \( \cos(n\theta + n\pi) = \cos(n\pi) \cos(n\theta) - \sin(n\pi) \sin(n\theta) = (-1)^n \cos(n\theta) \) for \( n \in \mathbb{N} \). Therefore, the following are equivalent

\[ \overset{\text{(2.5)}}{\iff} A_0 + \sum_{n=1}^{\infty} A_n \cos(n\theta) = -A_0 - \sum_{n=1}^{\infty} (-1)^n A_n \cos(n\theta) \]

\[ \iff 2 \left[ A_0 + \sum_{m=1}^{\infty} A_{2m} \cos(2m\theta) \right] = 0 \]

\[ \iff A_0 = A_{2m} = 0 \ \forall m \geq 1 \]

\(^1\)As we assume \( f(t, \theta) \) to be continuous and periodic, the functions we are dealing with are sufficiently smooth to split up summations and change the order of summation and integration.
by the Fourier Convergence Theorem. Hence, the pickup function is of the form

\[ m(\theta; \bar{\theta}) = \sum_{n=1}^{\infty} A_{2n-1} \cos((2n-1)(\theta - \bar{\theta})). \]

We calculate

\[ v_{bi}(t, \bar{\theta}) = \sum_{n=1}^{\infty} A_{2n-1} \int_{-\pi}^{\pi} f(t, \theta) \cos((2n-1)(\theta - \bar{\theta})) d\theta \]

\[ = \sum_{n=1}^{\infty} A_{2n-1} \left\{ a_0(t) \int_{-\pi}^{\pi} \cos((2n-1)(\theta - \bar{\theta})) d\theta \right. \]

\[ + \sum_{m=1}^{\infty} a_m(t) \int_{-\pi}^{\pi} \cos(m\theta) \cos((2n-1)(\theta - \bar{\theta})) d\theta \]

\[ + \sum_{l=1}^{\infty} b_l(t) \int_{-\pi}^{\pi} \sin(l\theta) \cos((2n-1)(\theta - \bar{\theta})) d\theta \right\} \]

\[ = \sum_{n=1}^{\infty} a_{2n-1}(t) \pi \cos((2n-1)\bar{\theta}) A_{2n-1} + b_{2n-1}(t) \pi \sin((2n-1)\bar{\theta}) A_{2n-1}. \]

We perceive that only the odd-numbered Fourier coefficients are incorporated in the voltage picked up by a bidirectional microphone. In contrast, the unidirectional pickup function does not ignore any particular information beforehand. And the omnidirectional microphone only encodes the zeroth Fourier coefficient. All three patterns can be found later in the investigations with simplified and modified pickup functions.

### 2.2 Simplified Pickup Function

The polar patterns in Figure 2.1 are generated by

\[ m(\theta; \bar{\theta}) = \alpha + (1 - \alpha) \cos(\theta - \bar{\theta}). \]

This function provides the omnidirectional case when \( \alpha = 1 \), bidirectional when \( \alpha = 0 \), hypercardioid when \( 0 < \alpha < \frac{1}{2} \) and cardioid when \( \alpha = \frac{1}{2} \). We refer to this special form as simplified pickup function.

**Omnidirectional microphones**

With \( \alpha = 1 \) the simplified form becomes

\[ m(\theta; \bar{\theta}) = 1. \]

That is precisely the exact pickup function with \( A_0 = 1 \) which results in

\[ v_{omni}(t) = a_0(t)2\pi. \]
2.2 Simplified Pickup Function

**Bidirectional microphones**

When $\alpha = 0$ we get the pickup function

$$m(\theta; \bar{\theta}) = \cos(\theta - \bar{\theta}).$$

By Lemma 2.1,

$$v_{bi}(t, \bar{\theta}) = \int_{-\pi}^{\pi} f(t, \theta) \cos(\theta - \bar{\theta}) d\theta$$

$$= a_0(t) \int_{-\pi}^{\pi} \cos(\theta - \bar{\theta}) d\theta + \sum_{n=1}^{\infty} \left\{ a_n(t) \int_{-\pi}^{\pi} \cos(n\theta) \cos(\theta - \bar{\theta}) d\theta + b_n(t) \int_{-\pi}^{\pi} \sin(n\theta) \cos(\theta - \bar{\theta}) d\theta \right\}$$

$$= 0 + \sum_{n=1}^{\infty} a_n(t) \pi \cos(\bar{\theta}) \delta_{1,n} + b_n(t) \pi \sin(\bar{\theta}) \delta_{1,n}$$

$$= a_1(t) \pi \cos(\bar{\theta}) + b_1(t) \pi \sin(\bar{\theta})$$

$$= \begin{pmatrix} \pi \cos(\bar{\theta}) \\ \pi \sin(\bar{\theta}) \end{pmatrix} \begin{pmatrix} a_1(t) \\ b_1(t) \end{pmatrix}.$$ 

**Cardioid microphones**

Here and in the following we only consider the simplified unidirectional pickup generated by $\alpha = \frac{1}{2}$, that is the cardioid microphone pickup. Accordingly we have

$$m(\theta; \bar{\theta}) = \frac{1}{2} + \frac{1}{2} \cos(\theta - \bar{\theta})$$

and therefore

$$v_{card}(t, \bar{\theta}) = \int_{-\pi}^{\pi} f(t, \theta) \left( \frac{1}{2} + \frac{1}{2} \cos(\theta - \bar{\theta}) \right) d\theta$$

$$= \frac{1}{2} \int_{-\pi}^{\pi} f(t, \theta) d\theta + \frac{1}{2} \int_{-\pi}^{\pi} f(t, \theta) \cos(\theta - \bar{\theta}) d\theta$$

$$= \frac{1}{2} a_0(t) 2\pi + a_1(t) \pi \cos(\bar{\theta}) + b_1(t) \frac{\pi}{2} \sin(\bar{\theta})$$

$$= \begin{pmatrix} \pi \cos(\bar{\theta}) \\ \frac{\pi}{2} \sin(\bar{\theta}) \end{pmatrix} \begin{pmatrix} a_0(t) \\ a_1(t) \\ b_1(t) \end{pmatrix}.$$ 

where the calculation of the second integral has been anticipated by the bidirectional case.

The crucial observation at this point is that a simplified microphone is only capable of picking up limited Fourier information. With the above calculations we provide basic tools for analyzing microphone arrays. For a specific microphone we have to determine its type given by the pickup pattern and its direction given by the angle $\bar{\theta} \in [-\pi, \pi)$. Then we are able to deal with the voltage functions in order to derive necessary conditions on the design of multiple-microphone arrays in terms of optimal Ambisonic recording.
2.2.1 Full Rank Encoding with \( N \) Microphones

In this section, we model the recording with \( N \in \mathbb{N} \) bidirectional or cardioid equidistant microphones within the unit circle. In subsection 2.2.2 we will see how common stereo microphone configurations such as Mid/Side and XY are related to this. For the first example with simplified bidirectional microphones we consider the angles \( \bar{\theta} = k\pi \) where \( k \) takes the integer values \( 0, \ldots, N - 1 \), because a bidirectional microphone is invariant to rotations of \( \pi \). For each microphone we obtain a voltage function

\[
v_{bi}(k)(t, \frac{k\pi}{N}) = a_1(t)\pi \cos \left( \frac{k\pi}{N} \right) + b_1(t)\pi \sin \left( \frac{k\pi}{N} \right) \\
= \left( \pi \cos \left( \frac{k\pi}{N} \right) \pi \sin \left( \frac{k\pi}{N} \right) \right) \left( a_1(t) \ b_1(t) \right).
\]

Since only the first part varies, we write these first vectors into the rows of a matrix that we then multiply by \( (a_1(t) \ b_1(t))^\top \in \mathbb{R}^2 \). Hence the associated matrix

\[
A_{bi}(N) = \begin{pmatrix}
\pi & 0 \\
\pi \cos \left( \frac{1\pi}{N} \right) & \pi \sin \left( \frac{1\pi}{N} \right) \\
\vdots & \vdots \\
\pi \cos \left( \frac{(N-1)\pi}{N} \right) & \pi \sin \left( \frac{(N-1)\pi}{N} \right)
\end{pmatrix} \in \mathbb{R}^{N\times2}
\]

contains the information on how the Fourier coefficients \( a_1(t), b_1(t) \) are encoded. We want to make sure that all possible information, that is the first Fourier coefficients, is picked up. On the other hand, it is not desirable to pick up a piece of information twice. A practical system will neither duplicate nor omit information. We demand an optimal array to include as much information as possible with as few microphones as possible. For this reason we conclude that the ideal number of microphones is given by the positive integer \( N \) for which the encoding matrix has full rank. Obviously, \( \text{rank}(A_{bi}(N)) \leq N \) and it is easy to check that \( A_{bi}(2) \) has full rank.

Analogously, we analyze a coincident array comprised of cardioid microphones. In this case we have \( \tilde{\theta} = \frac{2k\pi}{N} \) with \( k = 0, \ldots, N - 1 \), which equally apportions the unit circle. The voltage functions are

\[
v_{\text{card}}(k)(t, \frac{2k\pi}{N}) = a_0(t)\pi \cos \left( \frac{2k\pi}{N} \right) + a_1(t)\frac{\pi}{2} \cos \left( \frac{2k\pi}{N} \right) + b_1(t)\frac{\pi}{2} \sin \left( \frac{2k\pi}{N} \right) \\
= \left( \pi \frac{\pi}{2} \cos \left( \frac{2k\pi}{N} \right) \frac{\pi}{2} \sin \left( \frac{2k\pi}{N} \right) \right) \left( a_0(t) \ a_1(t) \ b_1(t) \right),
\]

which we write into the associated encoding matrix \( A_{\text{card}}(N) \) given below. The encoding matrix has full rank for \( N = 3 \), which can be verified by \( \det(A_{\text{card}}(3)) = \frac{3\sqrt{3}\pi^3}{8} \neq 0 \). Again
we conclude that this equals the ideal number of microphones.

\[
A_{\text{card}}(N) = \begin{pmatrix}
\pi & \pi \cos \left( \frac{1}{2} \frac{2\pi}{N} \right) & \frac{\pi}{2} \sin \left( \frac{1}{2} \frac{2\pi}{N} \right) \\
\pi & \pi \cos \left( \frac{(N-1) \cdot 2\pi}{N} \right) & \frac{\pi}{2} \sin \left( \frac{(N-1) \cdot 2\pi}{N} \right) \\
\vdots & \vdots & \vdots \\
\pi & \pi \cos \left( \frac{(N-1) \cdot 2\pi}{N} \right) & \frac{\pi}{2} \sin \left( \frac{(N-1) \cdot 2\pi}{N} \right)
\end{pmatrix} \in \mathbb{R}^{N \times 3}.
\]

Figure 2.2 displays the polar patterns of the coincident multiple-microphone arrays with the ideal number of bidirectional and cardioid microphones, respectively.

![Polar patterns](image)

(a) bidirectional array  
(b) cardioid array

Figure 2.2: Coincident arrays with ideal number of microphones

### 2.2.2 Analysis of Common Stereo Recording Techniques

With the above framework we can study common stereo recording techniques and evaluate if they would be optimal or at least acceptable in Ambisonic recording.

**Mid/Side Arrays** For a Mid/Side array, a bidirectional, and one other microphone (cardioid, omnidirectional, bidirectional) are combined. Note that we have already examined a Mid/Side array with two bidirectional microphones as part of subsection 2.2.1. We now want to study the use of different types of microphones. So we consider a Mid/Side arrangement with a bidirectional microphone pointed sidewards and a cardioid microphone pointed frontwards as shown in Figure 2.3a. The directional angles for this Type 1 are \( \bar{\theta}_{bi} = \frac{\pi}{2} \) and \( \bar{\theta}_{\text{card}} = 0 \). The encoding matrix

\[
\begin{pmatrix}
0 & \pi \cos(\bar{\theta}_{bi}) & \pi \sin(\bar{\theta}_{bi}) \\
\pi & \pi \cos(\bar{\theta}_{\text{card}}) & \frac{\pi}{2} \sin(\bar{\theta}_{\text{card}})
\end{pmatrix} = \begin{pmatrix}
0 & 0 & \pi \\
\pi & \frac{\pi}{2} & 0
\end{pmatrix}
\]
does have full rank, and all three Fourier coefficients are picked up when multiplying the matrix with the corresponding vector \((a_0(t) \ a_1(t) \ b_1(t))\). So we can conclude that this recording technique is optimal in the sense of Ambisonic recording. Studying Type 2 as shown in Figure 2.3b leads to a slightly different result. Again, we consider the same bidirectional microphone with \(\bar{\theta}_{bi} = \frac{\pi}{2}\), but additionally we have an omnidirectional microphone. Then, the encoding matrix given by

\[
\begin{pmatrix}
0 & \pi \cos(\bar{\theta}_{bi}) & \pi \sin(\bar{\theta}_{bi}) \\
2\pi & 0 & 0
\end{pmatrix}
= \begin{pmatrix} 0 & 0 & \pi \\ 2\pi & 0 & 0 \end{pmatrix}
\]

is to be multiplied with the same associated vector as above. The encoding matrix has full rank, but this time the Fourier coefficient \(a_1(t)\) is not picked up at all. That means that a possible piece of information is not extracted. But it is still acceptable in the following sense: We conclude that the Mid/Side arrays are optimal with regard to Ambisonic recording in the sense that two microphones pick up two linearly independent information.

XY Arrays Another stereo recording technique that we want to examine is the XY array shown in Figure 2.4. Again we record with two microphones, but this time we change the angle between the microphones from 90\(^\circ\) to 120\(^\circ\). Then the directional angles differ by \(\frac{2}{3}\pi\) and are given by \(\bar{\theta}_1 = \frac{\pi}{3}\) and \(\bar{\theta}_2 = \frac{5\pi}{3}\). By choosing two bidirectional microphones\(^2\), we obtain:

\[
\begin{pmatrix}
\pi \cos(\bar{\theta}_1) & \pi \sin(\bar{\theta}_1) \\
\pi \cos(\bar{\theta}_2) & \pi \sin(\bar{\theta}_2)
\end{pmatrix}
\begin{pmatrix} a_1(t) \\ b_1(t) \end{pmatrix}
= \begin{pmatrix} \frac{\pi}{2} & \frac{\sqrt{3}\pi}{2} \\ \frac{\pi}{2} & -\frac{\sqrt{3}\pi}{2} \end{pmatrix}
\begin{pmatrix} a_1(t) \\ b_1(t) \end{pmatrix}
\]

and observe that the encoding matrix has full rank. Hence, this recording technique is optimal in Ambisonic recording since it picks up all possible Fourier coefficients. By

\(^2\)Using two bidirectional microphones in an XY array is often called a Blumlein configuration[Cla58].

Figure 2.3: Common stereo recording technique: Mid/Side
choosing two cardioid microphones, we have

\[
\begin{pmatrix}
\pi & \frac{\pi}{2} & \cos(\bar{\theta}_1) & \frac{\pi}{2} & \sin(\bar{\theta}_1) \\
\pi & \frac{\pi}{2} & \cos(\bar{\theta}_2) & \frac{\pi}{2} & \sin(\bar{\theta}_2)
\end{pmatrix}
\begin{pmatrix}
a_0(t) \\
a_1(t) \\
b_1(t)
\end{pmatrix}
= 
\begin{pmatrix}
\pi & \frac{\pi}{4} & \sqrt{\frac{3\pi}{4}} \\
\pi & \frac{\pi}{4} & -\sqrt{\frac{3\pi}{4}}
\end{pmatrix}
\begin{pmatrix}
a_0(t) \\
a_1(t) \\
b_1(t)
\end{pmatrix}
\]

where the encoding matrix obviously does not have full rank. The Fourier coefficients \(a_0(t)\) and \(a_1(t)\) are encoded twice in exactly the same way, and there might be a more sophisticated way to pick up exactly those three coefficients with a different microphone array.

In the following, we focus on the full rank encoding with \(N\) microphones of the same type.

### 2.3 Modified Pickup Function

Real microphones do not behave in the simplified bidirectional and cardioid patterns described above. Their pickup patterns are more complex and must be described by functions with more nontrivial coefficients. So, in accordance to the exact pickup functions, we want to modify the previous assumptions to

\[
m_{bl}(\theta - \bar{\theta}) = \cos(\theta - \bar{\theta}) + \beta \cos(3(\theta - \bar{\theta})),
\]

\[
m_{card}(\theta - \bar{\theta}) = \frac{1}{2} + \frac{1}{2} \cos(\theta - \bar{\theta}) + \gamma \cos(2(\theta - \bar{\theta})).
\]

where \(\beta, \gamma\) are positive real numbers. The polar patterns are shown in Figure 2.5 and Figure 2.6 for representative values of \(\beta\) and \(\gamma\).
2.3 Modified Pickup Function

Repeating the process of the previous section, we have

\[ v_{bi}(t, \bar{\theta}) = \int_{-\pi}^{\pi} f(t, \theta) \cos(\theta - \bar{\theta})d\theta + \beta \int_{-\pi}^{\pi} f(t, \theta) \cos(3(\theta - \bar{\theta}))d\theta \]

\[ = a_1(t)\pi \cos(\bar{\theta}) + b_1(t)\pi \sin(\bar{\theta}) + a_0(t)\beta \int_{-\pi}^{\pi} \cos(3(\theta - \bar{\theta}))d\theta \]

\[ + \beta \sum_{n=1}^{\infty} a_n(t) \int_{-\pi}^{\pi} \cos(n\theta) \cos(3(\theta - \bar{\theta}))d\theta \]

\[ + \beta \sum_{m=1}^{\infty} b_m(t) \int_{-\pi}^{\pi} \sin(m\theta) \cos(3(\theta - \bar{\theta}))d\theta \]

\[ = a_1(t)\pi \cos(\bar{\theta}) + b_1(t)\pi \sin(\bar{\theta}) + a_3(t)\beta \pi \cos(3\bar{\theta}) + b_3(t)\beta \pi \sin(3\bar{\theta}) \]
2.3 Modified Pickup Function

\[ v_{\text{card}}(t, \bar{\theta}) = \frac{1}{2} \int_{-\pi}^{\pi} f(t, \theta) d\theta + \frac{1}{2} \int_{-\pi}^{\pi} f(t, \theta) \cos(\theta - \bar{\theta}) d\theta + \sum_{n=1}^{\infty} a_n(t) \int_{-\pi}^{\pi} \cos(n\theta) \cos(2(\theta - \bar{\theta})) d\theta \]

\[ = a_0(t) \pi + a_1(t) \frac{\pi}{2} \cos(\bar{\theta}) + b_1(t) \frac{\pi}{2} \sin(\bar{\theta}) + a_0(t) \gamma \int_{-\pi}^{\pi} \cos(2(\theta - \bar{\theta})) d\theta \]

\[ + \gamma \sum_{n=1}^{\infty} a_n(t) \int_{-\pi}^{\pi} \cos(n\theta) \cos(2(\theta - \bar{\theta})) d\theta \]

\[ + \gamma \sum_{m=1}^{\infty} b_m(t) \int_{-\pi}^{\pi} \sin(m\theta) \cos(2(\theta - \bar{\theta})) d\theta \]

\[ = a_0(t) \pi + a_1(t) \frac{\pi}{2} \cos(\bar{\theta}) + b_1(t) \frac{\pi}{2} \sin(\bar{\theta}) + a_2(t) \gamma \pi \cos(2\bar{\theta}) + b_2(t) \gamma \pi \sin(2\bar{\theta}). \]

Again, we apply \( \bar{\theta} = \frac{k\pi}{N} \) for \( N \) bi-directional microphones and \( \bar{\theta} = \frac{2k\pi}{N} \) for \( N \) cardioid microphones. The resultant matrix for the bi-directional microphones is

\[ A_{\text{bi}}(N) = \]

\[ \begin{pmatrix} \pi \cos \left( \frac{1\pi}{N} \right) & \pi \sin \left( \frac{1\pi}{N} \right) & 0 & 0 \\ \pi \cos \left( \frac{(N-1)\pi}{N} \right) & \pi \sin \left( \frac{(N-1)\pi}{N} \right) & 0 & 0 \\ \beta \pi \cos \left( \frac{1\pi}{N} \right) & \beta \pi \sin \left( \frac{1\pi}{N} \right) & \beta \pi \cos \left( \frac{1\pi}{N} \right) & \beta \pi \sin \left( \frac{1\pi}{N} \right) \\ \beta \pi \cos \left( \frac{(N-1)\pi}{N} \right) & \beta \pi \sin \left( \frac{(N-1)\pi}{N} \right) & \beta \pi \cos \left( \frac{(N-1)\pi}{N} \right) & \beta \pi \sin \left( \frac{(N-1)\pi}{N} \right) \end{pmatrix} \in \mathbb{R}^{N \times 4}. \]

We multiply the matrix with the vector \( (a_1(t) \ b_1(t) \ a_3(t) \ b_3(t)) \in \mathbb{R}^4 \). The ideal number of microphones is four since the matrix has full rank for \( N = 4 \). The corresponding matrix for cardioid microphones is given by \( A_{\text{card}}(N) \) below. It is to be multiplied with the vector \( (a_0(t) \ a_1(t) \ b_1(t) \ a_2(t) \ b_2(t)) \in \mathbb{R}^5 \). The ideal number of microphones is five as full rank is achieved for \( N = 5 \).

\[ A_{\text{card}}(N) = \]

\[ \begin{pmatrix} 2\pi & \pi \cos \left( \frac{1\pi}{N} \right) & \pi \sin \left( \frac{1\pi}{N} \right) & 0 & 0 \\ 2\pi & \pi \cos \left( \frac{1\pi}{N} \right) & \pi \sin \left( \frac{1\pi}{N} \right) & \gamma \pi \cos \left( \frac{1\pi}{N} \right) & \gamma \pi \sin \left( \frac{1\pi}{N} \right) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 2\pi & \pi \cos \left( \frac{1\pi}{N} \right) & \pi \sin \left( \frac{1\pi}{N} \right) & \gamma \pi \cos \left( \frac{1\pi}{N} \right) & \gamma \pi \sin \left( \frac{1\pi}{N} \right) \end{pmatrix} \in \mathbb{R}^{N \times 5}. \]

As one can guess, the aim is to model the pickup function \( m(\theta; \bar{\theta}) \) for any possible frequency. Therefore we generalize the ansatz and define

\[ m_{bi}(\theta; \bar{\theta}) = \sum_{i=1}^{M} \beta_i \cos((2i-1)(\theta - \bar{\theta})), \quad (2.6) \]

where \( \beta_1 = 1 \). We consider the angles \( \bar{\theta} = \frac{k\pi}{N} \) where \( k = 0, \ldots, N-1 \) and \( N \) is the number
of microphones. This leads to the voltage functions

\[ v_{bi}^{(k)}(t, \bar{\theta}) = \sum_{i=1}^{M} a_{2i-1}(t) \beta_i \pi \cos \left( \frac{(2i-1)\pi k}{N} \right) + b_{2i-1}(t) \beta_i \pi \sin \left( \frac{(2i-1)\pi k}{N} \right). \]

As before, we want to express these relations in the matrix \( A_{bi}(N) \). We use the denotations \( c_{i,j} = \cos \left( \frac{ij\pi}{N} \right) \) and \( s_{i,j} = \sin \left( \frac{ij\pi}{N} \right) \) for any integers \( i, j \). The encoding matrix is

\[ A_{bi}(N) = (\hat{C}_{1,1} \quad \hat{S}_{1,1} \quad \hat{C}_{3,2} \quad \hat{S}_{3,2} \quad \hat{C}_{2M-1,M} \quad \hat{S}_{2M-1,M}) \in \mathbb{R}^{N \times 2M}, \]

where the column vectors are given by

\[
\hat{C}_{i,j} = \begin{pmatrix} c_{0,j} \beta_j \pi \\ c_{1,j} \beta_j \pi \\ c_{2,j} \beta_j \pi \\ \vdots \\ c_{(N-1),j} \beta_j \pi \end{pmatrix} \quad \text{and} \quad \hat{S}_{i,j} = \begin{pmatrix} s_{0,j} \beta_j \pi \\ s_{1,j} \beta_j \pi \\ s_{2,j} \beta_j \pi \\ \vdots \\ s_{(N-1),j} \beta_j \pi \end{pmatrix},
\]

and the associated vector is

\[
(a_0(t) \quad b_1(t) \quad a_3(t) \quad b_3(t) \quad \cdots \quad a_{2M-1}(t) \quad b_{2M-1}(t))^\top \in \mathbb{R}^{2M}.
\]

Similarly, we generalize the pickup function for a cardioid microphone to

\[ m_{card}(\theta, \bar{\theta}) = \frac{1}{2} + \sum_{i=1}^{M} \gamma_i \cos(i(\theta - \bar{\theta})), \quad (2.7) \]

where \( \gamma_1 = \frac{1}{2} \). We consider the angles \( \bar{\theta} = \frac{2k\pi}{N} \) where \( k = 0, \ldots, N-1 \) and \( N \) is the number of microphones. This leads to the voltage functions

\[ v_{card}^{(k)}(t) = \pi a_0(t) + \sum_{i=1}^{M} a_i(t) \gamma_i \pi \cos \left( \frac{2i\pi k}{N} \right) + b_i(t) \gamma_i \pi \sin \left( \frac{2i\pi k}{N} \right). \]

Again, we want to capture these relations in the matrix \( A_{card}(N) \) where we use the notations from above. The corresponding matrix is given by

\[ A_{card}(N) = (\Pi \quad \hat{C}_{1,1} \quad \hat{S}_{1,1} \quad \hat{C}_{2,2} \quad \hat{S}_{2,2} \quad \cdots \quad \hat{C}_{M,M} \quad \hat{S}_{M,M}) \in \mathbb{R}^{N \times 2M+1} \]

where \( \Pi = (\pi \quad \cdots \quad \pi)^\top \in \mathbb{R}^{N} \) and the associated vector is

\[
(a_0(t) \quad a_1(t) \quad b_1(t) \quad a_2(t) \quad b_2(t) \quad \cdots \quad a_N(t) \quad b_N(t))^\top \in \mathbb{R}^{2M+1}.
\]

Definition 2.2 and Theorem 2.3 conclude this chapter with the general result for full rank encoding with modified pickup functions.
2.3 Modified Pickup Function

Definition 2.2 In a coincident multi-microphone array with modified pickup functions, we consider the number of microphones to be ideal if $k$ Fourier coefficients can be determined by $k$ signals.

Theorem 2.3 Given the pickup function (2.6), the ideal number of bidirectional microphones is $2M$. Given the pickup function (2.7), the ideal number of cardioid microphones is $2M + 1$.

Proof We want to show that the matrices $A_{bi}(N)$ and $A_{card}(N)$ have full rank for $N = 2M$ and $N = 2M + 1$, respectively. We do the proof for the matrix $A_{bi}(N)$. The proof for the matrix $A_{card}(N)$ is similar.

Without loss of generality we omit the coefficients $\pi, \beta, \gamma$. It suffices to prove that the columns of the matrix

$$
\hat{V} = \left( \begin{array}{cccccc}
1 & 0 & 1 & 0 & \ldots & \ldots & 1 & 0 \\
\omega_1 & \omega_0 & \omega_2 & \omega_3 & \ldots & \ldots & \omega_{2M} & \omega_{2M-1} \\
\omega_2 & \omega_3 & \omega_4 & \omega_5 & \ldots & \ldots & \omega_{2M-1} & \omega_{2M} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\
\omega_{2M} & \omega_{2M-1} & \omega_{2M-2} & \omega_{2M-3} & \ldots & \ldots & \omega_2 & \omega_1
\end{array} \right)
$$

build an orthogonal set in $\mathbb{R}^{2M}$. We let $i^2 = -1$ and define $\omega_n^m = \exp[i \frac{mn\pi}{2M}]$ as well as

$$
C_k = \left( \begin{array}{c}
\omega_0^k + \omega_1^k \\
\omega_0^k + \omega_{-1}^k \\
\omega_2^k + \omega_3^k \\
\vdots \\
\omega_{2M-2}^k + \omega_{2M-1}^k
\end{array} \right) \quad \text{and} \quad S_k = \left( \begin{array}{c}
i(\omega_0^k - \omega_1^k) \\
i(\omega_0^k - \omega_{-1}^k) \\
i(\omega_2^k - \omega_3^k) \\
\vdots \\
i(\omega_{2M-2}^k - \omega_{2M-1}^k)
\end{array} \right).
$$

Then the matrix $V = 2\hat{V} = (C_1 \ S_1 \ C_3 \ S_3 \ \ldots \ \ldots \ \ldots \ C_{2M-1} \ S_{2M-1}) \in \mathbb{R}^{2M \times 2M}$ resembles the Vandermonde matrix and we can easily calculate that its columns form an orthogonal set as required.

First, for every $n \in \mathbb{N}$ it holds

$$
\sum_{j=0}^{2M-1} \omega_j^n = \frac{1 - \omega_1^n}{1 - \omega_1^n} \cdot \sum_{j=0}^{2M-1} \omega_j^n = \frac{1}{1 - \omega_1^n} \cdot \sum_{j=0}^{2M-1} \omega_j^n - \omega_{j+1}^n = \frac{1}{1 - \omega_1^n} \cdot (\omega_0^n - \omega_{2M}^n) = 1 - \exp[\imath n\pi] = \frac{1 - 1}{1 - \omega_1^n} = 0.
$$

Let $1 \leq k, l \leq 2M - 1$ be odd and distinct. For the inner product of the column vectors it holds
\[ (C_k, S_k) = \sum_{j=0}^{2M-1} i(\omega_j^k + \omega_{-j}^k)(\omega_j^k - \omega_{-j}^k) \]
\[ = i \left\{ \left( \sum_{j=0}^{2M-1} \omega_j^{2k} \right) - 2M + 2M = \left( \sum_{j=0}^{2M-1} \omega_j^{-2k} \right) \right\} = 0, \]
\[ (C_k, C_l) = \sum_{j=0}^{2M-1} (\omega_j^k + \omega_{-j}^k)(\omega_j^l + \omega_{-j}^l) \]
\[ = \sum_{j=0}^{2M-1} \omega_j^{(k+l)} + \sum_{j=0}^{2M-1} \omega_j^{(-k-l)} + \sum_{j=0}^{2M-1} \omega_j^{(k-l)} + \sum_{j=0}^{2M-1} \omega_j^{(l-k)} = 0, \]
\[ (S_k, S_l) = \sum_{j=0}^{2M-1} i^2(\omega_j^k - \omega_{-j}^k)(\omega_j^l - \omega_{-j}^l) \]
\[ = -\sum_{j=0}^{2M-1} \omega_j^{(k+l)} - \sum_{j=0}^{2M-1} \omega_j^{(-k-l)} + \sum_{j=0}^{2M-1} \omega_j^{(k-l)} + \sum_{j=0}^{2M-1} \omega_j^{(l-k)} = 0, \]
\[ (C_k, C_l) = \sum_{j=0}^{2M-1} (\omega_j^k + \omega_{-j}^k)^2 = \sum_{j=0}^{2M-1} \omega_j^{2k} + 2\omega_j^k \omega_{-j}^k + \omega_{-j}^{2k} \]
\[ = 4M + 2 + \sum_{j=1}^{2M-1} \omega_j^{2k} + \omega_{-j}^{2k} \in [4, 8M], \]

since \(-2 \leq \omega_j^{2k} + \omega_{-j}^{2k} \leq 2\) for all \(1 \leq j \leq 2M - 1\). Therefore \((C_k, C_k) \neq 0\). And
\[ (S_k, S_k) = \sum_{j=0}^{2M-1} i^2(\omega_j^k - \omega_{-j}^k)^2 = \sum_{j=0}^{2M-1} \left( 2i \sin \left( \frac{jk\pi}{2M} \right) \right)^2 \]
\[ = 4M - \sum_{j=0}^{2M-1} 4 \cos \left( \frac{jk\pi}{M} \right) \in [4, 16M - 4] \]

since \(-4 \leq 4 \cos \left( \frac{jk\pi}{M} \right) \leq 4\) for all \(0 \leq j \leq 2M - 1\). Therefore \((S_k, S_k) \neq 0\).
3 Full-Sphere Ambisonic Recording

The aim of this chapter is to generalize the theory developed for horizontal-only recording to full-sphere, that is recording in three dimensions. We appoint pickup functions for omni-, uni-, and bidirectional microphones similar to the simplified pickup functions in chapter 2. We calculate their voltage functions and then analyze the Ambisonic B-Format of first order.

3.1 Spherical Harmonics

We extend the model by an upward pointing $z$-axis and let the $x$-axis be forward pointing and the $y$-axis be pointing to the right. Then we are in need of an additional angle, namely $\phi$. So $\Omega = (\theta, \phi)$ where $\theta \in [-\pi, \pi]$ is the azimuthal angle in the $xy$-plane and $\phi \in [0, \pi]$ is the elevation angle from the $z$-axis. Figure 3.1 elucidates the relations. This equals the common notation of a spherical coordinate system in mathematics, but one can find deviating notations in the publications related to sound reproduction.

An orthonormal basis of the system is given by the three vectors

\[
e_r(\theta, \phi) = \cos \theta \sin \phi \ e_1 + \sin \theta \sin \phi \ e_2 + \cos \phi \ e_3,
\]
\[
e_\phi(\theta, \phi) = \cos \theta \cos \phi \ e_1 + \sin \theta \cos \phi \ e_2 - \sin \phi \ e_3,
\]
\[
e_\theta(\theta) = -\sin \theta \ e_1 + \cos \theta \ e_2,
\]

Figure 3.1: Spherical coordinate system
where \( \{e_1, e_2, e_3\} \) is the standard orthonormal basis for \( \mathbb{R}^3 \). The spherical coordinates transformation of a point \((x, y, z)\) in Cartesian coordinates has the form

\[
\begin{align*}
x &= r \cos \theta \sin \phi, \\
y &= r \sin \theta \cos \phi, \\
z &= r \cos \phi.
\end{align*}
\]

As before, we want to write the sound field function in a general form. We need an analogue theory to the classical Fourier series to model full-sphere recording. It is provided by the spherical harmonics, which are the angular part of the solution of Laplace’s equation in spherical coordinates [Mac47]. They are defined as

\[
Y_n^l(\theta, \phi) = \begin{cases} 
\sqrt{\frac{2}{4\pi}} \frac{(n \geq 0)\, \sqrt{\frac{2}{4\pi}} \frac{(n < 0)}{\sqrt{\frac{2}{4\pi}} \frac{(n < 0)}} \frac{(n = 0)}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}{\sqrt{\frac{2}{4\pi}} \frac{(n = 0)}} \frac{(n < 0)}}/\]

where \( P_l^m(\cdot) \) are the associated Legendre polynomials

\[
P_l^m(x) = \frac{(-1)^n}{2^l l!} (1 - x^2)^{l/2} \frac{d^{l+n}}{dx^{l+n}} (x^2 - 1)^l
\]

and \( N_l^n \) are the normalization factors

\[
N_l^n = \sqrt{\frac{2l+1}{4\pi} \frac{(l-n)!}{(l+n)!}}.
\]

The spherical harmonics serve our purpose since they form a complete orthonormal set in the sense that

\[
\int_{\Omega \in S} Y_l^n Y_{l'}^{n'} dS = \delta_{l,l'} \delta_{n,n'} \quad \text{and} \quad \int_{\Omega \in S} |Y_l^n|^2 dS = 1,
\]

and for all \( g \in L^2(S) \), \( g(\Omega) = \sum_{l=0}^{\infty} \sum_{n=-l}^{l} g_l^n Y_l^n(\Omega) \) with coefficients \( g_l^n = \int_{\Omega \in S} g(\Omega) Y_l^n(\Omega) dS \) [Tak94]. Thus, we write the sound field function as a spherical harmonics expansion

\[
f(t, \theta, \phi) = \sum_{l=0}^{\infty} \sum_{n=-l}^{l} f_l^n(t) Y_l^n(\theta, \phi)
\]

or its truncated form

\[
f^L(t, \theta, \phi) = \sum_{l=0}^{L} \sum_{n=-l}^{l} f_l^n(t) Y_l^n(\theta, \phi),
\]
where \( f^m_n(t) \) are the spherical harmonics coefficients

\[
f^m_n(t) = \int_{\Omega \in S} f(t, \Omega) Y^m_n(\Omega) dS.
\]

The surface integral in three dimensions is given by

\[
\int_{\Omega \in S} dS = \int_{-\pi}^{\pi} \int_{0}^{\pi} \sin \phi \, d\phi \, d\theta,
\]

so we let \( \bar{\Omega} = (\bar{\theta}, \bar{\phi}) \) and define the voltage function as

\[
v(t, \bar{\Omega}) = \int_{\Omega \in S} f(t, \Omega) m(\Omega; \bar{\Omega}) dS = \int_{-\pi}^{\pi} \int_{0}^{\pi} f(t, \theta, \phi) m(\theta, \phi; \bar{\theta}, \bar{\phi}) \sin \phi \, d\phi \, d\theta.
\]

### 3.2 Simplified Pickup Function

The pickup function describes the pattern with which the information of the sound field function is extracted. Adjusted accordingly, we now have

\[ m = m(\theta, \phi; \bar{\theta}, \bar{\phi}). \]

In chapter 2 we had \( m(\theta; \bar{\theta}) = \cos(\theta - \bar{\theta}) \), that is cosine of the difference angle of \( \theta \) and \( \bar{\theta} \). We want to follow the same ansatz in three dimensions. Here, the difference angle of \((\theta, \phi)\) and \((\bar{\theta}, \bar{\phi})\) is given by \( \arccos(e_r(\theta, \phi) \cdot e_r(\bar{\theta}, \bar{\phi})) \). We let

\[
\begin{align*}
m_{omni}(\theta, \phi; \bar{\theta}, \bar{\phi}) &= 1, \\
m_{bi}(\theta, \phi; \bar{\theta}, \bar{\phi}) &= e_r(\theta, \phi) \cdot e_r(\bar{\theta}, \bar{\phi}), \\
m_{card}(\theta, \phi; \bar{\theta}, \bar{\phi}) &= \frac{1}{2} + \frac{1}{2} e_r(\theta, \phi) \cdot e_r(\bar{\theta}, \bar{\phi})
\end{align*}
\]

be the simplified pickup functions in the 3D model. By implication, we now compute

\[
e_r(\theta, \phi) \cdot e_r(\bar{\theta}, \bar{\phi}) = 
\begin{align*}
& \left( \cos \theta \sin \phi \, e_1 + \sin \theta \sin \phi \, e_2 + \cos \phi \, e_3 \right) \left( \cos \bar{\theta} \sin \bar{\phi} \, e_1 + \sin \bar{\theta} \sin \bar{\phi} \, e_2 + \cos \bar{\phi} \, e_3 \right) \\
& = \cos \theta \cos \bar{\theta} \sin \phi \sin \bar{\phi} + \sin \theta \sin \phi \sin \bar{\theta} \sin \bar{\phi} + \cos \phi \cos \bar{\phi} \\
& = \left( \frac{\cos(\theta - \bar{\theta}) + \cos(\theta + \bar{\theta})}{2} \right) \left( \frac{\cos(\phi - \bar{\phi}) - \cos(\phi + \bar{\phi})}{2} \right) \\
& \quad + \left( \frac{\cos(\theta - \bar{\theta}) - \cos(\theta + \bar{\theta})}{2} \right) \left( \frac{\cos(\phi - \bar{\phi}) - \cos(\phi + \bar{\phi})}{2} \right) \\
& \quad + \left( \frac{\cos(\phi - \bar{\phi}) + \cos(\phi + \bar{\phi})}{2} \right)
\end{align*}
\]

\[
= \cos(\theta - \bar{\theta}) \left( \frac{\cos(\phi - \bar{\phi}) - \cos(\phi + \bar{\phi})}{2} \right) + \left( \frac{\cos(\phi - \bar{\phi}) + \cos(\phi + \bar{\phi})}{2} \right)
\]

\[
= \frac{1}{2} \left( \cos(\phi - \bar{\phi})(1 + \cos(\theta - \bar{\theta})) + \cos(\phi + \bar{\phi})(1 - \cos(\theta - \bar{\theta})) \right).
\]
3.2.1 Voltage Function \(v_{\text{omni}}\)

The aim of this subsection is to calculate the voltage function for an omnidirectional microphone. It will come out to be useful to do this with the help of the spherical harmonics expansion of the pickup function

\[
m_{\text{omni}}(\Omega; \bar{\Omega}) = 1 = \sum_{l=0}^{\infty} \sum_{n=-l}^{l} a_l^n(\bar{\Omega}) Y_l^n(\Omega) = \sum_{l=0}^{\infty} \sum_{n=-l}^{l} a_l^n Y_l^n(\Omega).
\]

since an omnidirectional microphone is independent of the direction. Note that \(Y_0^0(\Omega) = (2\sqrt{\pi})^{-1}\), and therefore we can write 1 = \(2\sqrt{\pi} Y_0^0(\Omega)\). Moreover, the spherical harmonics are orthogonal. Thus, we have

\[
a_l^n = \int_{\Omega \in S} 1 \cdot Y_l^n(\Omega) dS = \int_{\Omega \in S} 2\sqrt{\pi} Y_0^0(\Omega) Y_l^n(\Omega) dS
\]

\[= 2\sqrt{\pi} \int_{\Omega \in S} |Y_0^0(\Omega)|^2 \delta_{n,0} \, dS = 2\sqrt{\pi} \delta_{n,0} \int_{\Omega \in S} \frac{1}{4\pi} dS
\]

\[= \frac{1}{2\sqrt{\pi}} \delta_{n,0} \delta_{l,0} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \sin \phi \, d\phi \, d\theta = \frac{1}{2\sqrt{\pi}} \delta_{n,0} \delta_{l,0} \left( \int_{-\pi}^{\pi} d\theta \right) \left( \int_{0}^{\pi} \sin \phi \, d\phi \right)
\]

\[= 2\sqrt{\pi} \delta_{n,0} \delta_{l,0}.
\]

The pickup function can be written as

\[
m_{\text{omni}}(\Omega; \bar{\Omega}) = \sum_{l=0}^{\infty} \sum_{n=-l}^{l} a_l^n(\bar{\Omega}) Y_l^n(\Omega) = a_0^0 Y_0^0(\Omega).
\]

We utilize this new version of the pickup function to find the voltage function. Using the truncated form of the sound field function allows us to change the order of the integral and the sums.

\[
v_{\text{omni}}^L(t) = \int_{\Omega \in S} m_{\text{omni}}(\Omega; \bar{\Omega}) f^L(t, \Omega) dS = \int_{\Omega \in S} a_0^0 Y_0^0(\Omega) \left( \sum_{l=0}^{L} \sum_{n=-l}^{l} f_l^n(t) Y_l^n(\Omega) \right) dS
\]

\[= \sum_{l=0}^{L} \sum_{n=-l}^{l} a_0^0 f_l^n(t) \int_{\Omega \in S} Y_l^n(\Omega) Y_l^n(\Omega) dS = a_0^0 f_0^n(t) \delta_{l,0} \delta_{n,0} = 2\sqrt{\pi} f_0^0(t)
\]

for every integer \(L\). So we have

\[
v_{\text{omni}}(t) = 2\sqrt{\pi} f_0^0(t).
\]
3.2.2 Voltage Function $v_{bi}$

In this subsection, we want to find the voltage function for a bidirectional microphone with the pickup function

$$m_{bi}(\theta, \phi; \bar{\theta}, \bar{\phi}) = \frac{1}{2} \left( \cos(\phi - \bar{\phi})(1 + \cos(\theta - \bar{\theta})) + \cos(\phi + \bar{\phi})(1 - \cos(\theta - \bar{\theta})) \right)$$

$$= \sum_{l=0}^{\infty} \sum_{n=-l}^{l} b_{l}^{n}(\bar{\Omega}) Y_{l}^{n}(\Omega).$$

Again we consider $n = 0, n > 0$ and $n < 0$ separately. For $n = 0$ it holds

$$b_{0}^{0}(\bar{\Omega}) = \int_{\Omega \in S} m_{bi}(\theta, \phi; \bar{\theta}, \bar{\phi}) Y_{0}^{0}(\Omega) dS$$

$$= \int_{-\pi}^{\pi} \int_{0}^{\pi} m_{bi}(\theta, \phi; \bar{\theta}, \bar{\phi}) N_{l}^{0} P_{l}^{0}(\cos \phi) \sin \phi \, d\phi \, d\theta$$

$$= \frac{1}{2} N_{l}^{0} \left( \int_{-\pi}^{\pi} (1 + \cos(\theta - \bar{\theta})) \, d\theta \right) \left( \int_{0}^{\pi} \cos(\phi - \bar{\phi}) P_{l}^{0}(\cos \phi) \sin \phi \, d\phi \right)$$

$$+ \frac{1}{2} N_{l}^{0} \left( \int_{-\pi}^{\pi} (1 - \cos(\theta - \bar{\theta})) \, d\theta \right) \left( \int_{0}^{\pi} \cos(\phi + \bar{\phi}) P_{l}^{0}(\cos \phi) \sin \phi \, d\phi \right)$$

$$= \frac{1}{2} N_{l}^{0} 2\pi \left( \int_{0}^{\pi} \cos(\phi - \bar{\phi}) P_{l}^{0}(\cos \phi) \sin \phi \, d\phi \right)$$

$$+ \frac{1}{2} N_{l}^{0} 2\pi \left( \int_{0}^{\pi} \cos(\phi + \bar{\phi}) P_{l}^{0}(\cos \phi) \sin \phi \, d\phi \right)$$

$$= N_{l}^{0} \pi \int_{0}^{\pi} \left[ \cos(\phi - \bar{\phi}) + \cos(\phi + \bar{\phi}) \right] P_{l}^{0}(\cos \phi) \sin \phi \, d\phi$$

$$= 2\pi N_{l}^{0} \cos \bar{\phi} \int_{0}^{\pi} P_{l}^{0}(\cos \phi) \cos \phi \sin \phi \, d\phi$$

$$= 2\pi N_{l}^{0} \cos \bar{\phi} I_{1}(l).$$

Similarly we obtain

$$b_{l}^{n}(\bar{\Omega}) = \delta_{n,1} \frac{2\pi}{\sqrt{2}} N_{l}^{1} \cos \bar{\theta} \sin \bar{\phi} \int_{0}^{\pi} P_{l}^{1}(\cos \phi) \sin^{2} \phi \, d\phi = \delta_{n,1} \frac{2\pi}{\sqrt{2}} N_{l}^{1} \cos \bar{\theta} \sin \bar{\phi} I_{2}(l)$$

for $n > 0$, and for $n < 0$ we have

$$b_{l}^{n}(\bar{\Omega}) = \delta_{n,-1} \frac{2\pi}{\sqrt{2}} N_{l}^{1} \sin \bar{\theta} \sin \bar{\phi} \int_{0}^{\pi} P_{l}^{1}(\cos \phi) \sin^{2} \phi \, d\phi = \delta_{n,-1} \frac{2\pi}{\sqrt{2}} N_{l}^{1} \sin \bar{\theta} \sin \bar{\phi} I_{2}(l).$$
It remains to calculate the integrals $I_1(l)$ and $I_2(l)$ for any integer $l \geq 0$.

\[
I_1(l) = \int_0^\pi \cos \phi \sin \phi P_0^l(\cos \phi) d\phi \\
= \frac{1}{2l!} \int_0^\pi \cos \phi \sin \phi \frac{d^l}{d(\cos \phi)^l} (\cos^2 \phi - 1)^l d\phi \\
= \frac{1}{2l!} \int_0^\pi \cos \phi \sin \phi \frac{d^l}{d(\cos \phi)^l} \left( \sum_{j=0}^{l} \binom{l}{j} (\cos^2 \phi)^{l-j} (-1)^j d\phi \right) \\
= \frac{1}{2l!} \int_0^\pi \cos \phi \sin \phi \sum_{j=0}^{l} (-1)^j \frac{l!}{j!(l-j)!} (\cos \phi)^{2l-2j} d\phi \\
= \frac{1}{2l!} \int_0^\pi \cos \phi \sin \phi \sum_{j=0}^{\lfloor l/2 \rfloor} (-1)^j \frac{l!}{j!(l-j)!} \frac{(2l-2j)!}{(l-2j)!} (\cos \phi)^{2l-2j} d\phi \\
= \frac{1}{2l!} \sum_{j=0}^{\lfloor l/2 \rfloor} (-1)^j \binom{l}{j} \frac{(2l-2j)!}{(l-2j)!} \frac{2}{l-2j+2}, \text{ if } l \text{ is odd.}
\]

By undertaking the same procedure for the second integral, we have

\[
I_2(l) = \int_0^\pi \sin^2 \phi P_1^l(\cos \phi) d\phi \\
= \int_0^\pi \sin^2 \phi \frac{-1}{2l!} \left( 1 - \cos^2 \phi \right)^{\frac{1}{2}} \frac{d^{l+1}}{d(\cos \phi)^{l+1}} (\cos^2 \phi - 1)^l d\phi \\
= \frac{(-1)}{2l!} \sum_{j=0}^{\lfloor l/2 \rfloor} (-1)^j \binom{l}{j} \frac{(2l-2j)!}{(l-2j-1)!} \int_0^\pi \sin^3 \phi (\cos \phi)^{2l-2j-1} d\phi \\
= \frac{(-1)}{2l!} \sum_{j=0}^{\lfloor l/2 \rfloor} (-1)^j \binom{l}{j} \frac{(2l-2j)!}{(l-2j-1)!} \frac{2(1 + (-1)^{l-2j-1})}{(l-2j)(2+l-2j)}, \text{ if } l \text{ is even.}
\]
This means that \( b_l^{-1}(\Omega) = b_l^0(\Omega) = b_l^1(\Omega) = 0 \) for every \( l \in \mathbb{N}_0 \) that is even. Here 0 is considered to be even. Indeed, the proof for \( l = 0 \) follows from the first line of the calculations. Moreover, we observe that \( b_l^n(\Omega) = 0 \) for \(|n| \geq 2\), thus

\[
m_{bi}(\Omega; \bar{\Omega}) = \sum_{l=0}^{\infty} \sum_{n=-l}^{l} b_l^n(\Omega) Y_l^n(\Omega) = \sum_{l=0}^{\infty} \sum_{n=-l}^{l} b_l^n(\Omega) Y_l^n(\Omega) = \sum_{k=0}^{\infty} \left( b_{2k+1}^{-1}(\Omega) Y_{2k+1}^{-1}(\Omega) + b_{2k+1}^0(\Omega) Y_{2k+1}^0(\Omega) + b_{2k+1}^1(\Omega) Y_{2k+1}^1(\Omega) \right).
\]

In the next step we want to see the impact of this result on the voltage function. Again, using the truncated forms both of the pickup and the sound field function allows us to change the order of the integrals and sums.

\[
v_{bi}(t, \Omega) = \int_{\Omega \in S} m_{bi}(\Omega; \bar{\Omega}) \left( \sum_{l=0}^{L} \sum_{n=-l}^{l} f_l^n(t) Y_l^n(\Omega) \right) dS = \sum_{k=0}^{L} \left[ \sum_{l=0}^{L} \sum_{n=-l}^{l} b_k^{-1}(\Omega) f_l^n(t) \int_{\Omega \in S} Y_k^{-1}(\Omega) Y_l^n(\Omega) dS \right] + \sum_{k=0}^{L} \left[ \sum_{l=0}^{L} \sum_{n=-l}^{l} b_k^0(\Omega) f_l^n(t) \int_{\Omega \in S} Y_k^0(\Omega) Y_l^n(\Omega) dS \right] + \sum_{k=0}^{L} \left[ \sum_{l=0}^{L} \sum_{n=-l}^{l} b_k^1(\Omega) f_l^n(t) \int_{\Omega \in S} Y_k^1(\Omega) Y_l^n(\Omega) dS \right]
\]

As \( L \rightarrow +\infty \), we obtain

\[
v_{bi}(t, \Omega) = \sum_{k=0}^{L} \left[ b_{2k+1}^{-1}(\Omega) f_{2k+1}^{-1}(t) + b_{2k+1}^0(\Omega) f_{2k+1}^0(t) + b_{2k+1}^1(\Omega) f_{2k+1}^1(t) \right].
\] (3.2)
3.2.3 Voltage Function $v_{\text{card}}$

Since the crucial part of the cardioid pickup function

$$m_{\text{card}}(\theta, \phi; \bar{\theta}, \bar{\phi}) = \frac{1}{2} + \frac{1}{4} \left( \cos(\phi - \bar{\phi})(1 + \cos(\theta - \bar{\theta})) + \cos(\phi + \bar{\phi})(1 - \cos(\theta - \bar{\theta})) \right)$$

$$= \sum_{l=0}^{\infty} \sum_{n=-l}^{l} c_l^n(\bar{\Omega}) Y_l^n(\Omega)$$

is the same as the bidirectional pickup function, the spherical harmonics coefficients $c_l^n(\bar{\Omega})$ are easily derived from the coefficients $b_l^n(\bar{\Omega})$. We have

$$c_l^n(\bar{\Omega}) = \frac{1}{\sqrt{2}} \int_{\Omega \in S} \frac{1}{2} + \frac{1}{2} m_{bi}(\Omega; \bar{\Omega}) Y_l^n(\Omega) dS = \frac{1}{2} \int_{\Omega \in S} d\Omega + \frac{1}{2} \int_{\Omega \in S} m_{bi}(\Omega; \bar{\Omega}) Y_l^n(\Omega) dS$$

$$= \frac{1}{2} \int_{-\pi}^{\pi} \sin \phi \, d\phi \, d\theta + \frac{1}{2} b_l^n(\bar{\Omega}) = 2\pi + \frac{1}{2} b_l^n(\bar{\Omega})$$

and therefore

$$m_{\text{card}}(\Omega; \bar{\Omega}) = \sum_{k=0}^{\infty} \sum_{m=-k}^{k} c_l^n(\bar{\Omega}) Y_l^m(\Omega)$$

$$= 2\pi \sum_{k=0}^{\infty} \sum_{m=-k}^{k} Y_l^m(\Omega) + \frac{1}{2} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} b_{2k+1}^l(\bar{\Omega}) Y_{2k+1}^l(\Omega).$$

Following the already familiar procedure and skipping the truncation and limiting process, this in turn yields

$$v_{\text{card}}(t, \bar{\Omega}) = \int_{\Omega \in S} m_{\text{card}}(\Omega; \bar{\Omega}) f(t, \Omega) dS$$

$$= 2\pi \int_{\Omega \in S} \left( \sum_{k=0}^{\infty} \sum_{m=-k}^{k} Y_l^m(\Omega) \right) \left( \sum_{l=0}^{\infty} f_l^n(\Omega) Y_l^n(\Omega) \right) dS$$

$$+ \frac{1}{2} \int_{\Omega \in S} \left( \sum_{k=0}^{\infty} \sum_{l=-1}^{1} b_{2k+1}^l(\bar{\Omega}) Y_{2k+1}^l(\Omega) \right) \left( \sum_{l=0}^{\infty} \sum_{n=-l}^{l} f_l^n(\Omega) Y_l^n(\Omega) \right) dS$$

$$= 2\pi \sum_{k=0}^{\infty} \sum_{m=-k}^{k} \sum_{l=0}^{\infty} f_l^n(t) \int_{\Omega \in S} Y_l^m(\Omega) Y_l^n(\Omega) dS + \frac{1}{2} \int_{\Omega \in S} Y_l^m(\Omega) Y_l^n(\Omega) dS$$

$$= 2\pi \sum_{k=0}^{\infty} \sum_{m=-k}^{k} f_{k}^{m}(t) + \frac{1}{2} v_{bi}(t, \bar{\Omega}).$$

Hence,

$$v_{\text{card}}(t, \bar{\Omega}) =$$

$$2\pi \sum_{k=0}^{\infty} \sum_{m=-k}^{k} f_{k}^{m}(t) + \frac{1}{2} \left[ b_{2k+1}^{-1}(\bar{\Omega}) f_{2k+1}^{-1}(t) + b_{2k+1}^{0}(\bar{\Omega}) f_{2k+1}^{0}(t) + b_{2k+1}^{1}(\bar{\Omega}) f_{2k+1}^{1}(t) \right].$$
3.3 Analysis of the First Order B-Format

The highest aim of surround sound systems would be to reproduce the whole original sound field in an exact way over a restricted listening area. Clearly this is not possible because of the tremendous amount of information. Hence we try to optimize the sound reproduction over a reasonable number of channels. For the First Order B-Format, this number is four. We saw in chapter 2 and chapter 3 that all we can actually do is to determine how much information we can capture with some combination of microphones that we characterize to be optimal in a prior defined way. Now we want to find out if the Ambisonic B-Format is optimal in the sense of our definition. Observe that in chapter 2 we included investigations about stereo recording techniques. In contrast, B-Format signals are not in any sense stereo compatible [Mal98].

In section 3.2 we found that the voltage functions of the simplified pickup functions are given by (3.1) for omnidirectional, (3.2) for bidirectional, and (3.3) for cardioid microphones. Examining First Order B-Format means to use their truncated forms with only the first order, that is having $k = 0$ instead of $k = 0, 1, 2, \ldots$. The voltage functions become

\[
v_{\text{omni}}(t, \Omega) = a_0^0 f_0^0(t),
\]
\[
v_{\text{bi}}(t, \Omega) = b_1^{-1}(\Omega)f_1^{-1}(t) + b_0^0(\Omega)f_0^0(t) + b_1^1(\Omega)f_1^1(t),
\]
\[
v_{\text{card}}(t, \Omega) = 2\pi f_0^0(t) + \frac{1}{2} [b_1^{-1}(\Omega)f_1^{-1}(t) + b_0^0(\Omega)f_0^0(t) + b_1^1(\Omega)f_1^1(t)].
\]

The B-Format array we want to study consists of one omnidirectional and three bidirectional microphones with the directional angles given in Figure 3.2.

<table>
<thead>
<tr>
<th>Microphone</th>
<th>Direction</th>
<th>Directional Angle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bidirectional 1</td>
<td>forward-pointing</td>
<td>$\Omega_1 = (0, \frac{\pi}{2})$</td>
</tr>
<tr>
<td>Bidirectional 2</td>
<td>leftward-pointing</td>
<td>$\Omega_2 = (\frac{3\pi}{4}, \frac{\pi}{2})$</td>
</tr>
<tr>
<td>Bidirectional 3</td>
<td>upward-pointing</td>
<td>$\Omega_3 = (0, 0)$</td>
</tr>
<tr>
<td>Omnidirectional</td>
<td>whole unit sphere</td>
<td>—</td>
</tr>
</tbody>
</table>

Figure 3.2: B-Format Microphone Types

An illustration of the setup is given in Figure 3.3, where the pictures show SoundField’s patented capsule array [Sou07]. We now calculate the particular voltage functions for the microphones. Note that we already calculated $a_0^0 = 2\sqrt{\pi}$ in subsection 3.2.1 and thus have

\[
v_{\text{omni}}(t) = 2\sqrt{\pi} f_0^0(t).
\]

Then, we need to calculate the voltage functions for the bidirectional microphones by plugging their directional angles into the truncated voltage function $v_{\text{bi}}$, as follows.
3.3 Analysis of the First Order B-Format

Figure 3.3: B-Format Microphone Array

\[ v_{b1}(t, \Omega_1) = b_0^{-1} \left(0, \frac{\pi}{2}\right) f_1^{-1}(t) + b_1^0 \left(0, \frac{\pi}{2}\right) f_1^0(t) + b_1^1 \left(0, \frac{\pi}{2}\right) f_1^1(t) \]
\[ = b_1^1 \left(0, \frac{\pi}{2}\right) f_1^1(t) = \frac{2\pi}{\sqrt{2}} N_1^1 \cos(0) \sin \left(\frac{\pi}{2}\right) I_2(1) f_1^1(t) \]
\[ = -2\sqrt{\frac{\pi}{3}} f_1^1(t) = \begin{pmatrix} 0 & 0 & -2\sqrt{\frac{\pi}{3}} \end{pmatrix} \begin{pmatrix} f_0^0(t) \\ f_1^{-1}(t) \\ f_1^0(t) \end{pmatrix}, \]
\[ v_{b2}(t, \Omega_2) = b_1^{-1} \left(\frac{3\pi}{4}, \frac{\pi}{2}\right) f_1^{-1}(t) + b_0^0 \left(\frac{3\pi}{4}, \frac{\pi}{2}\right) f_1^0(t) + b_1^1 \left(\frac{3\pi}{4}, \frac{\pi}{2}\right) f_1^1(t) \]
\[ = b_1^{-1} \left(\frac{3\pi}{4}, \frac{\pi}{2}\right) f_1^{-1}(t) + b_1^1 \left(\frac{3\pi}{4}, \frac{\pi}{2}\right) f_1^1(t) \]
\[ = -\sqrt{\frac{\pi}{6}} f_1^{-1}(t) + \sqrt{\frac{2\pi}{3}} f_1^1(t) = \begin{pmatrix} 0 & -\sqrt{\frac{\pi}{6}} & 0 \end{pmatrix} \begin{pmatrix} f_0^0(t) \\ f_1^{-1}(t) \\ f_1^0(t) \end{pmatrix}, \]
\[ v_{b3}(t, \Omega_3) = b_1^{-1}(0, 0) f_1^{-1}(t) + b_0^0(0, 0) f_1^0(t) + b_1^1(0, 0) f_1^1(t) \]
\[ = b_0^0(0, 0) f_1^0(t) = 2\pi N_1^0 \cos(0) I_1(1) \]
\[ = \sqrt{\frac{\pi}{3}} f_1^0(t) = \begin{pmatrix} 0 & 0 & \sqrt{\frac{\pi}{3}} \end{pmatrix} \begin{pmatrix} f_0^0(t) \\ f_1^{-1}(t) \\ f_1^0(t) \end{pmatrix}. \]
The associated matrix-vector multiplication with the encoding matrix $E \in \mathbb{R}^{4 \times 4}$ is

$$
\begin{pmatrix}
0 & 0 & 0 & -2\sqrt{\frac{3}{3}} \\
0 & -\sqrt{\frac{\pi}{6}} & 0 & \sqrt{\frac{2\pi}{3}} \\
0 & 0 & \sqrt{\frac{\pi}{3}} & 0 \\
2\sqrt{\frac{\pi}{2}} & 0 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
f_0^0(t) \\
f_0^{-1}(t) \\
f_1^0(t) \\
f_1^1(t)
\end{pmatrix}.
$$

The matrix $E$ has full rank. Thus, the First Order B-Format microphone array is optimal in terms of our definition regarding Ambisonic recording.
4 Conclusions

4.1 Overview of Results

In this thesis, we dealt with the encoding process of audio recording and reproduction in surround sound technology. We studied the recording of two dimensional (horizontal-only) and three dimensional (full-sphere) sound fields with coincident multiple-microphone arrays. We applied a simplified model that does not explicitly include physical components such as pressure and velocity, and which neglects the fact that true coincidence is not possible. Also, we used simplified assumptions about the pickup functions that describe the polar patterns of different types of microphones. We investigated Classical Fourier series representations in horizontal-only recording and spherical harmonics representations in full-sphere Ambisonic recording to describe the sound field.

Our first aim was to study the optimal use of multiple microphones on the plane. We obtained encoding matrices from which we derived necessary conditions for the design of the microphones for full rank encoding of the sound field information. For an array with $N$ equidistant bidirectional microphones with simplified pickup function we obtained the ideal number $N = 2$, and for the same setup with cardioid microphones we had $N = 3$. We defined the number of microphones to be ideal if $k$ Fourier or spherical harmonics coefficients are determined by $k$ recording signals. After modifying the pickup function by adding $M - 1$ terms, we concluded that $N = 2M$ and $N = 2M + 1$ are the ideal numbers of bidirectional and cardioid microphones, respectively.

The next aim was to extend the two dimensional theory to the three dimensional model. Analogous to the simplified pickup function in the 2D case, a simplified pickup function in three dimensions was derived. Eventually, the mathematical model was used to analyse the Ambisonic First Order B-Format with regard to full rank encoding. We concluded that it is optimal in the sense of Ambisonic recording since its corresponding encoding matrix has full rank.

4.2 Open Problems

There are a number of open problems and other approaches that we want to mention in this last subsection. First, we want to note that instead of modeling the sound field function with real Fourier and spherical harmonics expansions we could have used their
complex analogues. We chose the real representations since the sound field itself is real. But the complex approach would have been possible, too. The coherence of real and complex spherical harmonics is given in [Bat09].

In terms of full-sphere recording, we analyzed the First Order B-Format. Another microphone array that would be worth analyzing is the A-Format with four cardioid microphones pointing to the faces of a tetrahedron. In fact, one can consider any recording signals gained from a platonic solid. Regarding the basic Ambisonic recording formats, another open problem is the n-th Order B-Format. The aim would be to show that with our mathematical model, we would obtain an optimal number of \((n + 1)^2\) channels as Gerzon does with his dirac delta pickup [Ger73]. Gerzon also explains why we would want to increase the order of the encoding system: any desired degree of accuracy can be achieved by making \(n\) large enough.

In his paper [Ger73], Gerzon introduces the generalization of spherical harmonics in Ambisonic recording to spin harmonics. Spherical harmonics, in this context, are spin harmonics that have the same spin if and only if they have the same order. We only contemplated the special case of spherical harmonics in this thesis.

In chapter 3 we stated that B-Format signals are not stereo compatible [Mal98]. It could be part of future work to show the relation between Ambisonic and stereo signals by using the mathematical model constructed in this thesis.

Also, there have not been any tests on the described Ambisonic recording theory. There are numerous reports about tests on Ambisonic decoders, such as [BLH06], [LH07]. But the decoding process has not been part of this thesis at all. It would be an interesting task to study the optimal number of speakers, especially in relation to the ideal number of microphones. But it shall not be concealed that the decoding process is more complex than the encoding process.

A general open problem is given by psychoacoustics in its entirety. Constant progress in the understanding of how the human brain-ear combination localizes sound will help to incorporate all relevant components. Rather than neglecting the physical values hiding behind the Fourier and spherical harmonics coefficients, it would be an important step to include them and also try to regard such equally difficult and influential factors as reverberation.
Bibliography


[Inc09] INCORPORATED, Shure; SHURE (Hrsg.): Product Specifications SM58® Cardioid Dynamic Microphone. Shure, 2009


[Sou07] SOUNDFIELD ; TSL PROFESSIONAL PRODUCTS LTD. (Hrsg.): DSF-1 Performance Microphone System. TSL Professional Products Ltd., 2007

Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Erklärung

Hiermit versichere ich, dass ich diese Arbeit selbständig verfasst und keine anderen, als die angegebenen Quellen und Hilfsmittel benutzt, die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht und die Satzung des Karlsruher Instituts für Technologie zur Sicherung guter wissenschaftlicher Praxis in der jeweils gültigen Fassung beachtet habe.

Blacksburg, May 8, 2013