

# Implementierung eines 64 Kanal Mikrofonarrays

Daniel Hofer

Diplomarbeit durchgeführt am  
Institut für Elektronische Musik und Akustik  
Universität für Musik und Darstellende Kunst Graz  
&  
Technische Universität Graz



Beurteilung: Robert Höldrich, o.Univ.-Prof. Dr. techn. Mag. DI

Betreuung: Alois Sontacchi, Univ.-Ass. Dr. techn. DI

Franz Zotter, DI

Graz, im Juni 2008

*Gewidmet*

*meiner Frau und unseren drei Kindern Alina, Leticia und Gabriel*

*sowie meinen Eltern*

## Kurzfassung

Beamforming, oder auch räumliches Filtern, wird dazu eingesetzt, die Richtungsempfindlichkeit einer Mikrofonanordnung zu formen. Dazu werden mindestens zwei Mikrofone benötigt. Zur Verbesserung der räumlichen Auflösung bedarf es mehrerer Mikrofone. Dies stellt die Motivation dar, ein planares Mikrofonarray mit 64 Kanälen zu entwerfen. Die Implementierung eines kostengünstigen planaren Mikrofonarrays für 64 einzelne Kanäle ist der erste Schritt in dieser Arbeit. Die Grundkonfiguration der Mikrofone ist ein 8x8 Raster mit variablen Abständen der Rasterpunkte. Damit lassen sich unterschiedlichste Geometrien entwerfen. Die Vermessung der Geometrie stellt jedoch bei 64 Mikrofonen eine Herausforderung dar. Daher soll die Erfassung der Geometrie automatisiert erfolgen. Dazu werden Quellen an bekannten Raumpunkten benötigt. Als stabile räumliche Konstruktion wird in dieser Arbeit eine tetraederförmige Quellenanordnung eingesetzt. In den Ecken des Tetraeders befinden sich vier individuell bespielbare Lautsprecher. Mit diesen Lautsprechern werden Messsignale auf das Mikrofonarray gesendet. Die Laufzeiten zu den einzelnen Mikrofonen werden mittels *Generalized Cross Correlation Phasetransform* (GCC-PHAT) und linearer Regression bestimmt. Bei bekannter Schallgeschwindigkeit entsprechen die Laufzeiten den zurückgelegten Weglängen des Luftschalls. Aus diesen Weglängen wird mit numerischen Methoden (Simplex Methode) die Geometrie der Anordnung bestimmt. Das Resultat sind die Koordinaten der Mikrofone im Raum. Diese Koordinaten dienen in weiterer Folge als Grundlage zur Implementierung eines *Delay and Sum* Beamformers. Eine weitere Anwendung des planaren Mikrofonarrays ist die Nahfeldholographie.

## Abstract

Beamforming, or spatial filtering, is a useful technique for designing a directivity pattern with microphone arrays. At least two microphones are required for this application. To improve spatial resolution the number of microphones needs to be further increased. The motivation throughout this thesis is to implement a low cost planar microphone array with 64 channels. The basic configuration is an 8x8 grid with variable spacing inbetween the grid points. This offers a variety of different geometric layouts with 64 microphones. A challenge, however, is to determine the positions of the 64 microphones in space. Consequently, the geometry shall be captured automatically. For this purpose, loudspeakers are placed in space at known positions. This work employs a tetrahedral source arrangement as stable spatial construction. Four loudspeakers are mounted on the vertices of the tetrahedron. Every loudspeaker individually transmits a measurement signal towards the microphone array. The acoustic delay to every microphone is determined with generalized cross correlation phase transform (GCC-PHAT) and with linear regression. At the given speed of sound the acoustic delay corresponds to the length of the paths. From the path lengths, the coordinates of the microphones in space are determined with a numerical method (Simplex Method). The coordinates are required for an implementation of delay and sum beam formers or nearfield acoustical holography (NAH).

# Danksagung

Bedanken möchte ich mich bei Herrn Dr. Alois Sontacchi und Herrn DI Franz Zotter für die erstklassige Betreuung . Sie hatten immer ein offenes Ohr für meine Fragen. Sie gaben mir gute Anregungen und erteilten mir wertvolle Ratschläge, und vor allem nahmen sie sich Zeit für interessante Diskussionen. Ganz besonderer Dank gilt meiner Frau MMag. Evelin Kammerer, die mich das ganze Studium hindurch unterstützte und aufbaute und mir, so gut es ging, den Rücken freigehalten hat. Schließlich wäre das Studium ohne die finanzielle und tatkräftige Unterstützung meiner Eltern in dieser Form nicht möglich gewesen. Ihnen gebürt großer Dank. Danke !!!

# Inhaltsverzeichnis

<b>Kurzfassung</b>	<b>iii</b>
<b>Danksagung</b>	<b>v</b>
<b>Inhaltsverzeichnis</b>	<b>viii</b>
<b>1 Einleitung</b>	<b>1</b>
<b>2 Aufbau des Mikrofonarrays</b>	<b>3</b>
2.1 Stativ . . . . .	3
2.2 Anbindung an die Hardware . . . . .	4
2.3 Mikrofone . . . . .	4
<b>3 Vermessen der EMC8000 Mikrofone und Berechnung der Entzerrungsfilter</b>	<b>7</b>
3.1 Die Sweep-Messmethode . . . . .	7
3.1.1 Der Sweep . . . . .	7
3.1.2 Bestimmung der Impulsantwort . . . . .	8
3.1.2.1 Im Zeitbereich . . . . .	9
3.1.2.2 Im Frequenzbereich . . . . .	9
3.2 Messaufbau . . . . .	10
3.2.1 Messstrecke . . . . .	11
3.2.2 Messablauf . . . . .	12
3.3 Impulsantworten der Entzerrungsfilter . . . . .	12
3.3.1 Spektrale Division . . . . .	13
3.3.2 Cepstrum . . . . .	14

---

<b>4</b>	<b>Bestimmung der Mikrofonpositionen</b>	<b>15</b>
4.1	Time Delay Estimation . . . . .	16
4.1.1	Phasentransformation (PHAT) . . . . .	18
4.1.2	Experimente der Kreuzkorrelation und Phat . . . . .	18
4.1.3	Parabolische Interpolation . . . . .	21
4.2	Lineare Regression der Phase . . . . .	22
4.3	Gruppenlaufzeit . . . . .	23
4.4	Piezo-Hochtonhorn . . . . .	24
4.4.1	Gruppenlaufzeit und das akustische Zentrum . . . . .	25
4.5	Mathematische Modelle zur Auswertung der Mikrofonpositionen anhand der bestimmten Laufzeiten . . . . .	27
4.5.1	Lösung der quadratischen Gleichung . . . . .	27
4.5.2	Der Nelder-Mead Simplex-Algorithmus . . . . .	29
4.6	Die Implementierung . . . . .	32
4.7	Experimente und Ergebnisse . . . . .	33
<b>5</b>	<b>Beamformer</b>	<b>38</b>
5.1	Grundlagen . . . . .	38
5.2	Aperturen . . . . .	38
5.2.1	Kontinuierliche Aperturen . . . . .	38
5.2.2	Diskretisierte Aperturen . . . . .	41
5.3	Delay and Sum . . . . .	42
<b>6</b>	<b>Akustische Nahfeldholographie</b>	<b>45</b>
6.1	Ebene Wellen . . . . .	45
6.2	Evaneszente Wellen . . . . .	46
6.3	Homogene Lösung der Wellengleichung . . . . .	48
6.4	Wellenzahlspektrum . . . . .	49
6.5	Extrapolation des Schallfeldes . . . . .	49
6.6	Aliasing . . . . .	51
6.7	Simulation und Ergebnisse der Nahfeldholographie . . . . .	52
6.8	Schnelle und Druckverteilung auf einer Lautsprechermembran . . . . .	54

---

<b>7 Conclusio</b>	<b>58</b>
<b>Literaturverzeichnis</b>	<b>60</b>
<b>A Geräte und Datenblätter</b>	<b>63</b>
A.1 Mikrofon Behringer EMC8000 . . . . .	63
A.2 Vorverstärker Behringer Ultragain Pro-8 ADA8000 . . . . .	63
A.3 MADI (Multichannel Audio Digital Interface) . . . . .	64
A.4 RME Hammerfall DSP MADI . . . . .	64
A.5 Larson Davis CAL200 Sound Level Calibrator . . . . .	65
A.6 G.R.A.S Freifeld Mikrofon Typ 40AF . . . . .	65
A.7 Tetraeder . . . . .	65
A.8 Piezo-Hochtonhornwandler . . . . .	66



# Abbildungsverzeichnis

2.1	Mikrofonarray . . . . .	4
2.2	Signalfluss . . . . .	4
2.3	Frequenzgang des Messlautsprechers, gemessen mit EMC8000 Mikrofonen im Vergleich zur Messung mit dem Gras 40AF (rot-strichliert). . . . . .	5
2.4	Richtdiagramm EMC8000 Mikrofon; linker Halbkreis: obere 4 Oktaven, rechter Halbkreis: untere 4 Oktaven. . . . .	6
3.1	Tropfenlautsprecher . . . . .	8
3.2	Übertragungsstrecke des unbekanntes Systems . . . . .	9
3.3	Schematische Darstellung des Messaufbaus . . . . .	10
3.4	Impulsantworten der Behringer Mikrofone . . . . .	11
3.5	Bestandteile der Messstrecke . . . . .	12
3.6	Vermessung der Mikrofonfrequenzgänge in der Messkoje des IEM . . . . .	12
3.7	Blockschaltbild der Entzerrungsstrecke . . . . .	13
3.8	Impulsantworten der Entzerrungsfilter . . . . .	14
4.1	Aufbau der Messeinrichtung im Cube: im Vordergrund die tetraederförmige Lautsprecheranordnung, im Hintergrund das Mikrofonarray . . . . .	16
4.2	Standard Kreuzkorrelation von weißem Rauschen mit einer Verzögerung von 11 Samples . . . . .	19
4.3	Phat von weißem Rauschen mit einer Verzögerung von 11 Samples . . . . .	20
4.4	Standard Kreuzkorrelation der Impulsantwort eines Behringer EMC8000 mit einem Dirac-Impuls. . . . .	20

---

4.5	Phat der Impulsantwort eines Behringer EMC8000 mit einem Dirac-Impuls. . . . .	21
4.6	Parabolische Interpolation . . . . .	22
4.7	Frequenzgang des Piezo-Hochtonhorns . . . . .	25
4.8	Richtdiagramm des Hochtonlautsprechers . . . . .	25
4.9	Schematische Darstellung zur Gruppenlaufzeitmessung eines Lautsprechers . . . . .	26
4.10	Fluktuation des akustischen Zentrums, Gruppenlaufzeit des Piezo-Hochtonhorns . . . . .	27
4.11	Performance der Methode nach Reddi [Red93]. Die Lautsprecherpositionen sind in den Ecken des Tetraeders. In diesem Fall wird die 1. Lösung um den Tetraeder gespiegelt. Die beiden Lösungen können aber auch sehr eng nebeneinander zu liegen kommen. . . . .	28
4.12	Performance des Nelder-Mead Simplex-Algorithmus [LRWW98]. Die Lautsprecherpositionen sind in den Ecken des Tetraeders. . . . .	29
4.13	Gegenüberstellung der Messdaten von Vicon und der akustischen Messdaten, bestimmt über die Mittelung der Gruppenlaufzeiten. Die Messdaten der Gruppenlaufzeit sind um einen Offset verschoben, damit die Darstellung übersichtlicher wird. . . . .	33
4.14	Fehler der euklidischen Norm für das jeweilige Mikrofon in Bezug auf die Vicon-Messdaten. Der Fehler der Phat weist eine größere Streuung auf. . . . .	34
4.15	Varianzanalyse des euklidischen Fehlers der einzelnen Methoden zur Detektion der Laufzeiten. Die Phat weist einen Signifikanten Unterschied zu den beiden anderen Methoden auf. Lineare Regression und Gruppenlaufzeit weisen hingegen keinen signifikanten Unterschied auf. . . . .	35
4.16	Abweichung der Phat-Messdaten von Vicon-Messdaten . . . . .	35
4.17	Abweichung der Linearen-Regression-Messdaten von Vicon-Messdaten	36
4.18	Abweichung der Gruppenlaufzeit-Messdaten von Vicon-Messdaten . .	36
5.1	Direktive Systemantwort einer Apertur . . . . .	39

5.2	Directivity Pattern mit $N=8$ Sensoren, $d=0.1$ [m] und $f=1$ [kHz] $\alpha' = 45$	43
5.3	Directivity Pattern für 8 Sensoren, $d=0.1$ , $f=0.3-4$ kHz . . . . .	44
6.1	Wellenzahlraum der ebenen Wellen und der evaneszenten Wellen . .	48
6.2	K-Space-Filter für $\alpha = 0.1$ und $k_c = 1.2 \cdot k_r$ , ebene Wellen befinden sich im Durchlassbereich, evaneszente Wellen werden unterdrückt. . .	51
6.3	Wellenzahlraum einer ebenen Welle, $f=5000$ [Hz] Azimuth= $274^\circ$ , Elevation= $29.7^\circ$ . . . . .	53
6.4	Wellenzahlraum einer Punktquelle im Ferfeld, $f=5000$ [Hz] Azimuth= $274^\circ$ , Elevation= $29.7^\circ$ , die beiden Linien, die normal zueinander verlaufen, stellen die zweidimensionalen Sinc-Funktion dar. Die Krümmung der Wellenfront ist sehr schwach . . . . .	53
6.5	Wellenzahlraum einer Punktquelle im Nahfeld, $f=5000$ [Hz] Azimuth= $274^\circ$ , Elevation= $29.7^\circ$ . . . . .	54
6.6	Betrag vom Orts-Frequenz-Spektrum des Drucks, extrapoliert auf die Membranoberfläche bei $f=250$ Hz, und bereits mit dem k-Space Filter, den Anteil der evaneszenten Wellen reduziert. . . . .	55
6.7	Schnelleverteilung auf der Membranoberfläche bei $f=250$ Hz, mit zeropadding im Orts-Frequenz-Spektrum vor der Rücktransformation zur Erhöhung der Auflösung. . . . .	55
6.8	Nicht k-Space gefilterter Schalldruck im Wellenzahlraum dargestellt. Evaneszente Wellen liegen vor. . . . .	56
6.9	Schnelleverteilung auf der Membranoberfläche bei 500 Hz Sinuston im Nahfeld. Moden treten auf der Membran auf, aus denen die evaneszenten Wellen hervorgehen. . . . .	56
6.10	Schnelleverteilung bei $f=500$ Hz und Unterdrückung der Moden durch K-Space Filterung . . . . .	57
6.11	Schnelleverteilung auf der Membranoberfläche bei $f=1000$ Hz. Je mehr evaneszente Wellen vor der Rücktransformation zugelassen werden, desto mehr Moden werden sichtbar. . . . .	57
A.1	Behringer EMC8000 Mikrofon . . . . .	63

---

A.2	Behringer Ultragain Pro-8 ADA8000 . . . . .	64
A.3	RME ADI-648 . . . . .	64
A.4	RME HDSP MADI . . . . .	64
A.5	Tetraeder Lautsprecherarray . . . . .	66
A.6	Piezo-Hochtonhorn . . . . .	66

# Kapitel 1

## Einleitung

Das Anwendungsgebiet von Mikrofonarrays ist breit gefächert: Es reicht von räumlicher Abtastung des Schallfeldes für Wiedergabe über Ambisonics oder Wellenfeldsynthese, Quelllokalisierung, Verfolgung von bewegten Quellen bis hin zur Auswertung von Abstrahlungsverhalten schwingender Strukturen. Letzteres ist über die akustische Nahfeldholographie (NAH) möglich, die in der Theorie das erste Mal 1980 auftrat [WM80] und durch die steigende Rechnerkapazität in den letzten Jahren auch in der Industrie immer mehr zum Einsatz kommt. Für eine vernünftige Auflösung der NAH ist eine gewisse Anzahl an Mikrofonen nötig, je mehr Mikrofone desto schärfer kann das eintreffende Schallfeld abgebildet werden. Aus diesem Grund ist ein 64 Kanal Mikrofonarray implementiert worden. Die Anordnung der Mikrofone ist flexibel gestaltet und erlaubt diverse geometrische Anordnungen. Die Fragen, die sich in diesem Zusammenhang stellen sind: Wie kann das gesamte System kalibriert werden, und gibt es eine Möglichkeit, die einzelnen Positionen der Mikrofone schnell und einfach zu vermessen? Die Kalibrierung des gesamten Systems kann nur erfolgen, indem die Mikrofone und Vorverstärker einzeln vor jeder Messung mit einem Kalibrator kalibriert werden, da die verwendeten Vorverstärker keine Möglichkeit bieten, den ermittelten Gainwert zu speichern. Zur Positionsbestimmung der einzelnen Mikrofone am Stativ ist ein akustischer Messaufbau implementiert. Die zu bestimmende Größe ist die Laufzeit des Schalls von einem bekannten Referenzsystem aus zu den einzelnen Mikrofonen. Dazu werden drei Verfahren - die Phasentransformation, die lineare Regression der Phase und die Gruppenlaufzeit -

getestet und ausgewertet, wobei die Bestimmung der Laufzeit über die Gruppenlaufzeit einen neuen Ansatz darstellt. Die detektierten Laufzeiten werden schließlich einem Optimierungsverfahren übergeben, welches anhand der bekannten Lautsprecherpositionen des Referenzsystems die Koordinaten der 64 Mikrofone bestimmt.

Im letzten Kapitel wird die Umsetzung der akustischen Nahfeldholographie zur Messung von schwingenden Strukturen implementiert. Mit den heutigen Rechnerkapazitäten ist es möglich, die NAH in Echtzeit (alle 100 [ms] wird aktualisiert) zu betreiben. Die Implementierung erfolgt in PureData (PD). Eine Lautsprechermessung bei verschiedenen Frequenzen, das Zurückrechnen auf die Membranoberfläche sowie die Auswertung der Membranschnelle in Matlab, runden diese Diplomarbeit ab.

# Kapitel 2

## Aufbau des Mikrofonarrays

Das 64 kanalige Mikrofonarray am Institut für Elektronische Musik und Akustik in Graz (IEM) ist sehr ähnlich aufgebaut wie jenes am Institut für Schallforschung der Österreichischen Akademie der Wissenschaften in Wien(ÖAW) [PM04].

### 2.1 Stativ

Das Stativ besteht aus einer Aluminiumständer-Rahmenkonstruktion. Im Rahmen befinden sich 8 Aluminiumlatten, die vertikal verschiebbar sind. Auf diesen Aluminiumlatten sind die angebrachten Mikrofone horizontal verschiebbar. Weiters lässt sich der Rahmen auch aus dem Lot heraus bis  $35^\circ$  neigen. Die oberste Mikrofonreihe kann bis auf eine maximale Höhe von  $2m$ , die unterste Mikrofonreihe kann bis auf eine minimale Höhe von  $0,95m$  verstellt werden. Aufgrund der Befestigungsschellen können die Mikrofone auf einen Mindestabstand von  $0,045m$  zueinander verschoben werden. Diese Schellen sind mit der Aluminiumlatte fest verschraubt, siehe Abbildung 2.1. Vorteile dieser Aluminiumbauweise sind:

- Hohe akustische Transparenz
- Im Freien einsetzbar
- Geringe Anschaffungskosten
- Transportierbar
- Geringes Gewicht



Abbildung 2.1: Mikrofonarray

## 2.2 Anbindung an die Hardware

Verwendet werden 64 Mikrofone der Firma Behringer, Typ EMC8000, die an 8 Behringer Vorverstärkern, siehe Anhang 2 (A.2), Typ Ultragain Pro-8, angeschlossen werden. Die Ausgangssignale der Vorverstärker sind via ADAT mit dem MADI-Router von RME (A.3) verbunden und dieser wiederum mit dem PC via RME Hammerfall Soundkarte (A.4), siehe Abbildung 2.2. Die weitere Signalverarbeitung findet in PureData (PD) und Matlab<sup>®</sup> statt.

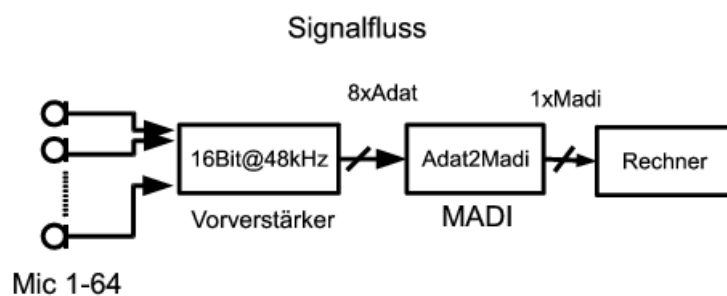


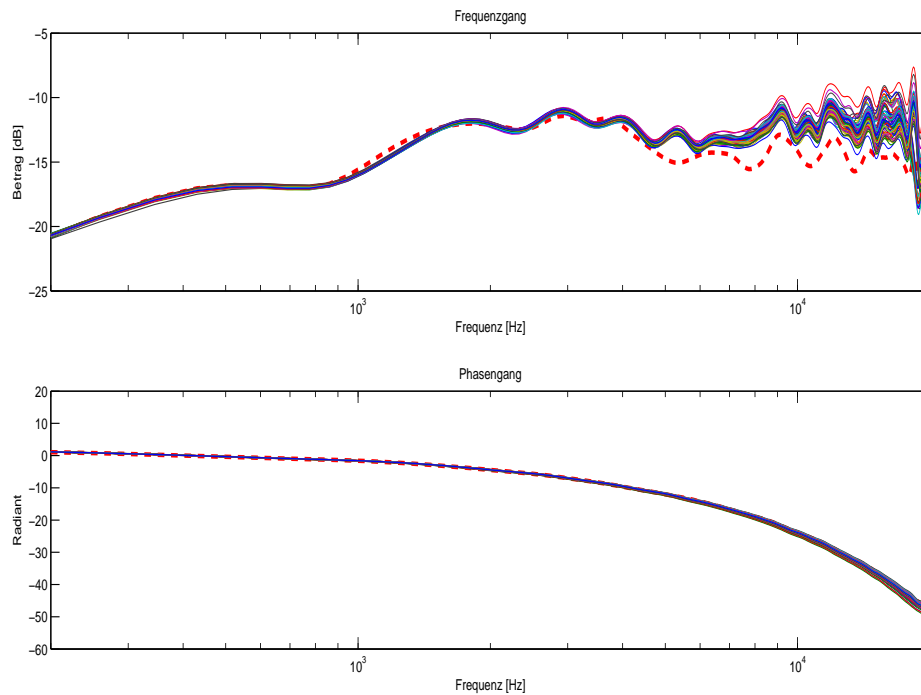
Abbildung 2.2: Signalfluss

## 2.3 Mikrofone

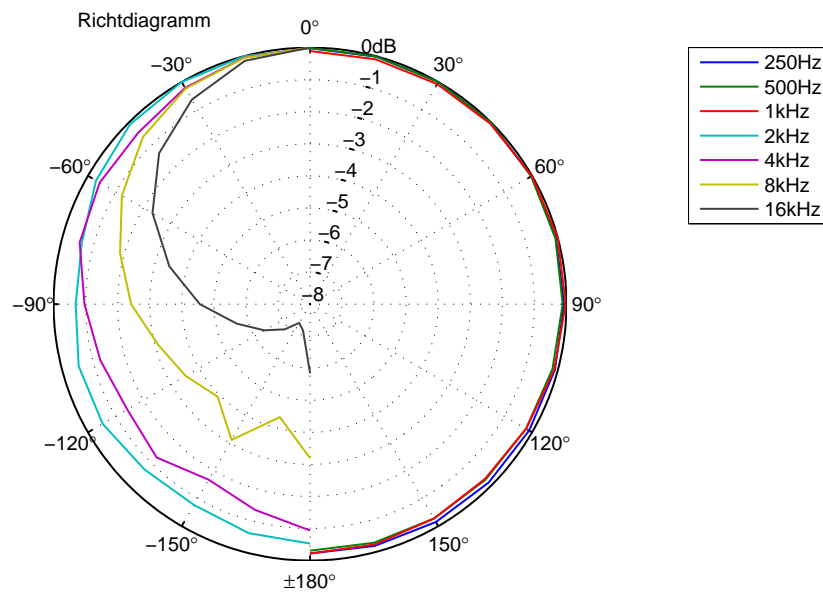
Die verwendeten Mikrofone sind omnidirektionale Elektret-Kondensator Mikrofone mit XLR-Anschluss, siehe (A.1). Vom Hersteller wird angegeben, dass sich diese Mi-



krofone besonders zur Raumeinmessung eignen, und dass sie einen glatten Frequenzgang aufweisen. Der angegebene Frequenzgang und die Omnidirektionalität wurden nachgemessen und in Abbildung 2.3 und Abbildung 2.4 dargestellt. Die Messungen wurden mit der Sweep-Methode nach Angelo Farina [Far00] in der Messkoje (3.6) am Institut für Elektronische Musik und Akustik (IEM) durchgeführt, siehe Kapitel 3. In Abbildung 2.3 sind die Frequenz- und Phasengänge aller 64 verwendeten Mikrofone im Vergleich zum qualitativ hochwertigen Messmikrofon GRAS 40AF dargestellt (rot strichliert). Auffällig ist, dass die EMC8000 Mikrofone einen 6 dB Anstieg zu hohen Frequenzen hin aufweisen. Daraus lässt sich folgern, dass sie Diffusfeld-entzerrt sind. In den Impulsantworten sind auch die Systemantworten der Messkoje und des verwendeten Lautsprechers (Tropfenlautsprecher, [War03]) enthalten.



**Abbildung 2.3:** Frequenzgang des Messlautsprechers, gemessen mit EMC8000 Mikrofonen im Vergleich zur Messung mit dem Gras 40AF (rot-strichliert).



**Abbildung 2.4:** Richtdiagramm EMC8000 Mikrofon; linker Halbkreis: obere 4 Oktaven, rechter Halbkreis: untere 4 Oktaven.

# Kapitel 3

## Vermessen der EMC8000 Mikrofone und Berechnung der Entzerrungsfilter

### 3.1 Die Sweep-Messmethode

#### 3.1.1 Der Sweep

Als Messsignal wird ein exponentieller Sweep verwendet, nach der Methode von Farina [Far00]. Ein Sweep ist ein periodisches Signal, das die Frequenz über eine gewisse Zeitdauer ändert, beginnend bei einer vordefinierten Startfrequenz bis zu einer definierten Stoppfrequenz. Die Frequenzänderung erfolgt bei dieser verwendeten Messung logarithmisch. Mathematisch kann der Sweep wie folgt beschrieben werden:

$$x(t) = \sin(f(t)). \quad (3.1)$$

Für einen exponentiellen Sweep gilt:

$$x(t) = \sin[K \cdot (e^{\frac{t}{L}} - 1)]. \quad (3.2)$$

Die Ableitung des Arguments der Sinusfunktion nach der Zeit liefert die Momentanfrequenz. Mit Hilfe der Startfrequenz  $\omega_1$  zum Zeitpunkt  $t = 0$  und der Stoppfrequenz  $\omega_2$  zum Zeitpunkt  $t = T$  lassen sich die unbekanntenen Konstanten  $K$  und  $L$  bestimmen.  $T$  ist die Sweeplänge.

$$K = \frac{T \cdot \omega}{\ln\left(\frac{\omega_2}{\omega_1}\right)} \quad (3.3)$$

$$L = \frac{T}{\ln\left(\frac{\omega_2}{\omega_1}\right)} \quad (3.4)$$

Somit ergibt sich die Beschreibung für den exponentiellen Sweep

$$x(t) = \sin\left[\frac{T \cdot \omega}{\ln\left(\frac{\omega_2}{\omega_1}\right)} \cdot e^{\frac{t}{T} \ln\left(\frac{\omega_2}{\omega_1}\right)} - 1\right]. \quad (3.5)$$

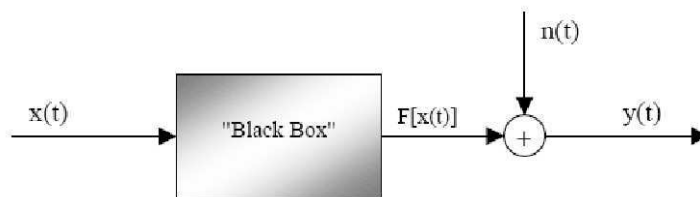
Der Vorteil der Sweeptchnik zur Messung von Impulsantworten ist, dass das zu messende System nicht absolut linear sein muss. Dies ist bei raumakustischen Messungen immer der Fall. Bei der verwendeten Messstrecke 3.2.1 weisen vor allem die verwendeten Tropfenlautsprecher, siehe Abbildung 3.1, Nichtlinearitäten auf. Näheres zum verwendeten Tropfenlautsprecher findet sich in der Diplomarbeit von Stefan Warum [War03]. Durch den exponentiellen Sweep werden die harmonischen Verzerrungen zeitlich vom Direktschall getrennt. Ursachen für die Zeitvarianz sind einerseits die Erwärmung der Lautsprecherspule und andererseits die Temperaturschwankungen des Raumes, welche die Ausbreitungsgeschwindigkeit des Schalls verändern.



Abbildung 3.1: Tropfenlautsprecher

### 3.1.2 Bestimmung der Impulsantwort

Die Messstrecke kann betrachtet werden als ein unbekanntes System (black box), siehe 3.2, das einen Eingang und einen Ausgang besitzt.



**Abbildung 3.2:** Übertragungsstrecke des unbekanntes Systems

Das Eingangssignal  $x(t)$  verursacht ein deterministisches Ausgangssignal  $y(t)$ , welches von Störungen  $n(t)$  überlagert wird,

$$y(t) = n(t) + F[x(t)]. \quad (3.6)$$

Für den Fall eines linearen zeitinvarianten Systems (LTI) entspricht die Funktion  $F[x(t)]$  der Faltung der Impulsantwort  $h(t)$  mit dem Eingangssignal  $x(t)$

$$y(t) = n(t) + x(t) \otimes h(t). \quad (3.7)$$

Zur Bestimmung von  $h(t)$  ist es notwendig, den Ausdruck  $x(t) \otimes h(t)$  zu „entfalten“. Diese Entfaltung ist auf zwei Arten lösbar.

### 3.1.2.1 Im Zeitbereich

Für die Lösung im Zeitbereich braucht man einen geeigneten inversen Filter  $f(t)$ , der als Faltungsergebnis mit dem Eingangssignal  $x(t)$  einen Dirac-Impuls  $\delta(t)$  liefert:

$$x(t) \otimes f(t) = \delta(t). \quad (3.8)$$

Die Impulsantwort lässt sich nun berechnen durch die Faltung des Ausgangssignals  $y(t)$  mit dem inversen Filter  $f(t)$ :

$$h(t) = y(t) \otimes f(t). \quad (3.9)$$

### 3.1.2.2 Im Frequenzbereich

Durch Division der Spektren  $X(k)$  und  $\hat{Y}(k)$  und durch Rücktransformation erhält man die Impulsantwort. Dieser Ansatz ist aber nur zulässig unter der Voraussetzung, dass genügend Mittelungen  $\hat{y}(t)$  des Ausgangssignals  $y(t)$  vorliegen. Denn, je mehr Messungen gemacht werden um so besser ist der SNR, genauer gesagt verbessert sich der SNR um 3 dB je Verdoppelung der Anzahl der Messungen. Ist dies gegeben, kann man auch davon ausgehen, dass der Rauschanteil  $n(t)$  sehr gering wird. Ansonsten kann bei der Entfaltung der Rauschanteil in Gleichung 3.7 überproportional verstärkt werden. Weiters ist darauf zu achten, dass Ausgangs- und Eingangssignal periodisch über die zirkuläre Faltung zusammenhängen. Dabei besteht die Gefahr von *Time-Domain-Aliasing*. Dies kann vermieden werden, indem die Signallänge

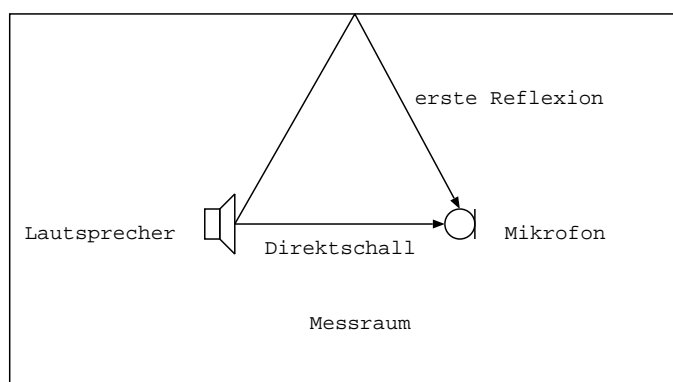
der Fourier-Transformationen für Messsignal und Systemantworten zumindest auf die Länge von  $M+N-1$  angepasst und mit Nullen aufgefüllt wird (*zero-padding*).  $N$  ist dabei die Länge des Messsignals  $x(t)$  und  $M$  die Länge der Systemantwort  $\hat{y}(t)$ .  $M+N-1$  gibt die Länge des Ergebnisses der linearen Faltung an. Die Dekonvolution erfolgt durch Division im Frequenzbereich. Die Impulsantwort ergibt sich durch die Rücktransformation des Quotienten:

$$h(t) = IFFT \left[ \frac{FFT(\hat{y}(t))}{FFT(x(t))} \right]. \quad (3.10)$$

## 3.2 Messaufbau

Die Impulsantworten der 64 Mikrofone wurden in der akustisch gedämpften Messkoje des (IEM) vermessen. Von fünf stichprobenartig gewählten Mikrofonen wurden zudem das Richtdiagramm aufgenommen, siehe Abbildung 2.4. Abbildung 3.3 zeigt den Messaufbau mit einem Mikrophon, das  $1m$  vor dem Tropfenlautsprecher positioniert ist. Alle Mikrofone wurden jeweils axial zur Lautsprechermembran in gleichem Abstand positioniert. Weiters wurden die Abstände zu den Wänden sowie zu Decke und Boden so gewählt, dass das zu messende Mikrophon genau in der Mitte der jeweiligen Strecken stand, um die ersten Reflexionen so spät wie möglich am Empfänger ankommen zu lassen.

Zur Vermessung der Omnidirektionalität wurden die Mikrofone, bei ähnlicher Positionierung, auf einen Drehteller gestellt. Die Ausrichtung des Drehtellers wurde in  $15^\circ$ -Schritten von  $0^\circ$  bis  $180^\circ$  variiert. Es wurde darauf geachtet, dass sich die Mikrophonkapsel im Zentrum befand.



**Abbildung 3.3:** Schematische Darstellung des Messaufbaus

Die Anordnung des Messaufbaus ist in Abbildung 3.3 gezeigt. Bei einer Distanz von  $1m$  zwischen Tropfenlautsprecher und zu messendem Mikrophon lässt sich das Eintreffen der ersten Reflexion mit dem Satz von Pythagoras berechnen. Der zeitliche Ver-

satz  $\Delta t$  zwischen Direktschall und erster Reflexion ergibt sich eine Verzögerungszeit von  $6ms$ . Für ein gültiges Ergebnis muss eine ganze Schwingungsperiode innerhalb des „Zeitfensters“  $\Delta t$  (gibt den Freifeldanteil an) am Mikrofon ankommen. Die freie Weglänge ist umso größer, je weiter reflektierende Flächen von dieser Anordnung entfernt sind. Somit ergibt sich eine untere Grenzfrequenz für diese Messanordnung, ausgehend von der Fernfeldbedingung eines Kugelstrahlers [Wes03]

$$kr \geq 2. \quad (3.11)$$

Bei  $r = 1$  ergibt sich für die untere Grenzfrequenz  $f_{ug} = 108 Hz$ , auf Samples bezogen bei  $f_s = 44100 Hz$  ein zeitliches Fenster von  $264 Samples$ , siehe Abbildung 3.4.

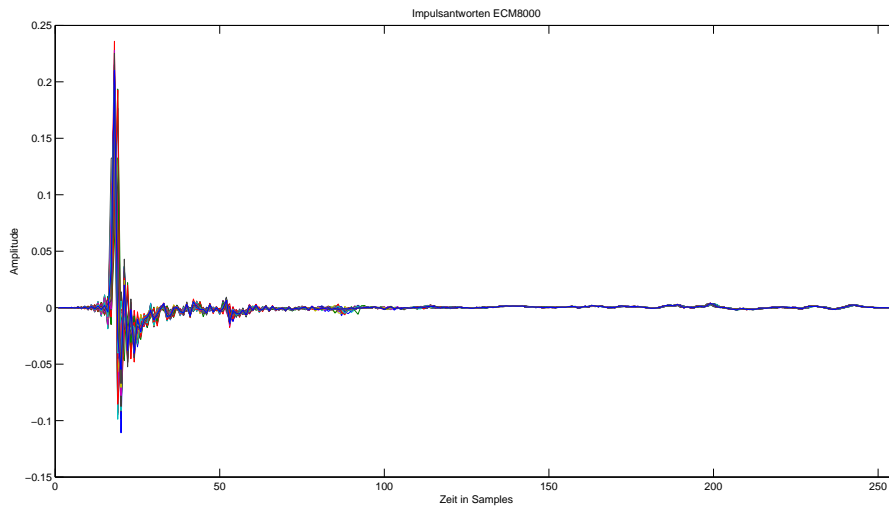


Abbildung 3.4: Impulsantworten der Behringer Mikrofone

### 3.2.1 Messstrecke

Die Messstrecke ist die serielle Verknüpfung der einzelnen Übertragungsfunktionen von DA-Wandler ( $H_{DA}$ ), Verstärker ( $H_{VS}$ ), Lautsprecher ( $H_{LS}$ ), Raum ( $H_R$ ), Messobjekt ( $H_{obj}$ ), Mikrofonvorverstärker ( $H_{MVS}$ ) und AD-Wandler ( $H_{AD}$ ). Das Übertragungsverhalten der Verstärker und der DA/AD-Wandler kann als linear angenommen werden. Nichtlinearitäten treten wie oben erwähnt bei der Übertragungsfunktion des Lautsprechers und des Raumes auf. Abbildung 3.5 zeigt die serielle Messstrecke.

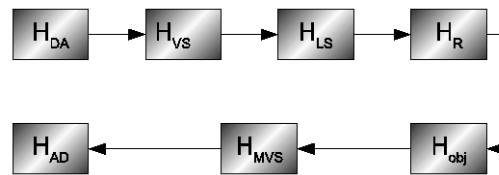


Abbildung 3.5: Bestandteile der Messstrecke

### 3.2.2 Messablauf

“Die gesamte Messstrecke” in Abbildung 3.2.1 wurde auf 94 dB bei  $1kHz$  kalibriert. Die Sweeplänge wurde mit  $T = 5s$  festgelegt. Das erzeugte Schallfeld wurde von dem zu messenden Mikrophon aufgenommen und im wav-Format à 16 Bit bei 44100 Hz Abtastfrequenz gespeichert. Mit dem “hochwertigen” Messmikrophon (A.6) wurde eine Referenzmessung der “gesamten Messstrecke” vorgenommen, da dieses Messmikrophon einen geradlinigeren Frequenzgang aufweist. Ziel ist es, damit den Frequenzgang der Mikrofone zu entzerren (3.3). In Abbildung 3.6 ist der Messaufbau der Messstrecke zu sehen.



Abbildung 3.6: Vermessung der Mikrofonfrequenzgänge in der Messkoje des IEM

## 3.3 Impulsantworten der Entzerrungsfiler

Mit Entzerrungsfilern soll das Übertragungsverhalten aller Mikrofone über den verwendeten Frequenzbereich so gut wie möglich kompensiert werden. Aufgrund der



Schwankungen im Frequenzgang der EMC8000 Mikrofone werden Entzerrungsfilter bestimmt, die eine geradlinige Frequenzantwort zu erhalten. Die Vorgehensweise ist in Abbildung 3.7 schematisch dargestellt und in (3.3.1) beschrieben. Der kausale Verzögerungsanteil der Übertragungsstrecke wird dabei ausgeklammert.

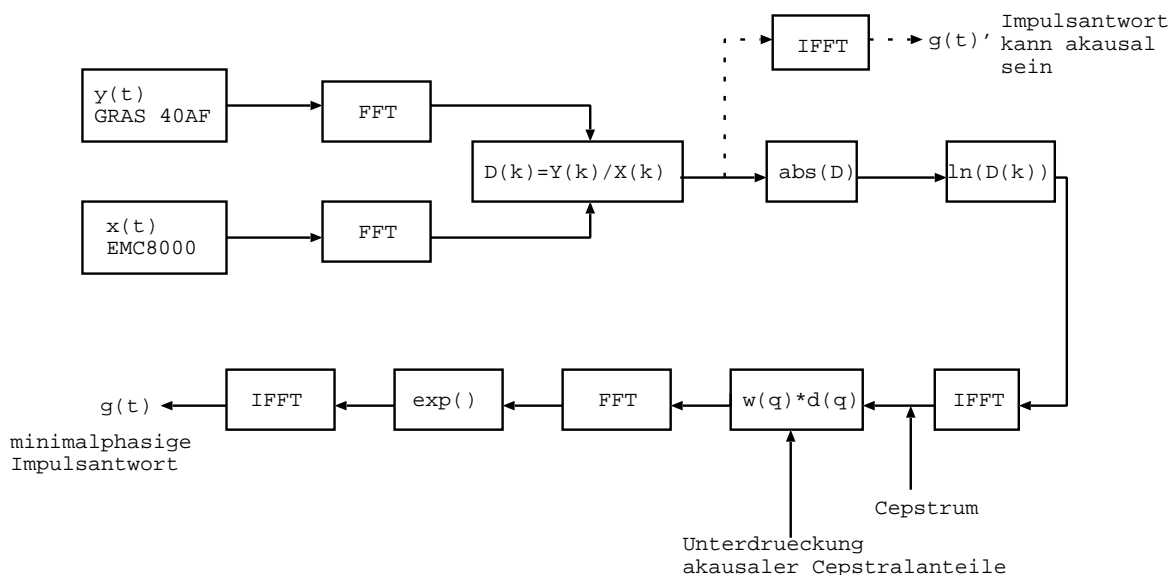


Abbildung 3.7: Blockschaltbild der Entzerrungsstrecke

### 3.3.1 Spektrale Division

Von der Impulsantwort des Messmikrofons wird das Betragsspektrum gebildet  $Y(k)$ , wie auch vom jeweiligen Array-Mikrofon  $X(k)$ . Danach wird die Division durchgeführt:

$$D(k) = Y(k)/X(k). \quad (3.12)$$

Zu beachten ist, dass nur der Freifeldanteil der Impulsantwort verwendet wird. Bei der Division können an Stellen, wo sehr kleine spektrale Beträge in  $X(k)$  vorliegen, Singularitäten auftreten. An diesen Stellen sollte mit einem additiven Zahlenwert regularisiert werden. Im Grunde führt die Rücktransformation des Quotienten  $IFFT\{D(k)\}$  bereits zur Impulsantwort eines Entzerrungsfilters. Dieser ist meist akausal, da die Positionierung der beiden Mikrofonkapseln am selben Raumpunkt nicht trivial ist. Somit tritt ein zeitlicher Versatz zwischen den beiden Messsignalen auf. Zur Bestimmung eines kausalen, minimalphasigen Filters wird nur der Betragsfrequenzgang des Quotienten  $D(k)$  verwendet. Da die Phaseninformation verloren ist, bedient man sich der cepstralen Domäne, um ein minimalphasiges System zum errechneten Amplitudengang zu bestimmen.

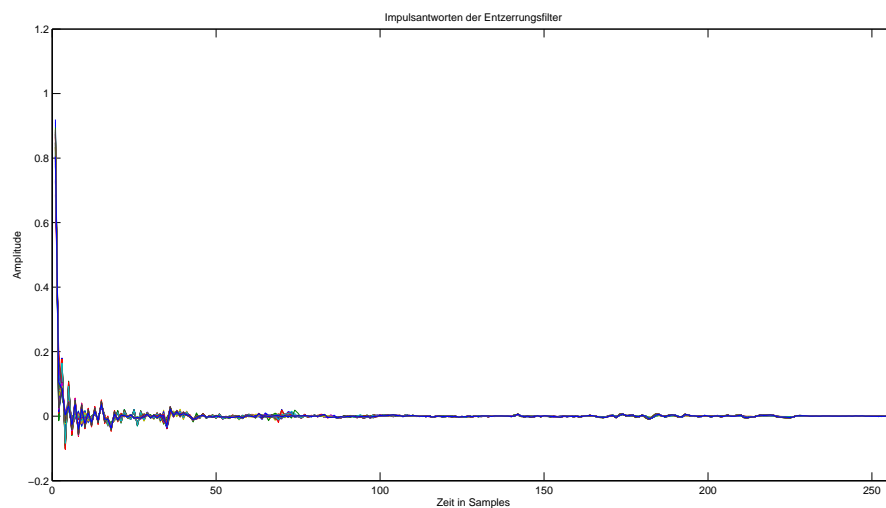
### 3.3.2 Cepstrum

Das Cepstrum ist die inverse Fourier-Transformierte des logarithmierten Spektrums. Eine Eigenschaft des Cepstrums ist, minimalphasige Komponenten von maximalphasigen Komponenten zu trennen. Durch Unterdrückung der maximalphasigen Cepstralanteile (akausal) und durch Rücktransformation der kausalen (minimalphasigen) Cepstralanteile in den Zeitbereich ist der Phasengang des minimalphasigen Systems eindeutig bestimmt. Nach Pei und Lin [PL06] errechnet sich der Phasengang mit der minimalphasigen Eigenschaft zu dem gefensterten Cepstrum des Betragsfrequenzgangs. Dieser ist im Allgemeinen gerade symmetrisch und reellwertig

$$h(n) = h_e(n) \cdot u_+(n) \quad (3.13)$$

$$u_+(n) = \begin{cases} 0 & n < 0 \\ 1 & n = 0 \\ 2 & n > 0 \end{cases} \quad (3.14)$$

Die somit erhaltenen Zeitsignale sind die Impulsantworten der Entzerrungsfiler sind in Abbildung 3.8 dargestellt.



**Abbildung 3.8:** Impulsantworten der Entzerrungsfiler

# Kapitel 4

## Bestimmung der Mikrofonpositionen

Die exakte Bestimmung der Mikrofonpositionen des Arrays ist von essentieller Bedeutung für die weitere Signalverarbeitung. Die naheliegendste Methode wäre, ein Maßband zu nehmen und die einzelnen Positionen von Hand zu vermessen. Diese Methode ist zeitaufwändig, im besonderen wenn unterschiedliche Mikrofonanordnungen untersucht werden, weshalb eine andere Lösung gesucht wird.

Die nahe liegendste Methode ist die Positionsbestimmung über die Auswertung der akustischen Flugzeiten mittels Time Delay Estimation (TDE). Für diese Methode benötigt es eine Referenz. Zur Bestimmung der drei Freiheitsgrade der Positionen im dreidimensionalen Raum muss das Referenzsystem mindestens einen Freiheitsgrad mehr aufweisen. Dazu wurde ein Tetraeder (A.7) konstruiert, an dessen Eckpunkten vier Lautsprecher (A.8) angebracht sind. Zusätzlich wurde ein fünfter, verschiebbarer, Lautsprecher an einer Kante des Tetraeders angebracht, um die Messungen zu verbessern. Die Positionen der montierten Lautsprecher am Tetraeder wurden mit dem *Vicon Motion Capture-System* im Cube des IEM vermessen. Die somit erhaltenen kartesischen Lautsprecherkoordinaten bilden das Referenzsystem für die akustische Vermessung der Mikrofonpositionen. Die angewandten Verfahren zur Bestimmung der Flugzeiten und die Auswertung der Daten sind in den folgenden Abschnitten ausgeführt.



**Abbildung 4.1:** Aufbau der Messeinrichtung im Cube: im Vordergrund die tetraederförmige Lautsprecheranordnung, im Hintergrund das Mikrofonarray

## 4.1 Time Delay Estimation

Die Kreuzkorrelation [HBC06] eignet sich zur Berechnung des zeitlichen Versatzes  $\tau$  zwischen zwei Signalen  $s_1$ , die von der gleichen Quelle stammen und an unterschiedlichen Orten aufgezeichnet werden. Mathematisch lässt sich der Vorgang wie folgt notieren:

$$x_1 = s_1(t) + n_1(t) \quad (4.1)$$

$$x_2 = s_1(t - D) + n_2(t), \quad (4.2)$$

wobei  $D$  die zeitliche Verzögerung,  $n_1(t)$  und  $n_2(t)$  unkorreliertes Rauschen mit den Signalen  $s_1(t)$  und  $s_2(t)$  repräsentieren. Die Kreuzkorrelationsfunktion ist definiert als

$$R_{x_1 x_2}(\tau) = \int_{-\infty}^{+\infty} x_1(t) \cdot x_2(t - \tau) dt. \quad (4.3)$$

Die Kreuzkorrelationsfunktion weist den größten Peak genau dann auf, wenn beide Signale am meisten Ähnlichkeit miteinander aufweisen. Die Fourier-Transformierte

von  $R_{x_1x_2}(\tau)$  wird als die Kreuz-Leistungsdichtespektrum-Funktion bezeichnet

$$G_{x_1x_2}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} R_{x_1x_2}(\tau) e^{j\omega t} dt. \quad (4.4)$$

Aufgrund der Eigenschaften der Fourier-Transformation entspricht die Faltung im Spektralbereich einer Multiplikation. Daraus folgt

$$G_{x_1x_2}(\omega) = E\{X_1(\omega) \cdot X_2^*(\omega)\} \quad (4.5)$$

$X_2^*$  steht für die komplex konjugierte Fourier-Transformierte von  $x_2(t)$  und  $E$  für den Erwartungswert. Beschreibt das Eingangssignal  $s_1$  einen stationären Zufallsprozess, und wird Ergodizität angenommen (Scharmittelwert=Zeitmittelwert), dann ist der Erwartungswert der Schätzgröße ein Zeitmittel, das in Ermangelung von Zeit über einen begrenzten Zeitraum erfolgt. Bei deterministischen Signalen, wie sie in der praktischen Ausführung dieser Arbeit verwendet wurden, ist diese Betrachtung überflüssig.

Die verallgemeinerte Kreuzkorrelationsfunktion (GCC) [KC76] fügt bewusst als Zusatz zwei Filter für die beiden Eingangssignale  $x_1$  und  $x_2$  mit  $H_1(\omega)$  und  $H_2(\omega)$  ein. Die Idee dahinter ist, dass beide Systeme weiß gefärbt werden, was bedeutet, dass die Bandbreite erhöht wird. Das führt zu einem klar ausgeprägten Peak im Korrelat. Wäre die Bandbreite der beiden Signale unendlich, würde sich im Korrelat ein idealer Dirac Delta ergeben

$$R_{x_1x_2}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} H_1(\omega) H_2^*(\omega) G_{x_1x_2}(\omega) e^{j\omega t} dt. \quad (4.6)$$

Die Filter  $H_1(\omega)$  und  $H_2(\omega)$  werden unter der frequenzabhängigen Gewichtsfunktion  $\Psi(\omega)$  zusammengefasst:

$$\Psi_{12}(\omega) = H_1(\omega) H_2^*(\omega) \quad (4.7)$$

$$R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Psi_{12}(\omega) G_{x_1x_2}(\omega) e^{j\omega t} dt. \quad (4.8)$$

Der zeitliche Versatz  $\tau$  zwischen den beiden Signalen  $x_1$  und  $x_2$  ist beschrieben durch

das Maximum von  $R_{x_1x_2}(\tau)$

$$\tau = \underset{\tau}{\operatorname{argmax}} R(\tau). \quad (4.9)$$

In der Theorie ist der zeitliche Versatz exakt beschrieben, allerdings ist der Sachverhalt in realen akustischen Umgebungen ein anderer. Die Standard Kreuzkorrelation kann mehrere lokale Maxima aufweisen, dadurch wird die Detektierung der Flugzeit des Schalls schwierig bis unmöglich. Lokale Maxima können entstehen durch erste Reflexionen und Umgebungsgeräusche sowie durch die Breitbandigkeit der übertragenen Signale. Diese erzeugen längere, eventuell periodische Faltungen. Eine Verbesserung schafft die Verwendung einer frequenzabhängigen Gewichtsfunktion, der Phasentransformation.

#### 4.1.1 Phasentransformation (PHAT)

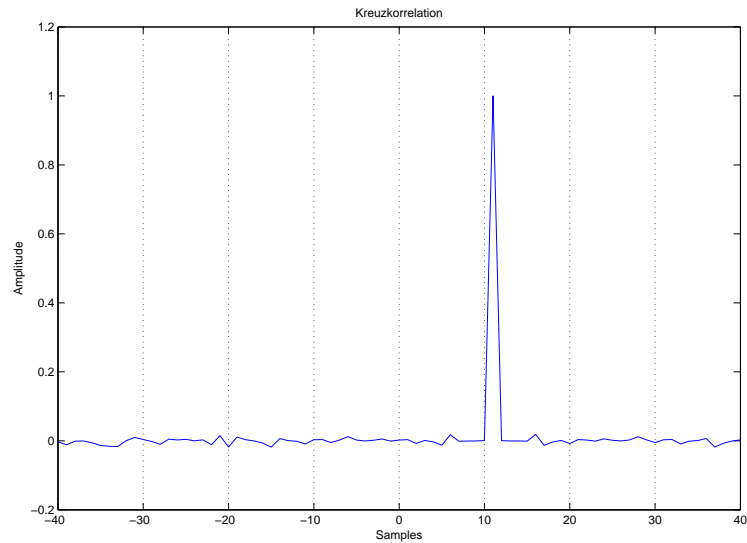
Unter mehreren Gewichtsfunktionen kristallisierte sich die Phasentransformation laut Literatur [KC76, OS96] für reale akustische Umgebungen als geeignet heraus:

$$\Psi_{12}(\omega) = \frac{1}{|X_1(\omega) \cdot X_2^*(\omega)|}. \quad (4.10)$$

Die Phasentransformation kann verstanden werden als ein *pre-whitening* der übertragenen Signale. Dabei wird jede Phasenkomponente des Kreuz-Spektrums gleich gewichtet. Beim dazugehörigen Delay tritt in der Phat ein besonders gut ausgeprägter Peak auf. Die Effizienz dieser Gewichtung hängt dennoch sehr stark von der Nachhallzeit sowie vom Signal-Rausch-Abstand und der Signalbandbreite ab. Es gibt auch Ansätze, die Phasentransformation weiter zu verbessern [BS97] [BW01]. Hierbei wird versucht, die Phase der Phat mittels least-Squares-Ansatz zu optimieren, siehe 4.2.

#### 4.1.2 Experimente der Kreuzkorrelation und Phat

Die Performance der Standard Kreuzkorrelation und der Phat werden hier gegenübergestellt. Als Erstes wird ein Beispiel angeführt unter idealen Voraussetzungen. In Abbildung 4.2 wird die Standard Kreuzkorrelation von weißem Rauschen bei



**Abbildung 4.2:** Standard Kreuzkorrelation von weißem Rauschen mit einer Verzögerung von 11 Samples

$f_s = 6000[Hz]$  und künstlich verzögert um 11 Samples dargestellt. Als Zweites folgt ein Beispiel unter realen Bedingungen. Dabei wird klar, dass die Phasentransformation alles nicht Relevante glättet und somit viel schärfer abbildet als die Standard Kreuzkorrelation.

In den Abbildungen 4.4 und 4.5 sind die Standard Kreuzkorrelation und Phat in realer akustischer Umgebung dargestellt, wobei die ersten Reflexionen aus der Impulsantwort herausgeschnitten sind. Die Faltung der beiden Signale - Referenzsignal mit empfangenem Signal - weist am Anfang einen geradlinigen Verlauf auf (=Null). Diese ist die Folge vom Referenzsignal, da dieser Impuls beinahe ideal ist.

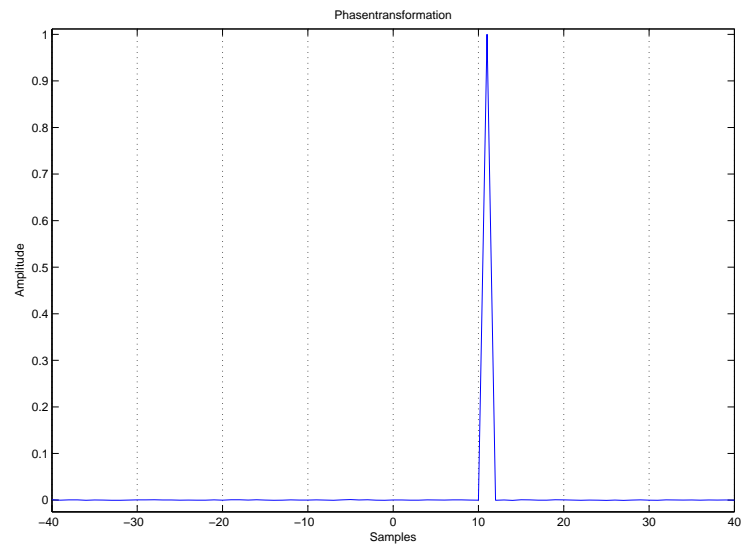


Abbildung 4.3: Phat von weißem Rauschen mit einer Verzögerung von 11 Samples

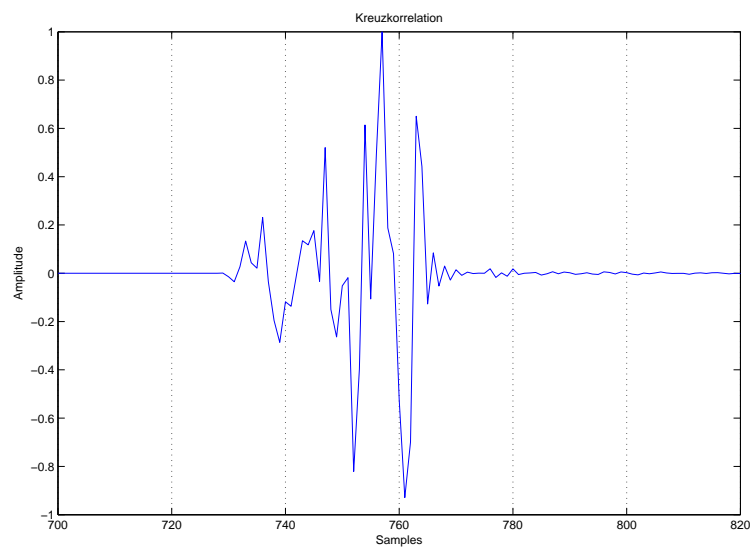
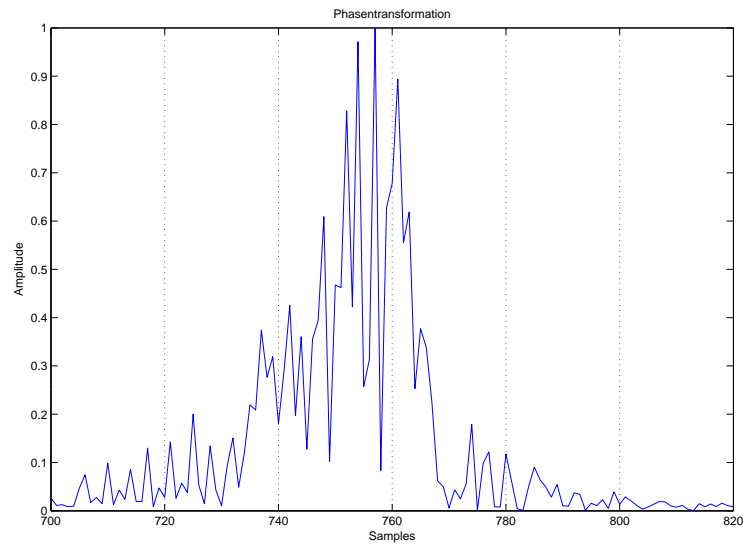


Abbildung 4.4: Standard Kreuzkorrelation der Impulsantwort eines Behringer EMC8000 mit einem Dirac-Impuls.





**Abbildung 4.5:** Phat der Impulsantwort eines Behringer EMC8000 mit einem Dirac-Impuls.

Das Pre-whitening der Phasentransformation ist in Abbildung 4.5 klar zu erkennen. Diese Eigenschaft der Phasentransformation kann das Ergebnis des Korrelats auch verschlechtern. Vergleicht man die Abbildung 4.4 mit 4.5 fällt auf, dass die Amplituden der Phat grössere Werte haben als jene in der Standard Kreuzkorrelation.

### 4.1.3 Parabolische Interpolation

Die Genauigkeit der TDE hängt natürlich von der Abtastfrequenz ab, denn wenn zum Beispiel  $f_s = 44100 [Hz]$  und  $c(25C) = 345 [m/s]$  ist, entspricht das einer räumlichen Abtastung des Schallfeldes von  $c/f_s = 0.72 [mm]$ . Eine wesentliche Verbesserung stellt die parabolische Interpolation dar [LT99]. Dabei wird eine Parabel in den maximalen Peak und dessen benachbarten Samples gelegt, siehe Abbildung 4.6. Das Maximum der Parabel gibt den neuen Verzögerungswert an, der den vorangegangenen Fehler um ein Vielfaches minimiert.

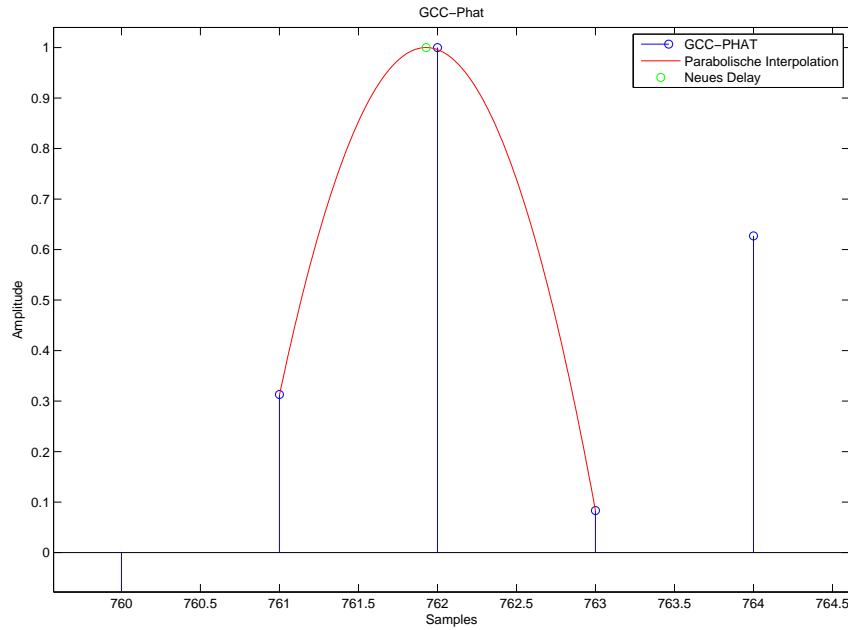


Abbildung 4.6: Parabolische Interpolation

## 4.2 Lineare Regression der Phase

Die Verzögerungszeit lässt sich auch über die Steigung der Phase bestimmen. Für einen idealen Impuls verläuft die Phase immer linear, und die Steigung ist direkt verknüpft mit der Position des Impulses, bezogen auf den Signalbeginn. Die lineare Regression stellt allgemein formuliert einen linearen Zusammenhang zwischen unabhängigen, stochastisch verteilten Messdaten und den Messstellen her. Bezogen auf die Phase eines realen Signales, mittelt die lineare Regression zwischen dem linearen Anteil und dem minimalphasigen Anteil des Phasengangs. Der lineare Anteil beschreibt die Laufzeit des Schalls in der Luft. Der minimalphasige Anteil wird durch das elektromechanische System verursacht und ist stets gekrümmt. Die lineare Regression kann, wie folgt, in vektorieller Schreibweise beschrieben werden:

$$\Delta N \cdot \vec{\omega} = -\vec{\varphi}, \quad (4.11)$$

wobei die Verzögerungszeit  $\Delta N$  der Faktor für die Steigung,  $\vec{\varphi}$  der Vektor des Phasengangs und  $\vec{\omega}$  der Vektor der Stützstellen auf der Frequenzachse ist,

$$\Delta N = \frac{\vec{\omega}^T \cdot \vec{\varphi}}{\vec{\omega}^T \cdot \vec{\omega}}. \quad (4.12)$$

Die Verzögerung des Signals berechnet sich folgendermaßen:

$$\tau = \frac{\Delta N}{f_s}, \quad (4.13)$$

dabei stellt die Normierung auf  $f_s$  den Bezug zur Abtastfrequenz her.

### 4.3 Gruppenlaufzeit

Ein weitere Möglichkeit, die Laufzeit zu messen, stellt die Auswertung der Gruppenlaufzeit dar. Durch Mittelung über die Bandbreite, in der Signale auftreten, ergibt sich ein mittlerer Verzögerungswert, welcher der Laufzeit entspricht. Die Gruppenlaufzeit ist definiert als die negative Ableitung der Phase nach der Frequenz  $-\frac{\partial \phi(\theta)}{\partial \theta}$ . Die Ableitung von zeitdiskreten Signalen kann über die Fourier-Transformation realisiert werden [FH02]. Dazu sind im Vorfeld die beiden Beziehungen nützlich:

$$H(e^{j\theta}) = \sum_{-\infty}^{\infty} h[n]e^{-j\theta n} \quad (4.14)$$

$$\frac{\partial H(e^{j\theta})}{\partial \theta} = \sum_{-\infty}^{\infty} (-jn) \cdot h[n]e^{-j\theta n}, \quad (4.15)$$

Aus dem Logarithmus einer komplexen Zahl  $\ln(r \cdot e^{j\phi}) = \ln(r) + i(\phi + 2\pi k)$  ergibt sich die Phase als Imaginärteil. Damit wird die gesuchte Ableitung zu

$$\phi(\theta) = \Im \{ \ln(H(e^{j\theta})) \} \quad (4.16)$$

$$\frac{\partial \phi(\theta)}{\partial \theta} = -\Im \left\{ \frac{1}{H(e^{j\theta})} \cdot \left[ \frac{\partial H(e^{j\theta})}{\partial \theta} \right] \right\}. \quad (4.17)$$

Die Gruppenlaufzeit  $gd$ , in Beziehung gesetzt zur zeitdiskreten Fourier-Transformation, lautet,

$$gd(\theta) = -\Im \left\{ \frac{\sum_{-\infty}^{\infty} (-jn) \cdot h[n] e^{-j\theta n}}{\sum_{-\infty}^{\infty} h[n] e^{-j\theta n}} \right\} = \Re \left\{ \frac{\sum_{-\infty}^{\infty} (n) \cdot h[n] e^{-j\theta n}}{\sum_{-\infty}^{\infty} h[n] e^{-j\theta n}} \right\}. \quad (4.18)$$

Der Mittelwert über die Bandbreite, welche die Signale beinhalten, ergibt die gesuchte Laufzeit

$$\tau_b(\theta) = \frac{1}{\Delta\theta} \int_{\theta}^{\Delta\theta} gd(\theta) d\theta. \quad (4.19)$$

## 4.4 Piezo-Hochtonhorn

Die Wahl eines Piezo-Lautsprechers zur Vermessung der Positionen ist deshalb von Vorteil, weil die Gruppenlaufzeit gering ausfällt, und das akustische Zentrum eine geringe Fluktuation aufweist. Allerdings verursacht das Exponentialhorn eine Verschlechterung der Gruppenlaufzeit, weil das Übertragungsverhalten eines Horns dem eines Hochpassfilters hoher Ordnung gleichkommt. Je steiler der Anstieg der Flanke dieses Übertragungsverhaltens ist, desto größere Werte nimmt die Gruppenlaufzeit an, siehe Abbildung 4.10. Der Frequenzgang und die Richtcharakteristik des Lautsprechers sind in Abbildung 4.7 und 4.8 dargestellt. Der Öffnungswinkel beträgt bei 8 kHz ca. 60°. Diese Darstellung stellt die Mittelung über die Oktavbänder mit den abgebildeten Mittenfrequenzen dar.

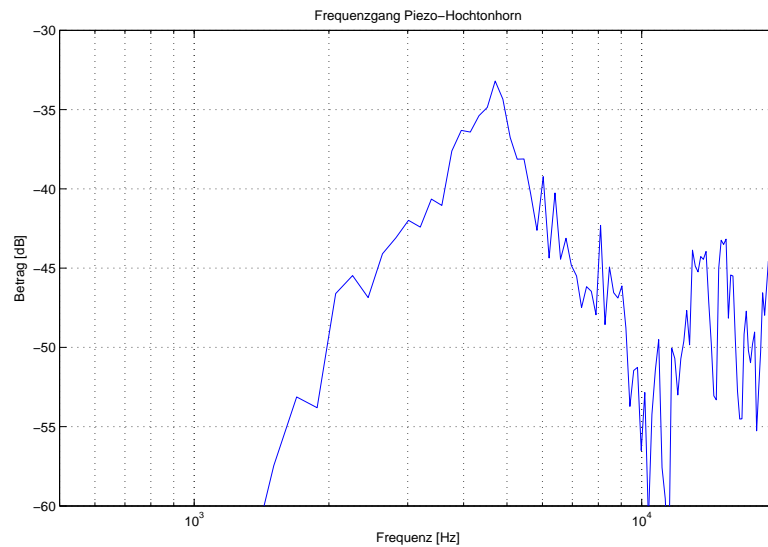


Abbildung 4.7: Frequenzgang des Piezo-Hochtonhorns

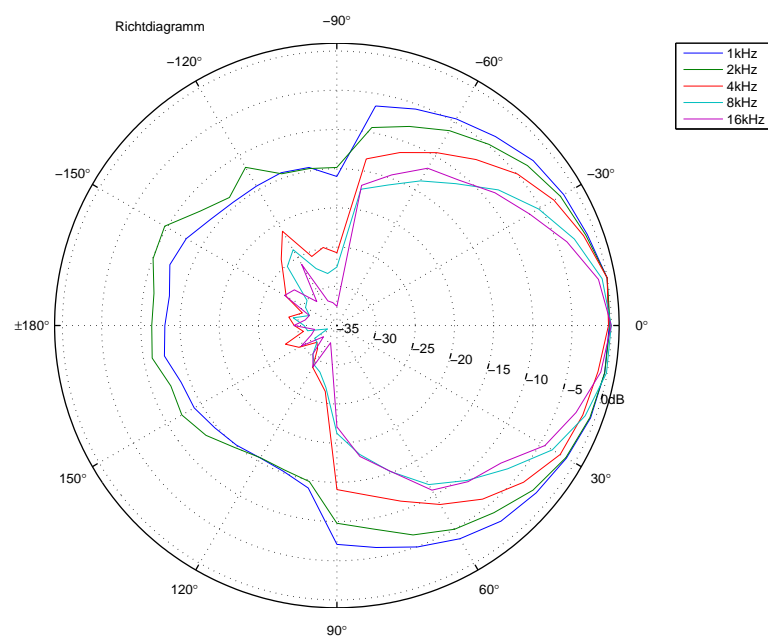
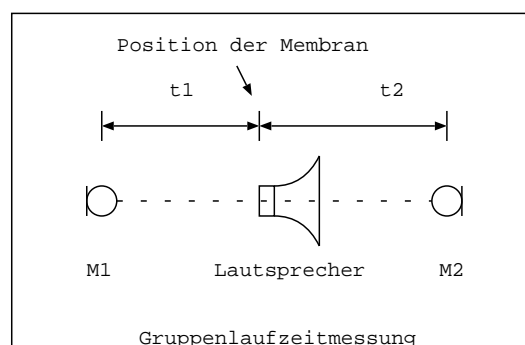


Abbildung 4.8: Richtdiagramm des Hochtonlautsprechers

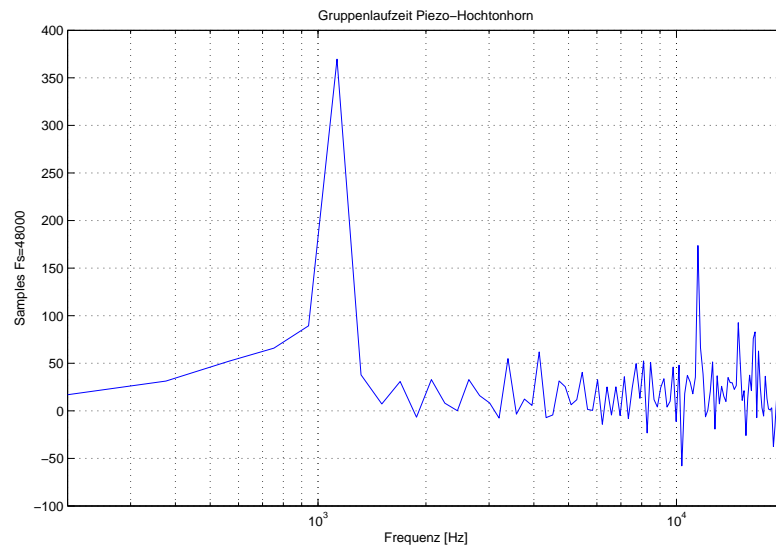
#### 4.4.1 Gruppenlaufzeit und das akustische Zentrum

Wie schon in 4.1 erwähnt, ist die Bandbreite eine bestimmende Größe wie scharf der Peak des Korrelats ausgebildet ist. Jeder Lautsprecher weist aber eine frequenzabhängige Fluktuation des akustischen Zentrums auf. Das heißt, ein breitbandig

abgestrahltes Signal weist schon bei der Ausbreitung im Raum frequenzabhängige Laufzeitunterschiede auf. Unter akustischem Zentrum versteht man jenen Punkt an dem die abgestrahlte Schallwelle im Lautsprecher entsteht, also wo sich der Ausgangspunkt befindet. Die Messung der Gruppenlaufzeit ist schematisch in Abbildung 4.9 dargestellt. Dabei wird auf einer horizontalen Linie ein Mikrofon vor und eines hinter dem Lautsprecher im Fernfeld positioniert. Bei dieser Messung wurden die Mikrofone in einem Abstand von  $0.8 [m]$  aufgestellt, also  $t_1 = t_2$ . Dem Lautsprecher wird ein exponentieller Sweep zugeführt und dieser von den Mikrofonen aufgezeichnet. Dann wird mit der Methode, wie in Abschnitt 4.3 beschrieben, die Gruppenlaufzeiten bestimmt und die beiden Ergebnisse voneinander abgezogen, da sie ja den selben Abstand zum Lautsprecher aufweisen. Das Resultat ist in Abbildung 4.10 dargestellt und beschreibt die Fluktuation des akustischen Zentrums für den verwendeten Piezo-Hochtonlautsprecher. Ab  $1 [kHz]$  tritt ein sehr scharf ausgeprägter Peak auf, der vom Horn abzuleiten ist. Das Exponentialhorn ist nämlich ein akustischer Hochpassfilter hoher Ordnung. Je steiler die Flanke des Hochpasses generell, desto größere Werte nimmt die Gruppenlaufzeit an. Außerdem ist ersichtlich, dass dieser Peak rasch wieder abfällt, und dass sich in den weiteren Bandbreite die Werte bei  $30 \text{ Samples}$  einpendeln. Oberhalb von  $10 [kHz]$  bilden sich nochmals ein Peak aus, der sich von einem ausgeprägtem Einbruch im Frequenzgang des Lautsprechers ableitet. Dieser sogenannte Notch entsteht durch die eingebaute Druckkammer.



**Abbildung 4.9:** Schematische Darstellung zur Gruppenlaufzeitmessung eines Lautsprechers



**Abbildung 4.10:** Fluktuation des akustischen Zentrums, Gruppenlaufzeit des Piezo-Hochtonhorns

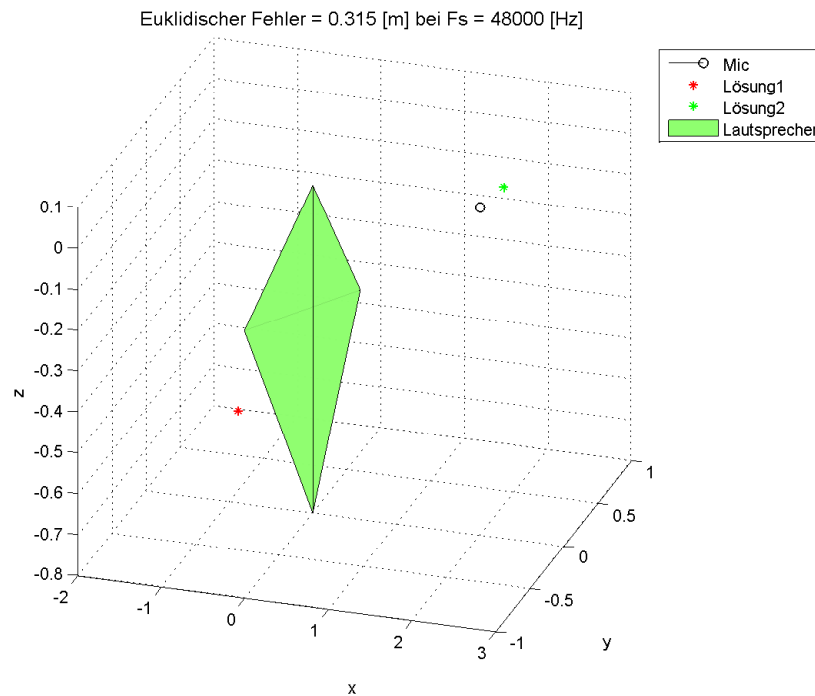
## 4.5 Mathematische Modelle zur Auswertung der Mikrofonpositionen anhand der bestimmten Laufzeiten

Geometrisch kann man sich die detektierten Laufzeiten als Radien von Kugeln vorstellen, die ihren Ursprung an der jeweiligen Lautsprecherposition haben. Der Schnittpunkt aller fünf Kugeln gibt die Position des jeweiligen Mikrofons wieder. Im Folgenden sind zwei Verfahren vorgestellt, die diese Position bestimmen.

### 4.5.1 Lösung der quadratischen Gleichung

Der folgende Lösungsweg ist von S.S.Reddi [Red93] übernommen, und im weiteren nur das Ergebnis vorgestellt (ohne Herleitung). Ausgehend von mindestens vier bekannten Schallquellen im  $n$ -dimensionalen Raum  $R_{n+1}$ , und vom gesuchten Abstand  $R_i$  des Mikrofons, mit der Voraussetzung, dass die Laufzeiten ermittelt sind, lässt sich die Lösung in Matrixschreibweise folgendermaßen anschreiben:

$$R_i^2(CC^T - 1) + R_i(C^TY + Y^TC) + Y^TY = 0, \quad (4.20)$$



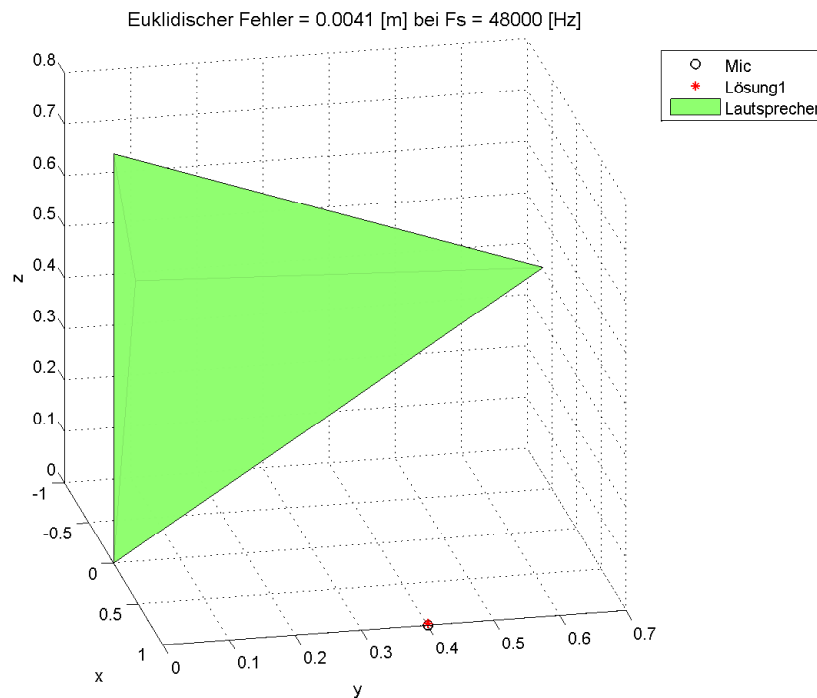
**Abbildung 4.11:** Performance der Methode nach Reddi [Red93]. Die Lautsprecherpositionen sind in den Ecken des Tetraeders. In diesem Fall wird die 1. Lösung um den Tetraeder gespiegelt. Die beiden Lösungen können aber auch sehr eng nebeneinander zu liegen kommen.

wobei  $C$  und  $Y$  Vektoren sind, mit den Koordinaten der einzelnen Quellen und den bekannten Laufzeiten. Durch Lösen dieser quadratischen Gleichung ergeben sich die Koordinaten für die gesuchte Position  $V$  zu

$$V = R_t C + Y. \quad (4.21)$$

Das Lösen der quadratischen Gleichung ergibt immer zwei Lösungen, allerdings stimmt nur eine für die gesuchten Koordinaten. Damit stellt sich ein neues Problem, denn die Wahl der richtigen Lösung kann in manchen Fällen nicht einmal visuell erfolgen, siehe [SGW07]. In Abbildung 4.11 ist das Ergebnis der Simulation dargestellt. Dabei wurden die Lautsprecherpositionen definiert und die jeweiligen Strecken zum Mikrofon auf die Abtastfrequenz bezogen, und die somit erhaltenen Laufzeiten wurden der quadratischen Gleichung übergeben. Der resultierende euklidische Fehler beträgt  $0.34 [m]$ , somit ist die Performance nicht sehr gut. Die selben Daten wurden auch dem Nelder-Mead Simplex-Algorithmus übergeben. Der eukli-





**Abbildung 4.12:** Performance des Nelder-Mead Simplex-Algorithmus [LRWW98]. Die Lautsprecherpositionen sind in den Ecken des Tetraeders.

dische Fehler ist in Abbildung 4.12 dargestellt und beträgt  $0.004 [m]$ . .

### 4.5.2 Der Nelder-Mead Simplex-Algorithmus

Der Nelder-Mead Simplex-Algorithmus [LRWW98], [TS05] zählt zu den Downhill-Simplex-Verfahren und ist ein Suchverfahren für nichtlineare Funktionen mit  $n$  Parametern. Die Besonderheit dieses Verfahrens ist, dass es ohne Ableitung der Parameter auskommt. Der Algorithmus startet mit einem Simplex, der das einfachste Volumen in einem  $n$ -Parameterraum darstellt. Im zweidimensionalen Raum entspricht dieses einem Dreieck und im dreidimensionalen Raum einen Tetraeder. Jeder Eckpunkt dieses Dreiecks entspricht einem Satz von Variablen, die in der zu optimierenden Funktion eingesetzt werden und einen Funktionswert wiedergeben. Dies wird für alle  $n + 1$  Eckpunkte durchgeführt. Als nächster Schritt werden die Funktionswerte verglichen und der "schlechteste" verworfen, stattdessen wird ein neuer Punkt nach den Kriterien von Nelder-Mead bestimmt. Der beste Punkt wird immer beibehalten. Diese Iteration konvergiert beinahe linear zum gesuchten Optimum der

Funktion.

Die Vorteile dieses Verfahrens sind, dass keine Ableitung vonnöten ist, somit keine Unstetigkeiten von Funktionen zu beachten sind, und dass sich wiederum daraus eine gewisse Robustheit ergibt. Andererseits ist im Vergleich zu anderen Optimierungsverfahren mit Ableitungen die Konvergenzzeit deutlich länger, und wie bei den meisten Optimierungsverfahren, kann der Algorithmus zu lokalen Nebenminima konvergieren. Weiters können sich auch ein oder mehrere Parameter zu schnell zusammenziehen, was den Iterationsvorgang zu früh stoppen kann.

Die von Nelder-Mead definierten Koeffizienten sind Reflexion  $\rho$ , Expansion  $\chi$ , Kontraktion  $\gamma$  und Reduktion  $\sigma$ , wobei gilt

$$\rho > 0, \quad \chi > 1, \quad \chi > \rho, \quad 0 < \gamma < 1, \quad 0 < \sigma < 1. \quad (4.22)$$

Die von Nelder-Mead vorgeschlagenen Werte, die in Standard Nelder-Mead Simplex Algorithmen verwendet werden, sind

$$\rho = 1, \quad \chi = 2, \quad \gamma = \frac{1}{2}, \quad \sigma = \frac{1}{2}. \quad (4.23)$$

Definitionen für den zweidimensionalen Raum,  $n = 2$ , seien  $x_b$  für den besten Punkt,  $x_z$  für den zweitbesten Punkt und  $x_s$  für den schlechtesten Punkt.  $x_s$  ist jener Punkt, der in jeder Iteration ersetzt wird. Zur Bestimmung eines neuen Punktes wird jeweils der Schwerpunkt  $x_0$  des Simplex ohne den schlechtesten Punkt  $x_s$  bestimmt. Im zweidimensionalen Fall ist dies die Halbierung der Strecke zwischen  $x_b$  und  $x_z$ . Die Eckpunkte des Simplex sind definiert als  $x_k, k = 1, \dots, n + 1$ :

$$x_0 = \frac{1}{n} \sum_{i \neq s} x_i. \quad (4.24)$$

- Der Reflexionspunkt  $x_r$  von  $x_s$

$$x_r = (1 - \rho)x_0 - \rho x_s. \quad (4.25)$$

- Der Expansionspunkt  $x_e$ , ausgehend von  $x_0$  in Richtung von  $x_r$

$$x_e = \chi x_r - (1 - \chi)x_0. \quad (4.26)$$

- Der Kontraktionspunkt  $x_t$ , der die Strecke zwischen  $x_0$  und  $x_s$  teilt,

$$x_t = \gamma x_s - (1 - \gamma)x_0. \quad (4.27)$$

- Der Reduktionspunkt  $x_k$  ersetzt alle Eckpunkte, abhängig von  $x_b$  und  $x_k$

$$x_k = \frac{1}{2}(x_k + x_b). \quad (4.28)$$

Der Algorithmus im Detail:

1. Schritt: Initialisierung eines Simplex mit  $n + 1$  Eckpunkten
2. Schritt: Berechnen der Funktionswerte für Eckpunkte und Klassifizierung nach bester und schlechtester  $f(x_1) \leq f(x_2) \leq \dots \leq f(x_{n+1})$ , daraus folgt  $x_b$ ,  $x_z$ ,  $x_s$  und  $x_0$ .
3. Schritt: Berechnung des Reflexionspunktes  $x_r$ . Wenn gilt  $f(x_r) > f(x_b)$ , gehe zu 4. Schritt, ansonsten 5. Schritt.
4. Schritt: Berechnung des Expansionspunktes  $x_e$ . Wenn  $x_e$  besser als  $x_b$ , ersetze  $x_z$  mit  $x_e$ , ansonsten ersetze  $x_z$  mit  $x_r$  und gehe zu 2. Schritt.
5. Schritt: Wenn  $x_r$  besser oder gleich als  $x_z$ , ersetze  $x_s$  mit  $x_r$  und gehe zu 2. Schritt. Ansonsten gehe zu 6. Schritt.
6. Schritt: Wenn  $x_r$  besser als  $x_s$ , ersetze  $x_s$  mit  $x_r$ .
7. Schritt: Berechnung des Reduktionspunktes  $x_t$ . Wenn  $x_t$  besser als  $x_s$ , ersetze  $x_s$  mit  $x_t$ . Ansonsten werden alle Punkte mit den Reduktionspunkten ersetzt, bis auf  $x_b$ , und gehe zu 2. Schritt.

Der Algorithmus verlagert den Simplex iterativ hin zum Optimum, um welches er sich schließlich zusammenzieht bis das Konvergenzkriterium erfüllt ist.

## 4.6 Die Implementierung

Voraussetzung ist, dass die Koordinaten der Lautsprecher bekannt sind. Der Tetraeder wird möglichst normal in einem "beliebigen" Abstand zum Mikrofonarray aufgestellt. Je größer der Abstand ist, um so schlechter wird das Ergebnis, da Unterschiede der Flugzeiten in der Unendlichkeit gegen Null gehen. Eine zu nahe Positionierung bringt auch keinen Vorteil, da der Öffnungswinkel des Lautsprechers  $30^\circ$  beträgt, und somit der SNR der Signale, die am Rand des Arrays aufgenommen werden, schlechter wird.

Allen Lautsprechern wird separat ein exponentieller Sweep zugespielt, den alle 64 Mikrofone aufzeichnen. Weiters werden diese aufgezeichneten Signale entfaltet und die resultierenden Impulsantworten um den Hauptimpuls auf wenige Samples gefensterter. Dabei bleibt natürlich die Information vom Anfang der Signale bis zum Impuls erhalten. Die Laufzeit wird mit Phat, linearer Regression und Gruppenlaufzeit bestimmt. Für jedes Mikrofon sind nun  $n$  Flugzeiten berechnet, die als Eingangsgrößen dem Nelder-Mead Simplex Algorithmus [LRWW98] mit adequate Startwerten übergeben werden. Der Erwartungswert des quadratischen Fehlers wird minimiert

$$E(\vec{\alpha}(j)) \equiv \sum_{i=1}^I (f(i, \vec{\alpha}(j)) - \hat{t}(i, j))^2, \quad (4.29)$$

wobei  $I$  die Anzahl der Lautsprecher ist,  $j$  die Anzahl der Mikrofone,  $\hat{t}$  die geschätzte Zeit und  $\vec{\alpha}$  der Vektor mit den gesuchten kartesischen Koordinaten der Mikrofone.

$$f(i, \vec{\alpha}(j)) \equiv \frac{\sqrt{(\alpha_x(j) - s_x(i))^2 + (\alpha_y(j) - s_y(i))^2 + (\alpha_z(j) - s_z(i))^2}}{c(\vartheta)} \quad (4.30)$$

Hierbei gibt  $\vec{s}$  die kartesischen Koordinaten der Lautsprecherpositionen und  $c(\vartheta)$  die Schallgeschwindigkeit, abhängig von der Temperatur in Grad Celsius an. Die Minimierung des Erwartungswertes übernimmt die Simplex-Methode:

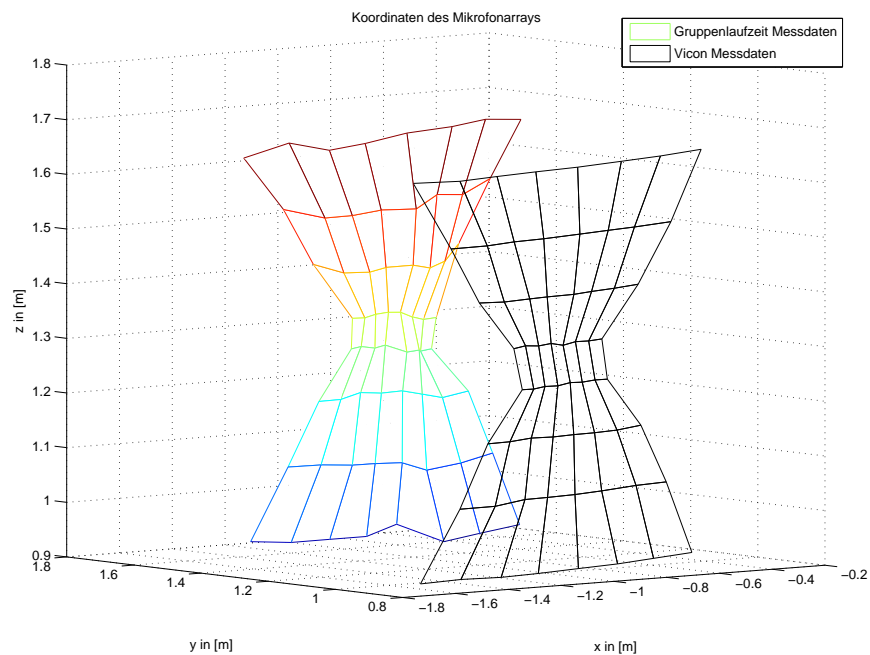
$$\vec{e}(j) \equiv \underset{\vec{\alpha}(j)}{\operatorname{argmin}}(E(\vec{\alpha}(j))). \quad (4.31)$$

Die geschätzten kartesischen Koordinaten der einzelnen Mikrofone des Arrays sind

in  $\vec{e}(j)$  enthalten.

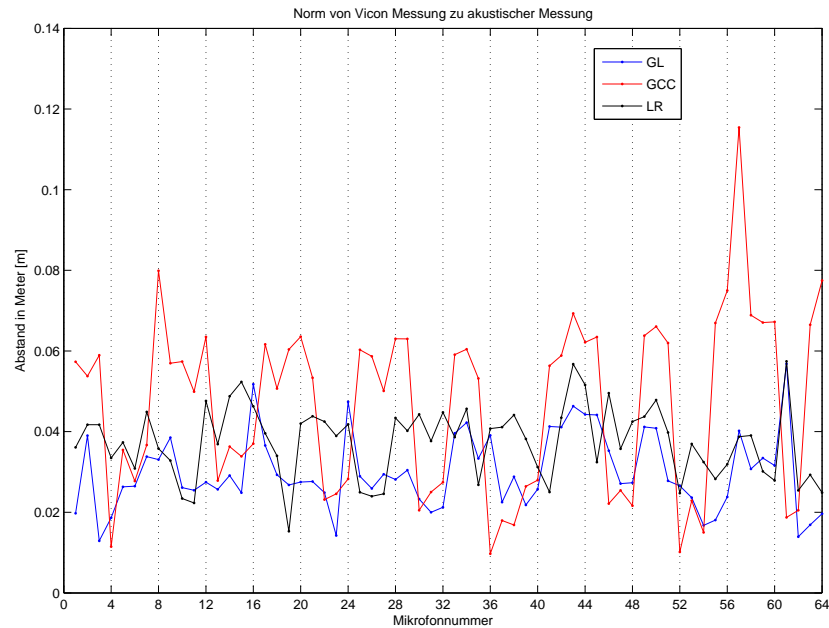
## 4.7 Experimente und Ergebnisse

Die Messergebnisse für die einzelnen Verfahren zur Bestimmung der Laufzeit [Phat, Lineare Regression (LR) und Gruppenlaufzeit (GL)] wurden in Matlab ausgewertet und sind in den nun folgenden Abbildungen dargestellt. In Abbildung 4.13 sind die Messdaten der Gruppenlaufzeit bewusst um einen Offset in y-Richtung verschoben, damit die Darstellung übersichtlicher wird. Messfehler sind klar ersichtlich, wie z.B. dass die horizontale Linienführung im Vergleich zur Vicon-Messung zum Teil stark von einer Geraden abweicht.



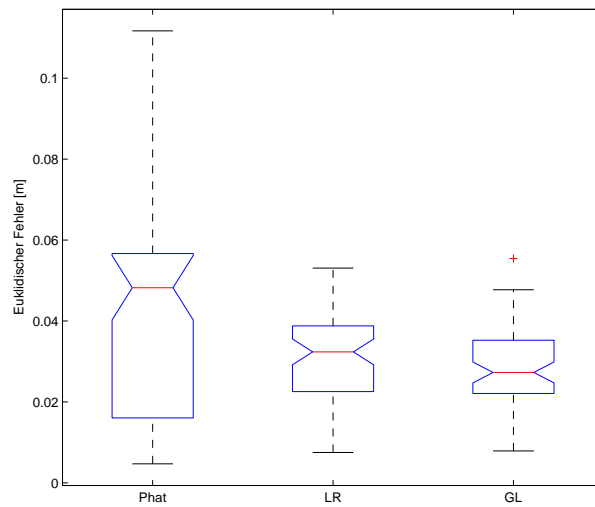
**Abbildung 4.13:** Gegenüberstellung der Messdaten von Vicon und der akustischen Messdaten, bestimmt über die Mittelung der Gruppenlaufzeiten. Die Messdaten der Gruppenlaufzeit sind um einen Offset verschoben, damit die Darstellung übersichtlicher wird.

In Abbildung 4.14 ist die euklidische Norm für das jeweilige Verfahren abgebildet. Die Detektionen der Laufzeiten mit linearer Regression und mit Gruppenlaufzeit weisen in etwa die gleiche Streuung auf. Die Messdaten der Phat weisen eine größere Streuung auf. In den Abbildungen 4.18, 4.17, 4.16 sind die Abweichungen der einzelnen Messungen von der Vicon-Messung dargestellt. Die euklidische Norm zwi-

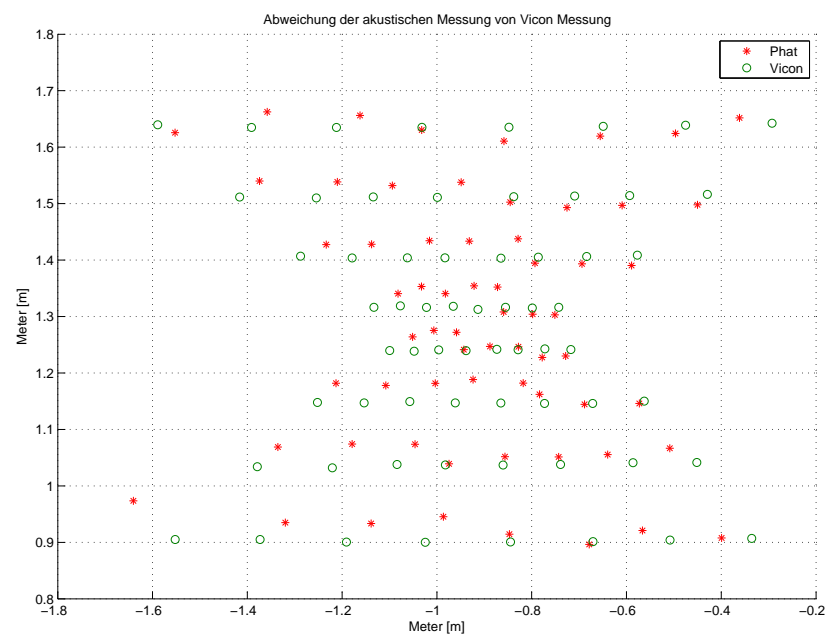


**Abbildung 4.14:** Fehler der euklidischen Norm für das jeweilige Mikrofon in Bezug auf die Vicon-Messdaten. Der Fehler der Phat weist eine größere Streuung auf.

schen wahren Wert (Vicon-Messdaten) und den akustisch vermessenen Daten stellt das Fehlermaß dar. Die einzelnen Verfahren wurden mit der Varianzanalyse (ANOVA) ausgewertet und sind in Abbildung 4.15 dargestellt. Das Konfidenzintervall der Phat-Messdaten ist um einiges größer als die der beiden anderen Messungen, das heißt die Streuung der Messdaten ist größer. Die Auswertung hat überdies ergeben, dass zwischen den Messdaten der linearen Regression und der Gruppenlaufzeit kein signifikanter Unterschied besteht, hingegen diese beiden im Vergleich zur Phat einen signifikanten Unterschied aufweisen.



**Abbildung 4.15:** Varianzanalyse des euklidischen Fehlers der einzelnen Methoden zur Detektion der Laufzeiten. Die Phat weist einen Signifikanten Unterschied zu den beiden anderen Methoden auf. Lineare Regression und Gruppenlaufzeit weisen hingegen keinen signifikanten Unterschied auf.



**Abbildung 4.16:** Abweichung der Phat-Messdaten von Vicon-Messdaten

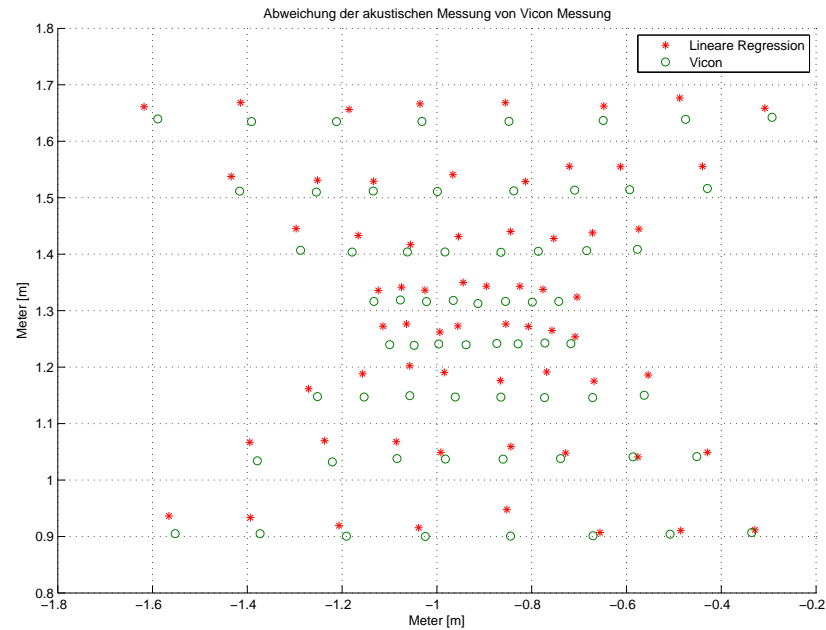


Abbildung 4.17: Abweichung der Linearen-Regression-Messdaten von Vicon-Messdaten

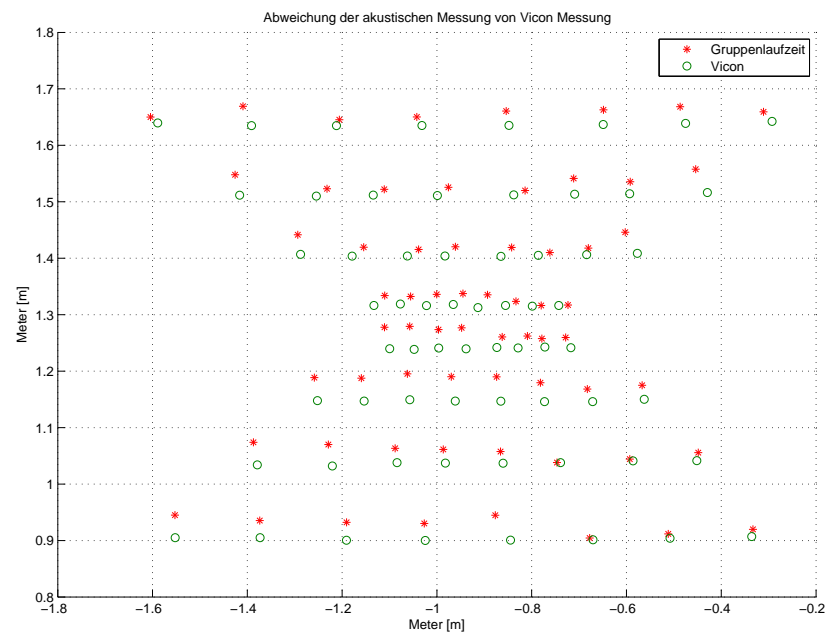


Abbildung 4.18: Abweichung der Gruppenlaufzeit-Messdaten von Vicon-Messdaten

In Abbildung 4.17 und 4.18 fällt ein systematischer Versatz der Messdaten in Richtung der Ordinate auf. Dieser Systematische Fehler konnte bis zum heutigen Tage nicht eindeutig zugeordnet werden. Darüber kann ich nur mutmaßen. Eine



Hypothese dazu wäre: Das Referenzsystem, der selbst konstruierte Tetraeder, weist eine Ungenauigkeit auf. Die Lautsprecher beschreiben nicht ein und die selbe Ebene, sie sind zueinander leicht verdreht. Die Messspitze konnte die einzelnen Membran nicht direkt berühren, da sich im Exponentialhorn Phasenausgleichskeile befinden. Der fehlende Zentimeter wurde ausgeglichen, indem die aufgezeichneten Signale um einen Sample weniger beschnitten wurden. Dieser Ausgleich um einen Sample kommt einer Verschiebung des Messpunktes entlang einer Raumachse, in diesem Fall  $y$ , gleich. Der Membranmittelpunkt befindet sich aber nicht auf dieser gedachten Linie und somit tritt ein Fehler auf der alle Messdaten systematisch verschiebt. Ein möglicher Lösungsvorschlag wäre die Raumkoordinaten der Lautsprecher nicht kartesisch auszuwerten, sondern in Polarkoordinaten.

# Kapitel 5

## Beamformer

### 5.1 Grundlagen

Mikrofonarrays ermöglichen die räumlich Abtastung eines Schallfeldes. Mehrere Sensoren (Mikrofone) sind im Schallfeld nach einer definierten Geometrie angebracht. Beim einfachsten Fall spricht man von einem *Linearray*, bei dem die Sensoren in einer Linie angebracht sind. Räumliches Filtern ist mit der Array-Technik gut realisierbar und wird vor allem dazu genutzt, verschiedene Quellen voneinander zu trennen. Das heißt, es entsteht eine “akustische Kamera”, die den Fokus, den Beam, in eine Raumrichtung richten kann und alle Signale (Störgeräusche), die aus einer anderen Richtung eintreffen, so gut wie möglich unterdrückt. Eine einfache mögliche Implementierung dieses Prinzips stellt der *Delay and Sum Beamformer* dar.

### 5.2 Aperturen

#### 5.2.1 Kontinuierliche Aperturen

Unter Apertur versteht man den Bereich, in dem Wellen entweder (aktiv) aus- oder eintreten (passiv) [Zio95] [Pap06]. Bei einem Mikrofonarray handelt es sich um eine passive Apertur. Die Aperturfunktion lässt sich für das Fernfeld wie folgt beschreiben: In einem infinitesimalen kleinen Volumen  $\delta V$  an einem Ort  $r$  tritt zu einer Zeit  $t$  das Signal  $x(r, t)$  ein. Das Volumen  $\delta V$  an der Stelle  $\vec{r}$  stellt einen linearen

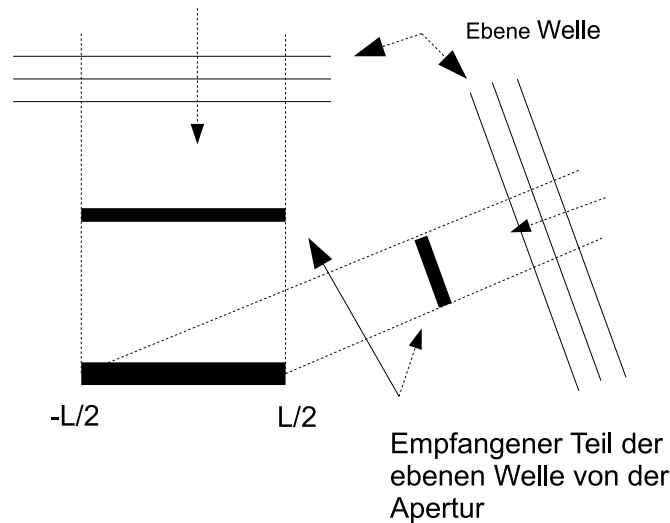
Filter mit der Impulsantwort  $a(\vec{r}, t)$  dar. Somit ist das empfangene Signal  $x_e(\vec{r}, t)$  definiert als Faltung des Signals  $x(\vec{r}, t)$  mit  $a(\vec{r}, t)$

$$x_e(\vec{r}, t) = \int_{-\infty}^{\infty} x(\vec{r}, \tau) a(\vec{r}, t - \tau) \delta\tau \quad (5.1)$$

und im Frequenzbereich

$$X_e(f, \vec{r}) = X(f, \vec{r}) \cdot A(f, \vec{r}). \quad (5.2)$$

$A(f, \vec{r})$  ist die Aperturfunktion/Empfindlichkeitsfunktion, welche die Systemantwort als eine Funktion der räumlichen Position auf der Apertur beschreibt. Die Systemantwort einer Apertur ist direktiv, das heißt, nur ein gewisser Anteil der Schallwelle, je nach Einfallswinkel, trifft auf die Apertur auf, siehe Abbildung 5.1.



**Abbildung 5.1:** Direktive Systemantwort einer Apertur

Die Systemantwort der Apertur als Funktion von der Frequenz wird "Directivity Pattern" oder auch "Beam Pattern" genannt. Das Directivity-Pattern und die Aperturfunktion sind über die Fourier-Transformation verknüpft. Das Fernfeld Directivity-Pattern  $D(f, \alpha)$  mit der Aperturfunktion  $A(f, \vec{r})$  berechnet sich zu

$$D(f, \alpha) = F\{A(f, \vec{r})\} = \int_{-\infty}^{\infty} A(f, \vec{r}) e^{j2\pi r \alpha} \delta r, \quad (5.3)$$

wobei

$$\vec{r} = \begin{Bmatrix} x_a \\ y_a \\ z_a \end{Bmatrix} \text{ und } \alpha = \frac{f \cdot k}{\omega} = \frac{1}{\lambda} [\cos\theta \cos\phi \cos\theta \sin\phi \sin\theta] \quad (5.4)$$

ist. Der Vektor  $\vec{r}$  beschreibt einen Punkt auf der Apertur,  $\alpha$  gibt den Einfallswinkel an und  $\lambda$  die Wellenlänge, womit die Frequenzabhängigkeit aufscheint. Um das Frequenzverhalten zu beschreiben, wird die Apertur mit einer endlichen Länge  $L$  entlang der x-Achse angenommen:

$$\vec{r} = \begin{Bmatrix} x_a \\ 0 \\ 0 \end{Bmatrix} \implies D(f, \alpha) = \int_{-L/2}^{L/2} A(f, \vec{r}) e^{j2\pi x_a \alpha} \delta x_a. \quad (5.5)$$

Für  $\alpha$  ergibt sich

$$\alpha = \frac{\cos\theta \cos\phi}{\lambda} \quad (5.6)$$

und der Directivity Pattern

$$D(f, \theta, \phi) = \int_{-L/2}^{L/2} A(f, x_a) e^{j2\pi x_a \frac{\cos\theta \cos\phi}{\lambda}} \delta x_a. \quad (5.7)$$

Ausgehend davon, dass die Aperturfunktion frequenzunabhängig ist, lässt sich diese folgendermaßen definieren

$$A(x_a) = \text{rect}\left(\frac{x_a}{L}\right) \equiv \begin{cases} 1 & \|x\| \leq \frac{L}{2} \\ 0 & \|x\| > \frac{L}{2} \end{cases}, \quad (5.8)$$

daraus folgt aus Gleichung 5.3

$$D(f, \alpha_x) = F\left\{\text{rect}\left(\frac{x_a}{L}\right)\right\} = L \text{sinc}(\alpha_x L). \quad (5.9)$$

Normalisiert mit  $\frac{1}{L}$ , auf Winkelform gebracht und nur unter Betrachtung des Azi-

miths  $\phi$  ergibt sich Gleichung 5.9 zu

$$D(f, \frac{\pi}{2}, \phi) = \text{sinc}\left(\frac{L \cdot \cos\phi}{\lambda}\right). \quad (5.10)$$

Dies entspricht einem räumlichen Filter. Die Nullstellenbreite der Sinc-Funktion ist frequenzabhängig und ist das Maß für die Beam Width. Je höher die Frequenz, desto mehr Nullstellen gibt es und damit verbunden mehr Nebenkeulen, wie in Abbildung 5.3 zu erkennen ist.

### 5.2.2 Diskretisierte Aperturen

Die Diskretisierung der Apertur auf  $N$  Punkte  $x_n$  lässt sich als räumliche Abtastung der Apertur verstehen. Dabei ist wichtig, dass jeder diskrete Punkt, oder auch Sensor, eine wiederum kontinuierliche Apertur beschreibt. Die gesamte Systemantwort des Sensorarrays ergibt sich aus der Superposition der einzelnen Sensoren. Diese haben eine eigene komplexe Frequenzantwort  $e_n(f, x_n)$  und können auch gewichtet sein:  $w_n(f)$ . Demnach lässt sich die Aperturfunktion wie folgt darstellen

$$A(f, x_a) = \sum_{n=\frac{N-1}{2}}^{\frac{N-1}{2}} w_n(f) \cdot e(f, x_a - x_n), \quad (5.11)$$

und über die räumliche Fourier-Transformation ergibt sich das Directivity Pattern zu

$$D(f, \alpha_x) = \sum_{n=\frac{N-1}{2}}^{\frac{N-1}{2}} w_n(f) \cdot E_n(f, \alpha_x) e^{j2\pi x_n \alpha_x}. \quad (5.12)$$

Nimmt man an, dass alle Sensoren gleich gewichtet sind,  $E_n(f, \alpha_x) = 1$ ,

$$D(f, \alpha_x) = \sum_{n=\frac{N-1}{2}}^{\frac{N-1}{2}} w_n(f) \cdot e^{j2\pi x_n \alpha_x}, \quad (5.13)$$

unter Betrachtung des Azimuths ergibt sich

$$D(f, \phi) = \sum_{n=\frac{N-1}{2}}^{\frac{N-1}{2}} w_n(f) \cdot e^{j2\pi x_n \frac{\cos\phi}{\lambda}}. \quad (5.14)$$

### 5.3 Delay and Sum

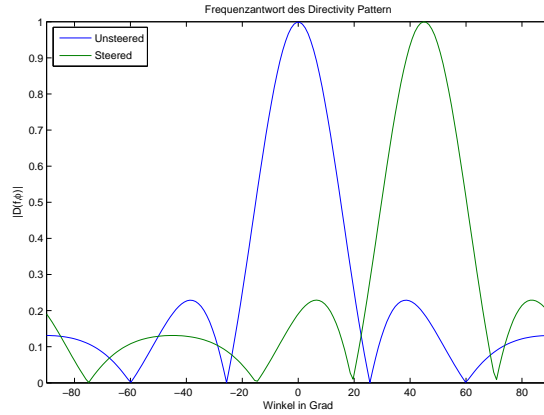
Wie der Name besagt, werden die eintreffenden Signale je nach Richtung, in die der Beam gerichtet werden soll, verzögert (Delay) und dann aufsummiert (Sum). Dadurch entsteht konstruktive und destruktive Interferenz zwischen den einzelnen Sensorsignalen, und so ergibt sich eine starke Richtwirkung und eine gute Unterdrückung für unkorrelierte Störungen. Dabei werden zwei Randbedingungen vorausgesetzt: (1) die Annahme, dass das eintreffende Signal schmalbandig ist, und (2) dass es sich dabei um ebene Wellen handelt (= Fernfeldannahme). Diese Annahme ist in den meisten Fällen gültig. Probleme ergeben sich bei der Annahme von schmalbandigen Signalen, da jedes Array durch die Abstände der Sensoren zueinander nur eine gewisse Bandbreite auflösen kann.

Der minimale Abstand ist ein Parameter für die obere Grenzfrequenz. Diese Begrenzung ist analog zu dem Nyquistkriterium für das Sampletheorem im Zeitbereich. Räumliches Aliasing tritt dann auf, wenn der kleinste Abstand  $\Delta x$  der Sensoren zueinander größer ist als die halbe Wellenlänge der höchsten auftretenden Frequenz.

$$\Delta x = \frac{c}{2 \cdot f_{max} \cdot \sin(\alpha_{max})}, \quad (5.15)$$

mit  $c$  für die Schallgeschwindigkeit,  $\alpha_{max}$  der größtmögliche Einfallswinkel und  $f_{max}$  die höchst auftretende Frequenz. Die Richtcharakteristik des Arrays hängt von der Apertur des Arrays ab, soll heißen, die Richtwirkung hin zu tiefen Frequenzen nimmt ab, die Hauptkeule weitet sich auf, siehe Abbildung 5.3.

Die Gewichtsfunktion  $w_n(f)$  kann nun komplex entworfen werden (Gl. 5.16), wobei mit dem Betrag die Form des Beam Pattern variiert wird und mit unterschiedlichen Phasenverzögerungswerten die Richtung der Hauptkeule gesteuert wird (*Steering*), siehe Abbildung 5.2.



**Abbildung 5.2:** Directivity Pattern mit  $N=8$  Sensoren,  $d=0.1$ [m] und  $f=1$ [kHz]  $\alpha' = 45$

Unter der Voraussetzung

$$w_n(f) = a_n(f)e^{j\varphi(f)} \quad (5.16)$$

ergibt sich für das Directivity Pattern

$$D(f, \alpha_x) = \sum_{n=\frac{N-1}{2}}^{\frac{N-1}{2}} e^{j2\pi x_n \alpha_x + \varphi(f)} \quad (5.17)$$

$$a_n(f) = 1 \quad \varphi(f) = 2\pi x_n \alpha'_x.$$

Der Steeringwinkel ist  $\alpha'$

$$\alpha'_x = \frac{\sin\theta' \cos\phi'}{\lambda} \implies D(f, \alpha') = \sum_{n=\frac{N-1}{2}}^{\frac{N-1}{2}} e^{j\frac{2\pi}{\lambda} x_n (\alpha_x - \alpha'_x)} = D(f, \alpha_x - \alpha'_x). \quad (5.18)$$

Der negative Phasenshift im Frequenzbereich entspricht einer zeitlichen Verzögerung im Zeitbereich. Diese Verzögerung wird nun auf einen Referenzsensor bezogen und ist äquivalent zur Laufzeit einer ebenen Welle zum  $n$ -ten Sensor:

$$\tau_n = \frac{\varphi_n}{2\pi f} = \frac{2\pi x_n \alpha'_x}{2\pi f} = \frac{x_n \cos\phi'}{f \cdot \lambda} = \frac{x_n \cos\phi'}{c}. \quad (5.19)$$

Durch Verzögern der einzelnen Signale und anschließende Addierung erhält man ein räumlich gefiltertes Monosignal, siehe Gl. 5.23. Ausgehend davon, dass die Sensoren

alle im gleichen Abstand angebracht sind, lässt sich das Directivity Pattern mit  $x_n = n \cdot d$  wie folgt darstellen,

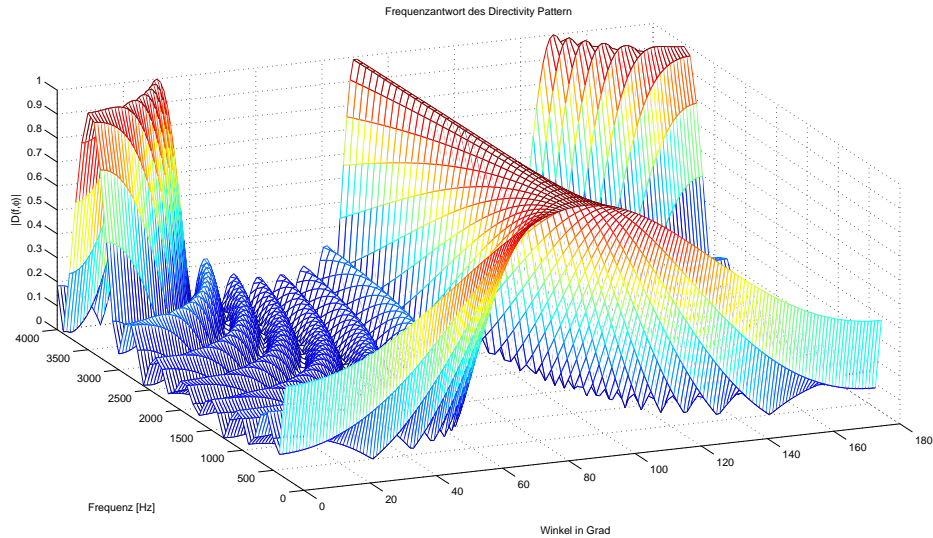
$$D(f, \phi) = \frac{1}{N} \sum_{n=1}^N e^{j \frac{2\pi(n-1)d(\cos\phi - \cos\phi')}{c}} \quad (5.20)$$

$$\tau_n = \frac{\cos\phi'(n-1)d}{c} \quad (5.21)$$

$$Y(f) = \frac{1}{N} \sum_{n=1}^N X(f) e^{j \frac{2\pi}{c} (n-1)d \cos\phi'} \quad (5.22)$$

$$y(t) = \frac{1}{N} \sum_{n=1}^N x_n(t - \tau_n). \quad (5.23)$$

In Abbildung 5.3 ist das Directivity Pattern für die Frequenzen von 300-4000 Hz abgebildet. Gut erkennbar ist das räumliche Aliasing ab 3000 Hz aufwärts.



**Abbildung 5.3:** Directivity Pattern für 8 Sensoren,  $d=0.1$ ,  $f=0.3-4$  kHz

Das in Abbildung 5.3 gezeigte Directivity Pattern bildet das Verhalten eines Line Arrays ab. Für ein planares Array entsteht eine zweidimensionale Sinc-Funktion. Diese Sinc-Funktion tritt auch bei der Nahfeldholographie auf, da es sich dabei, genau wie beim Delay and Sum Beamformer, um eine räumliche Transformation handelt über eine endliche Apertur, siehe Abbildung 6.4.



# Kapitel 6

## Akustische Nahfeldholographie

Akustische Nahfeldholographie [Wil99] stellt ein Analysewerkzeug dar, mit dem die Abstrahleigenschaften einer Schallquelle analysiert werden können. Die Schalldruckverteilung wird in einer Ebene im Raum gemessen, und zwar nahe zur Quelle. Durch 2D Fourier-Transformation wird die Schalldruckverteilung in ebene und evaneszente Wellen zerlegt, wodurch man das Wellenzahlspektrum erhält. Mit dem Verschiebungssatz der Fourier-Transformation läßt sich das Schallfeld an einer anderen Stelle im Raum und zur einer anderen Zeit auswerten. Die zugrunde liegenden Eigenschaften dieser Transformation werden nun im Detail vorgestellt.

### 6.1 Ebene Wellen

Für die Beschreibung der ebenen Welle wird die Wellengleichung in einem homogenen Schallfeld ohne Quellen für den Schalldruck  $p(x, y, z, t)$  gelöst

$$\nabla^2 p = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2}, \quad (6.1)$$

wobei der Operator  $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ ,  $c$  die Schallgeschwindigkeit und  $t$  die Zeit ist. Eine mögliche homogene Lösung für diese Differentialgleichung zweiter Ordnung ist der Produktansatz aus Linearkombination von ebenen Wellen

$$p_h(x, y, z, t) = e^{ik_x x} e^{ik_y y} e^{ik_z z} e^{-i\omega t} = e^{i\vec{k}^t \vec{r} - i\omega t}, \quad (6.2)$$

mit dem Wellenzahlvektor  $\vec{k} = (k_x, k_y, k_z)^t$  und dem Aufpunktvektor  $\vec{r} = (x, y, z)^t$ . Die Wellenzahl  $k$  kann mit unterschiedlichen Größen in Zusammenhang gestellt werden

$$k = \omega/c, \quad (6.3)$$

weilers auch mit der Wellenlänge  $\lambda$

$$k = 2\pi/\lambda. \quad (6.4)$$

Durch Einsetzen des Lösungsansatzes

$$\frac{\partial^2}{\partial x^2} p_h = -k_x^2 \cdot p_h \text{ und } \frac{\partial^2}{\partial t^2} p_h = -\omega^2 \cdot p_h \quad (6.5)$$

erhält man die charakteristische Gleichung

$$k_x^2 + k_y^2 + k_z^2 = \frac{\omega^2}{c^2} = k^2 \quad (6.6)$$

Sind die Komponenten des Wellenzahlvektors rein reellwertig  $k \in \mathbb{R}^3$  und liegen diese auf der von Gleichung 6.6 aufgespannten Kugel, ist dies die Lösung für die ebenen Wellen. Die ebene Welle lässt sich auch mit einem Richtungsvektor  $\vec{s}$  anschreiben  $s \in \mathbb{S}^2 \subset \mathbb{R}^3$  auf der Einheitskugel  $\mathbb{S}^2$  in Richtung  $(\varphi_s, \theta_s)$  und mit der Wellenzahl  $k$  als Betrag.

$$k_s = \begin{pmatrix} k_x \\ k_y \\ k_z \end{pmatrix} = k \cdot \vec{s} = k \cdot \begin{pmatrix} \cos(\varphi_s) \sin(\theta_s) \\ \sin(\varphi_s) \sin(\theta_s) \\ \cos(\theta_s) \end{pmatrix} \quad (6.7)$$

## 6.2 Evaneszente Wellen

Evaneszent ist gleichbedeutend mit "dahinschwindend". Werden für die Einträge des Wellenzahlvektors auch imaginäre Werte zugelassen  $k \in \mathbb{S}^2 \subset \mathbb{C}^3$ , treten evaneszente Wellen auf. Dieser Fall tritt auf, wenn z.B.  $k_x$  oder  $k_y > k$  ist. Gleichung 6.6 lässt sich umformen zu

$$k_z = \pm i \sqrt{-k^2 + k_x^2 + k_y^2} = \pm i k'_z. \quad (6.8)$$

Diese Erweiterung des Raumes auf  $\mathbb{C}$  erfüllt immer noch die Wellengleichung. Für die homogene Lösung der Wellengleichung ergibt sich

$$p = e^{-k'_z z} e^{i(k_x x + k_y y)}, \quad (6.9)$$

wobei  $k_z$  auch negative Werte annehmen kann, was in der Ausbreitung von Wellen zu örtlich exponentiell anklingenden Wellen führt und somit eher rechnerisch interessant ist. Für den physikalisch sinnvolleren Fall,  $k_z \geq 0$  ergibt sich, dass evaneszente Wellen exponentiell in  $z$ -Richtung abklingen. Dieses Phänomen tritt in der Nähe von Grenzflächen auf, wenn sich auf der Grenzfläche fein strukturierte Wellen/Schwingungsstrukturen bei sehr großen Wellenzahlen ausbilden, die in der Luft nicht ausbreitungsfähig sind. In der Optik wird dieser Effekt auch bei der Nahfeldmikroskopie genutzt. In der Akustik tritt dieser Effekt z.B. bei Materialschwingungen auf (biegeweich, schwach gespannte Membran), deren Wellenlängen nicht in das Medium Luft abgestrahlt werden können, da es keine koinzidente (dazu passende) Welle gibt. Im Wellenzahlraum ist das Wertepaar dann außerhalb des von  $k$  aufgespannten Kreises und befindet sich in der Region der evaneszenten Wellen, siehe Abbildung 6.1. Wichtig ist, dass die Lösung der Wellengleichung immer noch gegeben.

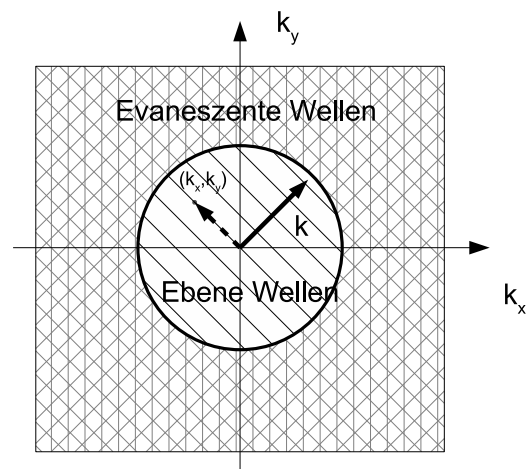


Abbildung 6.1: Wellenzahlraum der ebenen Wellen und der evaneszenten Wellen

### 6.3 Homogene Lösung der Wellengleichung

Die charakteristische Gleichung 6.6 auf Seite 46 zeigt, dass die einzelnen Komponenten des Wellenzahlvektors  $\vec{k} = (k_x, k_y, k_z)^t$  in einem exakten Zusammenhang mit  $\omega$  stehen. Somit kann die homogene Lösung auch vereinfacht angeschrieben werden, mit Dirac Delta [Zot08]

$$p_h(r, \omega) = e^{j(\vec{k}^t \vec{r} - \omega t)} \cdot \delta\left(\frac{\omega^2}{c^2} - \|\vec{k}\|^2\right). \quad (6.10)$$

Im Frequenzbereich läßt sich diese Beziehung noch einfacher notieren:

$$p_h(r) = e^{j\vec{k}^t \vec{r}} \cdot \delta(k - \|\vec{k}\|). \quad (6.11)$$

Ist die Wellengleichung im eingeschwungenen Zustand, kann die Summe der homogenen Lösungen im Frequenzbereich mit dem Integranden  $P(k)$  als Lösung für  $p(r)$

dargestellt werden,

$$p(r) = \int \int \int_{\mathbb{C}^3} P(k) \cdot e^{j\vec{k}^t \vec{r}} \cdot \delta(k - \|\vec{k}\|) dk \quad (6.12)$$

im komplexen dreidimensionalen Raum. Mit der Bedingung, dass  $\forall k \in \mathbb{C}^3 : \|\vec{k}\| = k$ , befindet man sich im Wellenzahlraum (k-space):

$$p(r) = \int \int \int_{\forall k \in \mathbb{C}^3 : \|\vec{k}\|=k} P(k) e^{j\vec{k}^t \vec{r}} dk. \quad (6.13)$$

## 6.4 Wellenzahlspektrum

Aus Gleichung 6.7 auf Seite 46 geht hervor, dass die zwei Variablen  $k_x, k_y$  unabhängig sind und die Variable  $k_z$  abhängig ist, und somit auch die Dimension des Raumes reduziert werden kann. Folglich gilt für  $p(r)$

$$p(\vec{r}) = \frac{1}{4\pi^2} \int dk_x \int_{\mathbb{S}^2 \subset \mathbb{C}^3} dk_y P(k_x, k_y) e^{j(k_x x + k_y y + k_z z)}. \quad (6.14)$$

Die Normierung auf  $\frac{1}{4\pi^2}$  geht aus der zweidimensionalen Fourier-Transformation hervor. Das Wellenzahlspektrum ist die Fourier-Transformation des Schalldrucks über die reellen Koordinaten  $r \in \mathbb{R}^3$

$$P(k_x, k_y) = \int dk_x \int dk_y p(x, y, 0) e^{-j(k_x x + k_y y)} \quad (6.15)$$

## 6.5 Extrapolation des Schallfeldes

Dieses Ergebnis ist das Rückgrat der Fourier-Akustik. Ist  $P(k_x, k_y)$  in der Ebene  $z = 0$  bestimmt, ist es ein Leichtes, mit Gleichung 6.14 das Schallfeld an einer anderen unendlichen Raumebene zu errechnen.

$$P(k_x, k_y, z) = P(k_x, k_y, z') e^{jk_z(z-z')} \quad (6.16)$$

Diese Beziehung zeigt, dass die Amplituden der ebenen Wellen nur eine Phasenänderung von einer Ebene zur anderen vollziehen. Dies gilt auch für evaneszente Wellen

und schreibt sich dann zu,

$$P(k_x, k_y, z) = P(k_x, k_y, z') e^{-|k_z|(z-z')}. \quad (6.17)$$

Gleichung 6.16 ist das Werkzeug der Nahfeldholographie, denn diese ermöglicht, das Schallfeld zurück bis zur Quelle zu rechnen und die Schwingungen der Quellstruktur zu analysieren. Da evaneszente Wellen sehr schnell abklingen, muß die Messung nahe zum Objekt erfolgen, ansonsten würde das Schallfeld nur in ebene Wellen zerlegt und es lägen keine Informationen über evaneszente Wellen vor. Wenn starke evaneszenten Wellen vorliegen, treten numerische Probleme auf, vor allem im Zusammenhang mit Rauschen in den Messwerten, aber auch durch nicht exakt ausgerichtete und kalibrierte Mikrofone.

Über die Eulersche Bewegungsgleichung

$$i\rho_0 \frac{\partial \vec{v}}{\partial t} = -\vec{\nabla} p$$

ist die Schallschnelle mit dem Schalldruck verknüpft. Durch Transformation über die Zeit, und Ableiten der Schnelle  $\vec{v}$  nach der Zeit, siehe 4.15 auf Seite 23, lässt sich die Eulersche Bewegungsgleichung anschreiben zu

$$i\omega\rho_0\vec{v} = \vec{\nabla} p, \quad (6.18)$$

wobei  $\omega = k \cdot c$ ,  $\rho_0$  die Dichte des Mediums Luft bei 20°C angibt. Dadurch ist es auch möglich die Schnelle des Objektes zu einem Zeitpunkt auszuwerten. Dazu wird die Ableitung des zweidimensionalen fouriertransformierten Drucks nur in z-Richtung angeschrieben

$$\mathcal{F}_x \mathcal{F}_y \left[ \frac{\partial p(x, y, z)}{\partial z} \right] = \frac{\partial P(k_x, k_y, z)}{\partial z} = ik_z P(k_x, k_y, z') e^{ik_z(z-z')}. \quad (6.19)$$

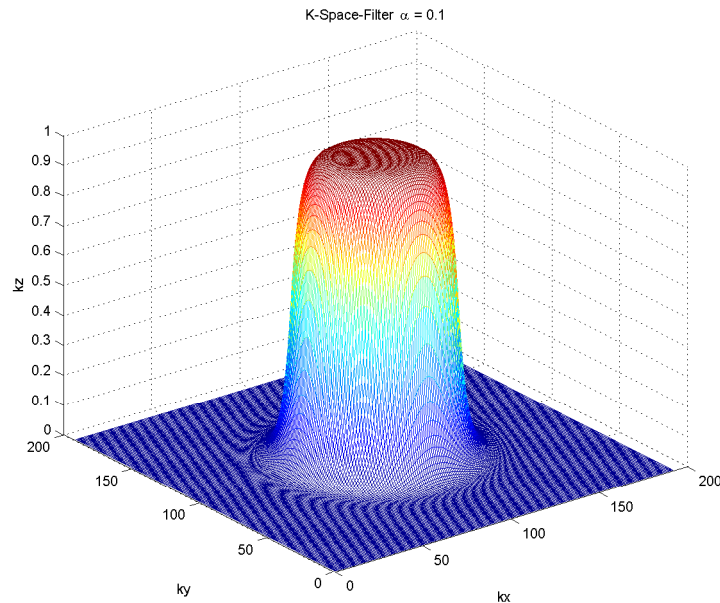
Dies führt zu der wichtigen Beziehung für die Schnelle normal auf die Ebene  $xy$  dargestellt im Wellenzahlraum

$$V_z(k_x, k_y, z) = -\frac{k_z}{\omega\rho_0} P(k_x, k_y) e^{ik_z(z-z')}. \quad (6.20)$$

Bei der Extrapolation ist zu beachten, dass Rauschen, schlechte Positionierung und Nichtkalibrierung der Mikrofone, die Extrapolation instabil wird. Abhilfe schafft Filtern der Spektralanteile mit einem so genannten *k-space filter*. Im Buch von Earl G. Williams [Wil99] wird der Filter mit  $k_c$  (Grenzfrequenz) und  $k_r = \sqrt{k_x^2 + k_y^2}$  wie folgt,

$$\Pi\left(\frac{k_r}{2k_c}\right) = \begin{cases} 1 - \frac{1}{2}e^{-(1-|k_r|/k_c)/\alpha} & |k_r| < k_c \\ \frac{1}{2}e^{1-|k_r|/k_c/\alpha} & |k_r| > k_c \end{cases}, \quad (6.21)$$

wobei  $\alpha$  die Flankensteilheit des Filters angibt. Typisch Werte dafür sind zwischen 0.2–0.05. Die Grenzfrequenz  $k_c$  sollte Werte in der Nähe von  $k$  annehmen um störende numerische Einflüsse zu unterdrücken. In Abbildung 6.2 ist der k-space Filter für  $\alpha = 0.1$  dargestellt. Der Anteil der ebenen Wellen befindet sich im Durchlassbereich, der Anteil der evaneszenten Wellen wird exponentiell abklingend unterdrückt.



**Abbildung 6.2:** K-Space-Filter für  $\alpha = 0.1$  und  $k_c = 1.2 \cdot k_r$ , ebene Wellen befinden sich im Durchlassbereich, evaneszente Wellen werden unterdrückt.

## 6.6 Aliasing

Räumliches Aliasing tritt dann auf, wenn der kleinste Abstand  $\Delta x$  der Sensoren zueinander größer ist als die halbe Wellenlänge der höchsten auftretenden Frequenz,

siehe 5.3.

$$f_{og,x} = \frac{c}{2 \Delta x}. \quad (6.22)$$

Über das Wellenzahlpaar  $k_x, k_y$  lässt sich auch Azimuth und Elevation der eintreffenden Welle bestimmen.

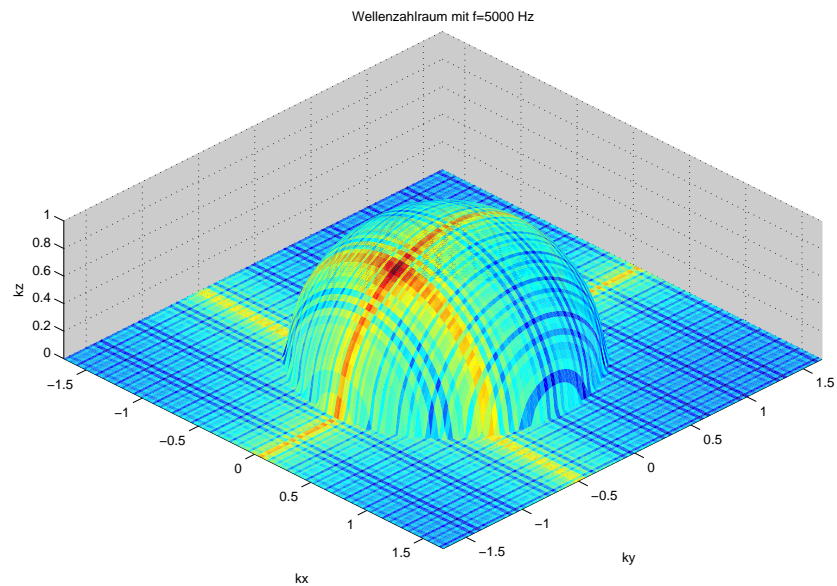
$$\varphi(k_x, k_y) = \arctan\left(\frac{k_y}{k_x}\right) \quad (6.23)$$

$$\theta(k_x, k_y) = \arccos\left(\frac{k_z}{k}\right) \quad (6.24)$$

## 6.7 Simulation und Ergebnisse der Nahfeldholographie

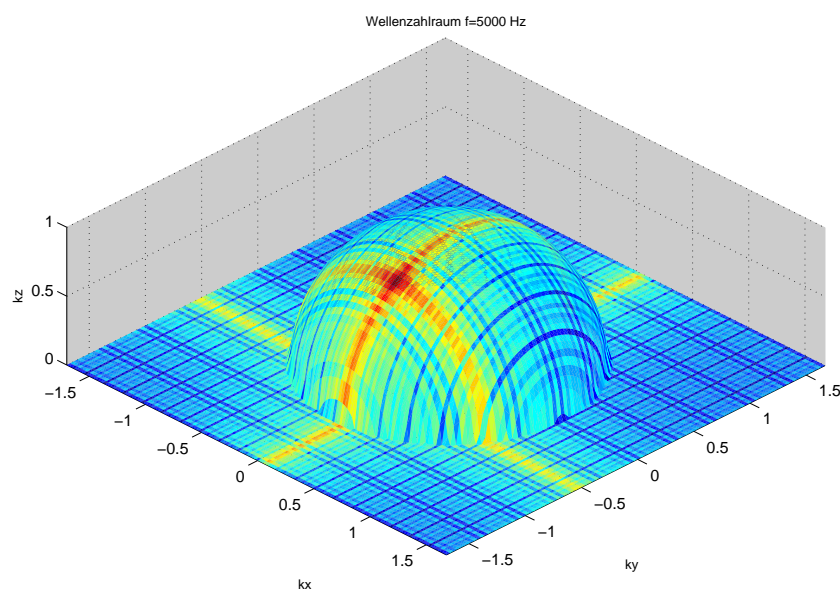
Die Simulation der Nahfeldholographie erfolgte in Matlab. Als diskretes Array wurde eine Fläche entworfen mit 100x50 Punkten. Dies legt die Grenzfrequenz auf 8500[Hz] fest. Als Schallquelle wurde eine ebene Welle sowie eine Punktquelle simuliert. Die Einfallsrichtung der Quelle ist über die Raumwinkel Azimuth und Elevation steuerbar. In Abbildung 6.3 ist der Wellenzahlraum für eine ebene Welle dargestellt mit den Einfallswinkeln Azimuth=274° und Elevation=29.7°. Die dunkelrot eingefärbte Fläche auf der Hemisphäre gibt die Einfallswinkel der ebenen Welle wieder. Die Größe der Fläche gibt die Genauigkeit der Winkelauflösung an. Je größer die Fläche ist, umso ungenauer kann die Einfallsrichtung der Schallwelle angegeben werden. Die beiden Linien, die sich in dieser Fläche kreuzen, repräsentieren die zweidimensionale Sinc-Funktion. Diese Funktion entsteht da die Aperturfunktion endlich ist.





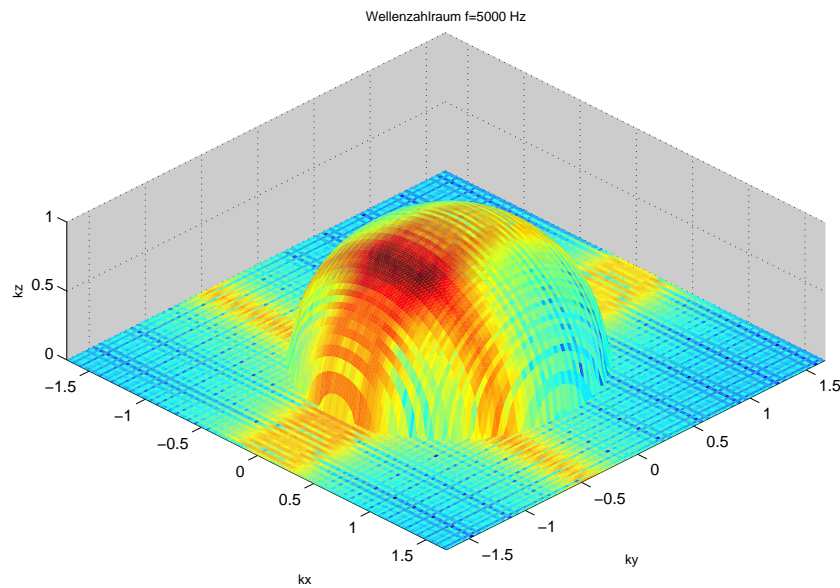
**Abbildung 6.3:** Wellenzahlraum einer ebenen Welle,  $f=5000$  [Hz] Azimuth= $274^\circ$ , Elevation= $29.7^\circ$

In Abbildung 6.4 ist der Wellenzahlraum für eine Punktquelle im Fernfeld dargestellt, mit gleicher Einfallsrichtung wie in Abbildung 6.3. Die Wellenfront ist sehr schwach gekrümmt.



**Abbildung 6.4:** Wellenzahlraum einer Punktquelle im Fernfeld,  $f=5000$  [Hz] Azimuth= $274^\circ$ , Elevation= $29.7^\circ$ , die beiden Linien, die normal zueinander verlaufen, stellen die zweidimensionalen Sinc-Funktion dar. Die Krümmung der Wellenfront ist sehr schwach

Die Genauigkeit der Winkelauflösung ist die gleiche wie für die ebene Welle. Interessant wird es wenn die Punktquelle bis zur Apertur geführt wird. Die Fläche auf der Hemisphäre vergrößert sich und breitet sich über die ganze Hemisphäre aus, wenn der Abstand auf Null reduziert wird. In Abbildung 6.5 ist die Punktquelle sehr nahe zur Apertur.



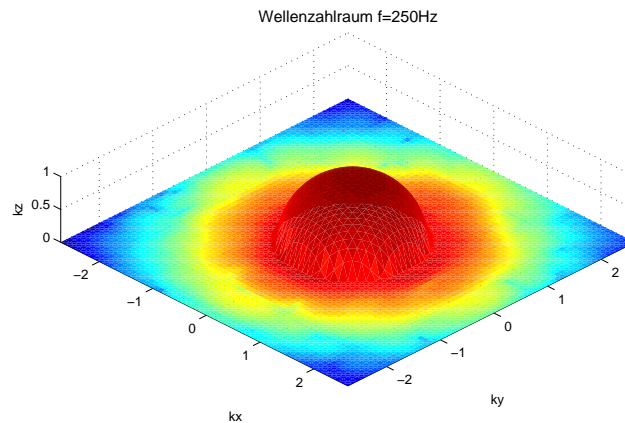
**Abbildung 6.5:** Wellenzahlraum einer Punktquelle im Nahfeld,  $f=5000$  [Hz] Azimuth= $274^\circ$ , Elevation= $29.7^\circ$

Die aufweitung der Fläche kommt daher, dass die eintreffende Wellenfront gekrümmt ist, und sich diese Krümmung nur mit Hinzufügen von ebenen Wellen beschreiben lässt. Dazu kann man sich eine Parabel vorstellen, die mit Tangenten approximiert wird. Das Gleiche passiert bei der 2D Fourier-Transformation: je stärker die eintreffende Welle gekrümmt ist, um so mehr ebene Wellen werden hinzugefügt.

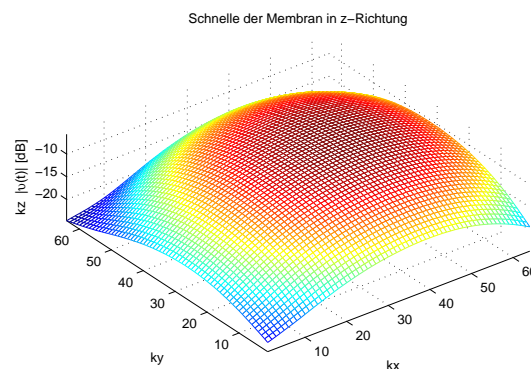
## 6.8 Schnelle und Druckverteilung auf einer Lautsprechermembran

Für diese Messung wurden die Mikrofone äquidistant und planar mit einem Abstand von  $0.1$  [m] angeordnet. Die Mikrofone wurden neu kalibriert, ein Tannoy Lautspre-

cher (System 6 NFM) in einen Abstand zum Mikrofonarray von  $0.1 [m]$  aufgestellt und jeweils ein Sinuston mit  $250 [Hz]$ ,  $500 [Hz]$  und  $1000 [Hz]$  abgestrahlt. Die Messergebnisse wurden in Matlab ausgewertet und sind in den folgenden Abbildungen dargestellt. In Abbildung 6.6 ist der Betrag des Orts-Frequenz-Spektrums logarithmiert und bereits mit dem k-Space-Filter gefiltert.



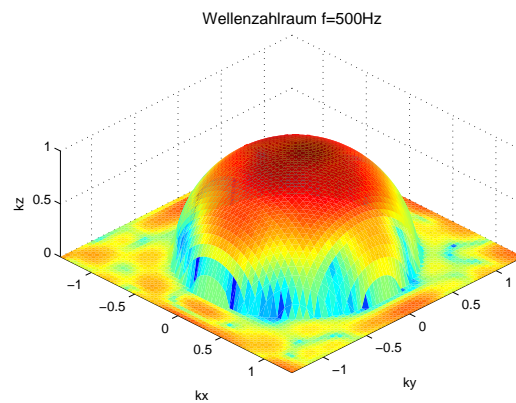
**Abbildung 6.6:** Betrag vom Orts-Frequenz-Spektrum des Drucks, extrapoliert auf die Membranoberfläche bei  $f=250Hz$ , und bereits mit dem k-Space Filter, den Anteil der evaneszenten Wellen reduziert.



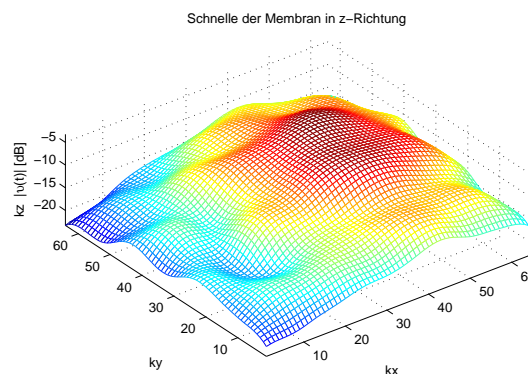
**Abbildung 6.7:** Schnelleverteilung auf der Membranoberfläche bei  $f=250Hz$ , mit zeropadding im Orts-Frequenz-Spektrum vor der Rücktransformation zur Erhöhung der Auflösung.

Der Lautsprecher wurde nicht genau im Zentrum des Mikrofonarrays positioniert, dies geht auch aus Abbildung 6.7 klar hervor, da das Maximum (dunkelrot

eingefärbt) nicht im Zentrum der Fläche zu liegen kommt. Kommt der k-Space-Filter nach der Extrapolation hin zur Membran nicht zum Einsatz, bildet die Schnelleverteilung Moden ab, aus denen die evaneszenten Wellen resultieren, siehe Abbildung 6.8 und 6.9.

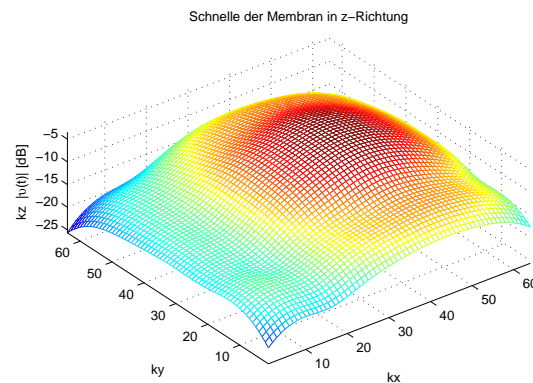


**Abbildung 6.8:** Nicht k-Space gefilterter Schalldruck im Wellenzahlraum dargestellt. Evaneszente Wellen liegen vor.



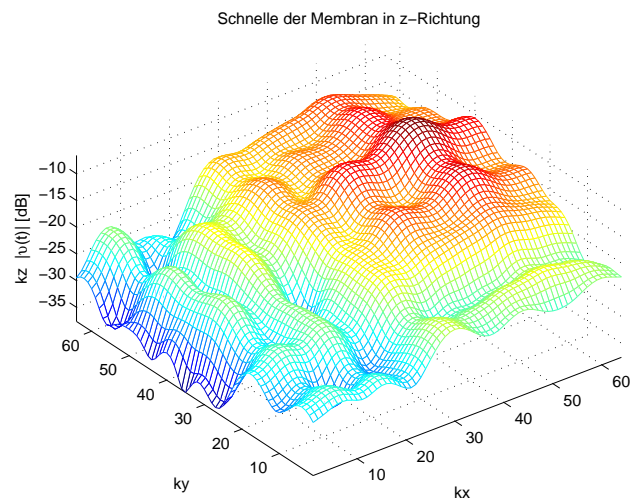
**Abbildung 6.9:** Schnelleverteilung auf der Membranoberfläche bei 500 Hz Sinuston im Nahfeld. Moden treten auf der Membran auf, aus denen die evaneszenten Wellen hervorgehen.

Durch Unterdrücken der evaneszenten Wellen mit dem k-Space-Filter bildet die Schnelleverteilung wieder nur die Grundschwingung der Membran ab, in diesem Fall  $500 [Hz]$ , siehe Abbildung 6.10.



**Abbildung 6.10:** Schnelleverteilung bei  $f=500$  Hz und Unterdrückung der Moden durch K-Space Filterung

Bei  $1000$  [Hz] für die gleiche Messanordnung treten schon sehr stark ausgeprägte evaneszente Wellen auf, die eine Regularisierung benötigen. Die Grenzfrequenz des Filters ist mit  $k_c = 1.1 \cdot k$  festgelegt. Die Schnelleverteilung beschreibt immer hochfrequenzere Moden, je mehr evaneszente Wellen, vor der Rücktransformation in den Zeitbereich, zugelassen werden.



**Abbildung 6.11:** Schnelleverteilung auf der Membranoberfläche bei  $f=1000$  Hz. Je mehr evaneszente Wellen vor der Rücktransformation zugelassen werden, desto mehr Moden werden sichtbar.

# Kapitel 7

## Conclusio

Der auftretende mittlere Fehler von 2 [cm] bei der Positionsbestimmung setzt sich aus mehreren Ursachen zusammen. Als Primäre Ursache ist die Abtastfrequenz und die exakte Vermessung des Lautsprecher-Referenzsystems (Tetraeder). Als sekundäre Fehlerquellen sind ist die Fluktuation des akustischen Zentrums und die Messung der Raumtemperatur, von der die auftretende Schallgeschwindigkeit abhängt, anzuführen. Mit dem konstruierten Messapparat für die Bestimmung der Lautsprecherpositionen mittels Viconsystem ist es nicht möglich bis zur Membran vorzudringen, da sich im Exponentialhorn sogenannte Phasenausgleichskeile befinden. Diese Messunsicherheit nimmt auch Einfluss auf das Messergebnis. Überdies ist anzuführen, dass die Lautsprecher eine leichte Verdrehung zueinander aufweisen, dies den Messfehler vergrößert. Eine Minimierung des Fehlers ist bei Verwendung anderer Lautsprecher, die eine geringere Gruppenlaufzeit aufweisen, möglich.

Die akustische Nahfeldholographie ist eine hochinteressante Theorie und lässt sich wunderbar implementieren. Zu untersuchen und bildlich darzustellen wie sich die Schalldruckverteilung und die Schnelleverteilung auf einer Lautsprechermembran verhalten, ist faszinierend, und trägt viel dazu bei, die Vorstellung und das Wissen über Schwingungsverhalten zu vertiefen. Auch die Zerlegung des Schallfeldes in ebene und evaneszente Wellen ermöglicht es, einen tieferen Einblick in die Komplexität von Schallfeldern und Schallquellen zu erlangen. Die erreichbare Auflösung, die mit den 64 Mikrofonen erreichbar ist, ist nicht unbedingt zufriedenstellend. Dazu wären noch weitere Experimente, oder sogar Erweiterungen des Verfahrens sinnvoll. Die

gezeigten Schnellverteilungen auf der Membranoberfläche wurde nur in  $z$ -Richtung ausgewertet. Auswertungen in  $x$  und  $y$  Richtung ist auch möglich und würde genaueren Aufschluss über das hydrodynamische Feld auf der Membranoberfläche ergeben.

# Literaturverzeichnis

- [BS97] M. Brandstein and H. Silverman, *A robust method for speech signal time-delay estimation in reverberant rooms*, 1997, vol. 1.
- [BW01] M. Brandstein and D. Ward, *Microphone Arrays: Signal Processing Techniques and Applications*. Springer, 2001.
- [Far00] A. Farina, “Simultaneous measurement of impulse response and distortion with a swept-sine technique,” *108th AES Convention, Paris*, pp. 18–22, 2000.
- [FH02] K. Fitz and L. Haken, “On the use of time-frequency reassignment in additive sound modeling,” *Journal of the Audio Engineering Society*, vol. 50, no. 11, pp. 879–893, 2002.
- [HBC06] Y. Huang, J. Benesty, and J. Chen, *Acoustic MIMO Signal Processing*. Springer, 2006.
- [KC76] C. Knapp and G. Carter, “The generalized correlation method for estimation of time delay,” *Acoustics, Speech, and Signal Processing [see also IEEE Transactions on Signal Processing]*, *IEEE Transactions on*, vol. 24, no. 4, pp. 320–327, 1976.
- [LRWW98] J. Lagarias, J. Reeds, M. Wright, and P. Wright, “Convergence properties of the nelder-mead simplex algorithm in low dimensions,” *SIAM Journal on Optimization*, vol. 9, no. 1, pp. 112–147, 1998.
- [LT99] X. Lai and H. Torp, “Interpolation methods for time-delay estimation using cross-correlation method for blood velocity measurement,” *IE-*



- EE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 46, no. 2, pp. 277–290, 1999.
- [OS96] M. Omologo and P. Svaizer, “Acoustic source location in noisy and reverberant environment using csp analysis,” *Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings., 1996 IEEE International Conference on*, vol. 2, 1996.
- [Pap06] L. Pape, “Vergleich robuster Mikrofonarrays,” Master’s thesis, Institut für Elektronische Musik und Akustik, 2006.
- [PL06] S. Pei and H. Lin, “Minimum-phase fir filter design using real cepstrum,” *Circuits and Systems II: Express Briefs, IEEE Transactions on [see also Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on]*, vol. 53, no. 10, pp. 1113–1117, 2006.
- [PM04] H. W. P. Majdak, M. Mihocic, “Aufbau und auswertung von vielkanalmessungen zur lokalisation von schallquellen,” *Presented at the Meeting of Austrian Section of Audio Engineering Society*, 2004.
- [Red93] S. Reddi, “An exact solution to range computation with time delay information for arbitrary array geometries,” *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]*, vol. 41, no. 1, 1993.
- [SGW07] A. Sontacchi, H. Gamper, and T. Walder, “Location determination of acoustical sources using a microphone array,” *3rd Congress of the Alps Adria Acoustics Association*, 2007.
- [TS05] T. Takahama and S. Sakai, “Constrained optimization by applying the/spl alpha/constrained method to the nonlinear simplex method with mutations,” *Evolutionary Computation, IEEE Transactions on*, vol. 9, no. 5, pp. 437–451, 2005.
- [War03] S. Warum, “Messung von Außenohrübertragungsfunktionen und direktionalen Raumimpulsantworten als Basis der Auralisation realer Räu-

- me,” Diplomarbeit, Institut für Elektronische Musik und Akustik der Universität für Musik und darstellende Kunst, Graz, 2003.
- [Wes03] G. Weselak, W.und Graber, *Raumakustik Skriptum*, v.2 ed., Institut für Nachrichtentechnik und Wellenausbreitung, Technische Universität Graz, 2003.
- [Wil99] E. Williams, *Fourier Acoustics: sound radiation and nearfield acoustical holography@bookmoser1988asa, title=Analyse und Synthese akustischer Spektren, author=Möser, M., year=1988, publisher=Springer-Verlag . Elsevier*, 1999.
- [WM80] E. Williams and J. Maynard, “Holographic imaging without the wavelength resolution limit,” *Physical Review Letters*, vol. 45, no. 7, pp. 554–557, 1980.
- [Zio95] L. Ziomek, *Fundamentals of Acoustic Field Theory and Space-Time Signal Processing*. CRC Press, 1995.
- [Zot08] F. Zotter, “Akustische Messtechnik Labor 2, LU, Institut für Elektronische Musik und Akustik, Graz,” 2008. [Online]. Available: [www.iem.at](http://www.iem.at)

# Anhang A

## Geräte und Datenblätter

### A.1 Mikrofon Behringer EMC8000

Bei diesem Mikrofon handelt es sich um ein omnidirektionales Elektret-Mikrofon.



Abbildung A.1: Behringer EMC8000 Mikrofon

Datenblatt und Spezifikation:

[http://www.behringerdownload.de/ECM8000/ECM8000\\_C\\_Specs.pdf](http://www.behringerdownload.de/ECM8000/ECM8000_C_Specs.pdf)

### A.2 Vorverstärker Behringer Ultragain Pro-8 ADA8000

Hierbei handelt es sich um einen 8-Kanal A/D-D/A @ 24 Bit Wandler mit Vorverstärker bei einer Samplefrequenz von wahlweise 44,1 kHz oder 48 kHz. Eingänge sind XLR und Line, Ausgänge ADAT optisch und Line. Phantomspannung von 48 V für jeden Kanal und Rasterpotentiometer für die Einstellung der Verstärkung.

Datenblatt und Spezifikation:



Abbildung A.2: Behringer Ultragain Pro-8 ADA8000

[http://www.behringerdownload.de/ADA8000/ADA8000\\_SPECS\\_Rev\\_C.pdf](http://www.behringerdownload.de/ADA8000/ADA8000_SPECS_Rev_C.pdf)

## A.3 MADI (Multichannel Audio Digital Interface)

Hierbei handelt es sich unter anderem um einen 8-Kanal 16x16 Router, der in diesem Fall dazu verwendet wird, die 8 ADAT Signal auf Madi zusammenzufassen.



Abbildung A.3: RME ADI-648

Datenblatt und Spezifikation:

[http://www.rme-audio.de/products\\_adi\\_648.php](http://www.rme-audio.de/products_adi_648.php)

## A.4 RME Hammerfall DSP MADI

Die HDSP MADI Pci stellt die Schnittstelle zum PC zur Verfügung.



Abbildung A.4: RME HDSP MADI

Datenblatt und Spezifikation:

[http://www.rme-audio.de/products\\_hdsp\\_madi.php](http://www.rme-audio.de/products_hdsp_madi.php)

## A.5 Larson Davis CAL200 Sound Level Calibrator

Batterie betriebener Kalibrator der Firma Larson Davis. Der Kalibrator strahlt  $1kHz$  wahlweise bei 94 dB SPL oder 114 dB SPL ab. <http://www.larsondavis.com/ModelCAL200.htm>

## A.6 G.R.A.S Freifeld Mikrofon Typ 40AF

Dabei handelt es sich um ein Freifeld-entzerrtes Präzisionsmikrofon der Firma G.R.A.S mit einem sehr guten Übertragungsverhalten.

<http://www.gras.dk/00012/00013/00028/00041/>

## A.7 Tetraeder

Der Tetraeder wurde in Eigenregie gebaut und besteht aus Hohl-Eisenstangen und Ahornholz in denen die Lautsprecher eingelassen sind. Der Körper ruht auf einem Flansch der auf ein  $1\frac{3}{8}$ " Stativ angebracht wird. Der Tetraeder ist auch frei nach Ikea Vorbild zerlegbar in seine Einzelteile. Sollte ihn jemand einmal zerlegen wollen: Einfach die Schrauben in den Ecken heraus schrauben mit einem sechskant Schlüssel.

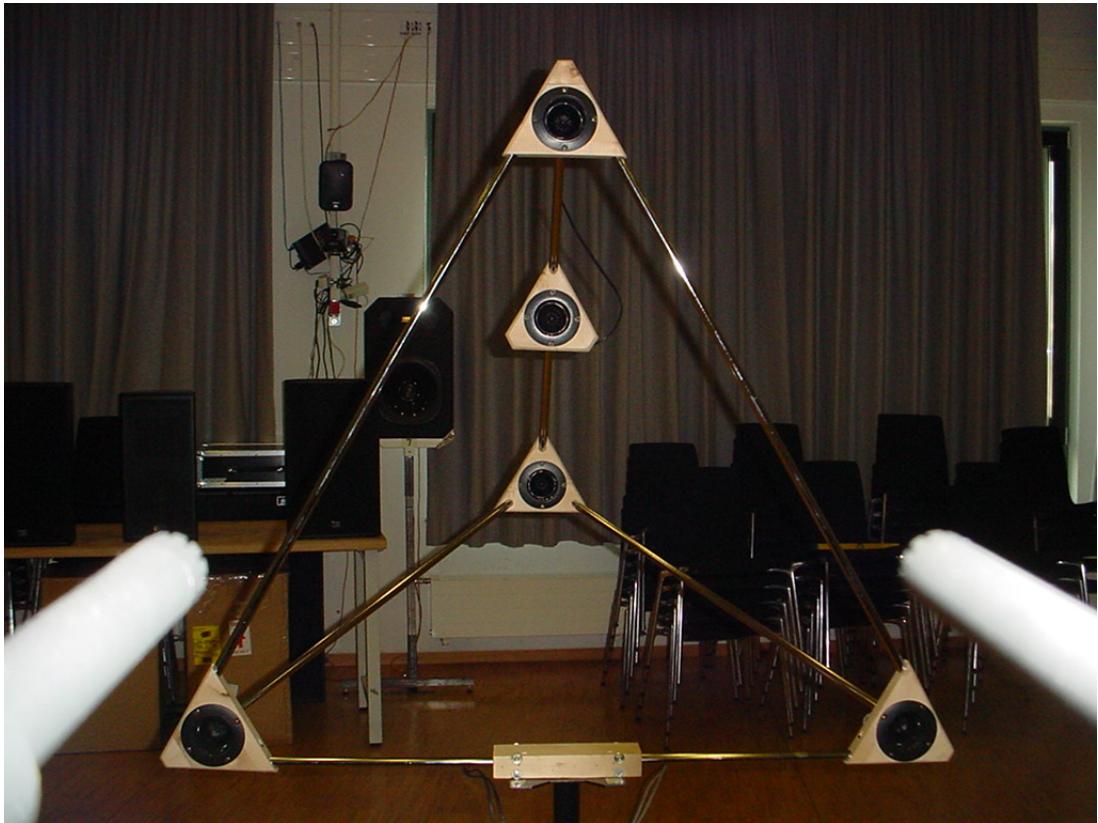


Abbildung A.5: Tetraeder Lautsprecherarray

## A.8 Piezo-Hochtonhornwandler

Dieser Wandler wurde bei Conrad bestellt. Es handelt sich um einen Piezo-Hochtonhorn Wandler TE-300. Den Hersteller konnte ich nicht ausfindig machen. Da er günstig war, färbt leider jeder Lautsprecher den Klang anders ein.



Abbildung A.6: Piezo-Hochtonhorn