

Masterarbeit

Steffen Johannes Zahn

Räumliche Filterung von Audiosignalen mit line-
aren Mikrofonarrays und Beamforming

Steffen Johannes Zahn

Räumliche Filterung von Audiosignalen mit linearen Mikrofonarrays und Beamforming

Masterarbeit eingereicht im Rahmen der Masterprüfung
im gemeinsamen Studiengang Mikroelektronische Systeme
am Fachbereich Technik
der Fachhochschule Westküste
und
am Department Informations- und Elektrotechnik
der Fakultät Technik und Informatik
der Hochschule für Angewandte Wissenschaften Hamburg

Betreuender Prüfer : Prof. Dr.-Ing. Hans-Peter Kölzer
Zweitgutachter : Prof. Dr.-Ing. Hans-Dieter Schütte
Abgegeben am 14. August 2015

Steffen Johannes Zahn

Thema der Masterarbeit

Räumliche Filterung von Audiosignalen mit linearen Mikrofonarrays und Beamforming

Stichworte

Räumliche Filterung, Beamforming, Breitband-Beamforming, Mikrofonarrays, Uniform Linear Array, Uniform Circular Array, Arraysignalverarbeitung, MVDR-Filter

Kurzzusammenfassung

In dieser Arbeit werden räumliche Filteralgorithmen für Audiosignale basierend auf linearen und zirkularen Mikrofonarrays entworfen. Nach einer Betrachtung verschiedener Einsatzumgebungen wird auf unterschiedliche Entwurfsverfahren eingegangen. Die durch Simulation verifizierten Ansätze werden in idealer und realer Umgebung getestet und auf ihre Leistungsfähigkeit geprüft. Abschließend werden die Ergebnisse gegenübergestellt um Stärken und Schwächen einzelner Ansätze hervorzuheben.

Steffen Johannes Zahn

Title of the master thesis

Spatial filtering of audio signals with linear microphone arrays and beamforming

Keywords

spatial filtering of audio signals, beamforming, broadband-beamforming, microphone arrays, uniform linear array, uniform circular array, audio signal processing, MVDR-beamformer

Abstract

This thesis is about spatial filtering of audio signals using linear and circular microphone arrays. At first operational environments are analyzed. This knowledge is used to present different development approaches. After simulation and verification the beamformers are evaluated and tested in ideal and real environments. Finally the results are being compared to emphasize individual weaknesses and strengths of each beamformer.

Inhaltsverzeichnis

Abbildungsverzeichnis	VI
1 Einleitung	1
1.1 Einführung	1
1.2 Aufgabenstellung	3
1.3 Aufbau der Arbeit	3
2 Stand der Technik	4
2.1 Bisherige Arbeiten zum Thema Mikrofonarrays	4
2.2 Modernes Beamforming	5
3 Entwicklungsumgebung und Hardware	9
3.1 DModule.C6713	9
3.2 DModule.PCM3003	10
3.3 Datenverarbeitung	11
3.4 Mikrofonarrays	12
4 Theoretische Grundlagen	14
4.1 Signal- und Arraymodell	14
4.2 Mikrofonabstand und räumliche Effekte	16
4.3 Evaluation von Beamformern	17
4.4 Rauschumgebung	22
4.5 Array Geometrie	25
5 Entwurf von Beamformern	31
5.1 Delay-and-Sum	31
5.2 Breitband-Beamforming mit Schmalband Entwurf	38
5.3 Breitband D&S-Beamformer	40
5.4 Breitband Minimum Variance Distortionless Response	42
5.5 Breitband Linear Constraint Minimum Variance	48
5.6 Nahfeldkompensation	52
6 Beamforming mit Subarrays	53
6.1 Beampattern	54
6.2 Array Gain	56
7 Implementierung in C	58

7.1	Programmablauf ULA	58
7.2	Programmablauf Subarray-ULA	59
8	Funktionalitätsmessungen in idealer Umgebung	60
8.1	Messumgebung und -aufbau	60
8.2	Durchführung	62
8.3	Messergebnisse ULA	66
8.4	Messergebnisse Subarray-ULA	71
9	Messungen mit Sprachsignal	73
9.1	Ideale Umgebung	74
9.2	Reale Umgebung	80
9.3	Bewertung	90
10	Zusammenfassung und Fazit	92
11	Ausblick	94
	Literaturverzeichnis	96
	Anhang	98

Abbildungsverzeichnis

1.1	Modellszenario einer Videokonferenz	1
1.2	Modellszenario eines Hörsaals	2
2.1	Active Electronically Steered Antenna (A-ESA) eines F-15 Kampffjets	6
2.2	Räumliches Multiplexverfahren durch Beamforming im LTE-Standard	7
3.1	Hardwareplattform	9
3.2	Blockschaltbild, D-Module.C6713 DSignT	10
3.3	Blockschaltbild, D-Module.PCM3003 DSignT	11
3.4	Arbeitsweise, EDMA Ping-Pong Verfahren	12
3.5	Lineares Mikrofonarray	13
4.1	Linear angeordnetes Mikrofonarray	15
4.2	Darstellungsbeispiel für ein Beampattern	19
4.3	Kohärenzfunktionen für ein diffuses Rauschfeld	23
4.4	Kohärenzfunktionen für ein zylindrisch isotropisches Rauschfeld	24
4.5	Array mit nested Subarrays	27
4.6	Zirkulares Mikrofonarray mit konstanten Mikrofonabständen (UCA)	28
5.1	Beampattern, 90° D&S-Beamformers für ein UCA	33
5.2	Beampattern, 90° D&S-Beamformers für ein UCA, Polardarstellung	34
5.3	Beampattern, 90° D&S-Beamformers für ein ULA	35
5.4	Beampattern, 90° D&S-Beamformers für ein ULA über Frequenz und Winkel	36
5.5	Beampattern, 0° D&S-Beamformers für ein ULA	37
5.6	Beampattern, 0° D&S-Beamformers für ein ULA über Frequenz und Winkel	37
5.7	Breitband Beamformer, Grundstruktur	39
5.8	Beampattern, 0° Breitband D&S-Beamformer	40
5.9	White Noise Gain, 0° Breitband D&S-Beamformers	41
5.10	Directivity Index, 0° Breitband D&S-Beamformers	42
5.11	Beampattern, 0° MVDR Beamformers für diffuses Rauschfeld	44
5.12	Beampattern, 0° MVDR Beamformers für zylindrisch isotropes Rauschfeld	45
5.13	WNG, 0° MVDR Beamformers für diffuses Rauschfeld	46
5.14	WNG, 0° MVDR Beamformers für zylindrisch isotropes Rauschfeld	46
5.15	Directivity Index, 0° MVDR Beamformers für diffuses Rauschfeld	47
5.16	Directivity Index, MVDR Beamformers für zylindrisch isotropes Rauschfeld	47
5.17	Beampattern, 0° / 50° LCMV Beamformers für unkorreliertes Rauschfeld, $\mu = 0.001$	49
5.18	Beampattern, 0° / 50° LCMV Beamformers für diffuses Rauschfeld, $\mu = 0.001$	50

5.19 White Noise Gain, 0° / 50° LCMV Beamformers für diffuses Rauschfeld	51
5.20 Directivity Index, 0° / 50° LCMV Beamformers für diffuses Rauschfeld	51
6.1 Arrayaufbau, 3 harmonische Subarrays	53
6.2 Beampattern, 90° Subarray-D&S-Beamformer	55
6.3 Beampattern, 0° MVDR-Subarray-Beamformer für diffuses Rauschfeld, $\mu = 0.001$	55
6.4 White Noise Gain verschiedener Subarray-Beamformer	56
6.5 Directivity Index verschiedener Subarray-Beamformer	57
7.1 Programmablaufplan, Beamformer Implementierung für ULA	58
7.2 Programmablaufplan, Beamformer Implementierung für Subarrays	59
8.1 Messaufbau im Schallmessraum der HAW Hamburg	61
8.2 Elektrischer Messaufbau im Schallmessraum der HAW Hamburg	62
8.3 Konfiguration des Audio Analyzers	63
8.4 Beispielergebnis der Amplitudengangsanalyse mit einem R&S Audio Analyzer	64
8.5 Amplitudengang des Übertragungssystems	65
8.6 Beampattern, 90° Delay and Sum Beamformer	66
8.7 Beampattern, 0° D&S-Beamformer	66
8.8 Beampattern, 90° MVDR-Beamformer für diffuse Rauschumgebung	67
8.9 Beampattern, 0° MVDR-Beamformer für diffuse Rauschumgebung	67
8.10 Beampattern, 90° MVDR-Beamformer für zylindrisch isotropische Rauschumgebung	68
8.11 Beampattern, 0° MVDR-Beamformer für zylindrisch isotropische Rauschumgebung	68
8.12 Beampattern, 90° 50° LCMV-Beamformer	69
8.13 Beampattern, Subarray D& S-Beamformer	71
8.14 Beampattern, Subarray MVDR-Beamformer für diffuses Rauschfeld	71
9.1 Referenzsignal an Mikrofon 3, weibliche Sprecherin	74
9.2 Spektrogramm, Referenzsignal an Mikrofon 3, weibliche Sprecherin	75
9.3 Ausgangssignal, D&S-Beamformer, weibliche Sprecherin	76
9.4 Spektrogramm, Ausgangssignal D&S-Beamformer, weibliche Sprecherin	77
9.5 Ausgangssignal, MVDR-Beamformer für diffuses Rauschfeld, weibliche Sprecherin	78
9.6 Spektrogramm, Ausgangssignal MVDR-Beamformer für diffuses Rauschfeld, weibliche Sprecherin	79
9.7 Grundriss des Versuchsbüros im Raum 14.04 der HAW-Hamburg	81
9.8 Referenzsignal an Mikrofon 3, weibliche Sprecherin in Büroumgebung	82
9.9 Referenzsignal an Mikrofon 3, weibliche Sprecherin in Büroumgebung	83
9.10 Ausgangssignal, Subarray D&S-Beamformer, weibliche Sprecherin in Büroumgebung	84

9.11	Spektrogramm, Ausgangssignal Subarray D&S-Beamformer, weibliche Sprecherin in Büroumgebung	84
9.12	Ausgangssignal, Subarray MVDR-Beamformer für diffuses Rauschfeld, weibliche Sprecherin in Büroumgebung	86
9.13	Spektrogramm, Ausgangssignal Subarray MVDR-Beamformer für diffuses Rauschfeld, weibliche Sprecherin in Büroumgebung	86
9.14	Ausgangssignal, D&S-Beamformer, weibliche Sprecherin in Büroumgebung	87
9.15	Spektrogramm, Ausgangssignal D&S-Beamformer, weibliche Sprecherin in Büroumgebung	88
9.16	Ausgangssignal, MVDR-Beamformer für diffuses Rauschfeld, weibliche Sprecherin in Büroumgebung	89
9.17	Spektrogramm, Ausgangssignal MVDR-Beamformer für diffuses Rauschfeld, weibliche Sprecherin in Büroumgebung	89
11.1	Beamformer in „Generalized-Sidelobe-Canceller“-Struktur	94

Abkürzungen und Übersetzungen

Abkürzung	Englisch	Deutsch
ADC	Analog Digital Converter	Analog Digital Umsetzer
-	Beamformer	Räumlicher Filter
-	Beampattern	Räumliches Empfangsmuster
D&S	Delay-and-Sum	Verzögern und Addieren
DFT	Discrete Fourier Transformation	Diskrete Fourier Transformation
DI	Directivity Index	-
DOA	Direction of Attack	Einstrahlungswinkel, Angriffswinkel
EDMA-Controller	Enhanced Direct Memory Access Controller	Direkter Speicherzugriffs-Controller
FFT	Fast Fourier Transformation	Schnelle Fourier Transformation
LCMV	Fast Fourier Transformation	Schnelle Fourier Transformation
OFDM	Orthogonal Frequency Division Multiplex	Orthogonales Frequenzmultiplexverfahren
McBSP	Multichannel Buffered Synchronous Serial Port	-
MVDR Beamformer	Minimum Variance Distortionless Response Beamformer	Minimalvarianz Beamformer mit unverzerrter Antwort
SNR	Signal-to-Noise Ratio	Signal- zu Rauschabstand
UART	Universal Asynchronous Receiver Transmitter	-
UCA	Uniform Circular Array	Gleichverteiltes zirkulares Array
ULA	Uniform Linear Array	Gleichverteiltes lineares Array
WNG	White Noise Gain	Verstärkung von weißem Rauschen

Mathematische Notation

Darstellung		Bedeutung
kleiner Buchstabe	x	Skalar im Zeitbereich
großer Buchstabe	X	Skalar im Frequenzbereich
fetter Buchstabe	x, X	Vektor im entsprechenden Bereich
hochgestelltes T	$()^T$	Transponiert
hochgestelltes H	$()^H$	Komplex konjugiert transponiert
fettgedruckte 1	1	Spaltenvektor dessen Elemente 1 sind
fettgedrucktes I	I	Einheitsmatrix
Versor	\angle	Winkel des nachfolgenden Ausdrucks

Symbolverzeichnis

D	Directivity Faktor
\mathbf{d}	Steering-Vektor
d_{mic}	Mikrofonabstand
d_{SA}	Entfernung zwischen Quelle und Array
DI	Directivity Index [dB]
f	Frequenz
\mathbf{I}	Einheitsmatrix
J_0	Bessel-Funktion erster Art
L_{Array}	Arraylänge
N_{mic}	Mikrofonanzahl
P_{Nutz}	Nutzleistung
P_{Rausch}	Rauschleistung
\mathbf{p}_i	Mikrofon-Ortsvektor
s	Quellsignal
R	Radius des UCA
v	Geschwindigkeit (meist Schall)
v_i	Störsignal
w_i	Beamformer-Koeffizienten
WNG	White Noise Gain
x_i	Signal am Mikrofon i
y	Beamformer-Ausgangssignal
α_i	Dämpfungsfaktoren
$\mathbf{\Gamma}$	Kohärenzmatrix

ε	Elevation-Empfangswinkel
λ	Wellenlänge
ϕ	Azimuth-Empfangswinkel
Φ	Spektrale Leistungsdichte
Φ	Leistungsdichtematrix
Ω	Kreisfrequenz

1 Einleitung

1.1 Einführung

In der modernen Zeit spielt weltweite Kommunikation eine entscheidende Rolle. Video- und Telefonkonferenzen ersetzen Dienstreisen und lange E-Mail Korrespondenzen. In den letzten zehn Jahren hat sich hieraus ein neuer Wirtschaftszweig gebildet, der die Bedürfnisse internationaler Firmen weltweit bedient. Ein großes Problem in dieser Anwendung ist die Audio-Performance der verwendeten Systeme[14].

Ein typischer Anwendungsfall dieser Systeme sind Telefonkonferenzen mit mehreren Personen in einem durchschnittlich großen Büro. Die Teilnehmer befinden sich verteilt um ein entsprechendes Empfangsgerät herum und tragen abwechselnd Inhalte zum Gespräch bei. Um eine möglichst gute Sprachqualität zu erreichen, wäre es entsprechend von Vorteil den jeweils aktiven Sprecher zu identifizieren und seine Sprache zu verstärken. Weiterhin wäre es nützlich, parallel dazu äußere Störgeräusche wie Klimaanlage oder andere weiter entfernte Sprecher auszublenden.

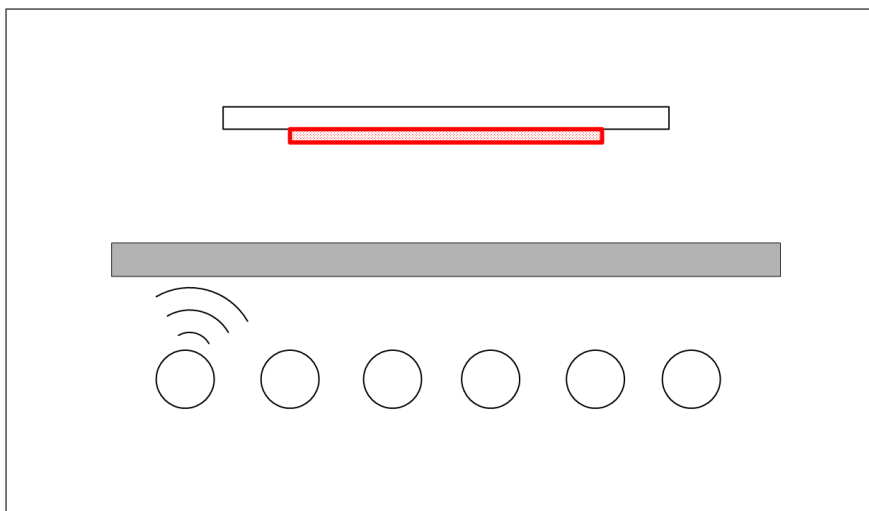


Abbildung 1.1: Modellszenario einer Videokonferenz

Ein weiterer denkbarer Anwendungsfall lässt sich z.B. in der VIP-Luftfahrt finden. Ein Sprecher befindet sich hierbei in einem privaten Abteil eines Flugzeuges. Zur optimalen Nutzung der z.T. langen Flugzeiten, werden viele Telefongespräche oder Videokonferenzen von Bord aus geführt. Um sich im Abteil frei bewegen zu können, wäre anstatt eines stationären Gerätes eine Freisprechanlage denkbar. Die Hauptaufgabe dieser Anlage wäre es, den Sprecher im Raum zu

„verfolgen“ , um jederzeit einen bestmöglichen Sprachempfang zu gewährleisten.

Ein letzter praxisnaher Fall findet sich beispielsweise bei größeren Vorträgen oder in Vorlesungen an Hochschulen. Ein Sprecher steht vor einer Gruppe von Zuhörern. Während seines Vortrages bewegt er sich vor dem Publikum, um Dinge an Tafeln oder Flipcharts zu erklären oder um Blickkontakt zu verschiedenen Teilnehmern herzustellen. Um nicht während des gesamten Vortrages ein Mikrofon mit sich herumtragen zu müssen, kommt auch hier eine stationäre Mikrofonlösung zum Einsatz. Die Hauptanforderung in diesem Fall ist wiederum, den Sprecher im Fokus zu halten und gleichzeitig die Geräusche hinter dem Mikrofon auszublenden.

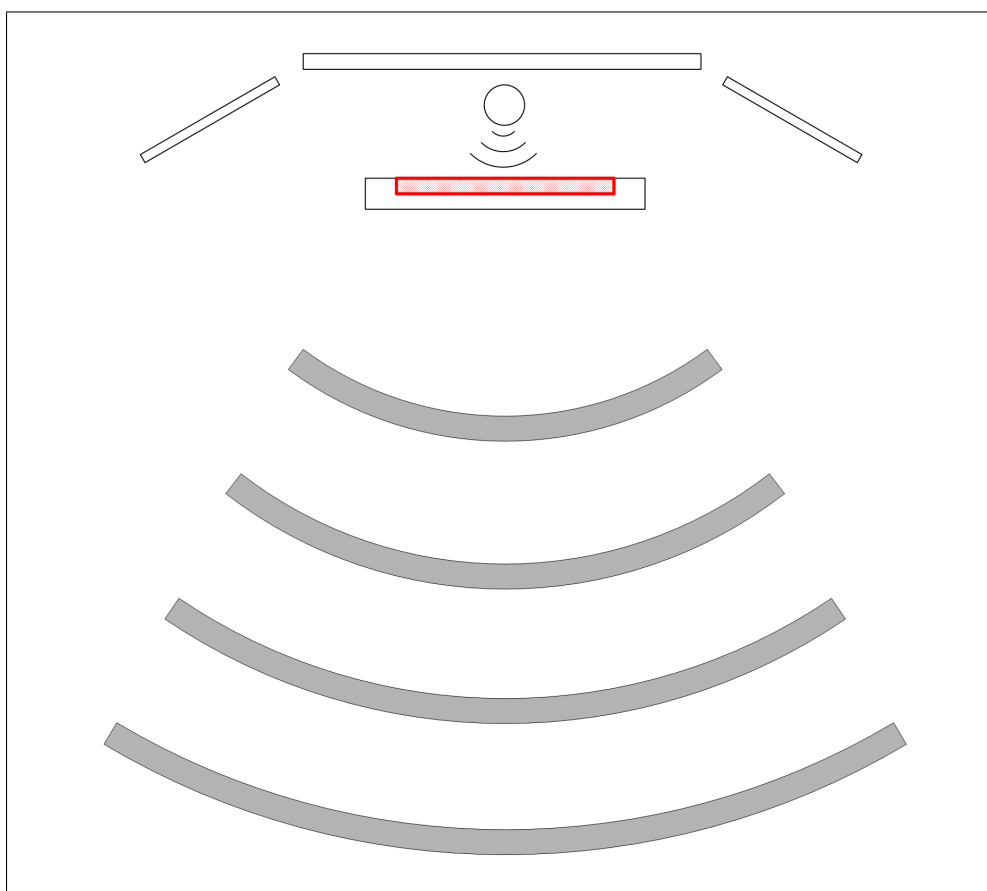


Abbildung 1.2: Modellszenario eines Hörsaals

Zusammengefasst erfordern die vorgestellten Szenarien eine Mikrofoneinheit, die in der Lage ist, eine adaptive räumliche Richtwirkung zu erzielen. Ein praxisnaher Ansatz für dieses Problem ist die Verwendung von Mikrofonarrays. Die Verwendung von mehreren räumlich getrennten Mikrofonen und durch geschickte Verrechnung der empfangenen Signale, kann die räumliche Empfangscharakteristik der Anordnung entscheidend beeinflusst werden. Diese Verfahren werden als „Beamforming“ bezeichnet und bilden das Kernthema dieser Arbeit.

1.2 Aufgabenstellung

Aufgabe ist die Entwicklung eines Algorithmus zur räumlichen Filterung von Audiosignalen. Basis des Projektes bilden die vorhandenen Mikrofonarrays und eine Hardwareplattform mit dem digitalen Signalprozessor TMS320C6713B der Firma Texas Instruments.

Ziel der Arbeit ist es, Audiosignale aus verschiedenen Richtungen zu gewichten, um z.B. die Audioqualität von Telefonkonferenzen zu verbessern.

1.3 Aufbau der Arbeit

Die Thesis ist in verschiedene Themenbereiche aufgliedert. Zunächst wird der aktuelle Stand der Technik beleuchtet. Es wird kurz auf bereits abgeschlossene Abschlussarbeiten im Bereich Mikrofonarray-Signalverarbeitung eingegangen und die generelle Forschungslage zum Thema betrachtet.

Nachdem die Hardwareplattform vorgestellt wird, folgt ein Kapitel zu theoretischen Grundlagen wie Signalmodell, Array Geometrie, Einfluss räumlicher Effekte und Rauschumgebungen. In Abschnitt 5 werden verschiedene Entwurfsansätze für Beamformer betrachtet und deren Eigenschaften beschrieben. Die Ergebnisse der Entwürfe werden an Hand von Simulationen dargestellt und erläutert. Neben den grundlegenden Schmalband-Entwurfsmethoden wird ein Verfahren zur Verarbeitung von Breitbandsignalen vorgestellt. Im Abschluss des Kapitels wird auf eine Anpassung des Filterverhaltens für Nahfeld-Empfang eingegangen. Im folgenden Abschnitt ist der Programmablauf der Hardwareimplementierung in C als Ablaufplan zu finden. Die Funktionalität der Beamformer wird in Kapitel 8 an Hand von Messungen im Schallmessraum der HAW-Hamburg getestet. Nachdem diese sichergestellt ist, finden Messungen verschiedener Beamformer in realer Umgebung statt. Den Abschluss der Arbeit bilden Zusammenfassung, Fazit und ein Ausblick auf zukünftige Verbesserungen und Anwendungsgebiete.

2 Stand der Technik

In diesem Kapitel wird kurz vorgestellt, welche Aspekte der Array-Signalverarbeitung bisher an der HAW-Hamburg untersucht worden sind. Nachfolgend wird beschrieben, in welchen Bereichen Beamforming in der heutigen Zeit angewendet wird.

2.1 Bisherige Arbeiten zum Thema Mikrofonarrays

In der Vergangenheit sind an der HAW-Hamburg diverse Arbeiten zum Thema Mikrofonarrays entstanden. Kernthemen der Arbeiten waren vor allem Audiolokalisierung und generelle Signalverarbeitung mit Mikrofonarrays. Unterteilen lassen sich die vorliegenden Untersuchungen in Art der untersuchten Algorithmen sowie Form des Arrays.

Im Bereich lineare Mikrofonarrays wurden vor allem Lokalisierungsalgorithmen auf Basis der Multichannel Crosscorrelation Coefficients (MCCC) untersucht, vor allem die Arbeiten von Kolja Pikora [12] und Markus Müller [10]. Dieses Verfahren nutzt die Zeitverzögerung zwischen den empfangenen Audiosignalen an den Mikrofonen, um über Korrelationsalgorithmen den zugehörigen Empfangswinkel abzuschätzen. Aus der Masterarbeit von Pikora geht z.B. hervor, dass es mit Hilfe dieses Verfahrens möglich ist, Sprecher in Azimuthrichtung mit einer Genauigkeit von bis zu $\pm 8^\circ$ zu lokalisieren [12, S.103]. Die entwickelten Algorithmen sind auf ihren jeweiligen Hardwareplattformen echtzeitfähig.

Im Bereich zirkulare Arrays existiert eine aktuelle Arbeit von Jens Reermann [13]. Der verwendete Ansatz nutzt einen unterraumbasierten Algorithmus zur Lokalisierung der Audioquellen, den sogenannten UCA-ESPRIT. Grundlage bildet ein zirkulares Mikrofonarray, das es erlaubt, neben der Azimuthlokalisierung auch eine Elevationslokalisierung vorzunehmen. Der Algorithmus beginnt mit der Abschätzung der Autokorrelationsmatrix der Mikrofon-signale, gefolgt von der Transformation in den sogenannten Phasen-Moden-Raum mit anschließender Eigenwertzerlegung. Die ermittelten Eigenwerte werden genutzt, um einen Signal- und Rauschunterraum zu bilden, dessen Eigenwerte wiederum Rückschlüsse auf den Empfangswinkel zulassen. Die durchgeführten Messungen zeigen eine hohe Genauigkeit in der Winkelbestimmung von 1° in Azimuthrichtung und 4° in Elevationsrichtung [13, S.172].

Die aus den beschriebenen Verfahren erhaltenen Richtungsinformationen sollen nun von räumlichen Filteralgorithmen genutzt werden, um eine auf den Sprecher ausgerichtete Empfangscharakteristik zu erhalten.

2.2 Modernes Beamforming

Beamforming ist eine weit verbreitete Methode zur räumlichen Filterung. Klassische Anwendungsfelder liegen im Bereich Mobilfunk und in der Radartechnik. Bedingt durch die für diesen Bereich im allgemeinen sehr hohen Frequenzen von mehreren Gigahertz können Sensor- bzw. Antennenarrays sehr klein und damit kompakt entworfen werden. Außerdem erlauben die kleinen Wellenlängen hochfrequenter Signale hochselektive Filtereigenschaften.

Eine der bekanntesten Anwendungen von Beamforming findet sich in der aktiven Radarerfassung z.B. von Flugzeugen. Im Buch „Radartechnik - Grundlagen und Anwendungen“ von Jürgen Göbel findet sich hierzu folgende Aussage:

„In vielen Radarsystemen sind (aus-)richtbare Antennen von großer Bedeutung, so beispielsweise bei der Luftraumüberwachung oder in Flugzeug-Bordradaren. Für die Realisierung kommen zwei Ansätze in Betracht: Neben den rein mechanisch richtbaren Antennen (Mechanically Steerable Antenna, MSA) erobern seit einigen Jahrzehnten elektronisch richtbare Antennen (Electronically Steerable Antenna, ESA) immer mehr Einsatzbereiche.“

[6, S.37]

Das Prinzip der elektronischen Ausrichtung beruht hauptsächlich auf Phasenänderung der Antennensignale, also dem Prinzip, das dem Beamforming zu Grunde liegt. Daraus leitet sich auch der Begriff Phased-Array-Antenna her[vgl. 6, S.37].

Die großen Vorteile der ESA liegen in der schnellen Anpassung des Empfangswinkels gegenüber der mechanischen Einstellung. Auf Grund dieser Eigenschaft werden Sie daher auch verstärkt im militärischen Bereich eingesetzt, Abbildung 2.1 zeigt z.B. einen F15 Kampfjet mit aktivem ESA¹. Die Methoden des in der Radartechnik als „digital Beamforming“ bezeichneten Technik [6, S.48] ermöglichen es weiterhin, die Richteigenschaften der Antenna -in gewissen Grenzen- beliebig zu beeinflussen.

¹Quelle: <http://aviationintel.com/the-great-radar-race-aesa-development-in-high-gear/>



Abbildung 2.1: Active Electronically Steered Antenna (A-ESA) eines F-15 Kampffjets

Einen weiteren großen Anwendungsbereich von Beamforming bilden kabellose Funkverbindungen, wie sie zum Beispiel im Mobilfunk oder in Wireless-Lan (W-LAN) verwendet werden. Es wird somit möglich, neben den „klassischen“ Multiplex Verfahren in Frequenz- oder Zeitschlitzern wie zum Beispiel bei OFDM (Orthogonal-Frequency-Division-Multiplex) auch räumliche Unterscheidungen einzubringen. Ein heute im Mobilfunk bekannter Standard verwendet diese Kombination. Es handelt sich um den sogenannten Long-Term-Evolution-Standard LTE. LTE ermöglicht Datenraten von bis zu 300 Megabit pro Sekunde [16]. Um dies zu erreichen, wird neben OFDM auch ein sogenanntes MIMO-System (Multiple Input Multiple Output) aus Antennen verwendet. Dieses Antennensystem kann entweder zur Erhöhung der Robustheit oder zur Erhöhung der Kanalkapazitäten durch räumlichen Multiplex verwendet werden. Abbildung 2.2 zeigt das Filterverhalten der LTE-Antennen:

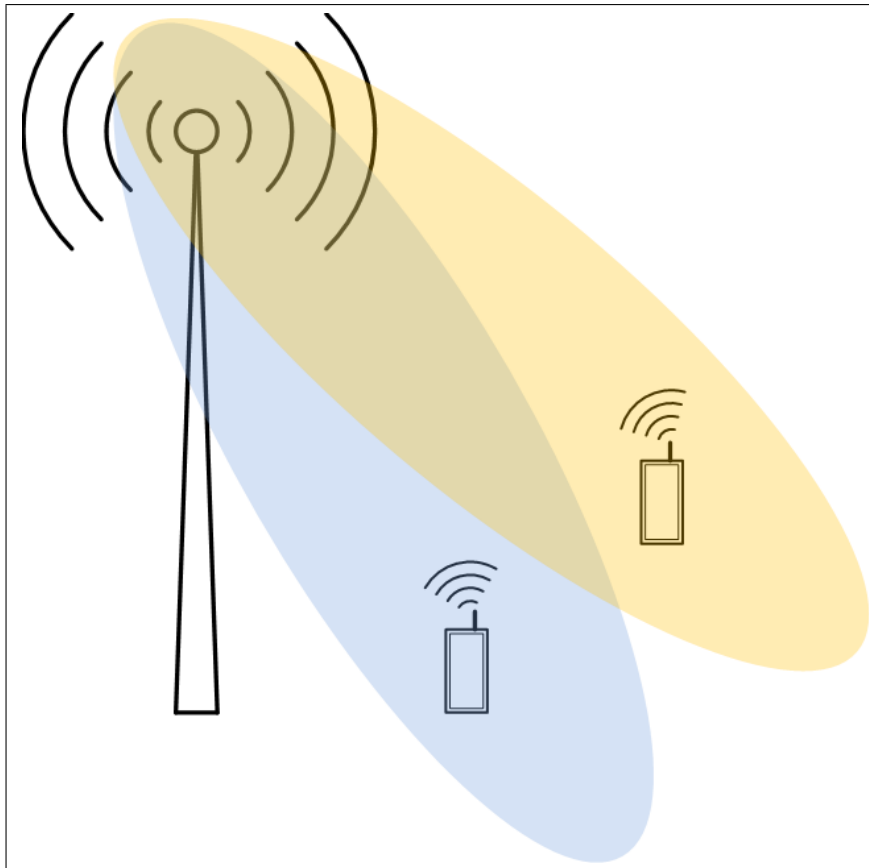


Abbildung 2.2: Räumliches Multiplexverfahren durch Beamforming im LTE-Standard

Die in der Radartechnologie und im Mobilfunk schon seit längerer Zeit genutzten Verfahren finden heutzutage nach und nach Einzug in andere Technologiebereiche, wie zum Beispiel der gerichteten Auto zu Auto [7] oder in-Auto Kommunikation [5, S.9].

Die Herausforderung in der Portierung der bestehenden Konzepte besteht vor allem in der Anpassung auf breitbandige Signale. Während für Radarantennen im Normalfall nur ein schmalbandiger aber hochfrequenter Bereich betrachtet werden muss, umspannen Audiosignale einen wesentlich größeren Bereich von mehreren Kilohertz im tiefen Frequenzband. Ein weiterer Unterschied liegt in den entsprechend großen Wellenlängen, die andere Arraydimensionen nötig machen.

Während Untersuchungen zum Thema Breitband Beamforming schon seit langem vorhanden sind (z.B. das Referenzwerk „*Adaptive Signal Processing*“ von Bernard Widrow aus dem Jahre 1985 [15]), so sind tatsächliche Implementierungen noch vergleichsweise selten zu finden. Viele der wissenschaftlichen Papiere beinhalten theoretische Aspekte sowie Software-Simulationen. Nur ein geringer Prozentsatz behandelt tatsächliche Messungen und praktische Aufbauten. Dies liegt möglicherweise darin begründet, dass z.T. sehr aufwendige Messsysteme und entsprechend ausgestattete Messräume sehr kostspielig sind.

Zusammenfassend lässt sich feststellen, dass das Thema Beamforming im Audibereich wach-

sendem Interesse ausgesetzt ist. Gerade in den letzten Jahren häufen sich Artikel, die sich auch mit Implementationen und Endprodukten beschäftigen. Hierzu zählt zum Beispiel der Artikel *New Insights Into the MVDR Beamformer in Room Acoustics* [8], erschienen bei der IEEE im Januar 2010 oder die schon angesprochenen Untersuchungen zur Autokommunikation.

3 Entwicklungsumgebung und Hardware

Für die Umsetzung der Verfahren wird eine Kombination aus verschiedenen Baugruppen verwendet. Grundlage bildet der digitale Signalprozessor (Digital Signal Processor, DSP) TMS320C6713B der Firma Texas Instruments. Die nötige Peripherie wird vom D.Module.C6713 bereitgestellt. Die Analog-Digital-Wandlung für bis zu 8 Kanäle übernimmt das Audio-Codec-Modul D.Module.PCM3003. Abbildung 3.1 [aus 13, S.25] zeigt die gesamte Hardwareplattform:

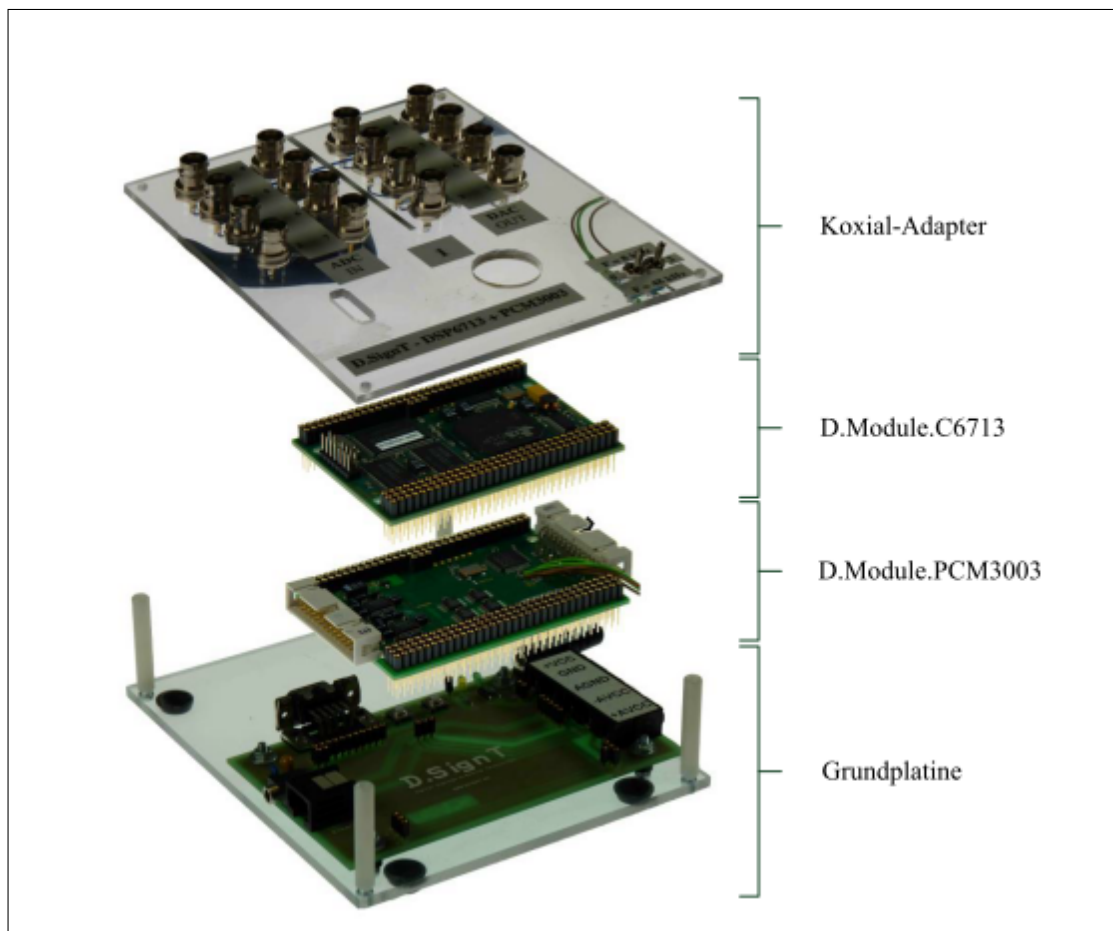


Abbildung 3.1: Hardwareplattform

In den folgenden Abschnitten wird genauer auf die Funktionsweise der Elemente eingegangen.

3.1 DModule.C6713

Das DModule.C6713 stellt den RISC Prozessor TMS320C6713 und die nötige Peripherie bereit. Der Prozessor stellt 8 parallele Instruktionseinheiten zur Verfügung welche bei 300 MHz

Systemtakt bis zu 1800 MFLOPS¹ Rechenleitung erzielen. Zur Datenanbindung stehen zwei „Multi-Channel Buffered Synchronous Serial Ports“ (McBSP) zur Verfügung. Speicherzugriffe werden über einen „Enhanced Direct Memory Access Controller“ (EDMA) gesteuert. Zur Kommunikation mit der Umgebung ist ein „Universal Asynchronous Receiver Transmitter“ (UART) mit RS232 Interface verbaut. Das folgende Blockschaltbild ist dem Datenblatt [3, S.1] entnommen und zeigt den internen Aufbau des Boards:

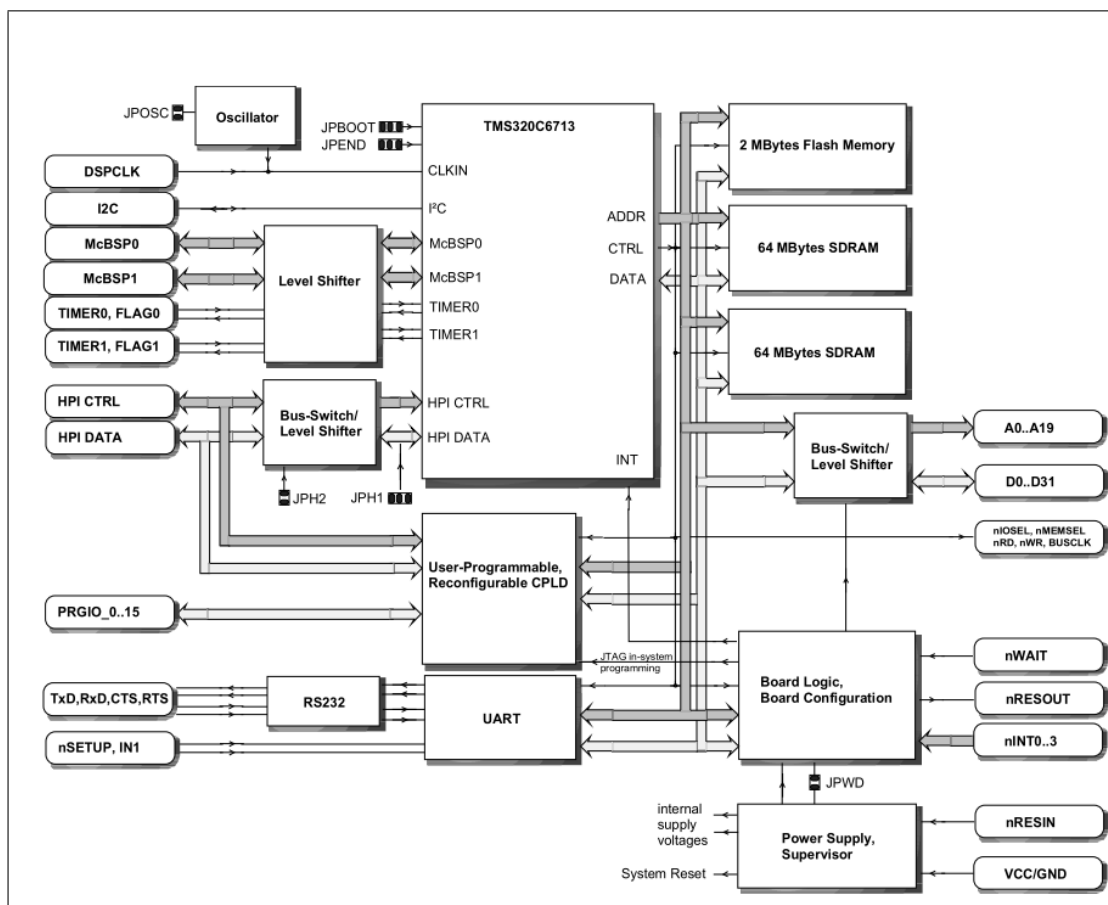


Abbildung 3.2: Blockschaltbild, D-Module.C6713 DSigT

3.2 DModule.PCM3003

Die Wandlung der analogen Mikrofonsignale übernimmt das PCM3003-Modul. Das Audio-Codex-Board ermöglicht die parallele Umsetzung von bis zu 8 Kanälen. Die verwendeten delta-Sigma ($\Delta\Sigma$) Analog-Digital-Konverter arbeiten mit wählbarer Abtastfrequenz von 8 kHz bis 48 kHz und 64-facher Überabtastung. Die Auflösung beträgt 16 Bit bei 1,8V_{pp} Vollaussteuerung. Die Kommunikation mit dem Prozessormodul erfolgt über ein 8-Kanal serielles Interface, wobei die Takt-

¹MFLOPS: Mega Floating Point Operations Per Second

und Synchronisationssignale vom PCM3003-Board selbst generiert werden. Der prinzipielle Aufbau kann dem Blockschaltbild des technischen Datenblattes [4, S.1] entnommen werden:

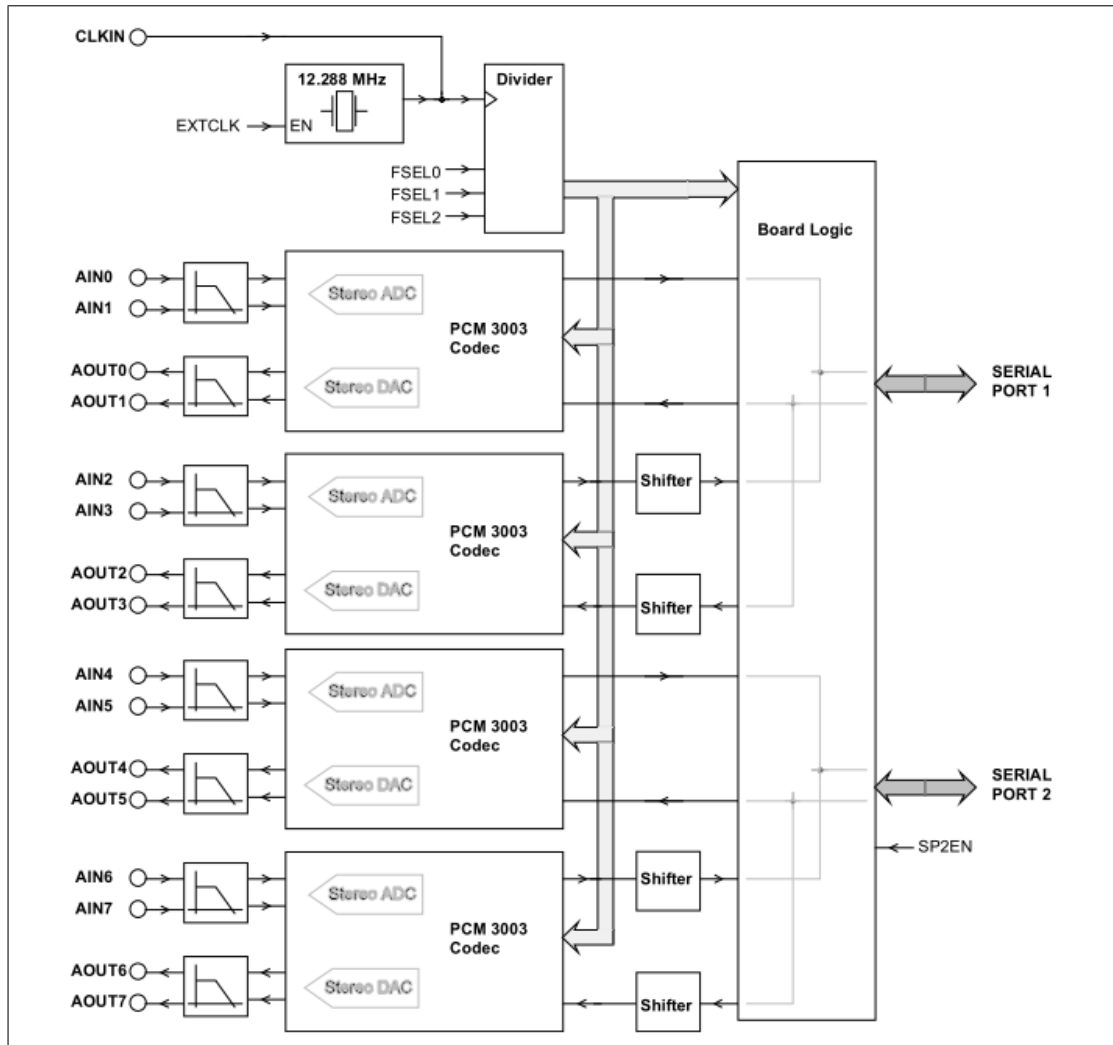


Abbildung 3.3: Blockschaltbild, D-Module.PCM3003 DSignT

3.3 Datenverarbeitung

Die über die serielle Schnittstelle des PCM3003-Boards bereitgestellten Daten werden auf dem C6713-Modul verarbeitet. Bei jedem Datenempfang an den McBSP-Schnittstellen wird ein Interrupt generiert, um die weitere Verarbeitung einzuleiten. Bedingt durch die Abtastfrequenz von 8 kHz würde das Hauptprogramm jedes Mal unterbrochen werden, wenn neue Daten zur Verfügung stehen. Diese Unterbrechungen können durch den vorhandenen „Enhanced Direct Memory Access-Controller“ (EDMA-Controller) vermieden werden. Der EDMA verwaltet eigenständig und parallel zur Ausführung des Hauptprogramms das Speichern der Daten. Anstatt bei jedem neuen Wert einen Interrupt zu generieren, wird bei diesem Verfahren ein Block von Daten in

einen Zwischenspeicher eingelesen und bei Erreichen einer definierten Blocklänge ein Interrupt ausgelöst. Dieser Interrupt ist das Signal für das Hauptprogramm, dass ein neuer Datenblock zur Verarbeitung bereit steht.

Ein Problem, das bei Verwendung des EDMA auftreten kann, ist das Überschreiben noch nicht verarbeiteter Datenblöcke. Der EDMA schreibt automatisch und direkt die neuen Daten in den Speicher. Findet die Datenverarbeitung auf denselben Speicherplätzen statt, auf denen der EDMA speichert, z.B. derselben Variablen, können notwendige Daten überschrieben werden. Um dieses Problem zu umgehen, wird ein zweiter gleichlanger Zwischenspeicher angelegt. Nach jedem Datenblock wird das Ziel für den nächsten Schreibvorgang auf den jeweils anderen Speicher gesetzt, somit ist sichergestellt, dass nie auf demselben Speicher geschrieben und gearbeitet werden kann. Folgende Grafik veranschaulicht dieses sogenannte Ping-Pong Verfahren¹:

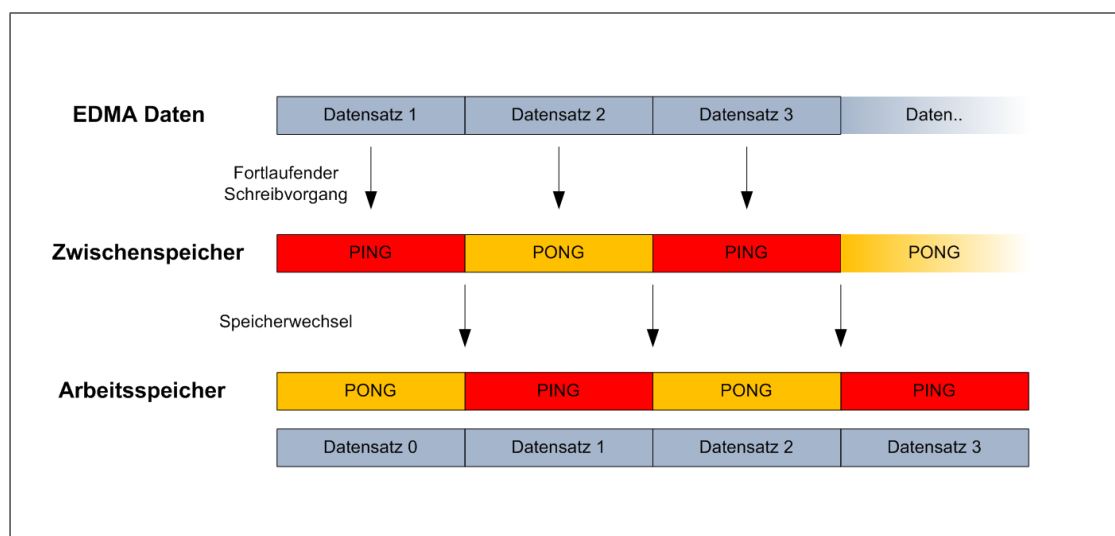


Abbildung 3.4: Arbeitsweise, EDMA Ping-Pong Verfahren

Diese Herangehensweise stellt sicher, dass zur Verarbeitung der Daten fast die gesamte Empfangszeit eines neuen Datenblocks zur Verfügung steht.

3.4 Mikrofonarrays

Zur Datengewinnung stehen zwei Arraysysteme zur Verfügung, ein lineares Arraysystem und ein zirkulares Array. Das zirkulare Array ist mit 8 festen Mikrofonen im Radius von 5 cm bestückt. Die Datenanbindung und Stromversorgung erfolgt über VGA durch ein Adaptermodul. Das lineare System besteht aus mehreren Elementen:

¹Arbeitsspeicher bezeichnet in diesem Fall das Zwischenspeicher-Element auf dem die Datenverarbeitung stattfindet.

- Ein Metallrahmen, in den bis zu 15 Mikrofonmodule im Abstand von 5 cm eingesetzt werden können
- Separate Mikrofonmodule zum Empfang und zur Verstärkung der Audiosignale
- Eine Rahmenhalterung, die gleichzeitig bis zu 8 Mikrofonensignale über Klinkenbuchsen annimmt und über einen VGA-Anschluss gebündelt weiterleitet
- Einem, zum zirkularen Array identischen, Adaptermodul sowohl zur Stromversorgung der Mikrofone über Klinkenkabel als auch zur Adaptierung des VGA-Anschlusses auf den Stecker des DSP-Boards

Die Signale werden beim linearen Array von den Mikrofonen über einen Operationsverstärker innerhalb des Mikrofonmoduls verstärkt und über die Halterung via VGA an den Versorgungsadapter weitergeleitet. Im zirkularen Array geschieht dasselbe, es ist allerdings keine Halterung notwendig. Die Schaltpläne beider Geometrien sind im Abschnitt [D] des Anhangs zu finden.

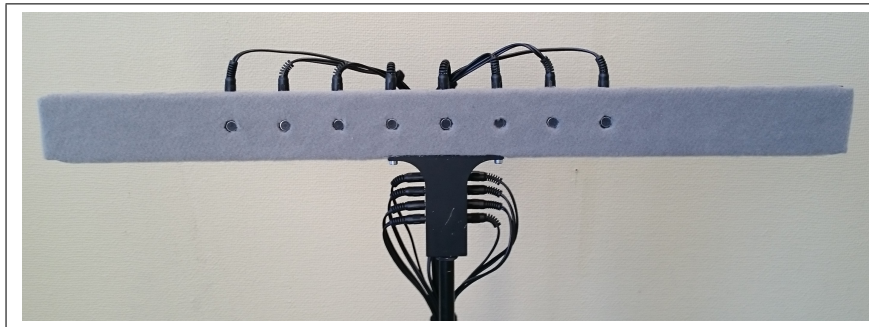


Abbildung 3.5: Lineares Mikrofonarray

4 Theoretische Grundlagen

Zunächst wird in diesem Kapitel das Signalmodell und die mathematische Beschreibung des Mikrofonarrays dargestellt. Im Folgenden wird anhand eines Beispiels die Funktionsweise von Beamforming erläutert.

Nachdem das Grundprinzip erklärt ist, wird auf unterschiedliche Rauschumgebungen eingegangen. Jede dieser Umgebungen modelliert mathematisch das Rauschverhalten in einer spezifizierten Umgebung wie z.B. in einem Büro. Diese Modellierung ist von entscheidender Bedeutung für den späteren Filterentwurf.

Um Filter entwerfen zu können, müssen weiterhin Faktoren bestimmt werden, die die Qualität des Beamformers beschreiben. Die drei wesentlichen Kenngrößen werden in Abschnitt 4.3 vorgestellt.

Die Auswirkungen verschiedener Arraygeometrien bzw. Parameter werden in Abschnitt 4.5 erläutert, bevor mit diesen Kenntnissen räumliche Filter (Beamformer) entwickelt und in Simulationen getestet werden.

4.1 Signal- und Arraymodell

Um den in den folgenden Abschnitten zu Grunde liegenden Modellaufbau zu verdeutlichen, wird zunächst ein Aufbau mit einer Quelle und einem linearen Array betrachtet. Die Umgebung wird als ideal reflexionsfrei angenommen.

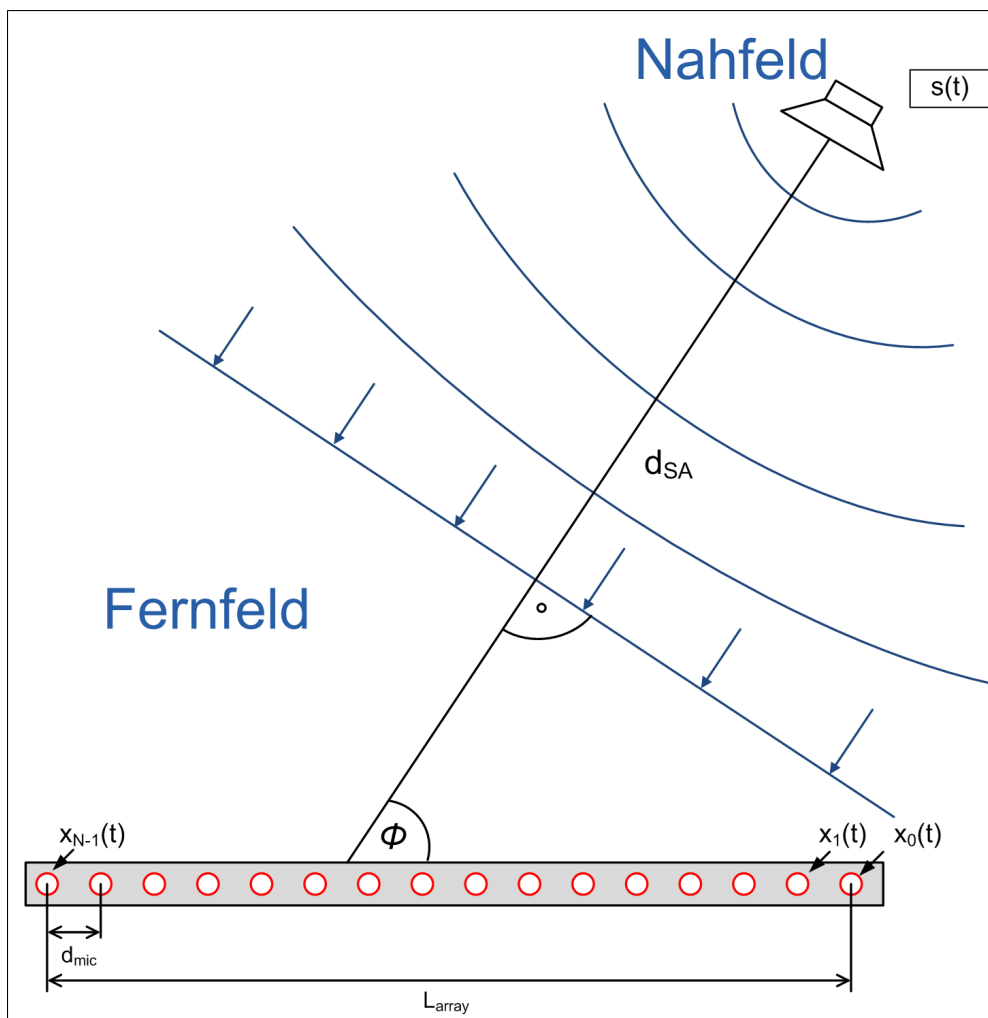


Abbildung 4.1: Linear angeordnetes Mikrofonarray

Das Quellsignal $s(t)$ breitet sich omnidirektional aus und wird je nach Form des Arrays zeitverzögert an den Mikrofonen empfangen. Die Signale an den Mikrofonen werden als $x_0(t)$ am Referenzmikrofon bis $x_{N-1}(t)$ am letzten Mikrofon bezeichnet. Der Zeitbezug wird an dieser Stelle auf Grund der digitalen Verarbeitung durch eine diskrete Indizierungsgröße ersetzt. Wie dem Bild zu entnehmen ist, wird zunächst davon ausgegangen, dass das Quellsignal als planare Wellenfront an den Mikrofonen ankommt. Diese Annahme setzt voraus, dass die Entfernung d_{AS} zwischen Quelle und Array so groß ist, dass sich das Array im Fernfeld der Quelle befindet. Hierfür muss die Fernfeldbedingung:

$$d_{AS} \gg \frac{L_{\text{Array}}^2}{v} f \quad (4.0a)$$

bzw.

$$d_{AS} \gg \frac{L_{\text{Array}}^2}{\lambda} \quad (4.0b)$$

erfüllt sein¹.

Um eine realere Nachbildung zu erreichen, wird nun ein zufälliges Rauschen $v_i(t)$ an jedem Mikrofon hinzugefügt. Zusammengefasst lässt sich das Verhalten des Modells in Matrixform wie folgt beschreiben [2, S. 20ff]:

$$\begin{pmatrix} x_0(k) \\ x_1(k) \\ \vdots \\ x_{N-1}(k) \end{pmatrix} = \begin{pmatrix} a_0 s(k - \tau_0(\phi)) \\ a_1 s(k - \tau_1(\phi)) \\ \vdots \\ a_{N-1} s(k - \tau_{N-1}(\phi)) \end{pmatrix} + \begin{pmatrix} v_0(k) \\ v_1(k) \\ \vdots \\ v_{N-1}(k) \end{pmatrix} \quad (4.1)$$

$$\mathbf{x}(k) = \boldsymbol{\alpha} \mathbf{s}(k - \boldsymbol{\tau}(\phi)) + \mathbf{v}(k) \quad (4.2)$$

Die Zeitverzögerungen mit Bezug auf das Referenzmikrofon x_0 werden durch τ_i beschrieben. Die Dämpfungsfaktoren α_i sind für eine ideale Fernfeldausbreitung ohne Reflexionen und mit omnidirektionalen Mikrofonen einheitlich.

4.2 Mikrofonabstand und räumliche Effekte

Die Betrachtung periodischer Signale beinhaltet eine Mehrdeutigkeit von Signalwerten im Sinne des Abtastzeitpunktes [sinngemäß 13, S.42]. Im räumlichen Sinn bedeutet dies, dass Signale die z.B. an einem Mikrofon eine Phase α aufweisen, denselben Wert hätten, wenn Sie die Phase $2\pi - \alpha$ aufweisen. Es ist also keine eindeutige räumliche Zuordnung möglich. Um diesen Effekt zu vermeiden, müssen die Phasenwinkel, die durch das Array erzeugt werden auf maximal π eingeschränkt werden. Für beliebige Arrays gilt entsprechend:

¹ v ist in diesem Zusammenhang die Ausbreitungsgeschwindigkeit der Schallwelle, also die Schallgeschwindigkeit. Sie beträgt bei Raumtemperatur $343 \frac{\text{m}}{\text{s}}$.

$$\Omega \frac{d_{mic}}{v} \varepsilon \leq \pi$$

ε beschreibt die Phasenänderungen die durch Azimuth- und Elevationswinkel auftreten können. Da diese sich aus Multiplikationen von Sinus- und Kosinusfunktionen zusammensetzen, ist der Betrag maximal 1. Mit $\Omega = 2\pi f$ ergibt sich:

$$\begin{aligned} 2\pi f \frac{d_{mic}}{v} &\leq \pi \\ d_{mic} &\leq \frac{v}{2f} \end{aligned} \quad (4.3a)$$

Oder bezogen auf die Wellenlänge mit $\lambda = \frac{v}{f}$:

$$d_{mic} \leq \frac{\lambda}{2} \quad (4.3b)$$

Der räumliche Alias-Effekt gibt harte Begrenzungen für die Arraygröße nach oben hin vor. Werden diese nicht eingehalten, ist keine sinnvolle Filterung mehr möglich. Auf der anderen Seite gibt es ebenfalls Beschränkungen, wie klein der Mikrofonabstand gewählt werden sollte. Generell gilt, je kleiner der Abstand zwischen den Mikrofonen, desto geringer die Zeitdifferenz der empfangenen Signale. Zieht man diesen Umstand in Betracht, ist es leicht zu sehen, dass bereits kleine Ungenauigkeiten in z.B. der Empfangswinkelvorgaben oder der später entworfenen Filterkoeffizienten große Fehler nach sich ziehen können. Weiterhin gibt der Abstand zwischen den Mikrofonen vor, wie gut das Array in der Lage ist, räumlich zu differenzieren bzw. aufzulösen. Im Beamforming bedeutet dies, dass bei großen Mikrofonabständen schmalere Empfangskeulen gebildet werden können als bei kleinen Abständen.

Zusammengefasst sollte der Mikrofonabstand also möglichst nah an der maximalen Grenze, die durch den Alias-Effekt vorgegeben wird, gewählt werden.

4.3 Evaluation von Beamformern

Bevor Beamformer entworfen werden können, ist es notwendig Qualitätsmaße bzw. Größen zu definieren. Mithilfe dieser Messgrößen kann während und nach dem Entwurf bestimmt werden, in welchem Umfeld der Filter wie performant arbeitet. Die folgenden Erläuterungen orientieren sich an Quelle [2, S. 20ff]

Gleichung 4.2 beschreibt das mathematische Verhalten des Arrays. Da im Beamforming vor allem die Frequenz und der Phasenversatz zwischen den Mikrofonsignalen eine Rolle spielt, bietet es sich an die Evaluation im Frequenzbereich vorzunehmen. Die Anwendung der Fourier Transformation ergibt ohne Einschränkung der Allgemeingültigkeit:

$$\mathbf{X}(e^{j\Omega}) = S(e^{j\Omega})\mathbf{d}(e^{j\Omega}, \tau(\phi)) + \mathbf{V}(e^{j\Omega}) \quad (4.4)$$

Der Vektor \mathbf{d} repräsentiert den Dämpfungs- und den Phasenfaktor des Arrays und wird immer als Kleinbuchstabe angegeben. Er wird im allgemeinen als Array-Response-Vektor oder steering-Vektor bezeichnet [1, S.7] und wie folgt mathematisch beschrieben:

$$\mathbf{d}(e^{j\Omega}, \phi) = [\alpha_0 e^{-j\Omega\tau_0(\phi)}, \alpha_1 e^{-j\Omega\tau_1(\phi)}, \dots, \alpha_{N-1} e^{-j\Omega\tau_{N-1}(\phi)}]^T \quad (4.5)$$

Die Zeitverzögerungen τ_i sind abhängig von der Form des Arrays. Im Abschnitt Geometrie wird hierauf näher eingegangen. Mit diesen Definitionen ergibt sich der gefilterte Ausgang des Arrays als

$$Y_b(e^{j\Omega}) = \sum_{n=0}^{N-1} W_n^*(e^{j\Omega}) X_n(e^{j\Omega}) = \mathbf{W}^H \mathbf{X} \quad (4.6)$$

W_n^* bzw. \mathbf{W}^H stellt hierbei die komplexwertigen Beamformer-Koeffizienten im Frequenzbereich dar. Es hat sich hierbei durchgesetzt, diese in komplex konjugierter Form anzugeben. Weiterhin wird im Folgenden, sowie in der meisten Literatur, die Abhängigkeit der einzelnen Größen von Winkel und Frequenz als gegeben vorausgesetzt und nur an besonders relevanten Stellen explizit mit angegeben.

4.3.1 Beampattern

Das sogenannte Beampattern ist die grundlegendste Art, das Verhalten eines Beamformers darzustellen. Es beschreibt das exakte Eingangs-/Ausgangsverhalten für jeden Winkel bei einer oder mehrerer Frequenzen¹. Allgemein lässt sich das Beampattern in zwei Formen angeben:

$$|R(\phi, \Omega)|^2 = |\mathbf{W}^H \mathbf{d}|^2 \quad (4.7)$$

¹Der Übersichtlichkeit halber wird nur eine Winkelabhängigkeit dargestellt. Für Arrays mit mehr als 2 Dimensionen besteht eine weitere Abhängigkeit vom Elevationswinkel.

[1, S.8]

oder als normierte Größe

$$|H(\phi, \Omega)|^2 = \frac{|\mathbf{W}^H \mathbf{d}|^2}{\mathbf{W}^H \Gamma_{\mathbf{v}} \mathbf{W}} \quad (4.8)$$

[2, S.22]

$\Gamma_{\mathbf{v}}$ ist die Kohärenzmatrix des Rauschens. Auf diese wird in Abschnitt 4.4 noch genauer eingegangen.

Betrachtet man das Beampattern bei einer festen Frequenz Ω_0 ist die Gleichung nur noch vom Angriffswinkel¹ Φ abhängig und man kann die Funktion in einem 2-dimensionalen Koordinatensystem darstellen. In vielen Fällen bietet es sich an, den Verlauf über 360° in einem Kreisdiagramm zu zeichnen. Ein Beispiel für beide Darstellungen zeigt Abbildung 4.2.

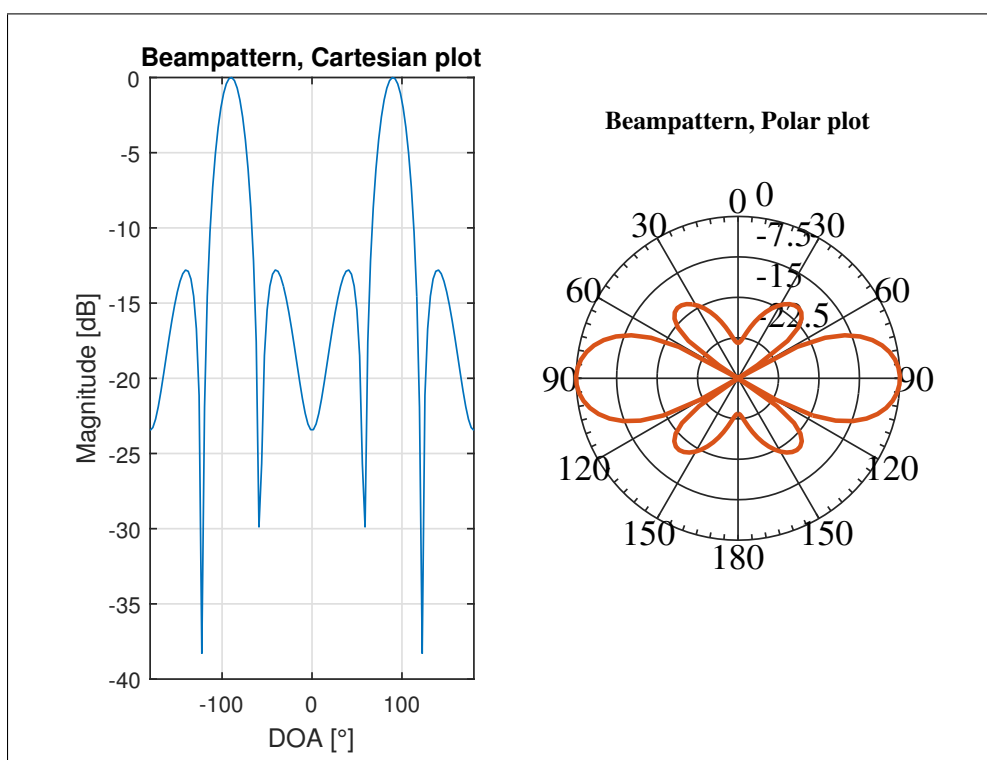


Abbildung 4.2: Darstellungsbeispiel für ein Beampattern

Das Beampattern hilft vor allem dabei, zu erkennen aus welchen Richtungen Störsignale effektiv unterdrückt und Nutzsignale optimal verstärkt werden. Keine offensichtliche Information liefert das Diagramm darüber, in welcher Weise dem Signal überlagertes Rauschen durch den räumlichen Filter beeinflusst wird. Diese Leistungsparameter müssen also über andere Größen ermittelt werden.

¹In der Literatur oft als Direction Of Attack, DOA bezeichnet

4.3.2 Array-Gain

Ein klassisches Maß für Signalqualität stellt das Signal zu Rausch Verhältnis (Signal-to-Noise Ratio, SNR) dar. Es beschreibt das Verhältnis von Signalleistung zu Rauschleistung und wird allgemein angegeben als¹

$$\text{SNR} = \frac{P_{\text{Nutz}}}{P_{\text{Rausch}}} \quad (4.9)$$

Der Array-Gain G setzt das SNR des gesamten Arrays ins Verhältnis zum SNR eines einzelnen Sensors²

$$G = \frac{\text{SNR}_{\text{Array}}}{\text{SNR}_{\text{Sensor}}} \quad (4.10)$$

Für stationäre Signale ergeben sich die Leistungen aus den spektralen Leistungsdichtespektren (LDS) von Nutzsignal und mittlerem Rauschen. Zunächst muss also das LDS am Ausgang des Arrays bestimmt werden:

$$\Phi_{Y_b Y_b} = \mathbf{W}^H \Phi_{\mathbf{X} \mathbf{X}} \mathbf{W} \quad (4.11)$$

$\Phi_{\mathbf{X} \mathbf{X}}$ ist die spektrale Leistungsdichtematrix der Mikrofon-signale und berechnet sich aus den einzelnen LDS:

$$\Phi_{\mathbf{X} \mathbf{X}} = \begin{pmatrix} \Phi_{X_0 X_0} & \Phi_{X_0 X_1} & \cdots & \Phi_{X_0 X_{N-1}} \\ \Phi_{X_1 X_0} & \Phi_{X_1 X_1} & \cdots & \Phi_{X_1 X_{N-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{X_{N-1} X_0} & \Phi_{X_{N-1} X_1} & \cdots & \Phi_{X_{N-1} X_{N-1}} \end{pmatrix} \quad (4.12)$$

Formel 4.11 lässt sich separat für zwei unterschiedliche Fälle betrachten. Im Fall, dass nur das Nutzsignal S vorhanden ist, vereinfacht sich der Ausdruck zu:

$$\Phi_{Y_b Y_b} |_{\text{Signal}} = \Phi_{SS} |\mathbf{W}^H \mathbf{d}|^2 \quad (4.13)$$

¹Die Abhängigkeit der Größen von Ω wird als aus Übersichtlichkeitsgründen nicht explizit mit angegeben

²Rechnung und Formeln anhand [2, S.21ff].

Im Fall, dass nur Rauschen an den Mikrofonen auftritt erhält man:

$$\Phi_{Y_b Y_b} |_{\text{Rausch}} = \Phi_{V_a V_a} \mathbf{W}^H \Phi_{\mathbf{V}\mathbf{V}} \mathbf{W} \quad (4.14)$$

$\Phi_{\mathbf{V}\mathbf{V}}$ beschreibt die normalisierte Kreuz-Leistungsdichtematrix des Rauschens. Die Normalisierung sorgt dafür, dass die Spur (Trace) der Matrix gleich der Anzahl der Mikrofone N ist.

Der Quotient beider Gleichungen ergibt den Array-Gain:

$$G = \frac{|\mathbf{W}^H \mathbf{d}|^2}{\mathbf{W}^H \Phi_{\mathbf{V}\mathbf{V}} \mathbf{W}} \quad (4.15)$$

Wobei sich der Quotient $\frac{\Phi_{SS}}{\Phi_{V_a V_a}}$ mit $\text{SNR}_{\text{Sensor}}$ aus der Gleichung eliminiert. Anhand der hergeleiteten Gleichung kann man erkennen, dass der SNR-Gewinn, der durch Beamforming möglich ist, neben den Filterkoeffizienten \mathbf{W}^* auch von der Geometrie und vor allem der Einsatzumgebung (Rauschumgebung) abhängt. Diese Abhängigkeit wird sich auch bei anderen Qualitätsmaßen zeigen. Daher wird im späteren Verlauf auf einige wichtige Rauschumgebungen eingegangen.

Unter der Annahme eines homogenen Rauschfeldes kann $\Phi_{\mathbf{V}\mathbf{V}}$ durch die Kohärenzmatrix¹ $\Gamma_{\mathbf{V}\mathbf{V}}$ dargestellt werden:

$$\Gamma_{\mathbf{V}\mathbf{V}} = \begin{pmatrix} 1 & \Gamma_{X_1 X_1} & \dots & \Gamma_{V_0 V_{N-1}} \\ \Gamma_{V_1 V_0} & 1 & \dots & \Gamma_{V_1 V_{N-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_{V_{N-1} V_0} & \Gamma_{V_{N-1} V_1} & \dots & 1 \end{pmatrix} \quad (4.16)$$

Für die Matrixelemente gilt:

$$\Gamma_{V_n V_m} (e^{j\Omega}) = \frac{\Phi_{V_n V_m} (e^{j\Omega})}{\sqrt{\Phi_{V_n V_n} (e^{j\Omega}) \Phi_{V_m V_m} (e^{j\Omega})}} \quad (4.17)$$

Eingesetzt in Gleichung 4.15 ergibt sich

$$G = \frac{|\mathbf{W}^H \mathbf{d}|^2}{\mathbf{W}^H \Gamma_{\mathbf{V}\mathbf{V}} \mathbf{W}} \quad (4.18)$$

¹Die Kohärenz entspricht einer normalisierten Form der Kreuz-Leistungsdichte [vgl. 9, S.21]

Da viele theoretische Rauschfelder über ihre Kohärenzfunktion beschrieben werden können, wird im weiteren Verlauf hauptsächlich diese Gleichung verwendet.

4.4 Rauschumgebung

Wie im vorherigen Abschnitt gezeigt, hängt die Leistung eines Beamformers essentiell von der Umgebung ab. Um verwendbare Simulations bzw. Entwurfsverfahren entwickeln zu können, müssen also Beschreibungen für unterschiedliche Rauschfelder gefunden werden.

Es gibt in der Literatur einschlägige Untersuchungen zu vielen verschiedenen Umgebungen. Eine solche ist z.B. das in [2, S.37] referenzierte Werk „Spatial-correlation functions for various noise models“ von B.F. Cron und C.H. Sherman aus dem Jahr 1962.

Wie in der Einleitung beschrieben, konzentriert sich diese Arbeit in erster Linie auf zwei unterschiedliche Anwendungsszenarien. Im Folgenden wird nach Modellen gesucht, die die dargestellten Situationen möglichst passend beschreiben.

4.4.1 Diffuse Noise Fields

Das erste Szenario beschreibt ein Büro, in dem eine Telefon- bzw. Videokonferenz stattfindet. Durch die begrenzte Größe des Büros kann davon ausgegangen werden, dass die Sensoren in dieser Umgebung Rauschsignale von in etwa gleicher Energie empfangen. Weiterhin treten Reflexionen an Wänden und Decken auf, sodass von mindestens schwacher Korrelation zwischen den Signalen ausgegangen werden muss.

Diese Arten von Umgebung, hierzu zählen zum Beispiel neben Büros auch Autos, werden „diffuse Rauschfelder“ („diffuse noise fields“) genannt¹ [9, S.21f]. Ihre Kohärenzmatrix ist durch die sinc-Funktion beschrieben:

$$\Gamma_{m,n}(e^{j\Omega})|_{\text{diffuse}} = \text{sinc}\left(\frac{\Omega d_{m,n}}{v}\right) \quad (4.19)$$

mit $d_{m,n}$ als Distanz zwischen den Mikrofonen m und n in Metern und v der Ausbreitungsgeschwindigkeit der Signale in Meter pro Sekunde.

Der Verlauf der Kohärenzfunktion ist in Abbildung 4.3 dargestellt, zu sehen ist die Kohärenzfunktion über der Frequenz f für Mikrofone mit den Abständen $d_1 = 0.05$ m, $d_2 = 0.10$ m und $d_3 = 0.20$ m.

¹In der Literatur findet sich in seltenen Fällen auch die Bezeichnung „spherical isotropic noise“

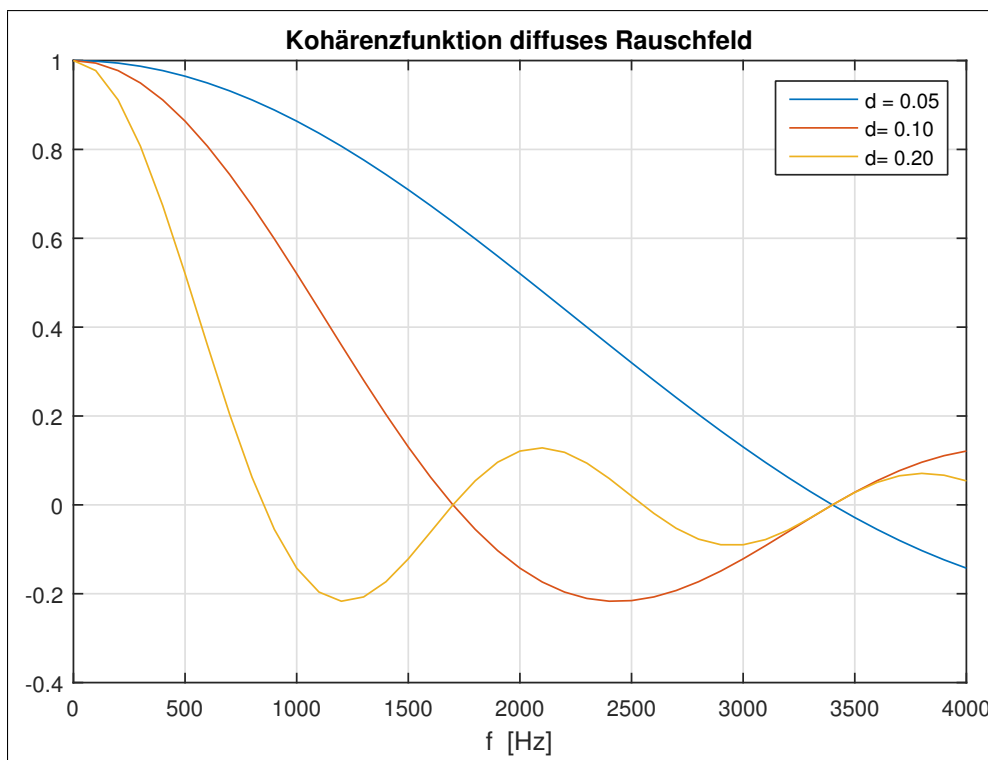


Abbildung 4.3: Kohärenzfunktionen für ein diffuses Rauschfeld

Man kann deutlich erkennen, dass die Kohärenz für nah beieinander liegende Mikrofone über einen großen Frequenzbereich hohe Werte liefert, während Sie für größere Abstände schnell absinkt.

Setzt man 4.19 in 4.18 ein, erhält man den wichtigen **Directivity Faktor**:

$$D = \frac{|\mathbf{W}^H \mathbf{d}|^2}{\mathbf{W}^H \Gamma_{\mathbf{v}\mathbf{v}}|_{\text{diffuse}} \mathbf{W}} \quad (4.20a)$$

In Dezibel dargestellt, wird er oft auch als **Directivity Index** beschrieben:

$$\text{DI [dB]} = 10 \log_{10}(D) \quad (4.20b)$$

Dieser gibt an, wie gut ein Mikrofonarray in der Lage ist, ein diffuses Rauschfeld zu unterdrücken.

4.4.2 Cylindrical Isotropic Noise

Das zweite Szenario behandelt einen großen (Hör-)Saal. In dieser Umgebung kann man davon ausgehen, dass Reflektionen von Wänden und Decken aufgrund der großen Entfernung zum Array minimal sind. Die Audiosignale werden also ohne Reflektion direkt empfangen, entsprechend

sind sie an jedem Mikrofon miteinander korreliert. Ein Modell, das dieser Situation entspricht ist nicht direkt in der betrachteten Literatur enthalten. Ein mögliches Modell behandelt eine sogenannte „cylindrical isotropic noise“ Umgebung oder auch „Cocktailparty Noise“. Die Annahmen sind hierbei keine Reflexionen von Wänden und Decken (open air) und viele Sprecher um das Array herum. Dieses Umfeld ist beschrieben durch die Besselfunktion erster Art[siehe 2, S.30]:

$$\Gamma_{m,n}(e^{j\Omega})|_{\text{cocktail}} = J_0\left(\frac{\Omega d_{m,n}}{c}\right) \quad (4.21)$$

Abbildung 4.4 zeigt erneut den Verlauf über der Frequenz für verschiedene Mikrofonabstände. Man kann erkennen, dass die Kohärenz für alle Abstände deutlich höhere Werte annimmt.

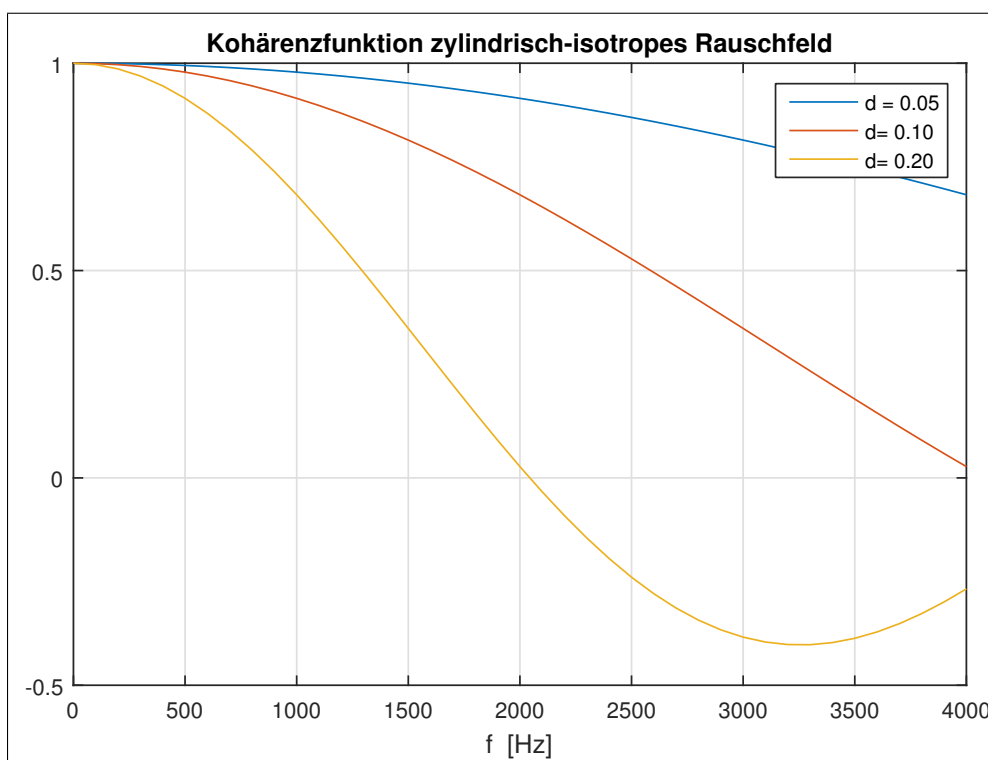


Abbildung 4.4: Kohärenzfunktionen für ein zylindrisch isotropes Rauschfeld

4.4.3 Spatially Uncorrelated Noise

Ein Rauschfeld welches in der realen Welt selten vorkommt aber bei SNR Betrachtungen sehr wichtig ist, ist das räumlich-unkorrelierte oder inkohärente Rauschfeld. Die zugehörige Kohärenzmatrix ist die Einheitsmatrix:

$$\Gamma_{m,n}(e^{j\Omega})|_{\text{uncorrelated}} = \mathbf{I} \quad (4.22)$$

In Kombination mit 4.18 erhält man den **White Noise Gain** [2, S.24]. Dieser Faktor spiegelt wider, wie gut ein Mikrofonarray weißes Rauschen unterdrückt. Ein hoher WNG Wert entspricht hierbei einer guten Unterdrückung während ein geringer Wert einer Verstärkung gleichkommt:

$$\text{WNG} = \frac{|\mathbf{W}^H \mathbf{d}|^2}{\mathbf{W}^H \Gamma_{\mathbf{v}\mathbf{v}}|_{\text{uncorrelated}} \mathbf{W}}$$

$$\text{WNG} = \frac{|\mathbf{W}^H \mathbf{d}|^2}{\mathbf{W}^H \mathbf{W}} \quad (4.23)$$

4.5 Array Geometrie

Verschiedene Anwendungsfälle erfordern unterschiedliche Formen von Mikrofonarrays. In diesem Kapitel werden zwei grundlegende Geometrien für Arrays betrachtet, das lineare und das zirkulare Array. Es wird darauf eingegangen, welche Arrays für welche Zwecke geeignet sind sowie deren amthematische Beschreibung hergeleitet.

4.5.1 Lineare Arrays

4.5.1.1 Uniform linear Array

Die einfachste Form von Mikrofonarrays sind Aufbauten mit linearer Mikrofonanordnung. Lineare Arrays sind aufgrund der 2-Dimensionalität nur in der Lage Filterung entlang des Azimuthwinkels ϕ umzusetzen¹.

Die Länge des Arrays L_{Array} gibt die maximale Ausdehnung des Arrays wieder. Ist der Abstand zwischen den Mikrofonen konstant d_{mic} , spricht man von einem „Uniform Linear Array (ULA)“. Das bereits in Abbildung 4.1 dargestellte Array erfüllt diese Beschreibung. Für ein ULA ergibt sich die Zeitverzögerung zwischen den einzelnen Mikrofonen x_i zu:

$$\tau_i = \frac{(i-1) d_{\text{mic}} \cos(\phi)}{v} \quad (4.24)$$

Es wird in dieser Gleichung das erste Mikrofon mit x_0 als Referenzmikrofon angesehen. Das bedeutet, dass das anliegende Signal x_0 die Bezugsphase Null hat und alle anderen Phasen an den Mikrofonen sich auf diese Nullphase beziehen.

¹Es ist möglich, lineare Arrays zu kombinieren, um Formen wie Quadrate und Dreiecke zu bilden. Diese können zusätzlich zum Azimuth auch Elevation auflösen. Arrays dieser Form sind jedoch nicht Thema der Arbeit.

Aus den beiden genannten Zusammenhängen lässt sich erkennen, dass die Zeitdifferenz zwischen den Mikrofonen nur abhängig von ϕ ist.

Ein Beispiel für $d_{mic} = 5 \text{ cm}$ und $i = 2$ ergibt den folgenden Zeitverzögerungsverlauf über ϕ :

Aus Gleichung 4.24 und 4.5 lässt sich nun der steering-Vektor eines ULA bestimmen:

$$\mathbf{d}_{\text{ULA}}(\Omega, \phi) = \left[1, e^{-j\Omega d_{mic} \frac{\cos(\phi)}{v}}, e^{-j\Omega 2 d_{mic} \frac{\cos(\phi)}{v}}, \dots, e^{-j\Omega (N-1) d_{mic} \frac{\cos(\phi)}{v}} \right]^T \quad (4.25)$$

Hierbei ist wichtig, dass der steering Vektor eine zusätzliche Abhängigkeit von der Signalfrequenz Ω aufweist. Diese Abhängigkeit erhält man durch die Darstellung der Zeitverzögerung als Phasenwinkel im Frequenzbereich. Die Dämpfungsfaktoren wurden in der Gleichung aufgrund der Fernfeldannahme und idealer Mikrofone zu 1 angenommen.

Weiterhin zeigt sich, dass gerade bei tiefen Frequenzen bzw. großen Wellenlängen die Phasenverschiebungen zwischen den einzelnen Mikrofonensignalen sehr gering werden können. Das bedeutet, dass die räumliche Auflösung bei tiefen Frequenzen deutlich schlechter ist, als bei hohen Frequenzen. Um dieses Problem zu adressieren, können sogenannte harmonisch-eingebettete Unterarrays verwendet werden.

4.5.1.2 Harmonisch-eingebettete Unterarrays

Wie im vorangegangenen Abschnitt beobachtet, hängt die räumliche Auflösung und somit die Breite der Hauptempfangskeule von der Frequenz bzw. Wellenlänge des Empfangssignals ab. Ist die Wellenlänge sehr groß im Verhältnis zum Mikrofonabstand, erhält man eine sehr schlechte räumliche Auflösung und entsprechend breite Empfangskeulen.

Ein Verfahren, um diese Auswirkungen abzuschwächen, ist kein Array mit konstantem Mikrofonabstand zu wählen, sondern mehrere Arrays mit unterschiedlichen Abständen ineinander zu „verschachteln“. Diese Art von Array wird im Allgemeinen als harmonisch-eingebettete Unterarrays („harmonically-nested subarrays“) bezeichnet¹. Die einzelnen Subarrays weisen hierbei meist einen logarithmischen Abstand auf. Oft wird der Faktor zwei als Basis verwendet. Abbildung 4.5 zeigt ein solches Array:

¹Im Sinne der Literatur wird im Folgenden von Subarrays gesprochen.

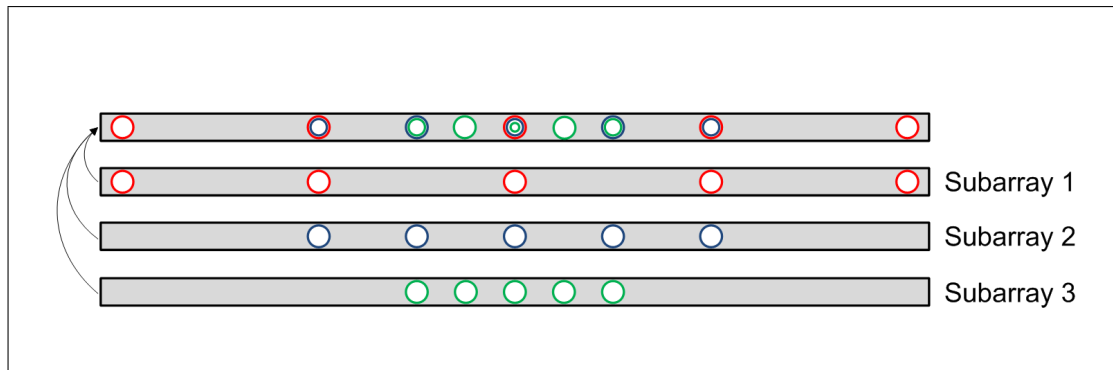


Abbildung 4.5: Array mit nested Subarrays

Dargestellt in der obersten Reihe ist das gesamte Array bestehend aus insgesamt 3 Unterarrays mit jeweils halbem Mikrofonabstand von Unter- zu Unterarray. Jedes Unterarray hat dieselbe Anzahl Mikrofone, außerdem werden Mikrofone doppelt z.T. sogar von allen Arrays genutzt. Das erste Subarray mit dem größten Mikrofonabstand verarbeitet die niederfrequenten Signalanteile, das Zweite die mittleren Frequenzen und das dritte Unterarray die hohen Frequenzanteile. Die Abgrenzung der Frequenzbereiche kann durch entsprechende Filter oder eine Frequenzerlegung z.B. durch DFT erfolgen.

4.5.2 Zirkulare Arrays

Eine weitere Form von Mikrofonarrays bilden zirkulare Anordnungen. Diese haben vor allem den Vorteil, dass sie kompakt sind und neben Azimuth auch Elevationsfilterung möglich machen. Bei zirkularen Mikrofonarrays sind die Elemente im Normalfall in gleichen Abständen auf einem Kreis mit Radius R angeordnet, Abbildung 4.6 zeigt ein solches „Uniform Circular Array (UCA)“:

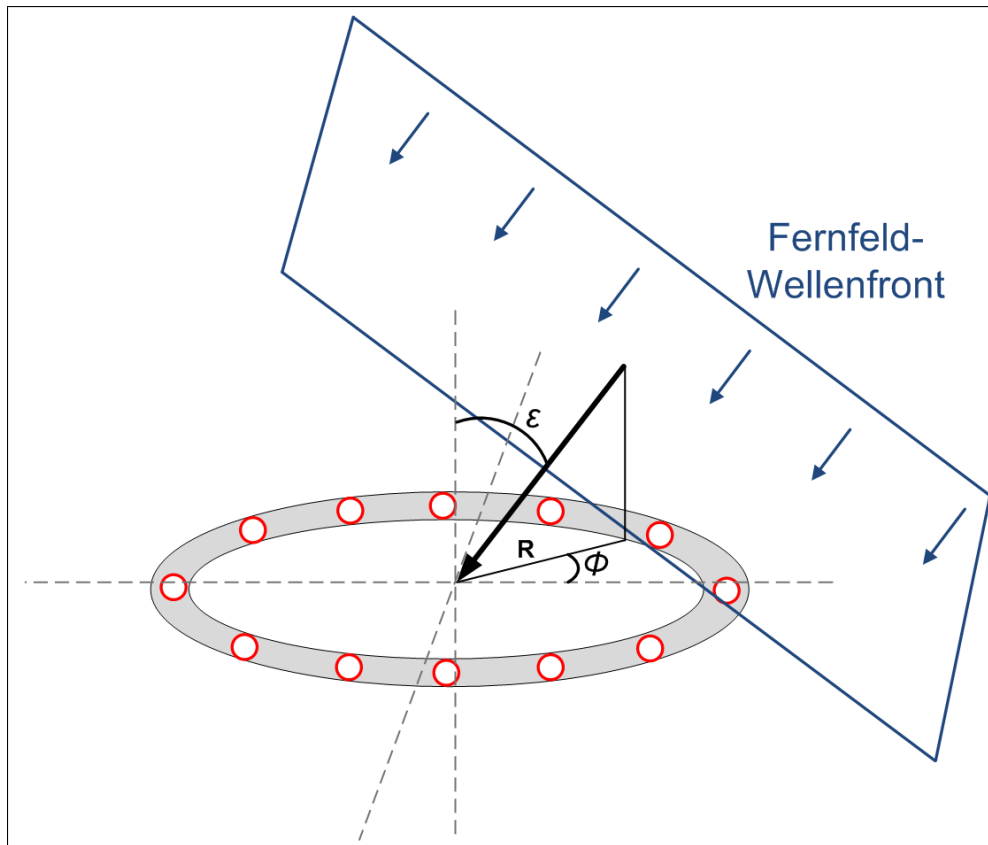


Abbildung 4.6: Zirkulares Mikrofonarray mit konstanten Mikrofonabständen (UCA)

Die mathematische Beschreibung muss entsprechend über drei Dimensionen aufgestellt werden. In Anlehnung an [13, S.69ff] ergibt sich zunächst für den Normalenvektor der Wellenfront aus den Kugelkoordinaten:

$$\mathbf{r}_{WF} = [\sin(\varepsilon) \cos(\phi), \sin(\varepsilon) \sin(\phi), \cos(\varepsilon)]^T \quad (4.26)$$

Die Mikrofonpositionen werden durch Ortsvektoren \mathbf{p}_i beschrieben:

$$\mathbf{p}_i = [R \cos(\phi_i), R \sin(\phi_i), 0]^T \quad (4.27)$$

Unter der Annahme, dass es sich um ein UCA handelt, also alle Mikrofone uniform angeordnet sind, lässt sich 4.27 vereinfachen zu:

$$\mathbf{p}_i = [R \cos(\Delta\phi(i-1)), R \sin(\Delta\phi(i-1)), 0]^T \quad (4.28)$$

mit

$$\Delta\phi = \frac{2\pi}{N} \quad (4.29)$$

Die Laufzeitunterschiede ergeben sich durch das Skalarprodukt aus Ortsvektor des Mikrofons und Normalenvektor der Wellenfront zu:

$$\tau_i = -\frac{1}{v} \mathbf{r}_{WF} \mathbf{p}_i \quad (4.30)$$

Durch Verwendung der Eulerschen Formeln:

$$\cos(a) = \frac{1}{2}(e^{ja} + e^{-ja}) \quad (4.30a)$$

$$\sin(a) = \frac{1}{2j}(e^{ja} - e^{-ja}) \quad (4.30b)$$

erhält man:

$$\tau_i = -\frac{R}{v} \sin(\varepsilon) \cos(\phi - \Delta\phi) \quad (4.31)$$

Diese Zeitdifferenz kann nun genutzt werden, um den steering-Vektor des UCA aufzustellen. Entsprechend Gleichung 4.5 für den 3-dimensionalen Fall ergibt sich:

$$\begin{aligned} \mathbf{d}_{UCA}(\Omega, \phi, \varepsilon) &= [a_0 e^{-j\Omega\tau_0(\phi, \varepsilon)}, a_1 e^{-j\Omega\tau_1(\phi, \varepsilon)}, \dots, a_{N-1} e^{-j\Omega\tau_{N-1}(\phi, \varepsilon)}]^T \\ \mathbf{d}_{UCA}(\Omega, \phi, \varepsilon) &= [a_0 e^{j\Omega \frac{R}{v} \sin(\varepsilon) \cos(\phi)}, a_1 e^{j\Omega \frac{R}{v} \sin(\varepsilon) \cos(\phi - \Delta\phi)}, \dots, a_{N-1} e^{j\Omega \frac{R}{v} \sin(\varepsilon) \cos(\phi - (N-1)\Delta\phi)}]^T \end{aligned} \quad (4.32)$$

Wie beim ULA werden die Dämpfungsfaktoren zunächst zu 1 angesetzt.

Es soll an dieser Stelle angemerkt sein, dass der räumliche Alias-Effekt bei zirkularen Arrays natürlich auch auftritt. Entsprechend gilt Gleichung 4.3a auch hier. Wie man am steering-Vektor erkennen kann, muss jedoch der Mikrofonabstand d_{mic} in der Formel durch den Array-Radius R ersetzt werden:

$$R \leq \frac{v}{2f} \quad (4.33)$$

4.5.3 Vergleich linear - zirkular

Nachdem für beide Arrays die mathematische Beschreibung hergeleitet ist, wird nun auf die wesentlichen Unterschiede zwischen den Arrays eingegangen. Um die Arrays sinnvoll vergleichen zu können, ist die Anzahl der Mikrofone beider Arrays gleich. Um den gleichen Frequenzbereich

ohne Alias-Effekt abzudecken, ist der Mikrofonabstand des linearen Arrays gleich dem Radius des zirkularen Arrays. Die Vorteile des zirkularen Arrays liegen vor allem in seiner Kompaktheit. Während ein zirkulares Array für Sprachsignale im Bereich bis ca. 3,5 kHz ohne Alias-Effekt entsprechend Gleichung 4.33 nur 10 cm Durchmesser hat, wäre ein lineares Array bei gleicher Mikrofonanzahl N_{mic} um den Faktor $\frac{N_{mic}}{2}$ länger. Zudem ist ein zirkulares Array in der Lage durch den drei-dimensionalen Aufbau auch eine Filterung in Elevationsrichtung zu erreichen, was mit gewöhnlichen linearen Arrays nicht möglich ist. Aufgrund der geringen vertikalen Ausdehnung, ist eine geringere Azimuthauflösung zu erwarten. Durch Simulation muss zunächst untersucht werden, ob sich das UCA unter diesen Voraussetzungen zum Beamforming eignet. Das lineare Array hat vor allem Vorteile in der optimalen Azimuthauflösung. Durch seine große vertikale Ausdehnung ist es möglich, wesentlich „schärfere“ Empfangskeulen zu bilden als das entsprechende zirkulare Array. Auch die einfache mathematische Beschreibung, welche im späteren Verlauf zur effizienteren Filterberechnung führt, spricht für das lineare Array. Von Nachteil kann die vertikale Ausdehnung sein, wenn Kompaktheit benötigt wird. Zudem entfällt bei linearen Arrays die Möglichkeit der Filterung in Elevationsrichtung.

5 Entwurf von Beamformern

In diesem Kapitel werden drei verschiedene Verfahren zum Entwurf von Beamformern vorgestellt. Zunächst wird auf den klassischen "Delay-and-Sum" Beamformer eingegangen. Im weiteren Verlauf werden anspruchsvollere Beamformer über Minimierungsansätze betrachtet.

Der Delay-and-Sum Beamformer stellt den Einstieg in die Entwurfsthematik dar. Es wird zunächst eine allgemeine Entwurfsmethode vorgestellt und deren Ergebnisse evaluiert. Diese Ergebnisse zeigen, warum Schmalbandansätze für die in der Einleitung beschriebenen Szenarien ungeeignet sind. Aufbauend auf der Schmalband Entwurfsmethodik wird ein Breitband-Ansatz vorgestellt, welcher im Folgenden für alle weiteren Beamformer ausschließlich angewandt wird. Da im realen Aufbau später maximal 8 Mikrofone mit einem Minimalabstand von 5 cm verwendet werden können, sind die Arrayparameter im Allgemeinen auf diese Werte eingestellt. Durch die Abstandsbeschränkung ergibt sich eine maximale Signalfrequenz an Hand von Gleichung 4.3a zu $\approx 3,5$ kHz.

Die Entwurfsmethoden machen keinen Unterschied zwischen Nahfeld oder Fernfeld Wellenausbreitung. Diese Information kann über das Array-steering \mathbf{d} eingebracht werden. Die gezeigten Simulationsergebnisse gehen zunächst von Fernfeldausbreitung aus.

5.1 Delay-and-Sum

5.1.1 Entwurfsmethodik

Die einfachste Möglichkeit einen gerichteten Empfang durch Mikrofonarrays zu erhalten, bildet der sogenannte „Delay-and-Sum“-Beamformer¹. Hierbei werden die an den Mikrofonen empfangenen Signale unterschiedlich lange verzögert und gewichtet addiert [vgl. 2, S.20].

Aus Gleichung 4.6 und 4.4 lässt sich der folgende Zusammenhang ableiten[2, S20f]:

$$Y_b(e^{j\Omega}) = \mathbf{W}^H [S(e^{j\Omega}) \mathbf{d} + \mathbf{V}(e^{j\Omega})] \quad (5.0a)$$

Da es sich bei $S(e^{j\Omega})$ um das Quellsignal handelt, also ein Skalar, lässt sich die Gleichung auch wie folgt schreiben:

¹Im Folgenden der Übersichtlichkeit halber als D&S-Beamformer Abgekürzt

$$Y_b(e^{j\Omega}) = \mathbf{W}^H \mathbf{d} S(e^{j\Omega}) + \mathbf{W}^H \mathbf{V}(e^{j\Omega}) \quad (5.0b)$$

An Gleichung 5.0b kann man leicht erkennen, dass das Produkt aus Gewichtsvektor \mathbf{W}^H und steering-Vektor \mathbf{d} die räumliche Filtercharakteristik ausmacht. Eine offensichtliche Wahl für Anpassung der Hauptempfangsrichtung ist, den Gewichtsvektor \mathbf{W} gleich einem auf $\frac{1}{N_{mic}}$ normierten steering-Vektor $\mathbf{d}_0(\Omega_0, \phi_0)$ zu setzen [1, S.11]:

$$\mathbf{W}_{D\&S} = \frac{1}{N_{mic}} \mathbf{d}_0 \quad (5.1)$$

Dies führt zu folgendem Ausgangssignal:

$$\begin{aligned} Y_b(e^{j\Omega}) &= \mathbf{W}_{D\&S}^H \mathbf{d} S(e^{j\Omega}) + \mathbf{W}_{D\&S}^H \mathbf{V}(e^{j\Omega}) \Big|_{\mathbf{w}=\frac{1}{N} \mathbf{d}_0} \\ Y_b(e^{j\Omega}) &= \frac{1}{N_{mic}} \mathbf{d}_0^H \mathbf{d} S(e^{j\Omega}) + \frac{1}{N_{mic}} \mathbf{d}_0^H \mathbf{V}(e^{j\Omega}) \end{aligned} \quad (5.2)$$

Die Filterwirkung kommt dadurch zustande, dass der Koeffizientensatz \mathbf{W} für eine feste Frequenz und einen festen Winkel entworfen wird, sich der steering-Vektor \mathbf{d} jedoch abhängig von Frequenz und Winkel ändert. Man erhält für diesen Fall die folgende Empfangscharakteristik für $\phi = \phi_0$ und $\Omega = \Omega_0$:

$$\begin{aligned} Y_b(e^{j\Omega})|_{\Omega=\Omega_0} &= \frac{1}{N_{mic}} \mathbf{d}_0^H \mathbf{d}_0 S(e^{j\Omega_0}) + \frac{1}{N_{mic}} \mathbf{d}_0^H \mathbf{V}(e^{j\Omega_0}) \\ Y_b(e^{j\Omega})|_{\Omega=\Omega_0} &= S(e^{j\Omega_0}) + \frac{1}{N_{mic}} \mathbf{d}_0^H \mathbf{V}(e^{j\Omega_0}) \end{aligned} \quad (5.3)$$

Gleichung 5.3 zeigt, dass für ϕ_0 bei Ω_0 am Ausgang des Beamformers das Quellsignal anliegt. Das Rauschen wird hierbei deutlich erkennbar um den Faktor $\frac{1}{N_{mic}}$ gedämpft.

5.1.2 Beampattern

Uniform Circular Array

Aus dem Vergleich der Arraygeometrien in Kapitel 4.5.3 bleibt die Frage offen, ob sich das vorhandene Zirkulare Array zum Beamforming eignet. Um eine sinnvolle Aussage hierüber treffen zu können, wird zunächst das vorhandene UCA mit Radius $R = 5$ cm und $N_{mic} = 8$ Mikrofonen simuliert. Für maximale Azimuthauflösung wird als Elevationswinkel im Entwurf $\phi_0 = 90^\circ$ angesetzt, der Azimuthwinkel unterliegt auf Grund der Zirkularität keinen Einschränkungen. Als Entwurfsfrequenz wird $f_0 = 2$ kHz angenommen, um das Verhältnis von Wellenlänge zu Arrayradius in einem guten Bereich zu halten. Das Beampattern zeigt folgendes Arrayverhalten:

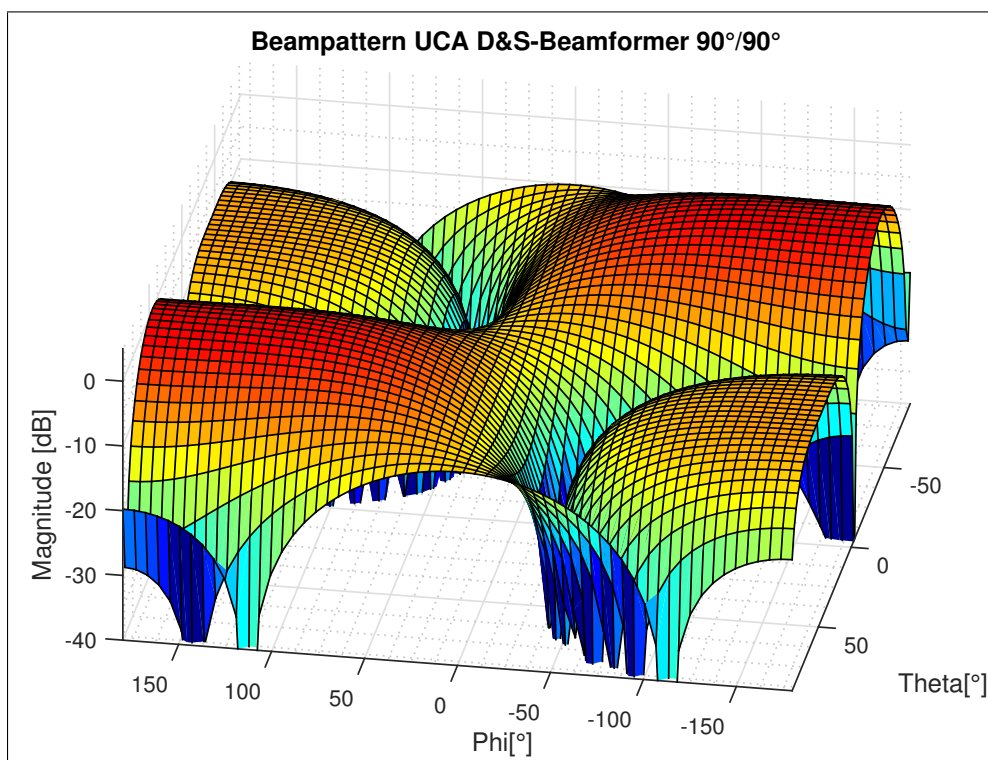


Abbildung 5.1: Beampattern, 90° D&S-Beamformers für ein UCA

Für einen festgesetzten Elevationswinkel von $\theta = \theta_0 = 90^\circ$ über ϕ :

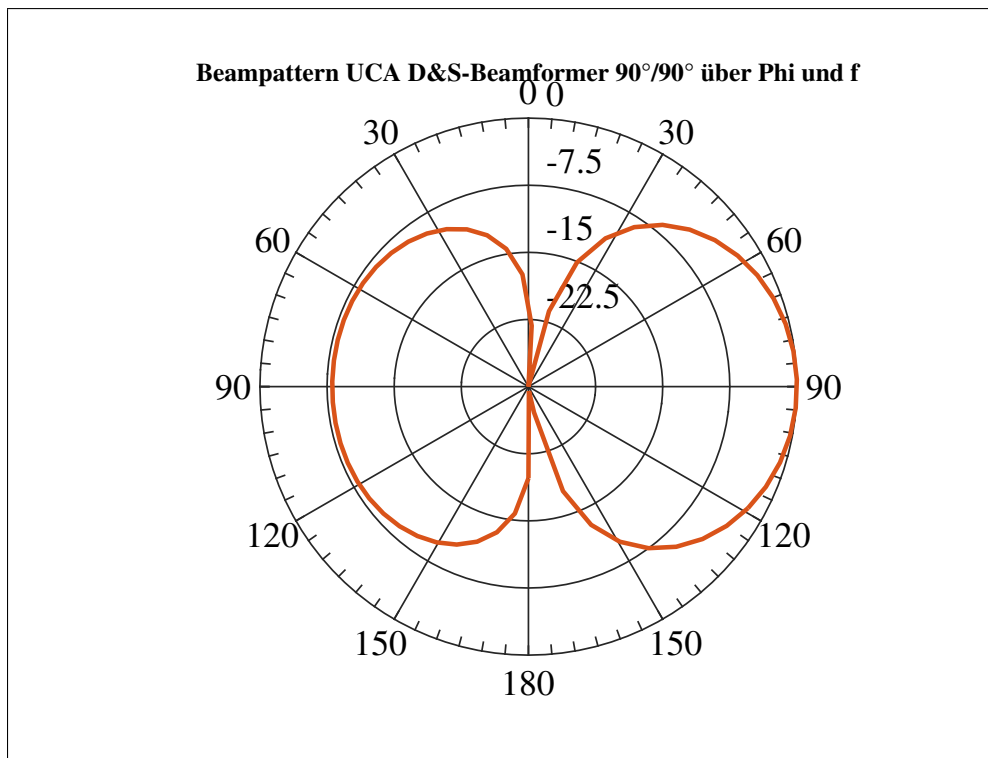


Abbildung 5.2: Beampattern, 90° D&S-Beamformers für ein UCA, Polardarstellung

Die Grafik zeigt, dass kein schmaler Empfangsbereich erzeugt werden kann. Die Beambreite beträgt ca. $\pm 60^\circ$ bei einer Randdämpfung von 9 dB. Die Nebenkeulen formen sich hinter dem Array und besitzen eine eher schwache Dämpfung von 7,5 dB. Da auch die folgenden Beamformer-Entwürfe die Beambreite bei Entwurfsfrequenz kaum beeinflussen, ist das vorhandene Array nicht zum Beamforming geeignet und wird im Folgenden nicht mehr betrachtet.

Uniform Linear Array

Der Vergleich der Arraygeometrien stellt als Haupteigenschaft des linearen Arrays die optimale Azimuthauflösung heraus. Entsprechend sollte das Beampattern des ULA deutlich besser als das des UCA sein. Das Beampattern des D&S-Beamformers ergibt sich aus 4.7 zu:

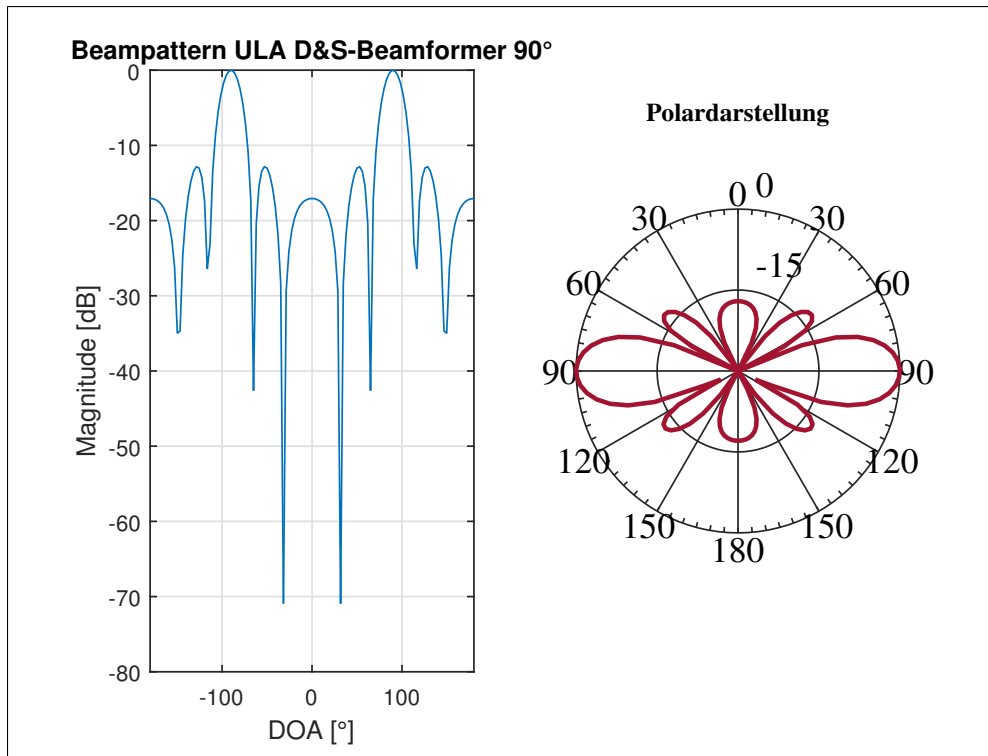


Abbildung 5.3: Beampattern, 90° D&S-Beamformers für ein ULA

Die Hauptkeulen des Beamformers sind auf exakt 90° ausgerichtet, die Verstärkung beträgt 0 dB. Die höchste Nebenkeule liegen für dieses Beispiel im Abstand von $\pm 40^\circ$ bei $-12,8$ dB.

Für die Verarbeitung von Audiosignalen sind allerdings nicht nur einzelne Frequenzen relevant, sondern ein breiter Bereich. Trägt man das Beampattern über den gesamtmöglichen Frequenzbereich bis $\approx 3,5$ kHz auf, erhält man folgendes Bild:

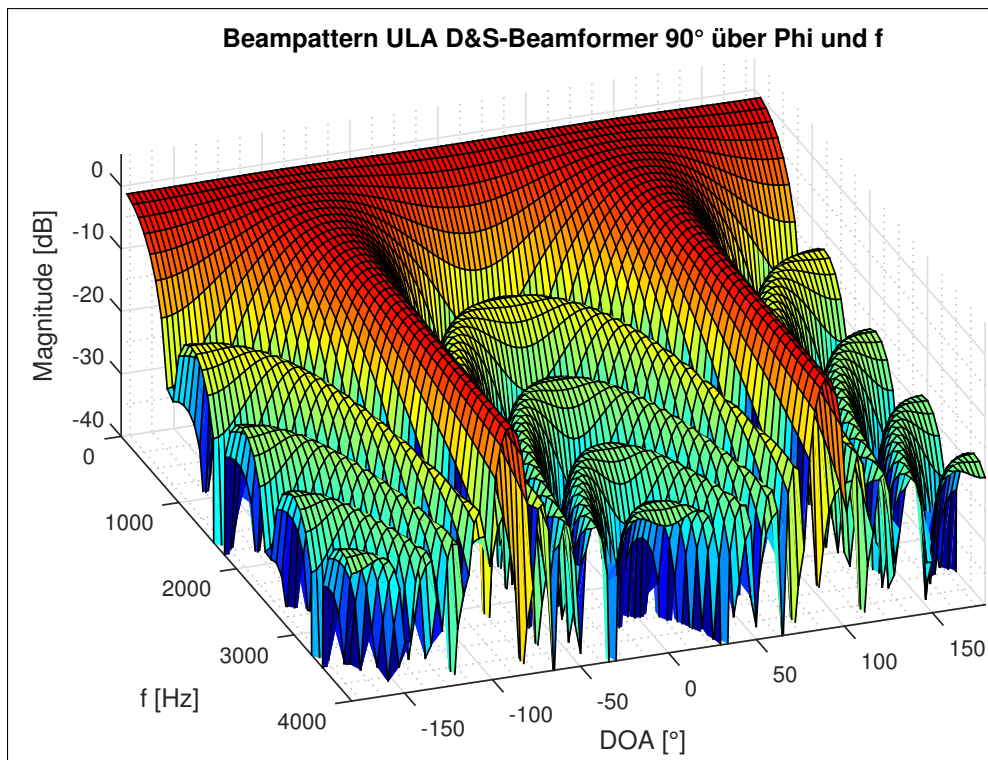


Abbildung 5.4: Beampattern, 90° D&S-Beamformers für ein ULA über Frequenz und Winkel

Man kann erkennen, dass vor allem im unteren Frequenzbereich sehr breite Empfangskeulen entstehen. Es ist somit kaum räumliche Filterwirkung erzielbar. Der Grund hierfür liegt in den sehr großen Wellenlängen im Verhältnis zum Mikrofonabstand. Ist dieses Verhältnis zu groß, ist kaum Phasenverschiebung zwischen den einzelnen Signalen vorhanden. Es ist somit nur schwer möglich, einen schmalen Empfangsbereich zu erzielen.

Für Entwurfswinkel um 0° laufen die Empfangskeulen aufgrund der Symmetrie zusammen. Für diesen Entwurfswinkel lassen sich einige kritische Schlussfolgerungen aus den korrespondierenden Beampattern ziehen:

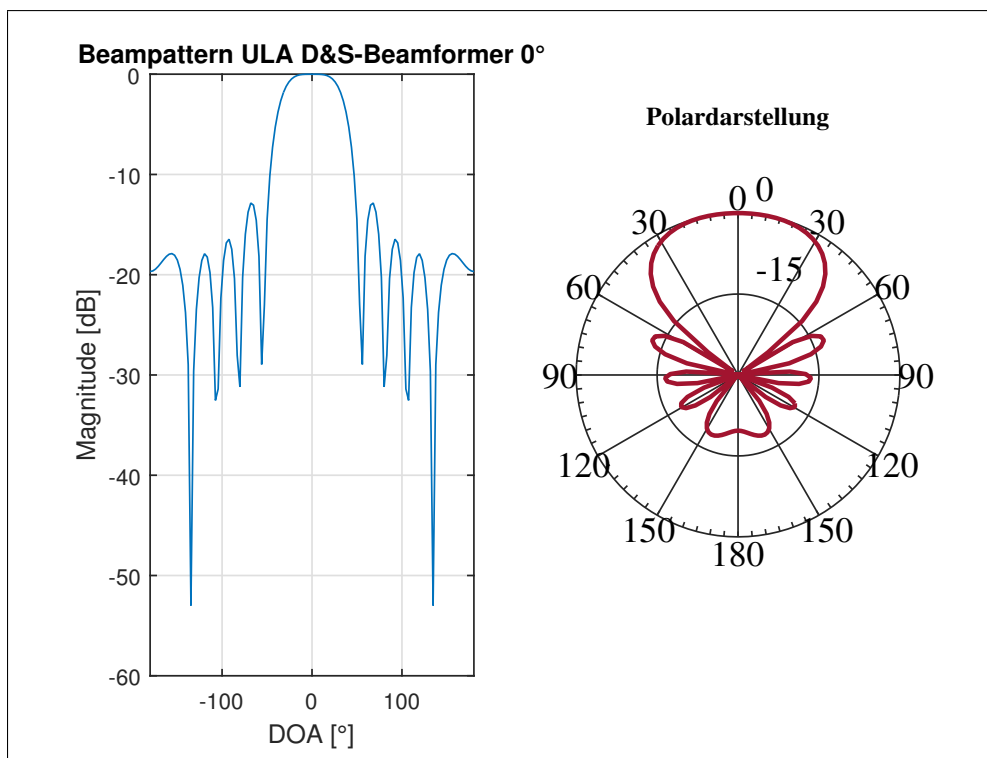


Abbildung 5.5: Beampattern, 0° D&S-Beamformers für ein ULA

Die Betrachtung bei der Entwurfshfrequenz $f_0 = 2$ kHz zeigt, dass die Hauptkeulen sich zusammengeschlossen haben. Weiterhin sind keine Auffälligkeiten festzustellen.

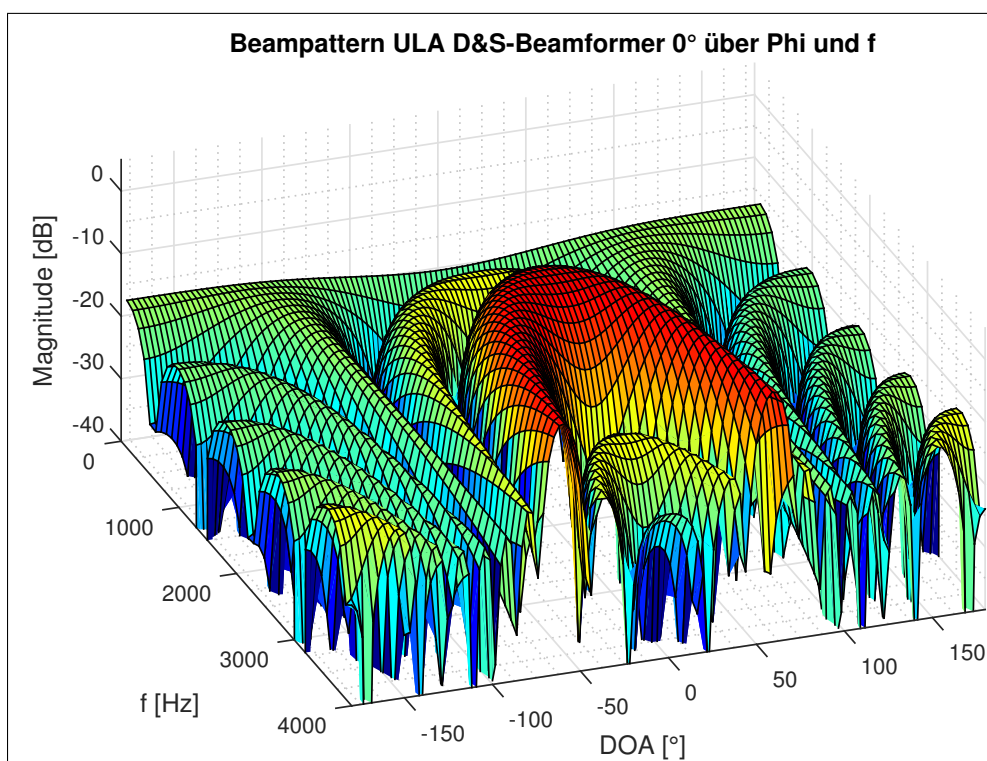


Abbildung 5.6: Beampattern, 0° D&S-Beamformers für ein ULA über Frequenz und Winkel

Das gesamte Beampattern zeichnet ein anderes Bild. Man kann deutlich erkennen, dass jenseits der Entwurfsfrequenz das Empfangsmuster willkürlich erscheint. Die Empfangsrichtung läuft von 2 kHz aufwärts wieder auseinander. Unterhalb von 2 kHz sind keine eindeutigen Hauptkeulen zu erkennen.

Die Ursache liegt in der unterschiedlichen Phasendifferenz. Während für $\phi_0 = 90^\circ$ Entwurfsfrequenz nie eine Phasendifferenz entsteht (vgl. 4.25), ist sie für $\phi_0 = 0^\circ$ maximal. Das bedeutet, je größer die Abweichung der Frequenz von f_0 , desto größer wird der Fehler zur angenommenen Phasendifferenz.

Es ist an diesem Beispiel deutlich erkennbar, dass der bis hierhin dargestellt Entwurf nur für schmalbandige Signale tauglich ist.

5.2 Breitband-Beamforming mit Schmalband Entwurf

Im vorherigen Abschnitt wurde gezeigt, dass für einen statischen Entwurf bei einer Frequenz kein ausreichend breitbandiger Beamformer entworfen werden kann. Abhilfe schafft eine Reihe von Beamformern, die jeweils auf einen bestimmten schmalbandigen Frequenzbereich angewendet werden. Die generelle Entwurfsmethodik wird somit beibehalten, jedoch muss das Signal in schmale Frequenzbereiche zerlegt werden. Die einfachste Möglichkeit dies zu erreichen ist die „Diskrete Fourier Transformation“ (Discrete Fourier Transformation, DFT) bzw. ihre effiziente Implementierung, die „Schnelle Fourier Transformation“ (Fast Fourier Transformation, FFT). Das Blockschaltbild eines Breitband Beamformers mit Frequenzerlegung durch DFT sieht wie folgt aus:

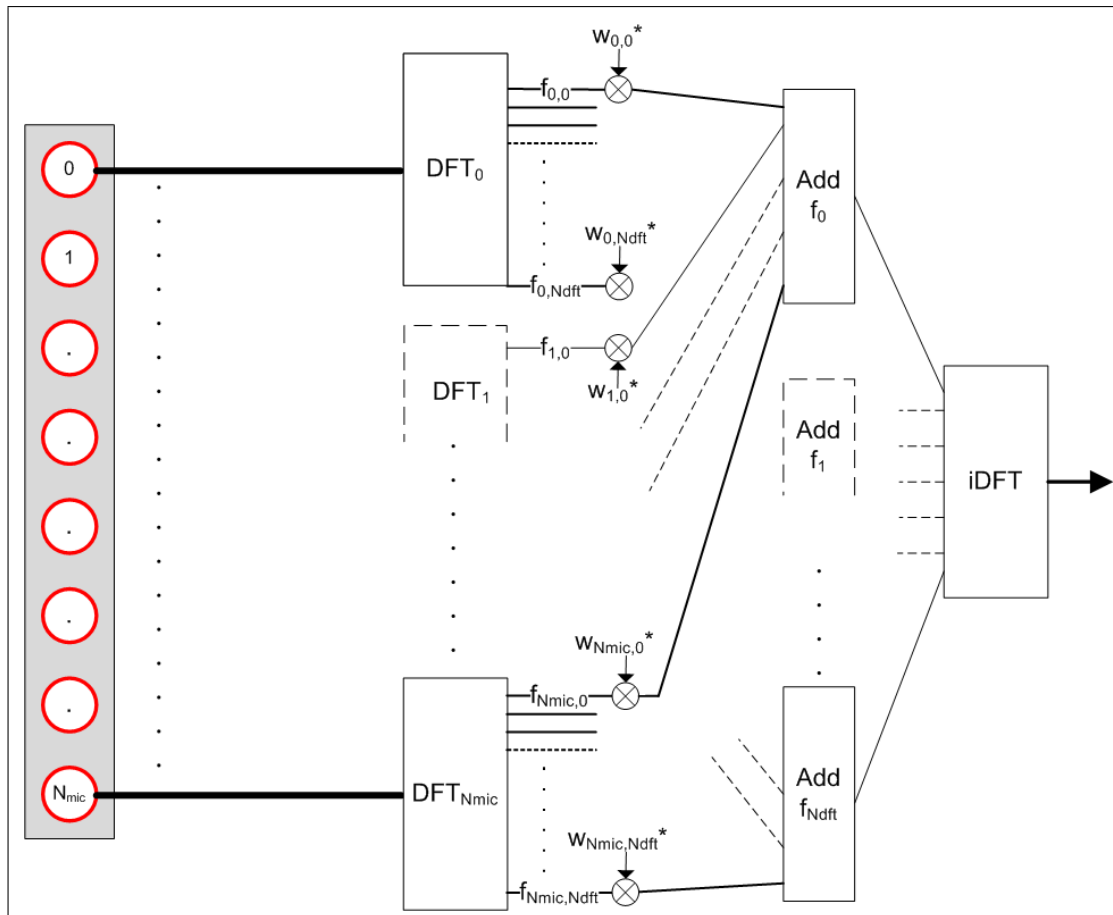


Abbildung 5.7: Breitband Beamformer, Grundstruktur

Die Signale der einzelnen Mikrofone werden durch DFT-Blöcke in den Frequenzbereich überführt. Für jedes Frequenzband f_0 bis f_{Ndft} wird jeweils ein Beamformer mit N_{mic} Koeffizienten w_i entworfen. Die Koeffizienten werden mit den jeweiligen frequenzzerlegten Signalen multipliziert und die zusammengehörigen Frequenzbänder addiert. Der iDFT-Block (inverse DFT) erzeugt letztendlich das gefilterte Zeitsignal.

Verglichen mit dem Schmalband-Ansatz steigt die Komplexität in diesem Verfahren massiv an, entsprechend sollte es nur verwendet werden, falls die zu verarbeitenden Signale nicht mehr als ausreichend schmalbandig betrachtet werden können.

Da die nun folgenden Breitband Beamformer in der späteren Implementierung verwendet werden, wird das Beampattern in der Draufsicht dargestellt. Die Farbskala am Rand gibt die zur „Heatmap“ gehörenden Signalpegel in dB wieder. Diese Darstellung ermöglicht eine bessere Übersicht über die Leistung des Beamformers.

5.3 Breitband D&S-Beamformer

5.3.1 Beampattern

Wendet man das bekannte D&S-Entwurfsverfahren für die Breitband Struktur an und wählt als Entwurfslfrequenz die jeweilige zum Frequenzbin der DFT korrespondierende Frequenz, erhält man einen Breitband D&S-Beamformer [1, S.9]. Das Beampattern verändert sich zu:

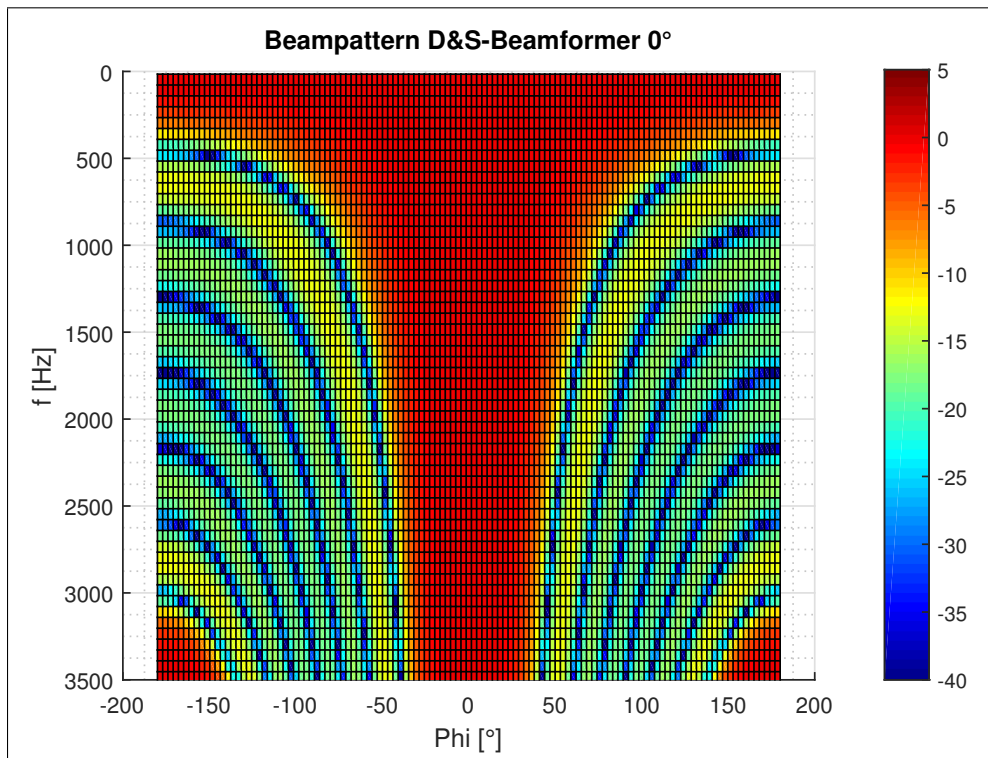


Abbildung 5.8: Beampattern, 0° Breitband D&S-Beamformer

Der Unterschied ist deutlich zu erkennen. Während im unteren Frequenzbereich aus gleichen Gründen weiterhin keine nennenswerte Filterwirkung erkennbar ist, beginnt sich ab ≈ 500 Hz eine deutliche Empfangskeule herauszubilden.

5.3.2 Array Gain

Zur Bestimmung des Array Gains muss entsprechend Gleichung 4.18 ein Rauschfeld angenommen werden.

Wird räumlich unkorreliertes Rauschen angenommen, erhält man als Array Gain den White Noise Gain. Für den beschriebenen D&S-Beamformer ergibt sich folgender Verlauf:

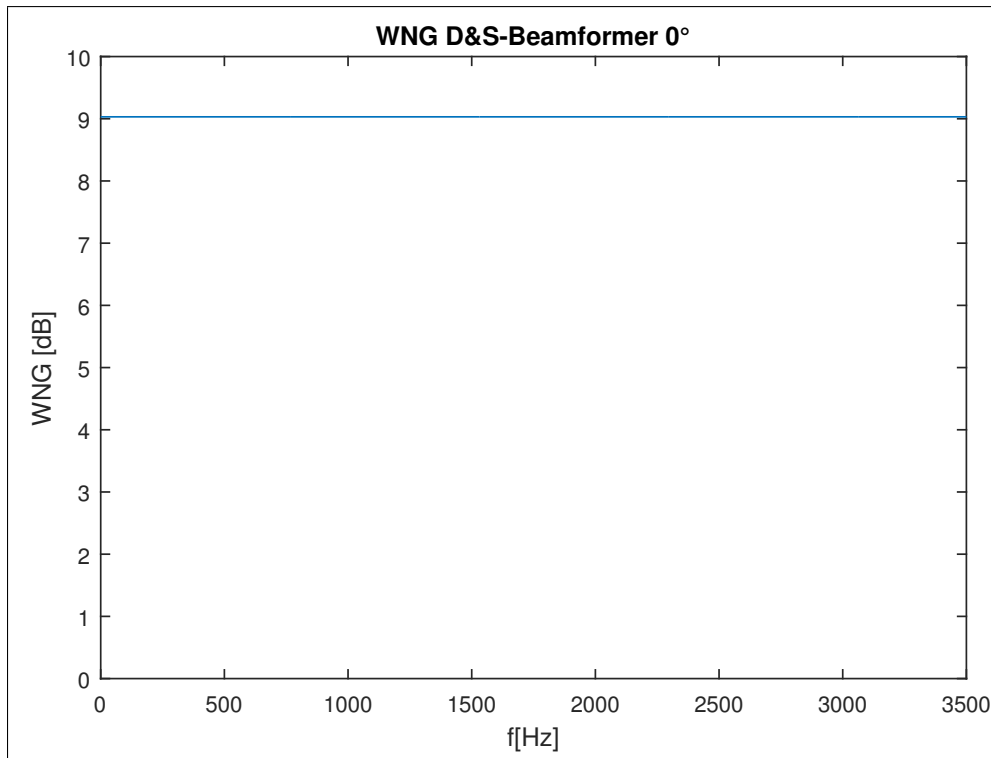


Abbildung 5.9: White Noise Gain, 0° Breitband D&S-Beamformers

Der White Noise Gain ist konstant über der gesamten Frequenz. Dass dies für alle D&S-Beamformer gelten muss, ist an Gleichung 5.3 erkennbar. Der WNG entspricht dem Dämpfungsfaktor $\frac{1}{N_{mic}}$ vor dem Rauschsignal \mathbf{V} . Entsprechend ist der erreichte WNG in diesem Fall:

$$\text{WNG}_{\text{D\&S,8Mic}} |_{\text{dB}} = 10 \cdot \log_{10}(8) \simeq 9 \text{ dB}$$

Für ein diffuses Rauschfeld erhält man die Directivity bzw. den Directivity Index, also die Fähigkeit des Beamformers ein diffuses Rauschfeld zu unterdrücken:

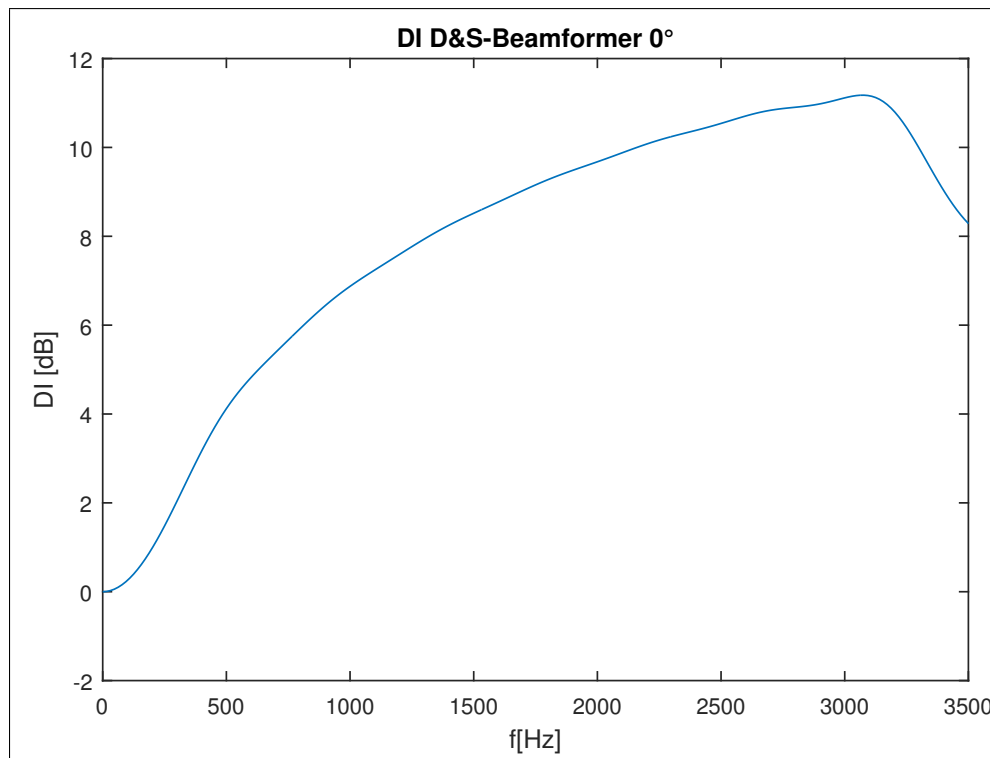


Abbildung 5.10: Directivity Index, 0° Breitband D&S-Beamformers

Wie man sehen kann, ist die Unterdrückung vor allem in höheren Frequenzen möglich, tiefe Frequenzen werden nicht unterdrückt.

5.4 Breitband Minimum Variance Distortionless Response

5.4.1 Entwurfsmethodik

Beim Entwerfen optimaler Beamformer werden Minimierungsverfahren eingesetzt, um definierte Eigenschaften zu erhalten bzw. zu vermeiden. Die zu optimierenden Parameter sind die Filterkoeffizienten $w_i(e^{j\Omega})$, der zu minimierende Parameter hängt vom gewünschten Ergebnis ab.

Das Ziel des Minimum Variance Distortionless Response (MVDR) Beamformers ist, die Leistung des Ausgangssignals zu minimieren (minimum variance) aber gleichzeitig eine unverzerrte Antwort für Signale aus einem festen Winkel Φ_0 zu erhalten (distortionless response). Diese Forderungen ergeben das folgende Minimierungsproblem¹:

$$\min_{\mathbf{W}} \mathbf{W}^H \Phi_{XX} \mathbf{W} \quad \text{unter Bedingung} \quad \mathbf{W}^H \mathbf{d} = 1 \quad (5.4)$$

¹Diese, sowie die in der folgenden Herleitung benötigten Formeln sind Quelle [2] ab Seite 24 entnommen.

Unter der Annahme, dass für den Entwurf Array-steering und Signalrichtung übereinstimmen, ist das Objektiv die optimale Unterdrückung des Rauschens. Somit kann im Ansatz 5.4 das Empfangs-LDS Φ_{xx} durch das Rausch-LDS Φ_{VV} ersetzt werden. Über das Verfahren der Lagrange-Multiplikatoren [A] erhält man die Lösung des Minimierungsproblems:

$$\mathbf{W} = \frac{\Phi_{VV}^{-1} \mathbf{d}}{\mathbf{d}^H \Phi_{VV}^{-1} \mathbf{d}} \quad (5.5)$$

Für ein homogenes Rauschfeld kann das Rausch-LDS Φ_{VV} durch die Kohärenzmatrix Γ_{VV} ersetzt werden:

$$\mathbf{W}_{MVDR} = \frac{\Gamma_{VV}^{-1} \mathbf{d}}{\mathbf{d}^H \Gamma_{VV}^{-1} \mathbf{d}} \quad (5.6)$$

Formel 5.6 beinhaltet Matrixinversionen, die unter Umständen numerisch instabil sein können. Zudem zeigt sich, dass MVDR-Beamformer die Eigenschaft haben, weißes Rauschen vor allem in tieferen Frequenzen tendenziell eher zu verstärken [2, S.27]. Aus diesem Grund wird die Berechnungsvorschrift leicht modifiziert, um eine Begrenzung für den WNG einzubringen und die Invertierbarkeit zu verbessern:

$$\mathbf{W}_{MVDR} = \frac{(\Gamma_{VV}^{-1} + \mu \mathbf{I}) \mathbf{d}}{\mathbf{d}^H (\Gamma_{VV}^{-1} + \mu \mathbf{I}) \mathbf{d}} \quad (5.7)$$

Zusammengefasst erzeugt der MVDR-Ansatz den Beamformer, der das gewählte Rauschfeld optimal unterdrückt.

Ein interessanter Sonderfall ist hierbei die räumlich-unkorrelierte Rauschumgebung. Setzt man 4.22 in 5.6 ein ergibt sich:

$$\mathbf{W}_{MVDR}|_{uncorrelated} = \frac{\mathbf{I}^{-1} \mathbf{d}}{\mathbf{d}^H \mathbf{I}^{-1} \mathbf{d}} \quad (5.8)$$

$$\mathbf{W}_{MVDR}|_{uncorrelated} = \frac{1}{N_{mic}} \mathbf{d} \quad (5.9)$$

Was dem D&S-Beamformer entspricht. Der D&S-Beamformer ist also der optimale Beamformer zur Unterdrückung von weißem Rauschen.

5.4.2 Beampattern

Wie im vorangegangenen Abschnitt beschrieben, kann über die MVDR Methode eine Vielzahl von Beamformern entworfen werden. Mit Rücksicht auf die betrachteten Anwendungsszenarien wird im folgenden das Beampattern für diffuse Rauschumgebung 4.4.1 und zylindrisch-isotrope

Umgebung 4.4.2 gezeigt.

Für ein diffuses Rauschfeld ergibt sich folgendes Beampattern¹

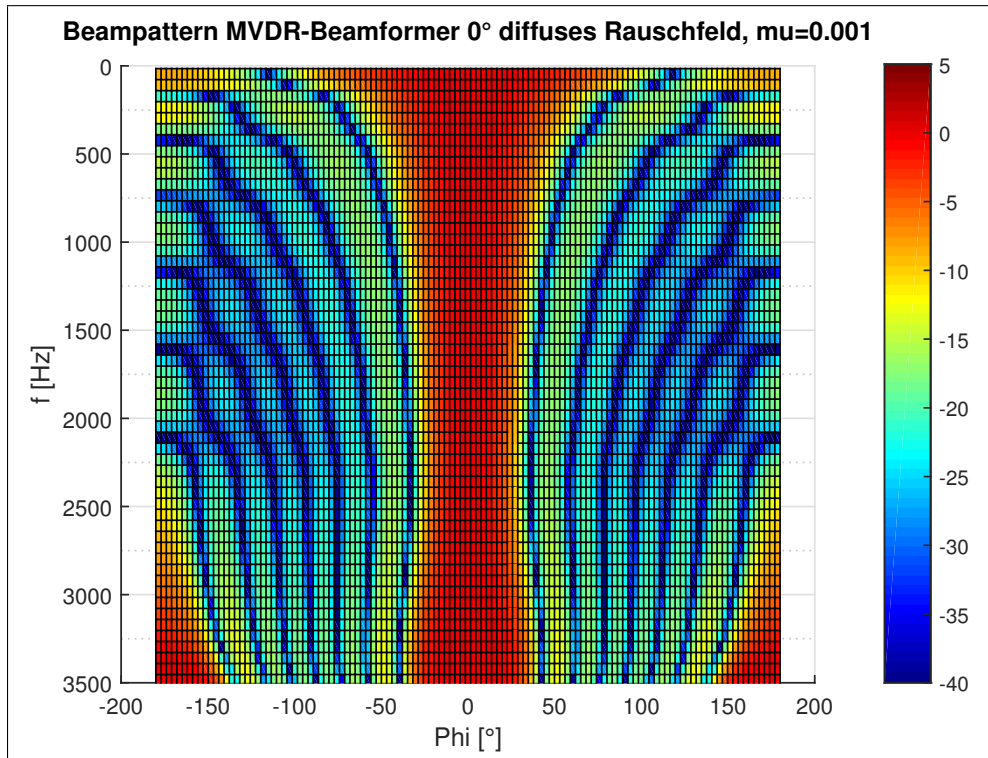


Abbildung 5.11: Beampattern, 0° MVDR Beamformers für diffuses Rauschfeld

Im Vergleich zum D&S-Beamformer lässt sich eine deutlich konstantere Hauptkeulenbreite über der Frequenz feststellen.

Das zylindrisch isotrope Rauschfeld erzielt den folgenden Verlauf:

¹Aus Gründen der Übersichtlichkeit wird hier nur ein Entwurfswinkel dargestellt. In Abschnitt [B] des Anhangs sind einige weitere Beampattern dargestellt.

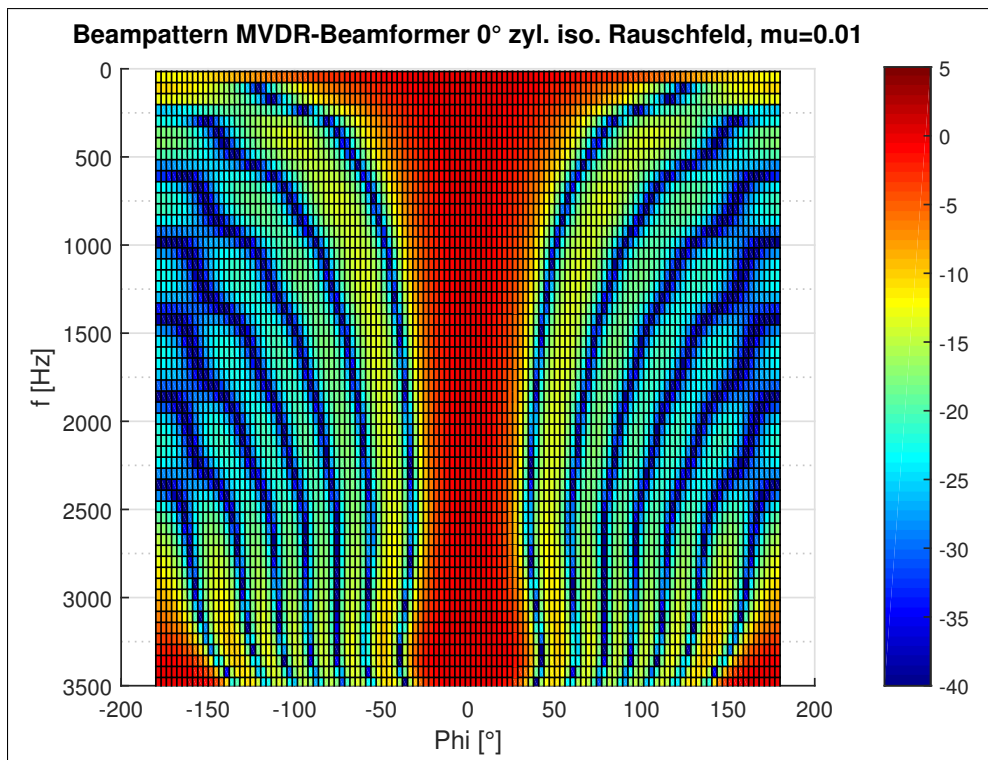


Abbildung 5.12: Beampattern, 0° MVDR Beamformers für zylindrisch isotropes Rauschfeld

Das Beampattern unterscheidet sich an dieser Stelle nicht wesentlich vom diffusen Entwurf. Ein Unterschied stellt z.B. die unterschiedliche Höhe der Nebenkeulen dar. Der Anpassungsfaktor μ hat dabei, solange er ausreichend klein gegenüber den eigentlichen Koeffizienten ist, nur marginale Auswirkung auf den Verlauf. Seine Auswirkungen werden erst später, in der Directivity und im WNG, sichtbar.

5.4.3 Array Gain

Der WNG ist stark von der gewählten Anpassung durch μ abhängig. Je nach gewähltem Parameter ergeben sich folgende Verläufe:

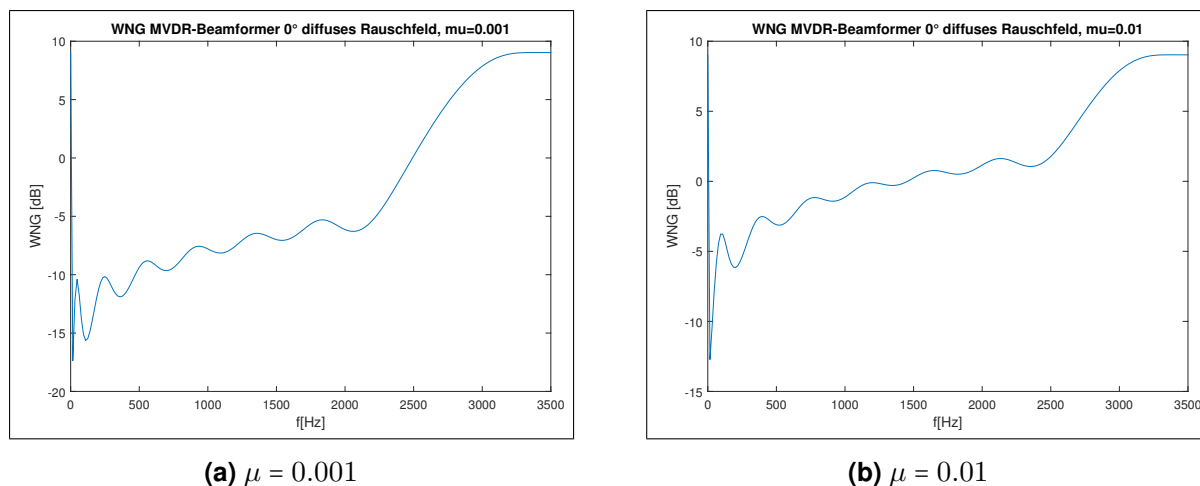


Abbildung 5.13: WNG, 0° MVDR Beamformers für diffuses Rauschfeld

Die Verbesserung im WNG bei größerem μ ist hier deutlich erkennbar. Man kann jedoch den Korrekturfaktor nicht beliebig groß wählen. Im Grenzübergang für sehr großes μ wird der Beamformer wieder zum Delay and Sum Beamformer:

$$\mathbf{W}_{\text{MVDR}} \Big|_{\mu \gg |w_i|} = \mathbf{W}_{\text{D\&S}} \quad (5.10)$$

Dieser erreicht dann erneut den optimalen WNG von in diesem Fall 9 dB. Der bessere WNG wird dabei gegen eine vergleichsweise breite Hauptkeule im tiefen Frequenzbereich getauscht.

Für das zylindrische Rauschfeld ergeben sich die folgenden Graphen:

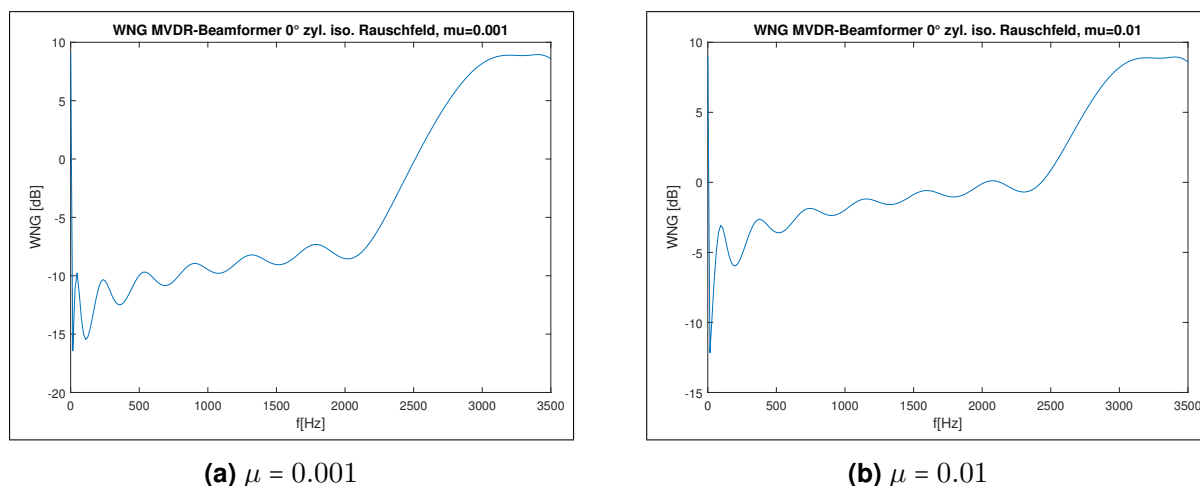


Abbildung 5.14: WNG, 0° MVDR Beamformers für zylindrisch isotropes Rauschfeld

Der Directivity Index zeigt folgende Verläufe:

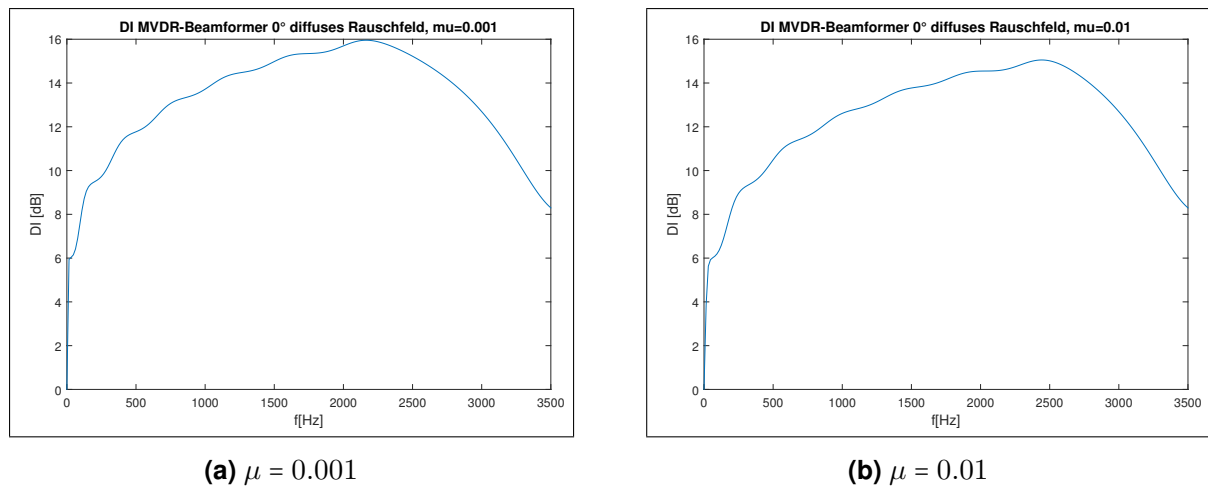


Abbildung 5.15: Directivity Index, 0° MVDR Beamformers für diffuses Rauschfeld

Erkennbar ist, dass durch Anheben von μ die Fähigkeit der Unterdrückung diffuser Rauschfelder sinkt. Dieser Zusammenhang wird auch dann offensichtlich, wenn der Grenzübergang 5.10 zum D&S-Beamformer durchgeführt wird.

Für das zylindrische Rauschfeld ergeben sich die folgenden Graphen:

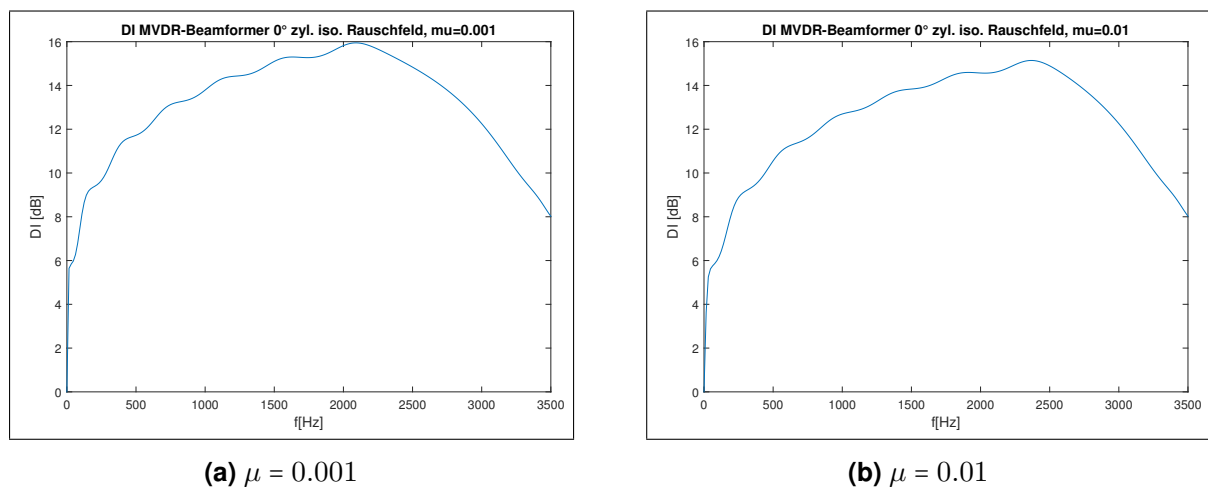


Abbildung 5.16: Directivity Index, MVDR Beamformers für zylindrisch isotropes Rauschfeld

Zusammenfassend lässt sich sagen, dass die Unterschiede im MVDR-Entwurf für diffuse und zylindrisch isotrope Rauschfelder minimal sind. Der eingebrachte Korrekturfaktor ermöglicht eine variable Anpassung des WNG an das jeweilige Umfeld. Eine Verbesserung im WNG erfolgt immer auf Kosten der Directivity sowie der Breite der Empfangskeule.

5.5 Breitband Linear Constraint Minimum Variance

5.5.1 Entwurfsmethodik

Der MVDR Ansatz 5.4 kann verallgemeinert werden, um multiple Nebenbedingungen zu erfüllen. Im allgemeinen Fall ergibt sich also das folgende Minimierungsproblem [9, S.27]:

$$\min_{\mathbf{w}} \mathbf{w}^H \Phi_{XX} \mathbf{w} \quad \text{unter Bedingung} \quad \mathbf{C}^H \mathbf{w} = \mathbf{g} \quad (5.11)$$

Die Bedingungs- bzw. Einschränkungsmatrix \mathbf{C} enthält hierbei die gewünschten Richtungsvektoren $[d_1, \dots, d_M]$ für welche der Antwortvektor \mathbf{g} den zugehörigen Gain enthält $[g_1, \dots, g_M]^T$. Die Lösung des oben genannten Problems ist der sogenannte „Linear Constraint Minimum Variance“ oder kurz LCMV-Beamformer. Unter denselben Annahmen wie im MVDR Fall ergibt sich inklusive der WNG-Anpassung und unter Abhängigkeit von Γ [aus 9, S.27]:

$$\mathbf{w}_{\text{LCMV}} = [(\Gamma + \mu \mathbf{I})^{-1} \mathbf{C} [\mathbf{C}^H (\Gamma + \mu \mathbf{I})^{-1} \mathbf{C}]^{-1} \mathbf{g}] \quad (5.12)$$

Dieser Filtertyp erlaubt es nun, bis zu $N_{mic} - 1$ Bedingungen beim Filterentwurf zu berücksichtigen. Er kann z.B. verwendet werden, um neben einer Hauptempfangsrichtung verschiedene Richtungen explizit mit Nullstellen zu behaften. Theoretisch ist es somit möglich, feste Rauschquellen aus bekannter Richtung optimal auszublenden.

5.5.2 Beampattern

Um das Prinzip zu veranschaulichen, werden zwei Beschränkungen im Entwurf berücksichtigt. Die Nutzquelle sendet aus $\Phi_{Src} = 0^\circ$, die Störquelle aus $\Phi_{Noise} = 50^\circ$, der Korrekturfaktor beträgt zunächst $\mu = 0.001$. Der Beamformer wird zunächst für räumlich unkorrelierte Rauschumgebung entworfen. Es ergibt sich folgendes Beampattern:

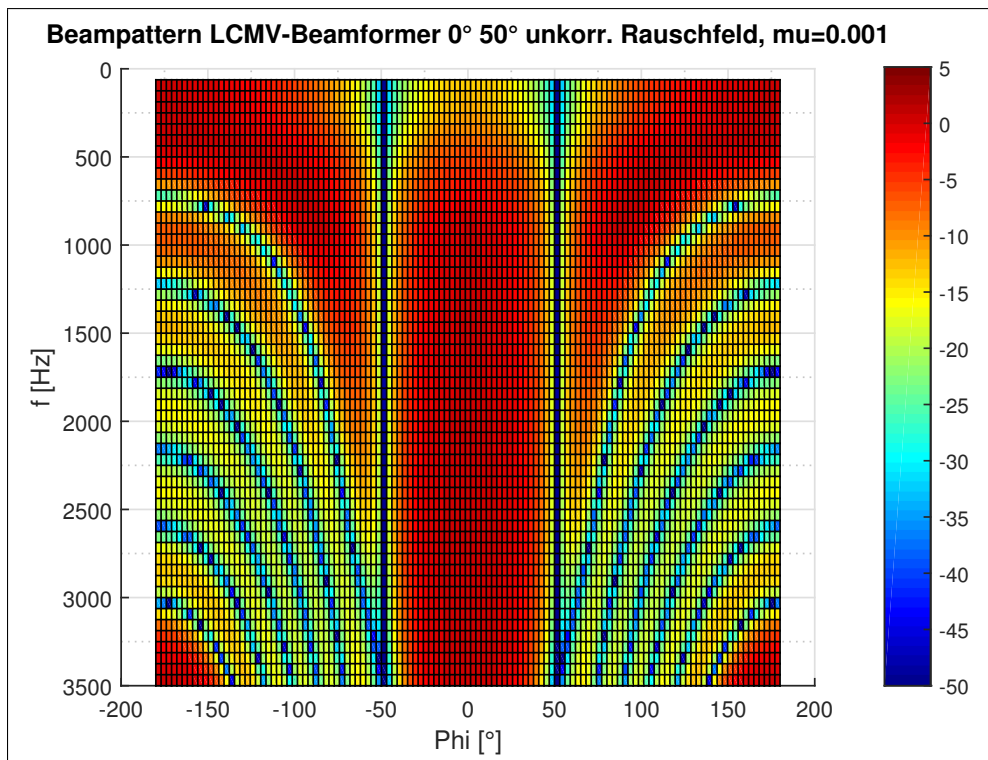


Abbildung 5.17: Beampattern, $0^\circ / 50^\circ$ LCMV Beamformers für unkorreliertes Rauschfeld, $\mu = 0.001$

Es ist deutlich die für unkorrelierte Rauschfelder erwartete D&S-Charakteristik zu erkennen. Neben dem erwarteten breiten Verlauf im tiefen Frequenzbereich ist jedoch durchgängig die tiefe Nullstelle um 50° entstanden.

Zum Vergleich wird derselbe Entwurf für ein diffuses Rauschfeld durchgeführt:

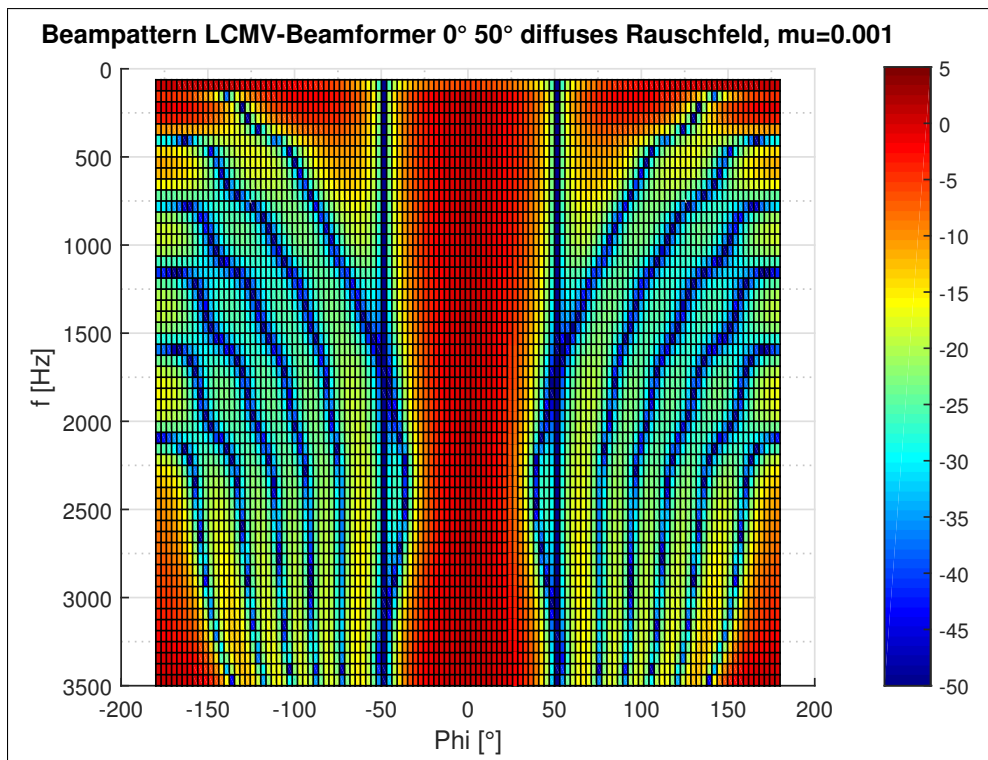


Abbildung 5.18: Beampattern, $0^\circ / 50^\circ$ LCMV Beamformers für diffuses Rauschfeld, $\mu = 0.001$

Analog zum vorherigen Bild bringt der Entwurf eine tiefe Nullstelle um 50° ein. Die Verbreiterung der Empfangskeule im tiefen Frequenzbereich ist weniger stark ausgeprägt als im unkorrelierten Entwurf.

5.5.3 Array Gain

Wie in den vorherigen Kapiteln wird der WNG als erstes betrachtet. Für den gewählten Korrekturfaktor von $\mu = 0.001$ ergeben sich folgende WNG-Verläufe für unkorreliertes und diffuses Rauschfeld:

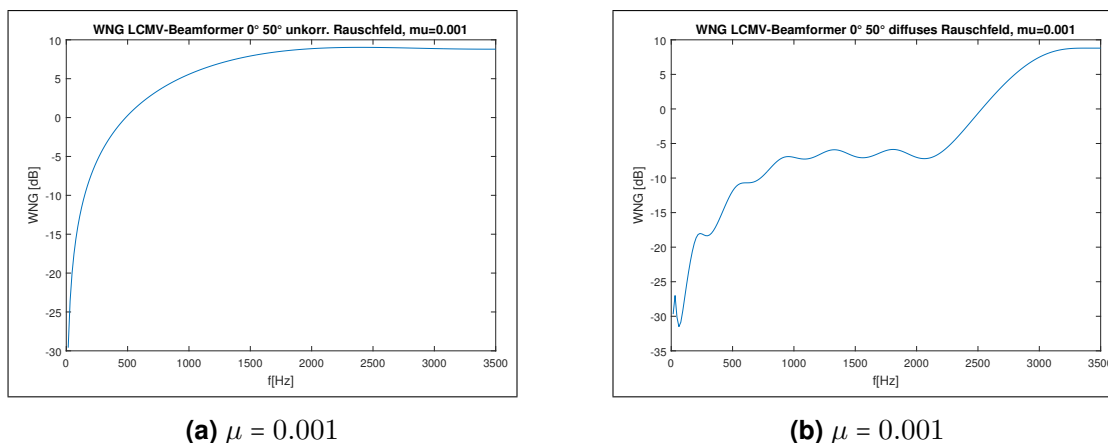


Abbildung 5.19: White Noise Gain, $0^\circ / 50^\circ$ LCMV Beamformers für diffuses Rauschfeld

Der Verlauf in Abbildung 5.19a ähnelt dem klassischen D&S-Verlauf. Das Einbringen der Nullstellen verschlechtert jedoch die Performance im niederfrequenten Bereich. Im diffusen Rauschfeld entspricht der LCMV-Verlauf in etwa dem MVDR Beamformer. Der WNG kann über Änderungen an μ entsprechend der Bedürfnisse angepasst werden.

Der Directivity Index beider Varianten stellt sich wie folgt dar:

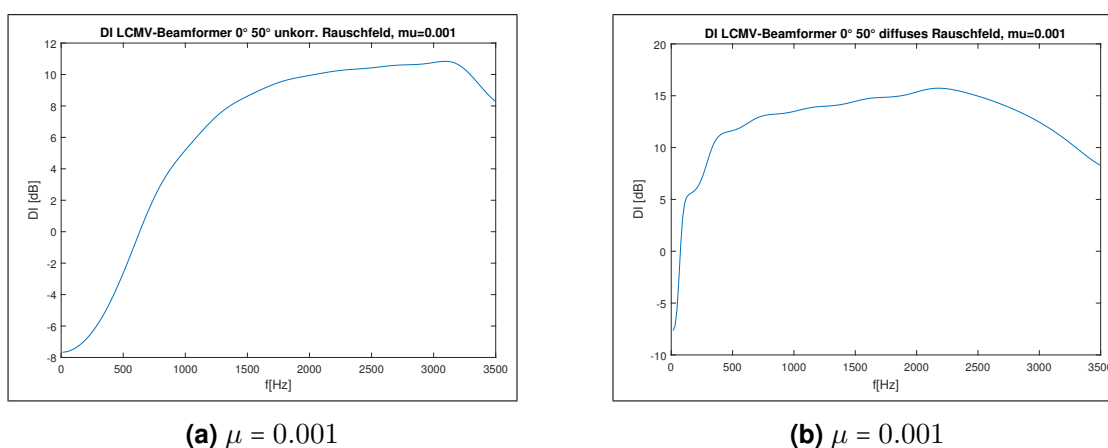


Abbildung 5.20: Directivity Index, $0^\circ / 50^\circ$ LCMV Beamformers für diffuses Rauschfeld

Wie erwartet, ist der Directivity Index des Entwurfs für diffuses Rauschfeld deutlich höher. Im Vergleich zur MVDR Lösung fällt auf, dass in einem kleinen Bereich bei den tiefsten Frequenzen eine Verstärkung des diffusen Rauschens anstatt einer Dämpfung auftritt.

5.6 Nahfeldkompensation

Die bisher vorgestellten Verfahren verwenden nicht explizit Nahfeld oder Fernfeldcharakteristika. Die Nahfeld- und Fernfeldeigenschaften sind im Array-steering d enthalten. Während im Fernfeld von planaren Wellenfronten ausgegangen wird, muss auf Grund der kugelförmigen Ausbreitung im Nahfeld ein Amplitudenunterschied sowie eine nicht konstante Phasendifferenz zwischen den Mikrofonen berücksichtigt werden.

Dr. Ian McCowan schlägt in [9] eine Variante vor, die diese Anpassung direkt an den über ein beliebiges Verfahren gewonnenen Koeffizienten \mathbf{w}_i implementiert:

$$\mathbf{w}_i(f) = \frac{d_n(d_{SA}, \phi)}{d_0(d_{SA}, \phi)} e^{j \frac{2\pi}{\lambda} (d_0(d_{SA}, \phi) - d_n(d_{SA}, \phi) + n d_{mic} \cos(\phi))} \quad (5.13)$$

mit d_i als Abstand zwischen Mikrofon und Quelle (ULA):

$$d_i(d_{SA}, \phi) = [d_{SA}^2 + 2d_{SA}id_{mic} \cos(\phi) + (id_{mic})^2]^{\frac{1}{2}} \quad (5.14)$$

sowie d_{SA} als Abstand von Mikrofon Null zur Quelle und d_{mic} als Abstand zwischen den Mikrofonen.

6 Beamforming mit Subarrays

Beim Entwurf von Beamformern konnte beobachtet werden, dass im niedrigen Frequenzbereich nur eine schwache räumliche Filterwirkung erzielbar ist. Da die Filterwirkung in erster Linie vom Verhältnis Wellenlänge zu Mikrofonabstand bestimmt wird, ist ein vielversprechender Ansatz dieses Problem zu lösen, die Verwendung der in Kapitel 4.5.1.2 vorgestellten harmonischen Subarrays.

Die verwendete Arraykombination enthält, um sinnvoll vergleichen zu können, erneut 8 Mikrofone. Es werden insgesamt 3 Unterarrays mit sich verdoppelnden Mikrofonabständen verwendet. Abbildung 6.1 zeigt den Aufbau:

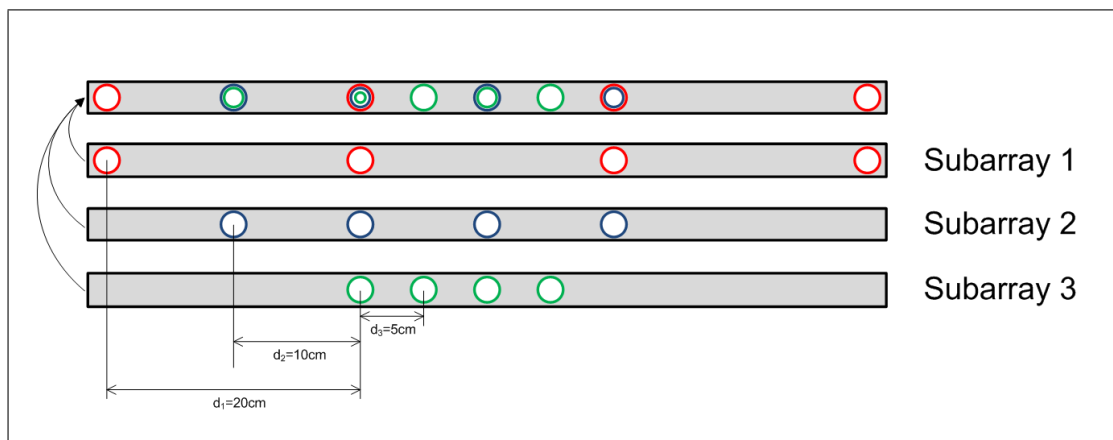


Abbildung 6.1: Arrayaufbau, 3 harmonische Subarrays

Auf Grund der physikalischen Gegebenheiten ist der Aufbau nicht symmetrisch. Für Subarray 1 ist der Arraymittelpunkt in der Mitte des Gesamtaufbaus. Für Subarray 2 und Subarray 3 ist der Mittelpunkt um $\Delta m_2 = 2,5 \text{ cm}$ bzw. $\Delta m_3 = 1,25 \text{ cm}$ nach links verschoben. Dies führt zu unterschiedlichen Auftreffwinkeln auf die Arrays. Die Winkelabweichung im Bezug auf den Gesamtmittelpunkt ergibt sich zu:

$$\Delta\phi_{\text{Subarray } 2} = \arctan\left(\frac{\Delta m_2}{d_{AS}}\right) = \arctan\left(\frac{2,5 \text{ cm}}{d_{AS}}\right) \quad (6.1)$$

$$\Delta\phi_{\text{Subarray } 3} = \arctan\left(\frac{\Delta m_3}{d_{AS}}\right) = \arctan\left(\frac{1,25 \text{ cm}}{d_{AS}}\right) \quad (6.2)$$

Für Entfernungen d_{AS} zwischen Array und Quelle von z.B. 1 m wäre die maximale Winkeldifferenz entsprechend $\Delta\phi_{\text{Subarray } 2} \simeq 2,9^\circ$. Im Vergleich zur bisher beobachteten Beambreite ist diese Abweichung nicht kritisch und somit vernachlässigbar. Wird das Array in der Anwendung jedoch in unmittelbarer Nähe zur Signalquelle betrieben, muss der Winkel im Entwurf entsprechend angeglichen werden.

Die vorgestellten Entwurfsmethoden können für diese Struktur ebenfalls verwendet werden. Im Gegensatz zum ULA wird jedoch für jedes Unterarray ein separater Entwurf durchgeführt, für den die Parameter entsprechend angepasst werden müssen.

Durch die gesteigerten Abstände zwischen den Mikrofonen, muss bei der Verarbeitung der Signale verstärkt auf den räumlichen Alias-Effekt geachtet werden. In diesem konkreten Anwendungsfall ergeben sich die folgenden Eckfrequenzen an Hand von Formel 4.3a:

$$f_{\text{Edge Subarray 1}} \leq \frac{v}{2d_1} = 875,5 \text{ Hz}$$

$$f_{\text{Edge Subarray 2}} \leq \frac{v}{2d_2} = 1715 \text{ Hz}$$

$$f_{\text{Edge Subarray 3}} \leq \frac{v}{2d_3} = 3430 \text{ Hz}$$

Das größte Subarray wird entsprechend für den tiefsten Frequenzbereich bis max 875,5 Hz eingesetzt. Das Mittlere von 875,5 Hz bis 1715 Hz und das Kleinste für den verbleibenden Bereich bis 3430 Hz. In die Praxis umgesetzt bedeutet dies, dass die einzelnen Frequenzbins der FFT je einem der Arrays bzw. einem der Beamformer zugeordnet werden.

6.1 Beampattern

Simuliert werden nun zwei Beamformer in der beschriebenen Struktur. Zum einen ein D&S-Beamformer für $\phi = 0^\circ$, zum anderen ein MVDR-Beamformer ($\mu = 0.001$) für diffuses Rauschfeld und gleichem Winkel. Die bestimmten Eckfrequenzen werden um 100 Hz abgesenkt, um Alias-Effekte sicher zu unterdrücken.

Der D&S-Beamformer erzielt folgendes Ergebnis:

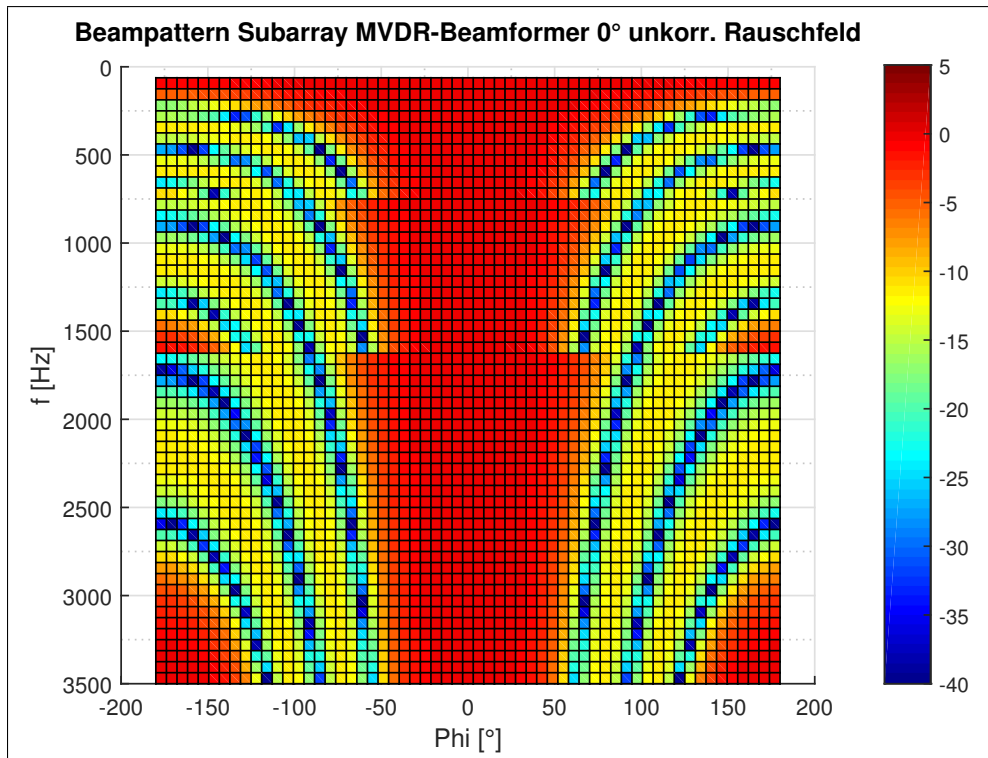


Abbildung 6.2: Beampattern, 90° Subarray-D&S-Beamformer

Für den MVDR-Beamformer im diffusen Rauschfeld erhält man:

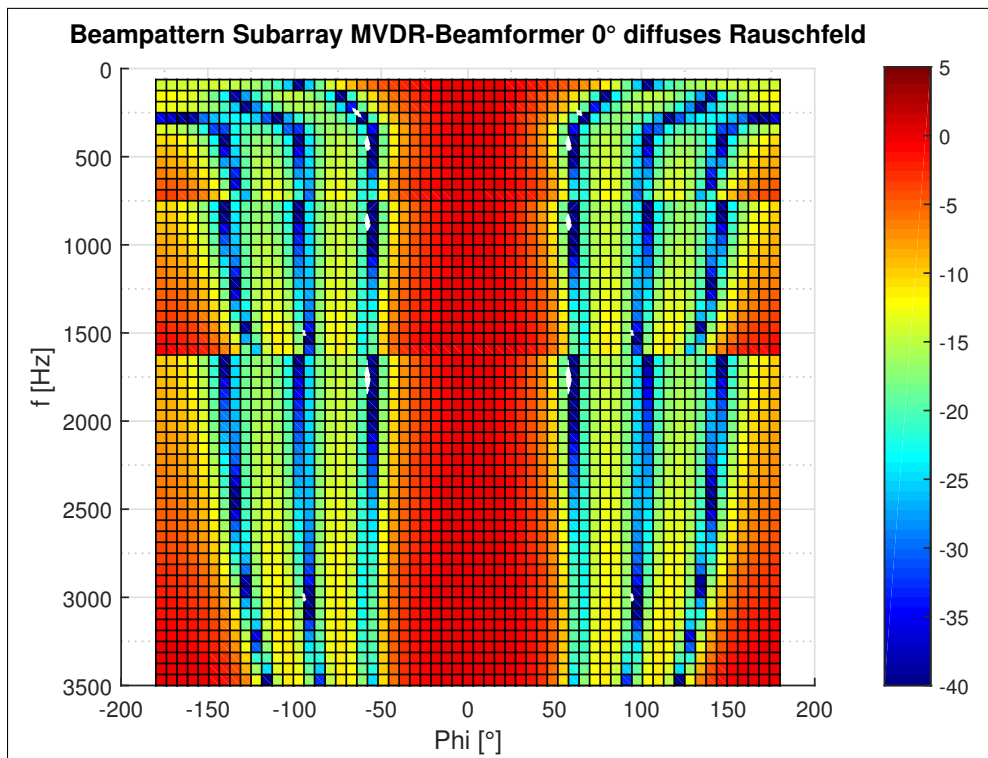
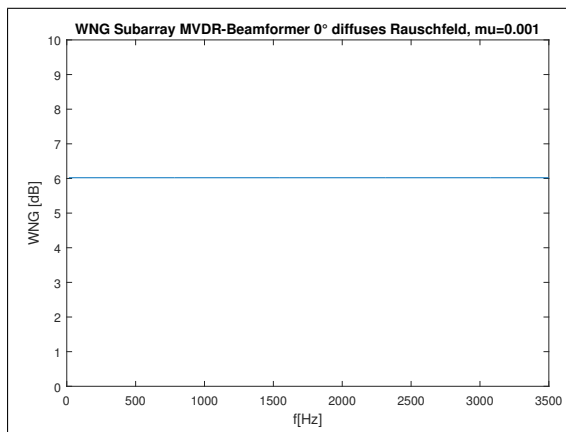


Abbildung 6.3: Beampattern, 0° MVDR-Subarray-Beamformer für diffuses Rauschfeld, $\mu = 0.001$

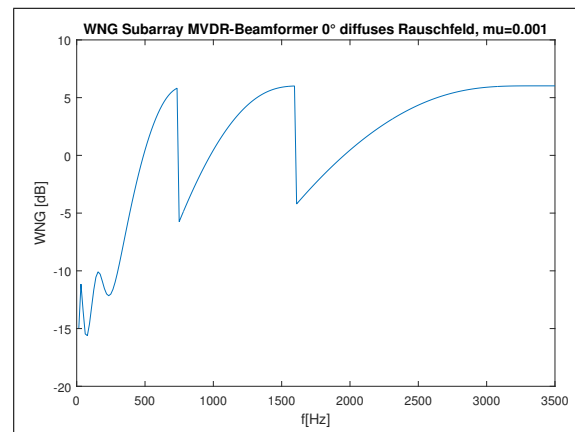
Die Verbesserung im Vergleich zum ULA ist deutlich erkennbar, die Empfangskeule läuft wesentlich weniger stark auseinander. Dadurch, dass pro Array jedoch nur 4 Mikrofone verwendet werden, sind weniger Nebenkeulen und entsprechend weniger Nullstellen vorhanden. Somit ist insgesamt die Dämpfung außerhalb der Hauptempfangsrichtung geringer. Die Höhe der größten Nebenkeule ist mit -12 dB jedoch nur geringfügig schlechter als die dB des ULA.

6.2 Array Gain

Wie in den vorherigen Simulationen gesehen, ist der WNG eines D&S-Beamformers ausschließlich von der Anzahl der Mikrofone abhängig. Entsprechend ergibt sich ein konstanter Verlauf von ≈ 6 dB:



(a) 0° Subarray-D&S-Beamformer

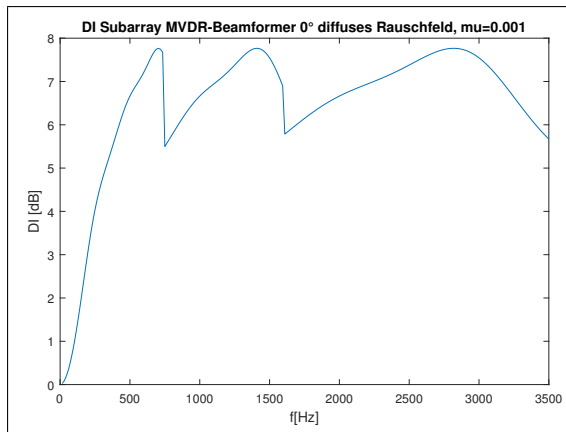


(b) 0° Subarray-MVDR-Beamformer für diffuses Rauschfeld, $\mu = 0.001$

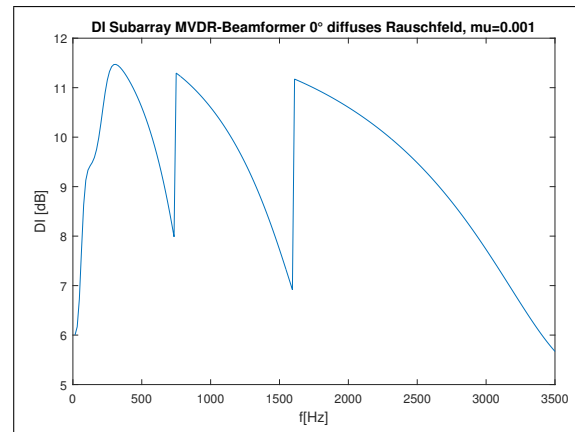
Abbildung 6.4: White Noise Gain verschiedener Subarray-Beamformer

Der Subarray MVDR-Entwurf zeigt einen anderen Verlauf mit Verstärkung des Rauschens im tiefen Frequenzbereich: Während der Subarray-D&S-Beamformer durchweg einen schlechteren WNG aufweist, ist der MVDR-Beamformer in Subarray Struktur insgesamt besser im WNG, als die einfache ULA Implementierung.

Der Directivity Index des Subarray-D&S-Beamformers zeigt eine ähnliche Tendenz wie im ULA. Der initiale Anstieg ist deutlich schneller, jedoch sind an den Übergangsstellen zwischen den Arrays sprunghafte Abfälle um einige dB zu erkennen. Insgesamt ist der Directivity Index über den gesamten Bereich besser als beim ULA, erreicht jedoch im hohen Frequenzbereich etwas schlechtere Werte.



(a) 0° Subarray Delay and Sum Beamformers

(b) 0° MVDR-Subarray-Beamformers für ein diffuses Rauschfeld, $\mu = 0.001$ **Abbildung 6.5:** Directivity Index verschiedener Subarray-Beamformer

Der Directivity Index des Subarray MVDR-Beamformers ist durchweg positiv. Während er insgesamt schneller ansteigt als beim ULA, ist die maximale Unterdrückung deutlich geringer.

7 Implementierung in C

Die Implementierungen der Beamformer in Hardware sind verhältnismäßig einfach. Der in Abschnitt 5.2 beschriebene Aufbau des Verfahrens wird in C übersetzt. Die zwei unterschiedlichen Algorithmen für das ULA und das Subarray sind in den nachfolgenden Abschnitten als Programmablaufplan dargestellt.

Die notwendigen Filterkoeffizienten wurden in Matlab vorberechnet und sind über Header-Files als mehrdimensionales Array hinterlegt. Für die FFT wird eine von Texas Instruments bereitgestellte hocheffiziente Assembler Implementierung verwendet. Das Grundgerüst der Datenverarbeitung über McBSP und EDMA liegt als Musterprogramm der Vertreiberfirma des Boards, DSignT, bereits vor.

7.1 Programmablauf ULA

Im Fall des ULA wird für den Beamformer jedes Mikrofonsignal verwendet. Die Eingangssignale werden nach Erhalt eines kompletten Datenblocks einer FFT mit 512 Punkten unterzogen. Die Frequenzbereichsdaten werden dann mit den jeweiligen Filterkoeffizienten entsprechend Abbildung 5.7 multipliziert. Die Frequenzbins der einzelnen Kanäle werden addiert und per iFFT zurück in den Zeitbereich transformiert. Ein ADC generiert aus diesen Daten das Ausgangssignal. Das Verfahren zeigt der abgebildete Programmablaufplan 7.1:

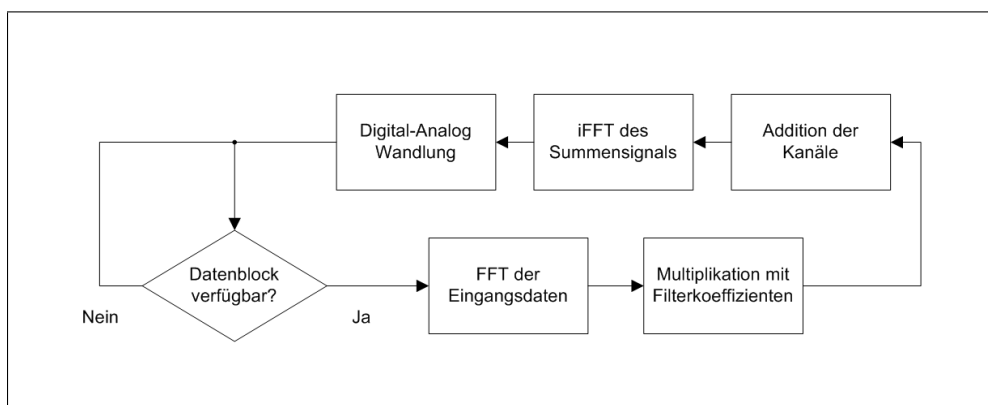


Abbildung 7.1: Programmablaufplan, Beamformer Implementierung für ULA

7.2 Programmablauf Subarray-ULA

Für die Subarrays muss vor der Anwendung des Beamformers die Zuordnung der Bins und Kanäle zu den zugehörigen Subarrays stattfinden. Das größte Unterarray arbeitet mit den Daten der Mikrofone 0, 1, 5 und 7, die verwendeten Frequenzbins sind 0-42. Das mittlere Subarray nutzt die Mikrofone 1, 3, 5 und 6 in den Frequenzbins 43-97. Die verbleibenden Frequenzbins werden vom kleinsten Subarray und den Mikrofonen 2, 3, 4 und 5 verarbeitet. Nachdem die Zuteilung stattgefunden hat, folgt das Verfahren prinzipiell dem der ULA-Implementierung nur dass bei der Addition entsprechend ausschließlich die jeweiligen Bins der Subarrays das Ausgangssignal generieren. Programmablaufplan 7.2 verdeutlicht das Vorgehen:

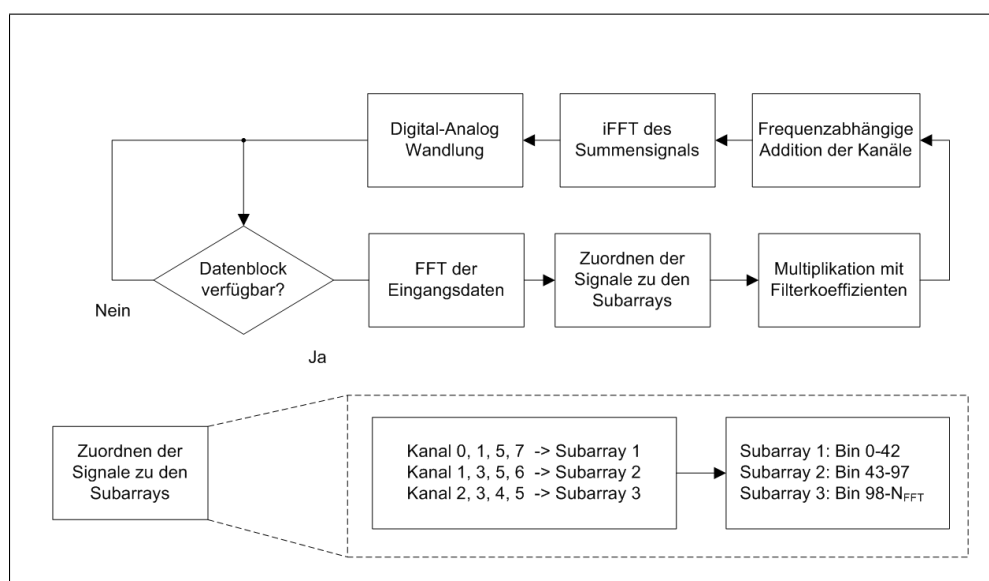


Abbildung 7.2: Programmablaufplan, Beamformer Implementierung für Subarrays

8 Funktionalitätsmessungen in idealer Umgebung

Die ersten Messreihen dienen der Funktionsanalyse der in C implementierten Beamformer. Es wird überprüft, ob die erstellten Algorithmen korrekt funktionieren und die Beamformer entsprechend der Simulation arbeiten. Um dies nachzuweisen, wird das Beampattern über einen relevanten Winkel- und Frequenzbereich gemessen.

Für diese Tests ist eine möglichst ideale Testumgebung notwendig. Die Umgebung sollte zum einen ausreichend groß sein, um Fernfeldausbreitung zuzulassen. Zum anderen sollten Reflexionen und äußere Störeinflüsse wie Rauschen bestenfalls nicht vorhanden sein.

8.1 Messumgebung und -aufbau

An der HAW-Hamburg steht für Messungen dieser Art ein Schallmessraum zur Verfügung. Der kubische Raum hat die Ausmaße von ca. 8m x 8m x 8m, die Wände sowie die Decke sind komplett mit schallabsorbierenden Schaumstoffspitzen eingekleidet. Der Boden selbst wird von einem gespannten Drahtseilnetz gebildet, unter welchem weitere Schaumstoffelemente Reflexionen von unten abfangen.

Eine zentrale Aufhängung erlaubt es, das Array mittig im Raum zu positionieren. Als Quelle wird ein Lautsprecher in einer der Ecken des Raumes aufgestellt um eine maximale Entfernung zum Array zu erreichen. Folgende Abbildung zeigt den Messaufbau:

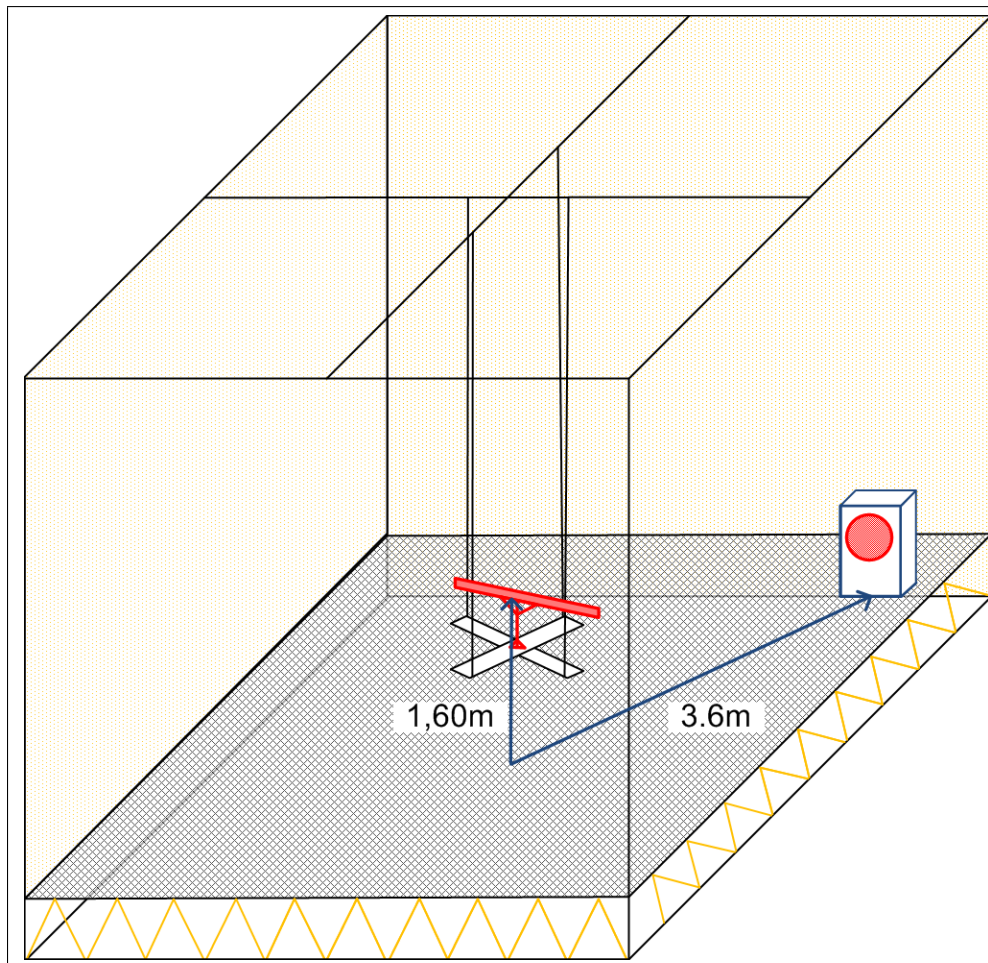


Abbildung 8.1: Messaufbau im Schallmessraum der HAW Hamburg

Als Referenz wird ein weiteres hochpräzises Messmikrofon in unmittelbarer Nähe zum Array aufgestellt. Mit diesem Mikrofon ist es möglich, Abweichungen vom erwarteten Ergebnis durch den Lautsprecher oder sonstige Effekte zu erkennen und aus dem Gesamtergebnis zu entfernen. Das Array wird auf die lange Kante gestellt, so dass die Mikrofone der Quelle zugewandt sind. Auf diese Weise soll möglichst omnidirektionales Verhalten für den positiven Winkelbereich erreicht werden.

Auf elektrischer Seite wird folgender Aufbau verwendet:

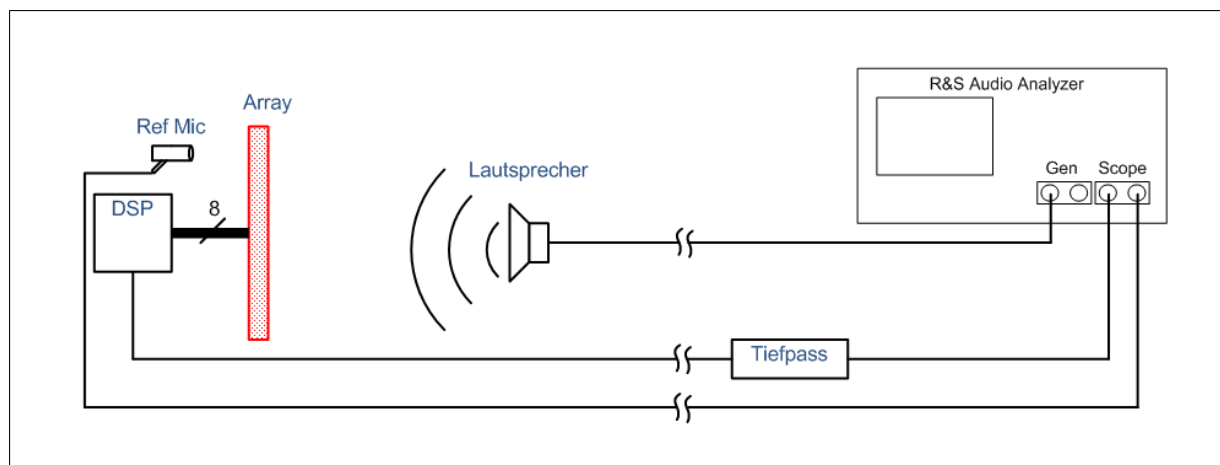


Abbildung 8.2: Elektrischer Messaufbau im Schallmessraum der HAW Hamburg

Die geschwungenen Linien im Leitungsverlauf kennzeichnen dabei den Übergang zwischen Messkammer und Vorraum. Verwendet werden folgende Geräte:

- **Generator:** Rhode & Schwarz UPV (Audio Analyzer)
- **Lautsprecher:** Neumann KH 120A
- **Referenzmikrofon:** Brüel & Kjær (Kapsel 4190, Vorverstärker 2669)
- **Testmikrofone (Array):** 8 Mikrofone vom Typ WM-52BM
- **Leitungen:** Es werden überwiegend 3-Pol XLR-Leitungen verwendet. Am direkten Übergang zum Tiefpass sowie zum DSP sind Coax-Leitungen notwendig.

Die entsprechenden Datenblätter können dem Anhang [E] entnommen werden.

8.2 Durchführung

Um die Funktion der implementierten Beamformer zu verifizieren, wird das reale Beampattern untersucht. Hierfür sind entsprechend Messungen über Winkel und Frequenz notwendig.

Um unterschiedliche Frequenzen abzudecken, kann die Amplitudengangsanalyse des Audio-Analyzers genutzt werden. Hierbei wird über den Generatorausgang ein Frequenzsweep (Chirp) ausgegeben und am Scope-Eingang mithilfe von Notch-Filtern die Energie im zugehörigen Frequenzband gemessen. Die Winkelabdeckung erfolgt von Hand durch Drehung des Array-Sockels in definierten Schritten. Da dieses Messverfahren insgesamt sehr zeitaufwendig ist, wird die Richtcharakteristik im Bereich von -90° bis 90° in 10° Schritten gemessen. Für jeden Winkel wird ein Frequenzsweep von 100 Hz bis 3500 Hz in 100 linearen Schritten von entsprechend 34 Hz ausgewertet. Pro Beamformer werden je zwei Entwurfswinkel ($\phi_0 = 0^\circ$ und $\phi_0 = 90^\circ$) getestet.

Zudem wird eine Messreihe mit Fernfeld- und eine mit Nahfeld-Koeffizienten durchgeführt. Es kommen ein D&S-Beamformer, ein MVRD-Beamformer für diffuses und ein MVDR-Beamformer für zylindrisch-isotropes Rauschfeld zum Einsatz.

Die Konfiguration des Audio Analyzers zeigt Abbildung 8.3:

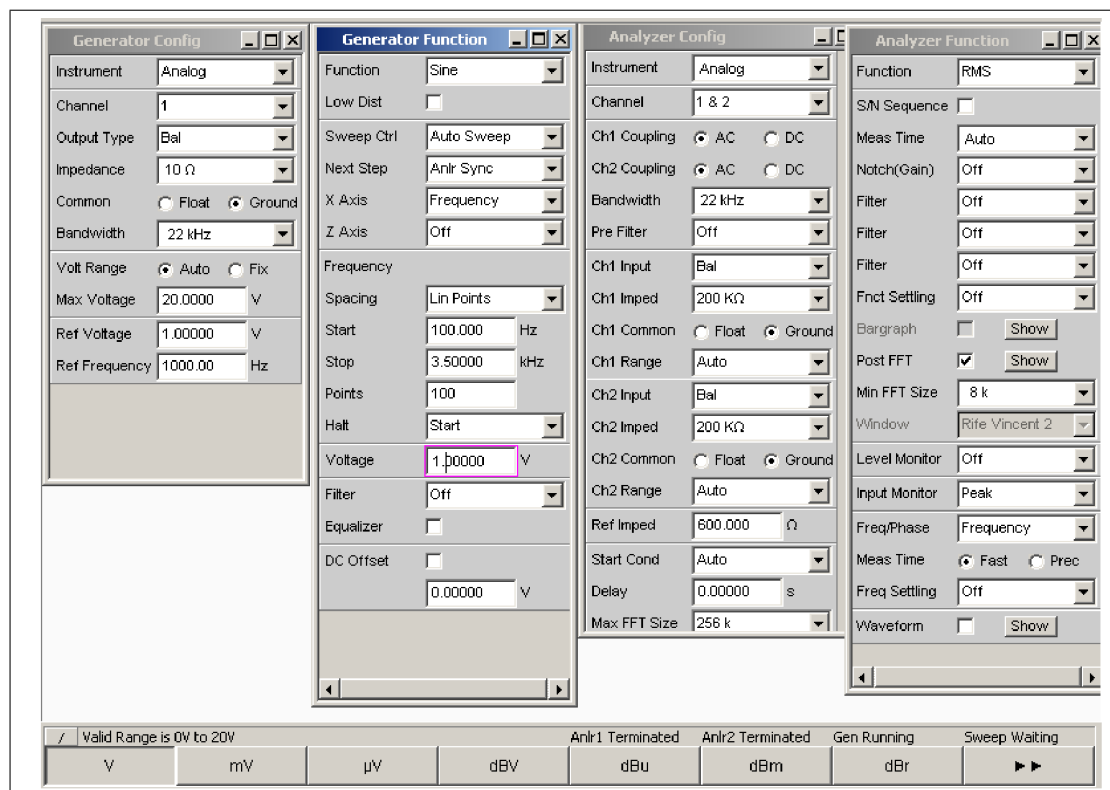


Abbildung 8.3: Konfiguration des Audio Analyzers

Das Ergebnis der Messung wird vom Audio Analyzer wie folgt dargestellt:

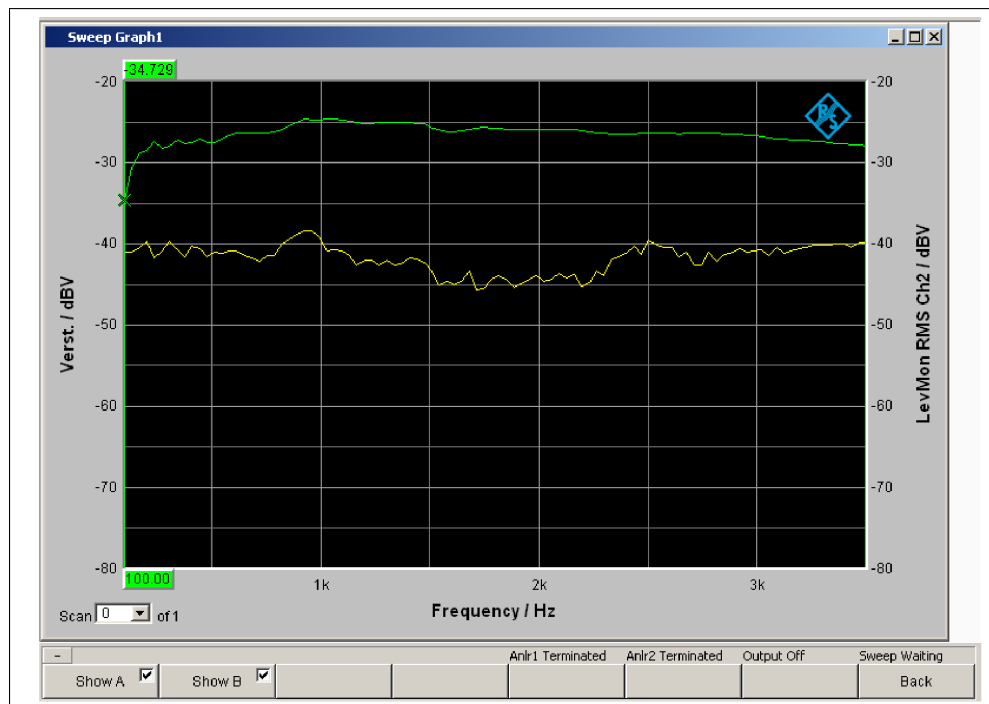


Abbildung 8.4: Beispielergebnis der Amplitudengangsanalyse mit einem R&S Audio Analyzer

Die grüne Linie entspricht in diesem Beispiel dem Beamformer-Output, die gelbe dem Referenzmikrofon.

Um die Interpretation der Ergebnisse zu vereinfachen und um unnötig viele Bilder zu vermeiden besteht die Möglichkeit, die Messreihen als sog. „Equalization-File“ zu speichern. Um diese Funktion zu nutzen, muss eine Norm-Frequenz angegeben werden, auf deren Pegel alle Messwerte normiert werden. Die Umwandlung der Equalization-Files in ein Beampattern übernimmt ein Matlab Skript. Zunächst werden hierbei alle Messwerte entnormiert und in dBV umgerechnet, im Anschluss werden alle Messreihen entsprechend des Messwinkels zusammengesetzt.

Bevor die Messreihen begonnen werden, wird der Amplitudengang des gesamten in Abbildung 8.2 dargestellten Systems aufgenommen. Betrachtet wird jedes verwendete Mikrofon separat, somit ergeben sich 8 Graphen:

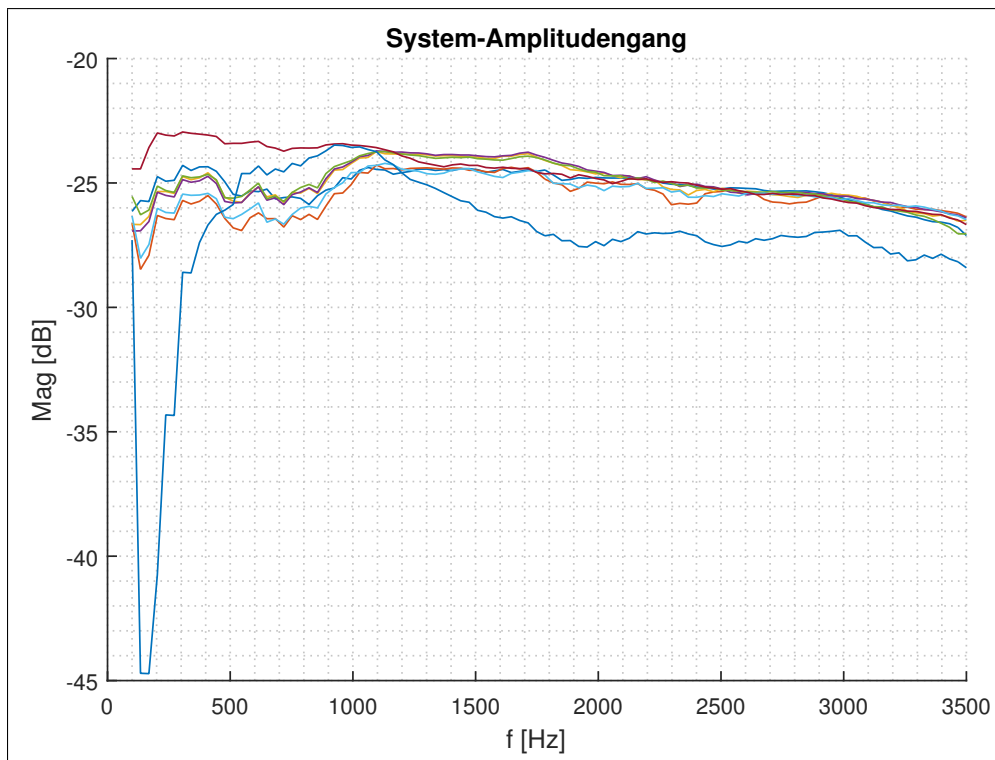
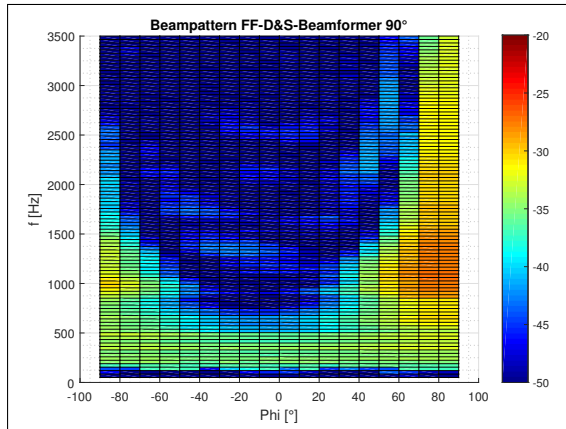


Abbildung 8.5: Amplitudengang des Übertragungssystems

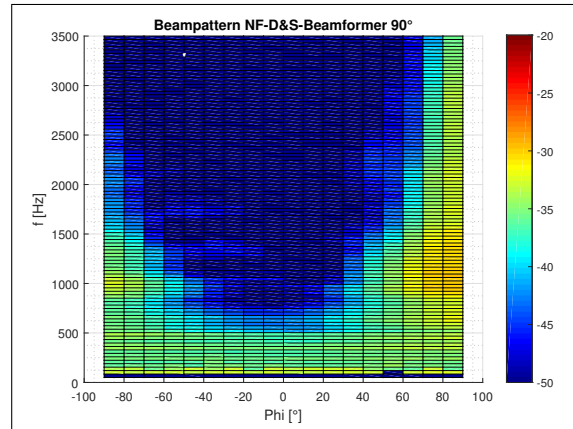
Die Amplitudengänge sind größtenteils sehr ähnlich, die Ausnahme bildet der blau dargestellte Pfad über Mikrofon 7. Im sehr niedrigen Frequenzbereich von ca. 0 – 250 Hz ist ein deutlicher Einbruch der Empfangsleistung erkennbar. Eventuelle Pegelunterschiede im Endergebnis des Beamformings bei tiefen Frequenzen könnten eventuell hierauf zurückgeführt werden.

8.3 Messergebnisse ULA

8.3.1 D&S-Beamformer

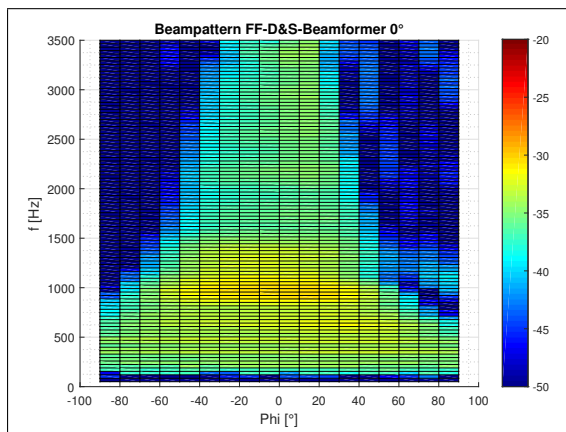


(a) Fernfeld-Koeffizienten

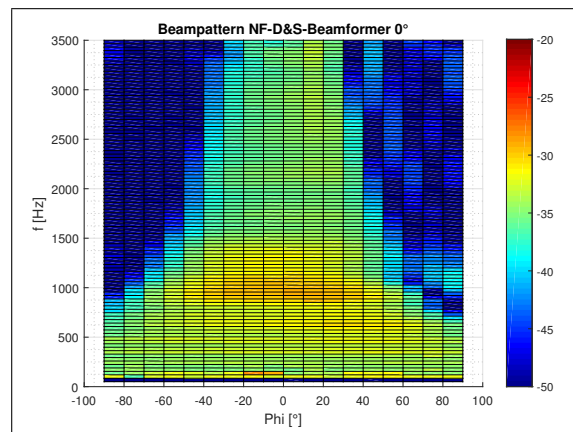


(b) Nahfeld-Koeffizienten

Abbildung 8.6: Beampattern, 90° Delay and Sum Beamformer



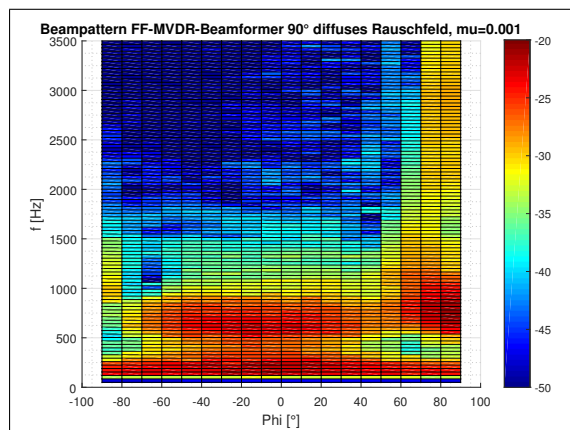
(a) Fernfeld-Koeffizienten



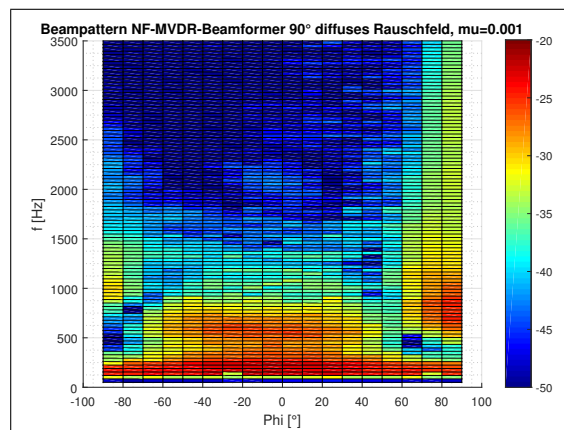
(b) Nahfeld-Koeffizienten

Abbildung 8.7: Beampattern, 0° D&S-Beamformer

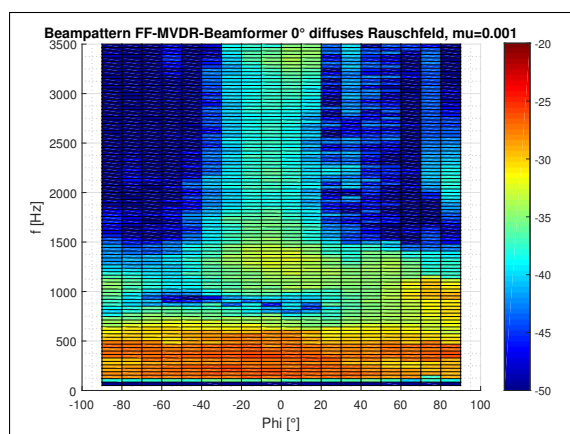
8.3.2 MVDR-Beamformer für diffuses Rauschfeld



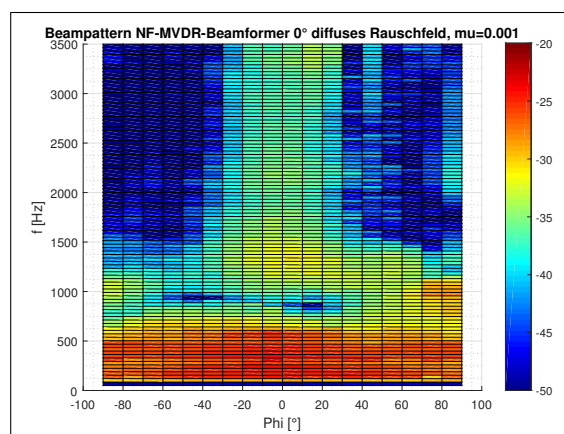
(a) Fernfeld-Koeffizienten



(b) Nahfeld-Koeffizienten

Abbildung 8.8: Beampattern, 90° MVDR-Beamformer für diffuse Rauschumgebung

(a) Fernfeld-Koeffizienten



(b) Nahfeld-Koeffizienten

Abbildung 8.9: Beampattern, 0° MVDR-Beamformer für diffuse Rauschumgebung

8.3.3 MVDR-Beamformer für zylindrisch isotropes Rauschfeld

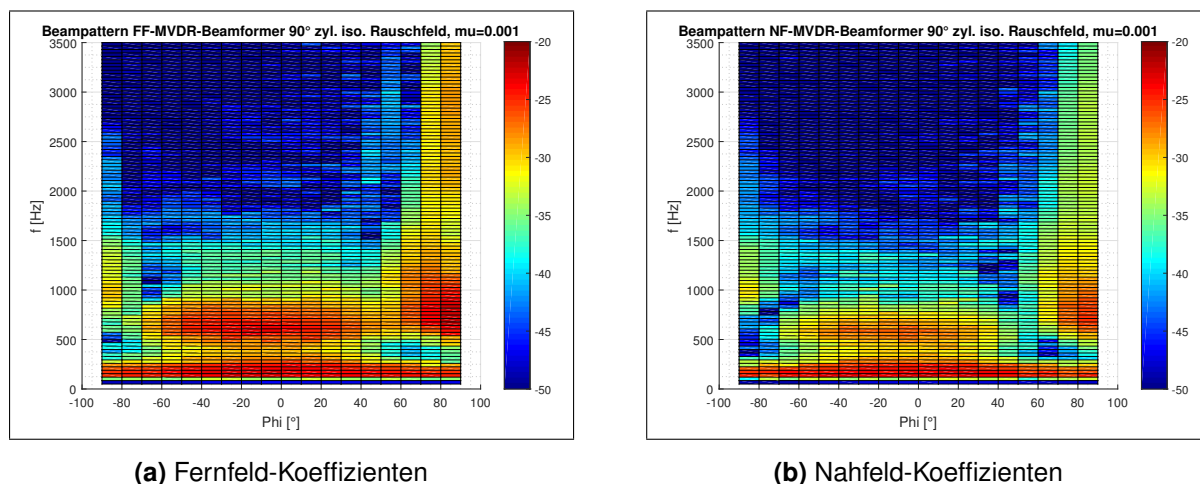


Abbildung 8.10: Beampattern, 90° MVDR-Beamformer für zylindrisch isotropische Rauschumgebung

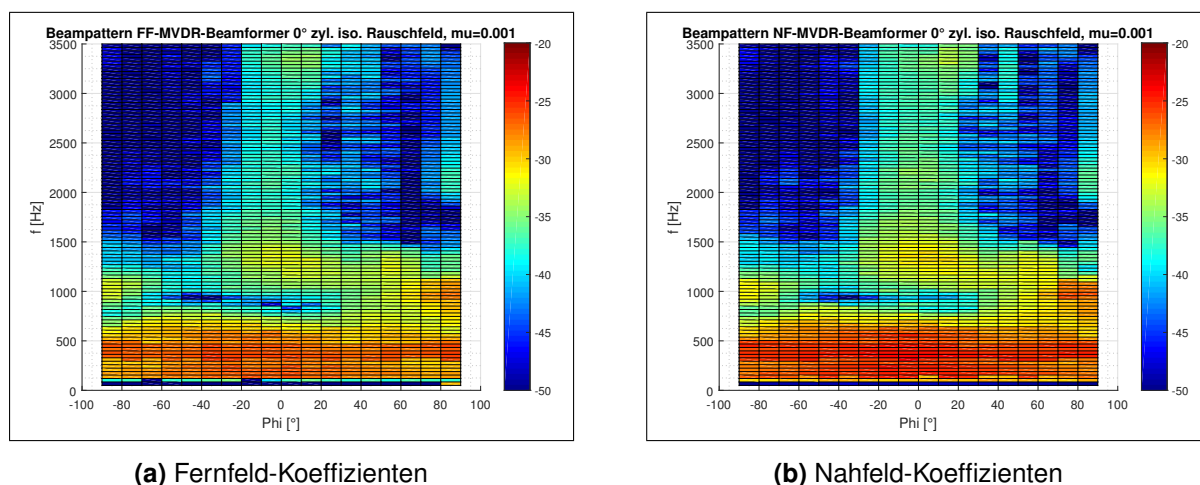


Abbildung 8.11: Beampattern, 0° MVDR-Beamformer für zylindrisch isotropische Rauschumgebung

8.3.4 LCMV-Beamformer

Da das LCMV-Beampattern insgesamt dem MVDR-Beampattern entspricht, wurde an dieser Stelle nur eine Messung vorgenommen. In dieser Messung liegt das Hauptaugenmerk darauf, zu beurteilen ob die Nullstellen, die zusätzlich eingebracht werden, erkennbar bzw. messbar sind.

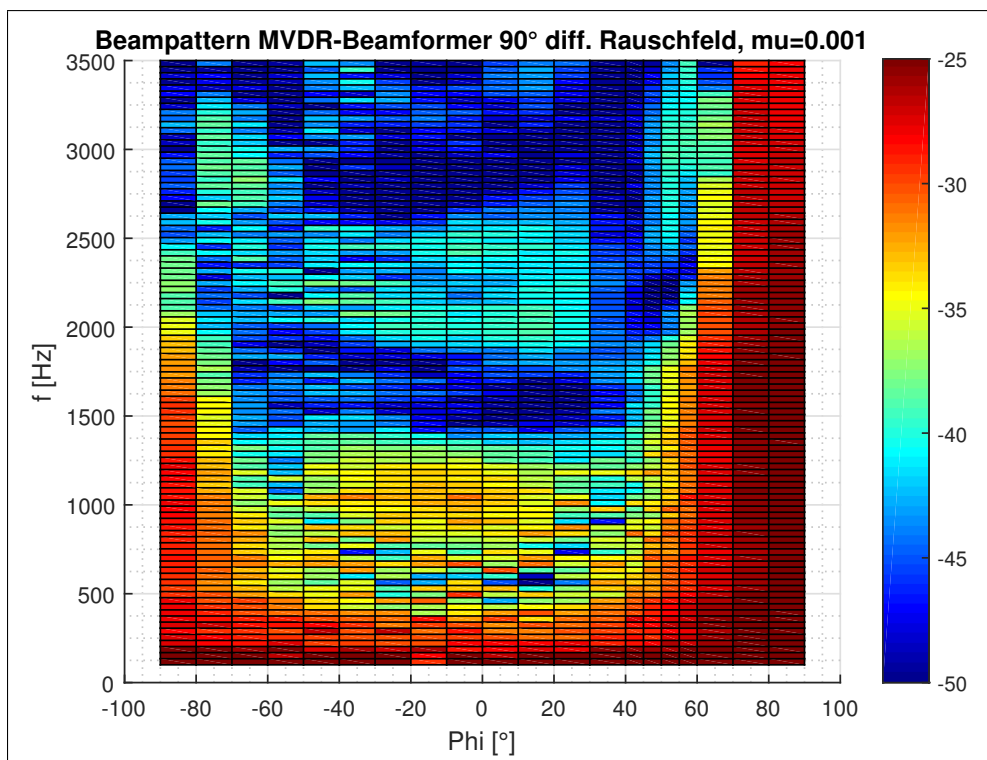


Abbildung 8.12: Beampattern, 90° 50° LCMV-Beamformer

Anhand der Grafik lassen sich die Nullstellen nicht erkennen. Die Auflösung der Winkelschritte wurde zwar im Bereich um 50° erhöht, dies führt jedoch nicht zu besseren Ergebnissen.

Ein Grund für das Fehlverhalten des Beamformers könnte in der eventuell ungenauen Beschreibung des Arrayverhaltens liegen. Die erzeugten Nullstellen sind sehr schmal und somit sehr von einer exakten mathematischen Beschreibung des Systems abhängig. Dies in Zusammenhang mit der nur groben Einstellung der Messumgebung könnte zum Verschwinden der Nullstellen führen.

Zur Überprüfung dieser Hypothese ist in Anhang [B] eine Simulation des Beampatterns für den beschriebenen LCMV-Beamformer abgebildet. Es wurde dabei eine zufallsverteilte Winkelabweichung von bis zu $\pm 2.5^\circ$ angenommen. Das Ergebnis zeigt, dass es theoretisch möglich ist, durch diese Art von Messfehlern Nullstellen zu „übersehen“.

8.3.5 Auswertung

Die Messergebnisse zeigen, dass alle Beamformer eine räumliche Filterwirkung erzielen. Allen voran der D&S-Beamformer, dessen Ergebnis fast exakt den Erwartungen entspricht. Auffällig ist jedoch, dass gerade im tiefen Frequenzbereich bis ca. 700 Hz von keinem Beamformer eine nennenswerte Filterwirkung zustande kommt. Während dieses Ergebnis im Fall der Delay and Sum Implementierung auch in der Simulation zu beobachten war, überrascht an dieser Stelle

das Ergebnis der MVDR-Beamformer.

Anhand der Simulation ist bekannt, dass im kritischen Frequenzbereich für Winkel von -90° bis 90° zwei Nebenkeulen mit ≈ -13 dB Pegel auftreten. Weiterhin sind die auftretenden Nullstellen sehr schmal gegenüber den Keulen. Dies kann dazu führen, dass auf Grund des Abstands der Winkelstützstellen von 10° die Nullstellen nicht erfasst werden. Dies ist im Vergleich der MVDR-Beamformer zwischen dem 0° und dem 90° Entwurf gut erkennbar.

Zusätzlich sind im tiefen Frequenzbereich die auftretenden Zeitdifferenzen und somit Phasenverschiebungen der Mikrofonsignale sehr gering, sodass es schwierig ist, Ungenauigkeiten im eingestellten Winkel gegenüber dem Sollwinkel zu erkennen.

Die Messungen im 90° Entwurf beider MVDR-Beamformer zeigen eine vermeintliche Überhöhung im Bereich von 0° . Tatsächlich ist diese auch in der Simulation zu sehen, da sich auf Grund der Symmetrie zwei Nebenkeulen treffen¹. Löst man die Farbskala feiner auf bzw. betrachtet die konkreten Werte, sieht man, dass der gemessene Pegelunterschied zur Hauptkeule von ca. -5 dB ziemlich genau dem simulierten Wert entspricht. Der 0° Entwurf zeigt bei feinerer Auflösung der Farbskala auch Pegelunterschiede von ca. 6 dB zwischen Hauptkeule und Nebenkeulenbereich. Dieser Wert liegt allerdings deutlich unter den Erwartungen von ≈ 12 dB.

Der verbleibende hohe Frequenzbereich aller MVDR-Beamformer zeigt in etwa das erwartete Bild. Die erreichte Dämpfung neben der Hauptkeule erreicht Werte im Bereich von -10 dB bis -15 dB, was dem Pegel der in der Simulation beobachteten Nebenkeulen entspricht. Die Symmetrie der Simulation ist in den Messungen nicht erkennbar. Dies ist jedoch leicht mit der Beschaffenheit des Arrays zu erklären. Während bei positiven Winkeln die Mikrofone dem Array zugewandt sind, werden sie bei negativen Winkeln mehr und mehr von hinten „angestrahlt“. Dies verschlechtert erwartungsgemäß den Empfang.

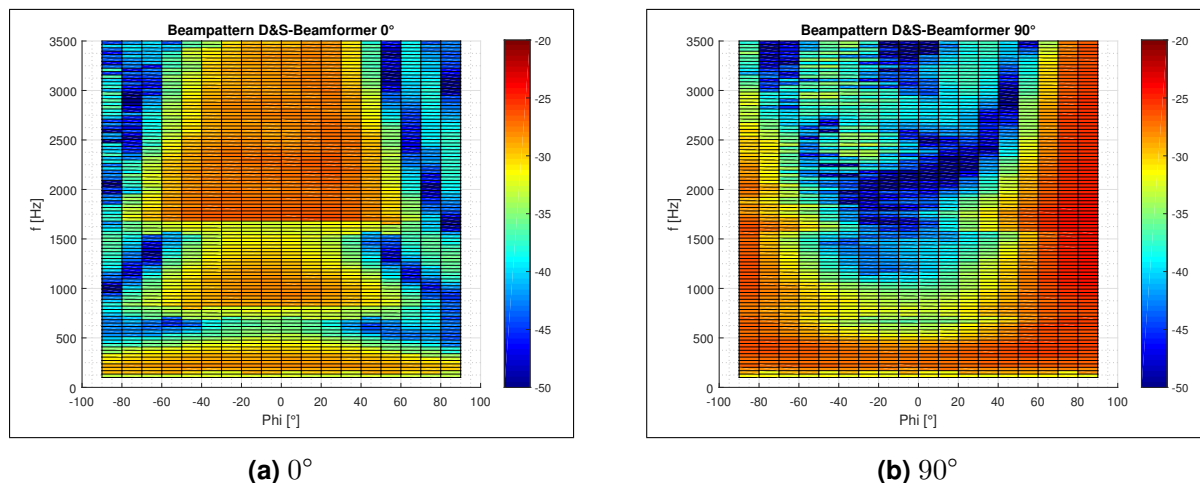
Der LCMV-Beamformer zeigt nicht die in der Simulation beobachteten Muster. Einige Ursachen hierfür sind im vorangegangenen Abschnitt bereits erläutert worden.

Zusammengefasst lässt sich sagen, dass die implementierten MVDR- und D&S-Beamformer funktionsfähig sind. Der LCMV-Beamformer funktioniert, zeigt aber nicht die gewünschten Nullstellen. Der D&S-Beamformer zeigt durchweg das erwartete Verhalten, während die MVDR-Beamformer vor allem im tiefen Frequenzband Leistungseinbußen erkennen lassen. Mit Verweis auf die geplanten Anwendungsszenarien, in denen vor allem Sprache verarbeitet werden soll, könnte dies zu Problemen führen.

¹Eine entsprechende Abbildung ist im Anhang [B] zu finden

8.4 Messergebnisse Subarray-ULA

8.4.1 SubarrayD&S-Beamformer

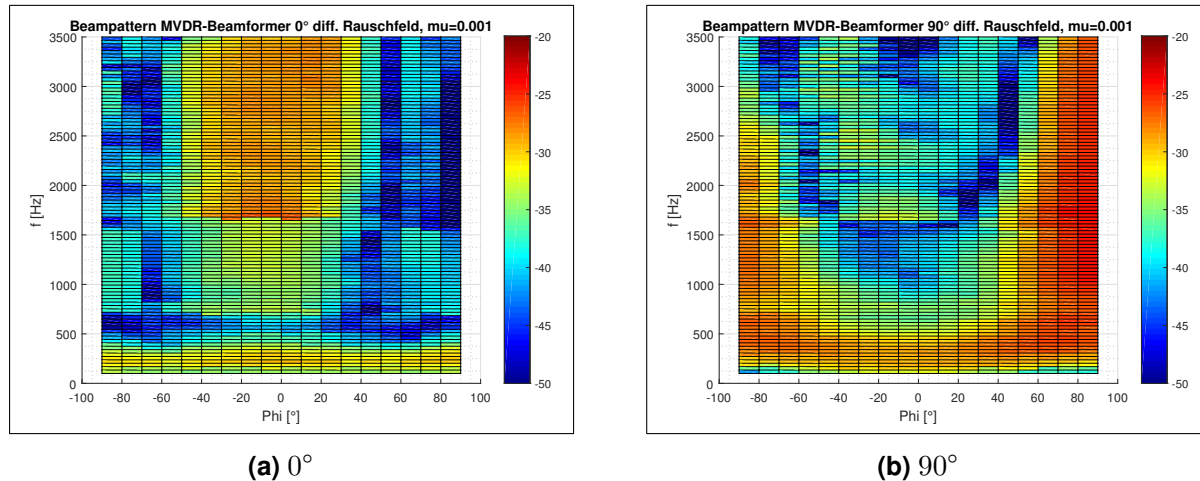


(a) 0°

(b) 90°

Abbildung 8.13: Beampattern, Subarray D & S-Beamformer

8.4.2 Subarray MVDR Beamformer



(a) 0°

(b) 90°

Abbildung 8.14: Beampattern, Subarray MVDR-Beamformer für diffuses Rauschfeld

8.4.3 Auswertung

Wie aus der Simulation erwartet, sorgt die Unterarray-Struktur für ein konstanteres Beamformer-Verhalten über der Frequenz. Besonders gut ist dies beim MVDR-Beamformer zu erkennen. Ein deutlich erkennbarer Nachteil liegt, bedingt durch die halbe Mikrofonanzahl pro Array, in der Breite der Empfangskeule. Während beim ULA eine Breite von ca $\pm 40^\circ$ für 0° Entwurfswinkel

entsteht, sind es beim Unterarray schon fast $\pm 60^\circ$. Allerdings sollte nicht außer Acht gelassen werden, dass nicht der gesamte Empfangsbereich von 360° gemessen wurde, somit erscheint die Keule breiter als sie unter Umständen ist.

Das Verhalten im kritischen Frequenzbereichs des ULA von 0 – 700 Hz ist beim Subarray-Beamforming deutlich verbessert. Allerdings ist für sehr niedrige Frequenzen um bis zu 300 Hz, immer noch eine nur schwache Filterwirkung erkennbar. Ein Grund hierfür könnte das „schlechte“ Mikrofon 7 sein. Wie in Abbildung 8.5 gezeigt, ist fast kein Empfang bis 200 Hz vorhanden. Da in diesem Bereich das größte Unterarray eingesetzt wird, welches auch Mikrofon 7 verwendet, fehlt unter Umständen bis zu ein Viertel der Signalleistung zum Ausgleich der anderen Signale.

Zusammengefasst kann man sagen, dass das Subarray dem ULA vorgezogen werden sollte, wenn weniger die Hauptkeulenbreite, als die gute Leistung über einen großen Frequenzbereich wichtig ist.

9 Messungen mit Sprachsignal

Nachdem die Funktionsfähigkeit der Beamformer sichergestellt ist, wird nun ein Beamformer ausgewählt und die Wirkung auf ein Sprachsignal untersucht. Die erste Messung wird erneut im Schallmessraum erfolgen, im zweiten Versuch wird unter realen Bedingungen in einem Büro gemessen. Hieraus soll eine Aussage getroffen werden, ob eine Unterdrückung unkorrelierter oder diffuser Rauschfelder sinnvoll ist.

Gemessen wird im Folgenden das Beamformer-Ausgangssignal aus 90° Hauptempfangsrichtung und aus 0° Dämpfungsrichtung. Um eine sinnvolle Referenz zu haben, wird nicht die originale Sprachdatei, sondern das Ausgangssignal eines einzelnen Mikrofons aufgezeichnet. Aus dem Vergleich des Beamformer-Resultats und dem Ergebnis, welches ein einzelnes Mikrofon erreichen würde, können Rückschlüsse auf die Leistungsfähigkeit der Implementierung ohne störende Einflüsse der System-Übertragungsfunktion gezogen werden.

Der Messaufbau im ideal reflexionsfreien Schallmessraum entspricht weiterhin dem in Abbildung 8.1 bzw. 8.2 dargestellten Verfahren. Das verwendete Sprachsignal ist ein Sample einer weiblichen Sprecherin, das bereits in der Arbeit von Herrn Reermann [13] verwendet wird. Die Datei liegt im wave-Format (.wav) vor und ist ca. 17 Sekunden lang. Der vorgetragene Text ist folgender:

To administer medicine to animals is frequently a very difficult matter and yet sometimes it is necessary to do so. The simplest method is to mix the medicine with butter or some other grease and smirr on the nose of the animal from time to time. Naturally it will lick the grease off.

Die Resultate werden mit dem Audio-Analyzer triggergesteuert mit einer Abtastrate von 48 kHz aufgenommen und gespeichert. Durch die hohe Abtastrate ist die maximale Aufnahmedauer auf 10 Sekunden begrenzt. Es wird also nicht der gesamte Text genutzt.

Da das Beamformerverhalten bisher in erster Linie in der Frequenz betrachtet worden ist, wird bei der Auswertung auch vor allem das Spektrogramm¹ der Signale betrachtet. Das verwendete Spektrogramm nutzt eine 256-Punkte STFT bei einer Überlappung der Signalausschnitte von 64 Samples. Als Fensterfunktion wird ein Hamming-Fenster verwendet. Um neben den Spektrogrammen und den subjektiven Hörproben der Beamformer-Signale eine feste Messgröße zu

¹Das Spektrogramm ist die Darstellung der Leistung eines Signals über Frequenz und Zeit. Durch Fourier Transformation einzelner sich überlappender Signalausschnitte wird ein Spektrum berechnet, dieses Verfahren nennt sich Kurzzeit-Fourier-Transformation (Short Time Fourier Transformation, STFT). Die Amplitude ist meist in Farbabstufungen codiert um eine 2-dimensionale Darstellung zu erhalten. Das Ergebnis ist abhängig von der Länge der STFT der Überlappung der einzelnen Signalausschnitte sowie den gängigen Parametern der Fourier Transformation

haben, wird die Gesamtenergie der Ausgangssignale betrachtet. Da die Beamformer im Allgemeinen im tiefen Frequenzband bis ca. 1000 Hz in den Simulationen schlechte Ergebnisse zeigen, wird die Energiebetrachtung in den Bereich kleiner 1000 Hz und größer 1000 Hz aufgeteilt. Betrachtet wird jeweils das Integral der Leistung, bezogen auf die Bandbreite und die Messzeit¹. Der resultierende Wert ist also Leistung pro Hertz und Sekunde.

9.1 Ideale Umgebung

9.1.1 Referenzsignal

Zunächst wird das Referenzsignal betrachtet. Im Zeitbereich ergibt sich folgender Signalverlauf:

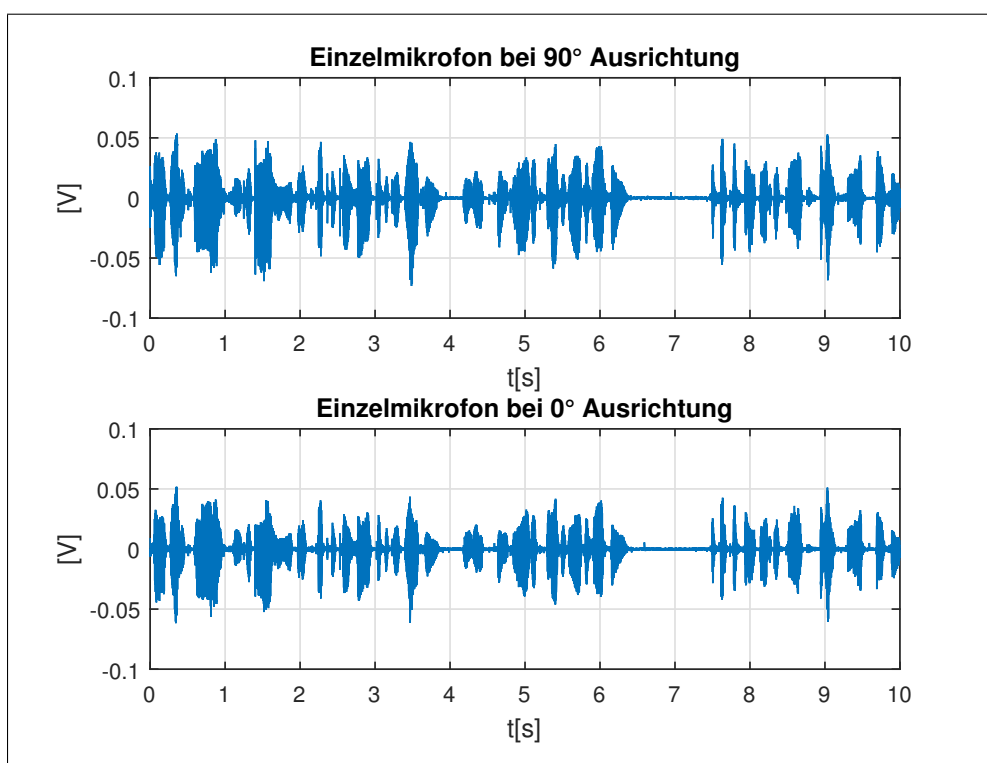


Abbildung 9.1: Referenzsignal an Mikrofon 3, weibliche Sprecherin

Das Referenzsignal hat folgendes Spektrogramm:

¹Es wird nicht über die gesamte Zeit integriert. Da das Mustersignal eine relativ große Sprachpause mit wenig Energie ab 6.4 s hat, endet das Integral hier um den Messwert exakter zu halten.

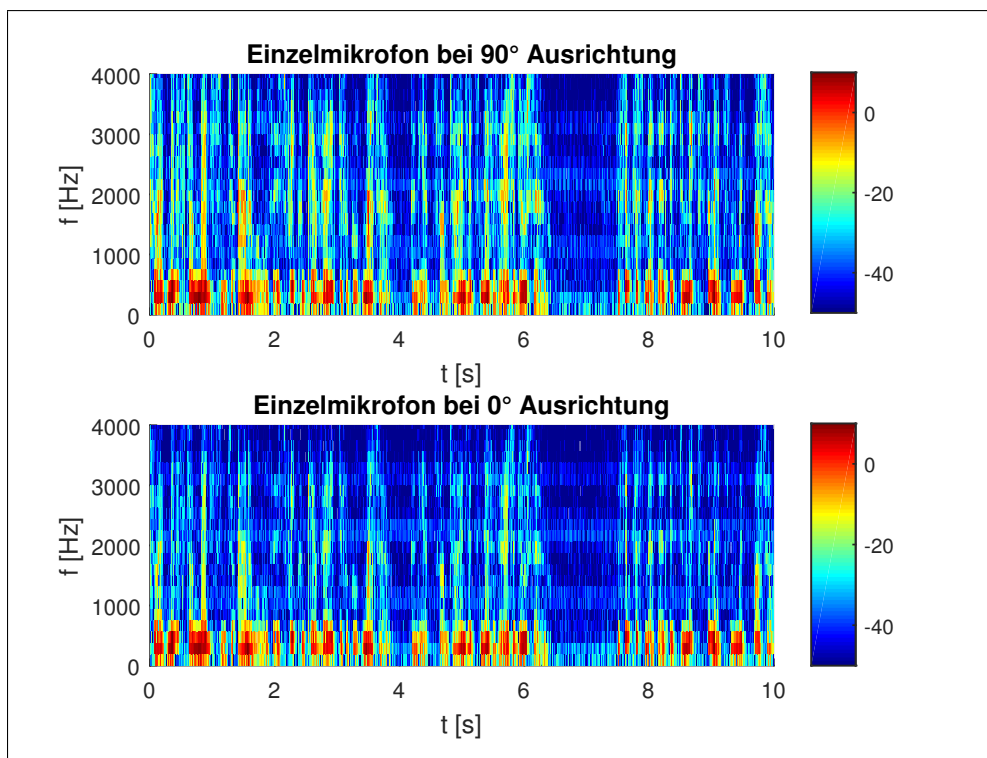


Abbildung 9.2: Spektrogramm, Referenzsignal an Mikrofon 3, weibliche Sprecherin

Die Durchschnittsleistungen ergeben folgende Werte:

90°	0 – 1 kHz	–36,35 dB
90°	1 – 3,5 kHz	–47,25 dB
0°	0 – 1 kHz	–37,8 dB
0°	1 – 3,5 kHz	–50,85 dB

Tabelle 9.1: Durchschnittsleistung Referenzsignal

Beide Darstellungsarten der Signale zeigen, dass der Unterschied zwischen 90° und 0° Ausrichtung vor allem Einfluss auf die hohen Frequenzen des einzelnen Mikrofons hat. Im Spektrogramm lässt sich gut erkennen, dass ein großer Teil der Signalenergie im niedrigen Frequenzbereich bis 1000 Hz auftritt. Da die ULA-Beamformer gerade in diesem Bereich Schwächen zeigen, könnte das Subarray im Realtest möglicherweise besser abschneiden.

9.1.2 Subarray D&S-Beamformer

Da der reflexionsfreie Schallmessraum am ehesten einem unkorrelierten Rauschfeld entspricht, wird zunächst der Subarray D&S-Beamformer verwendet. Die Messung liefert folgende Verläufe:

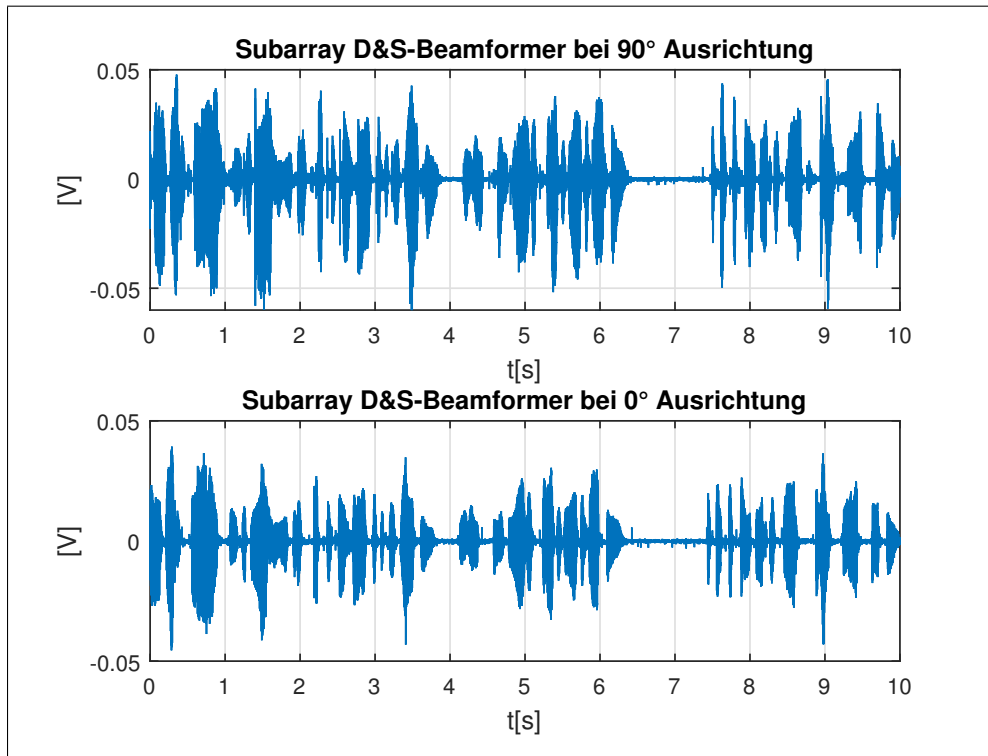


Abbildung 9.3: Ausgangssignal, D&S-Beamformer, weibliche Sprecherin

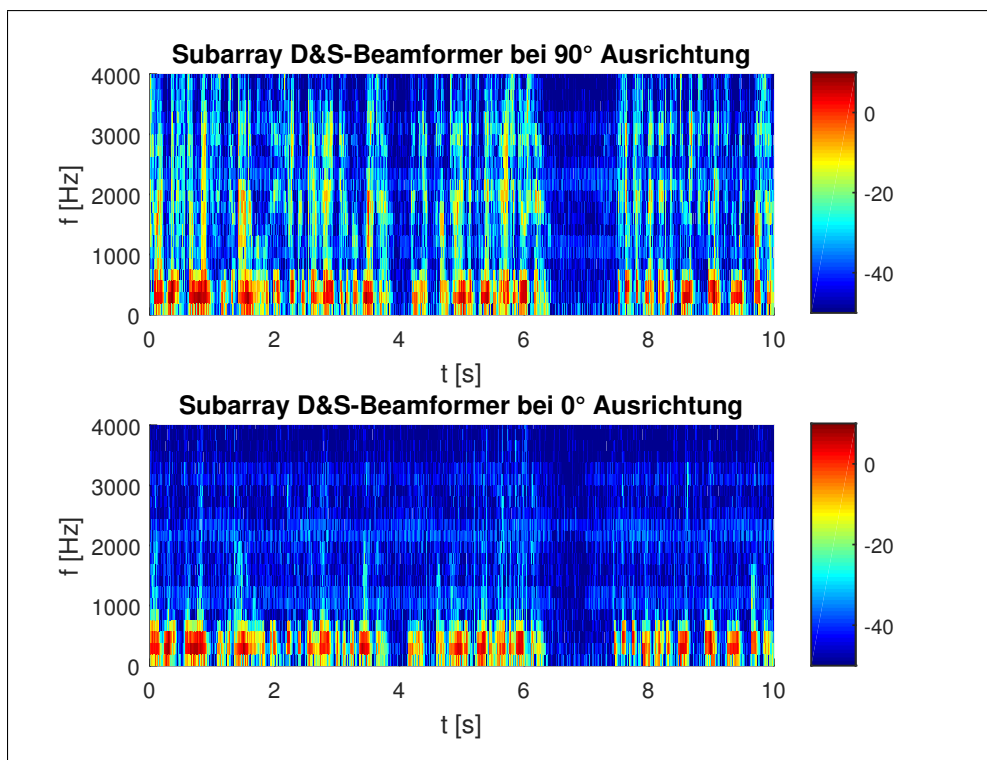


Abbildung 9.4: Spektrogramm, Ausgangssignal D&S-Beamformer, weibliche Sprecherin

Und folgende Durchschnittsleistungen:

90°	0 – 1 kHz	–38,78 dB
90°	1 – 3,5 kHz	–47,14 dB
0°	0 – 1 kHz	–2–42,9 dB
0°	1 – 3,5 kHz	–59,33 dB

Tabelle 9.2: Durchschnittsleistung Subarray D&-S-Beamformer

Anhand des Zeitsignals ist keine wirkliche Aussage über die Performance möglich. Man kann grob erkennen, dass die Gesamtenergie des Signals niedriger ist, außerdem sind die Pegel insgesamt kleiner.

Im Spektrogramm zeigt sich der Einfluss des Beamformers deutlicher. Die Durchschnittsdämpfung der hohen Frequenzanteile liegt bei ca. 12 dB, in manchen Bereich erreicht sie über 20 dB. Die Abschwächung im tiefen Frequenzbereich ist im Gegenzug deutlich geringer, die erreichte Dämpfung überschreitet nur selten mehr als 5 dB.

Insgesamt kann man sagen, dass das Signal räumlich betrachtet einer Tiefpassfilterung unterzogen wird. Hohe Frequenzanteile werden zuverlässig gedämpft, während niedrige Frequenzanteile nur schwach unterdrückt werden.

Auf der beiliegenden DVD sind die Ergebnisse des Beamformings als Hörprobe hinterlegt, siehe hierzu Anhang [C]. Anhand der aufgezeichneten Signale kann man die beschriebene Charakteristik gut „heraushören“. Das Signal aus 0° klingt deutlich dumpfer und erkennbar leiser als das Signal aus der Hauptempfangsrichtung von 90° .

9.1.3 Subarray MVDR-Beamformer

Im realen Szenario des Büros könnte der speziell auf diffuse Rauschfelder ausgelegte Subarray MVDR-Beamformer praktikabel sein. Als Referenz wird auch diese Variante im Schallmessraum getestet:

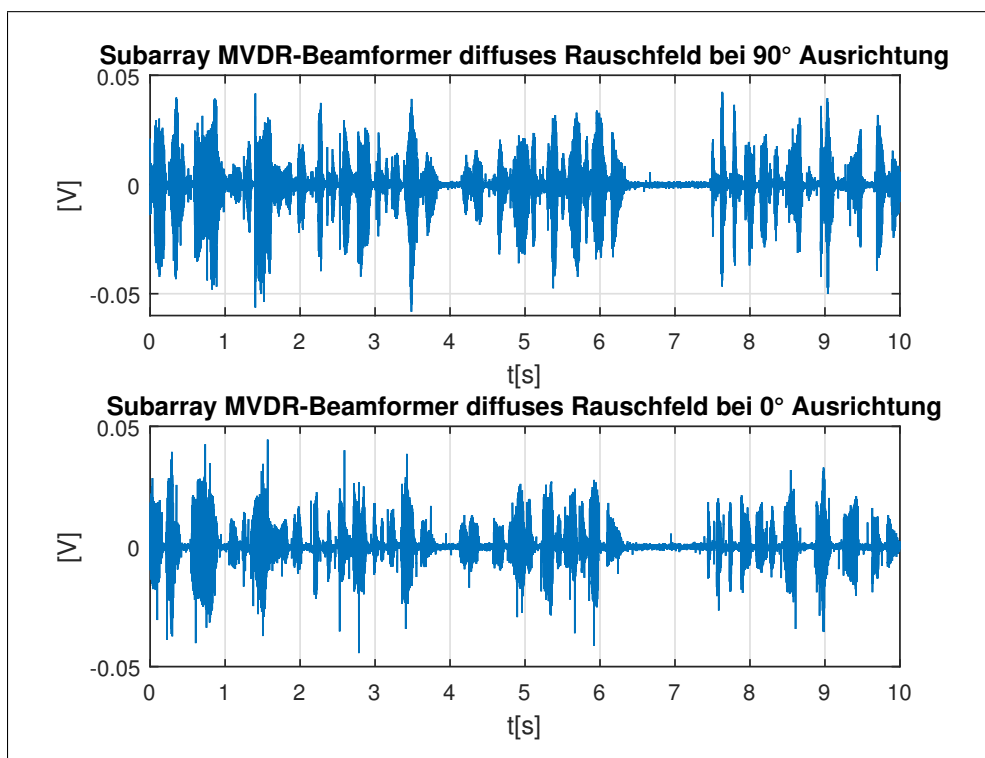


Abbildung 9.5: Ausgangssignal, MVDR-Beamformer für diffuses Rauschfeld, weibliche Sprecherin

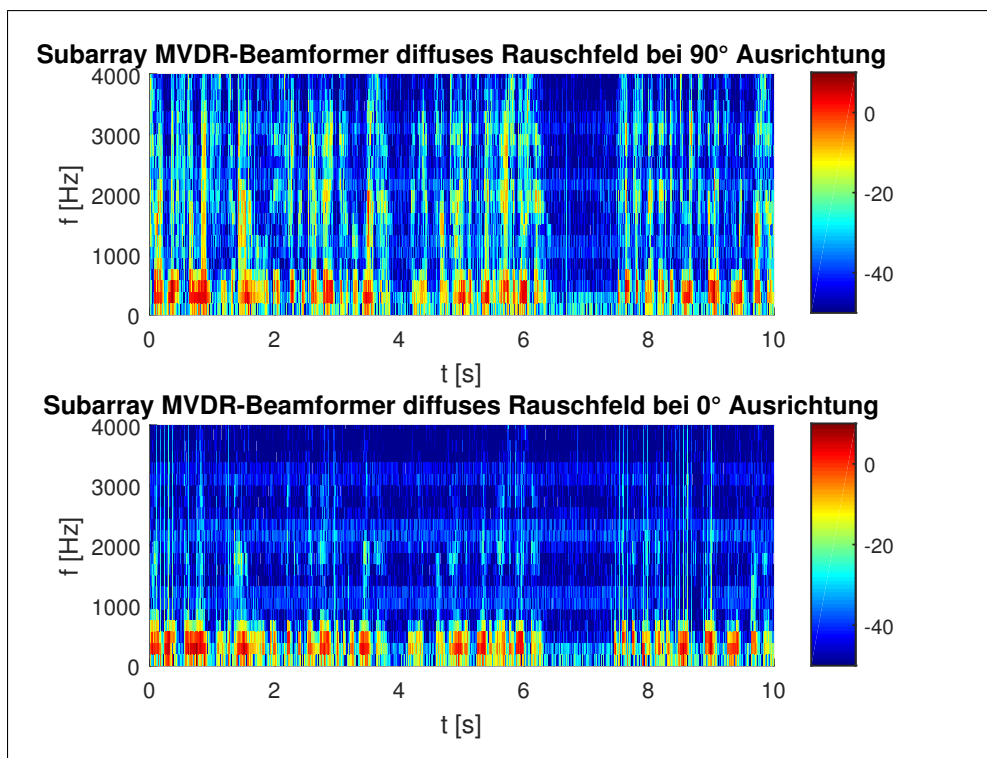


Abbildung 9.6: Spektrogramm, Ausgangssignal MVDR-Beamformer für diffuses Rauschfeld, weibliche Sprecherin

Mit den Durchschnittsleistungen:

90°	0 – 1 kHz	-39,8 dB
90°	1 – 3,5 kHz	-46,78 dB
0°	0 – 1 kHz	-43,03 dB
0°	1 – 3,5 kHz	-59,04 dB

Tabelle 9.3: Durchschnittsleistung MVDR-Beamformer für diffuses Rauschfeld

Auch in diesem Fall lässt das Zeitsignal nur wenige Rückschlüsse auf den Beamformer zu. Erkennbar ist, dass auch hier die Pegel etwas niedriger und generell „schmäler“ ausfallen.

Im Spektrogramm zeichnet sich ein ähnliches Bild mit gutem Dämpfungsverhalten bei hohen Frequenzen ab wie beim D&S-Beamformer. Ein wesentlicher Unterschied lässt sich dennoch erkennen: Lässt man die Unterdrückung der starken Frequenzen außer Acht, sieht man, dass das Hintergrundrauschen leicht verstärkt wird, dies zeigen auch die Durchschnittsleistungen. Besonders gut ist der Effekt im niedrigen Frequenzbereich zu erkennen: Während beim D&S-

Beamformer in den Sprachpausen dunkelblaue Bereiche im Spektrogramm sichtbar sind, werden diese im MVDR-Signal türkis gezeichnet.

Diese Beobachtung bestätigt das in der Simulation bereits beobachtete schlechte WNG Verhalten bei tiefen Frequenzen.

Wie beim D&S-Beamformer sind auch für diese Messungen Audiofiles hinterlegt. Das beschriebene Hintergrundrauschen ist im Dämpfungsfall bei 0° Ausrichtung sehr schön zu hören.

9.2 Reale Umgebung

Die Messung in realer Umgebung findet in einem Büro der HAW-Hamburg statt. Abbildung 9.7 zeigt den Grundriss des genutzten Raums 14.04:

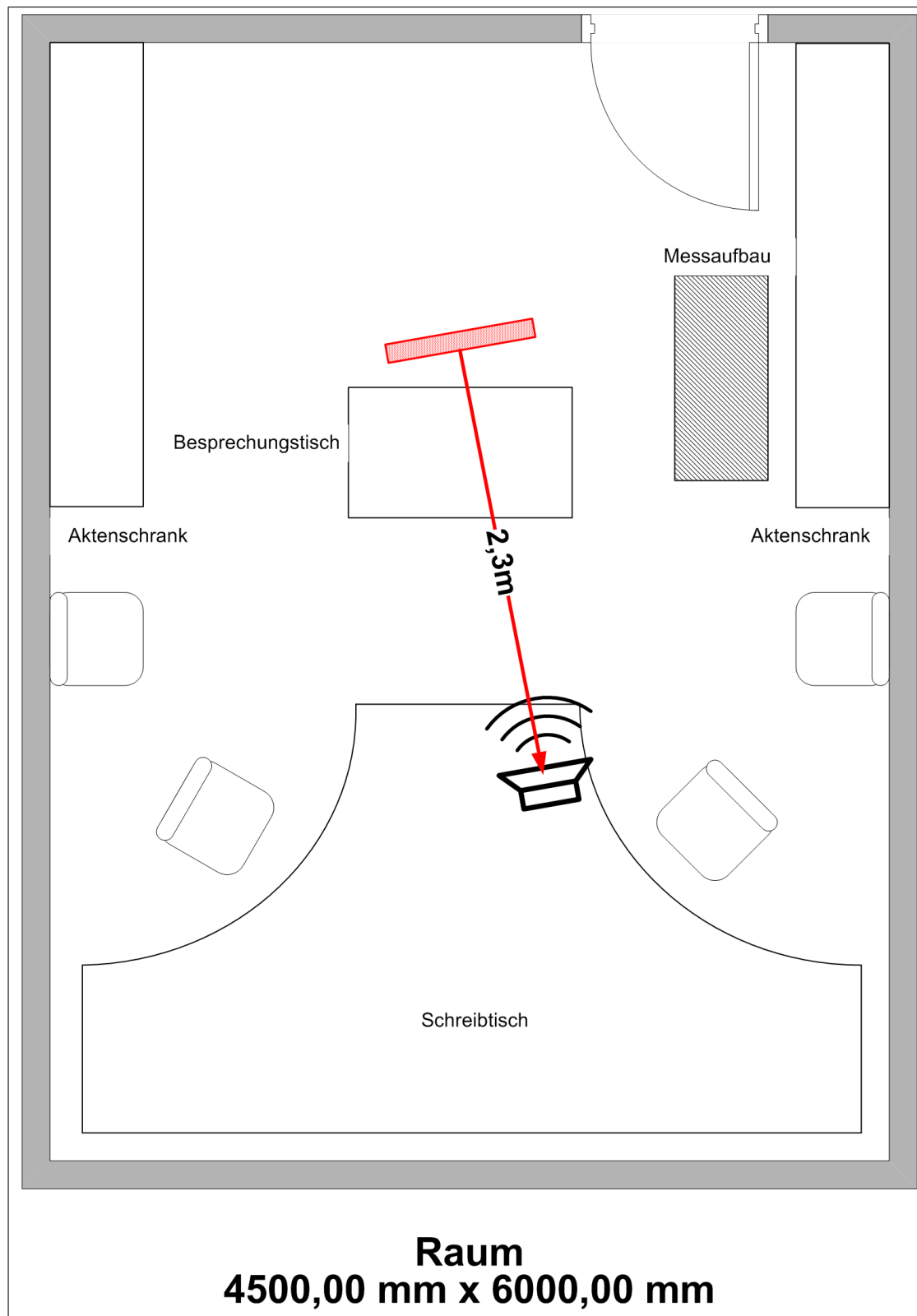


Abbildung 9.7: Grundriss des Versuchsbüros im Raum 14.04 der HAW-Hamburg

Der elektronische Messaufbau ist im Vergleich zum Schallmessraum, bis auf die Leitungslängen, unverändert.

Gemessen wird erneut zunächst ein einzelnes Mikrofonsignal als Referenz. Im Anschluss werden D&S- und MVDR-Beamformer zur räumlichen Filterung eingesetzt. Die Messungen werden sowohl mit dem ULA als auch mit dem Subarray-Aufbau durchgeführt.

9.2.1 Referenzsignal

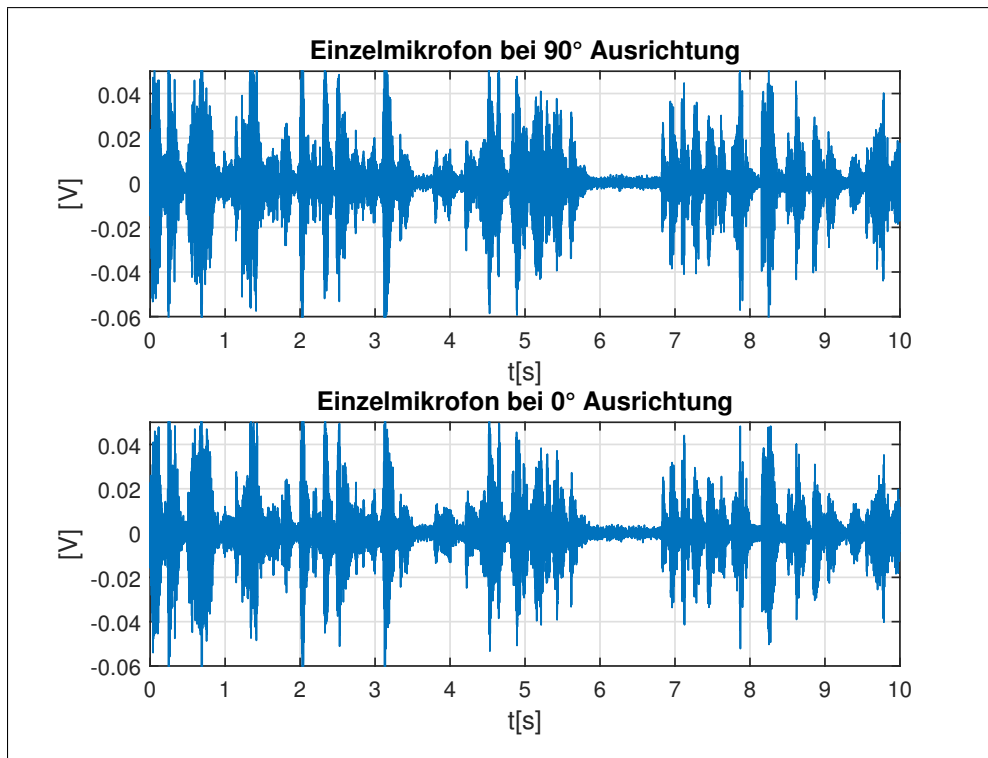


Abbildung 9.8: Referenzsignal an Mikrofon 3, weibliche Sprecherin in Bürourgebung

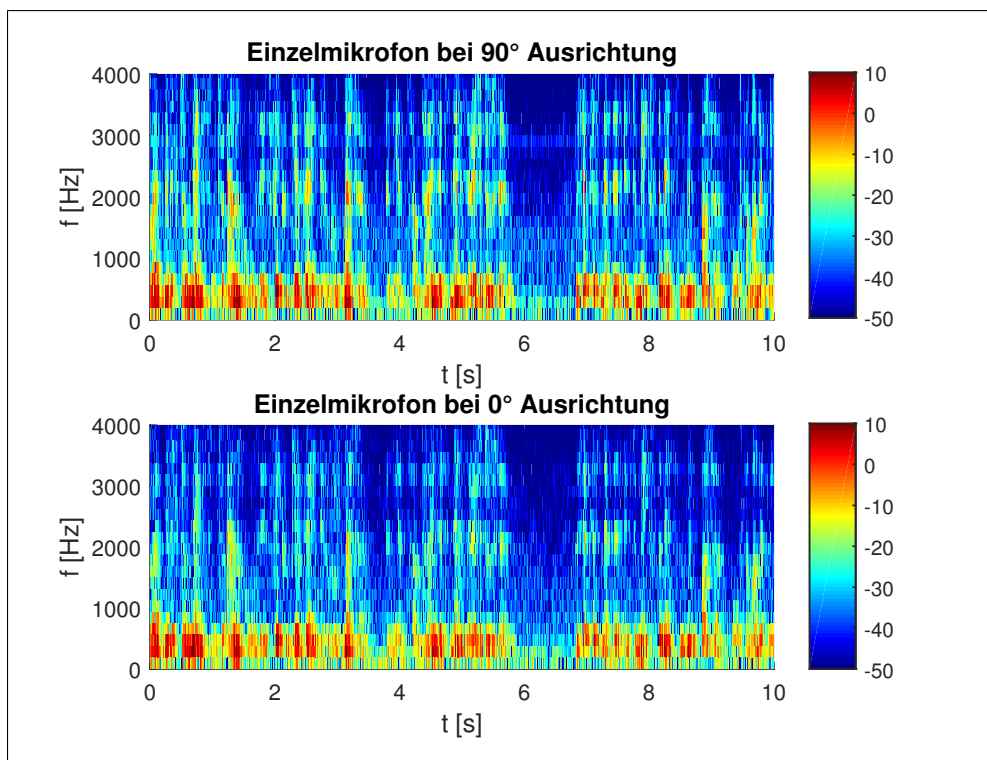


Abbildung 9.9: Referenzsignal an Mikrofon 3, weibliche Sprecherin in Büroumgebung

90°	0 – 1 kHz	-31,30 dB
90°	1 – 3,5 kHz	-48,72 dB
0°	0 – 1 kHz	-30,96 dB
0°	1 – 3,5 kHz	-51,56 dB

Tabelle 9.4: Durchschnittsleistung Referenzsignal

Das Referenzsignal zeigt keine wesentlichen Unterschiede zwischen 90° und 0° Ausrichtung. Im Vergleich zum Referenzsignal aus dem Schallmessraum ist zu erkennen, dass die Spektrogramme der Büromessung stärkere Anteile in den höheren Frequenzen aufweisen. Weiterhin ist insgesamt ein wesentlich stärkerer Rauschteppich als in der reflexionsfreien Umgebung erkennbar.

9.2.2 Subarray D&S-Beamformer

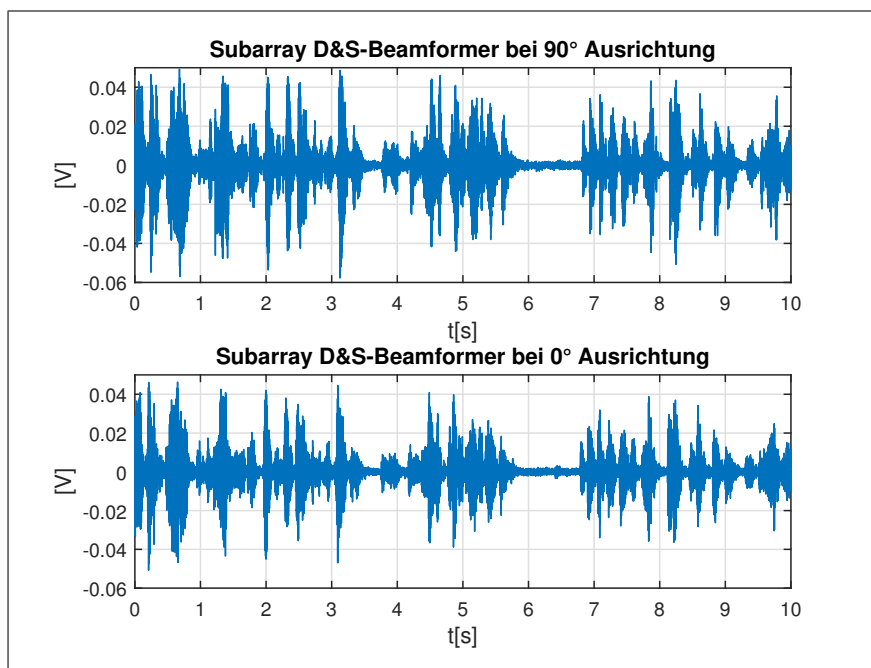


Abbildung 9.10: Ausgangssignal, Subarray D&S-Beamformer, weibliche Sprecherin in Büroumgebung

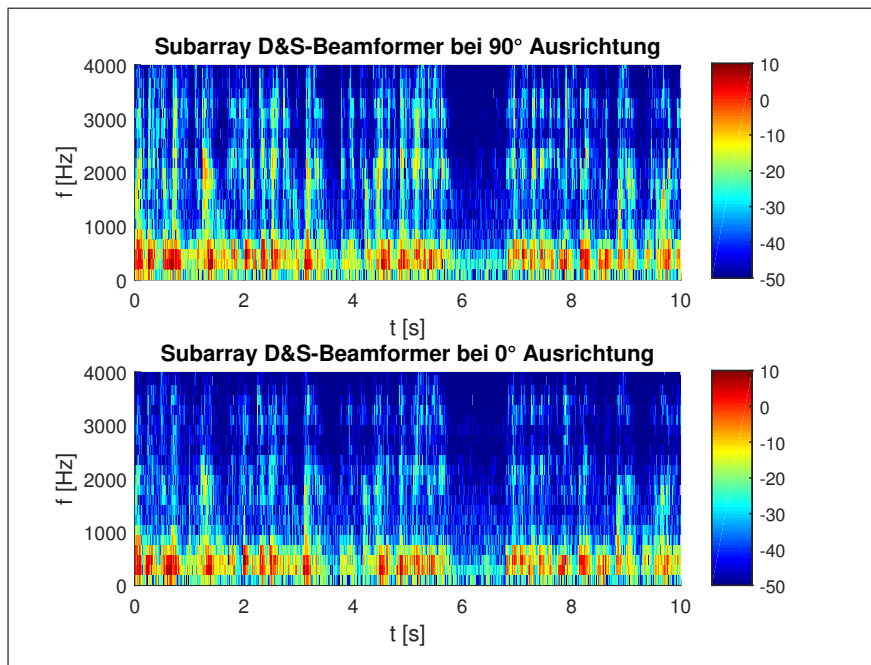


Abbildung 9.11: Spektrogramm, Ausgangssignal Subarray D&S-Beamformer, weibliche Sprecherin in Büroumgebung

90°	0 – 1 kHz	-34,44 dB
90°	1 – 3,5 kHz	-52,72 dB
0°	0 – 1 kHz	-35,70 dB
0°	1 – 3,5 kHz	-58,45 dB

Tabelle 9.5: Durchschnittsleistung Subarray D&S-Beamformer

Der Subarray D&S-Beamformer zeigt im Zeitbereich keine großen Unterschiede zwischen den beiden Ausrichtungen. Wie bereits im Schallmessraum beobachtet, ist die Höhe der Pegel insgesamt etwas geringer bei schmalerer „Signalbreite“. Die aus den Messungen in idealer Umgebung auftretende Dämpfung hoher Frequenzen lässt sich auch im realen Anwendungsfall beobachten. Vor allem in den höheren Frequenzbändern ab 750 Hz ist wieder eine Filterwirkung erkennbar. Die Dämpfung beträgt im Schnitt 6 dB, es werden an manchen Stellen Werte bis 12 dB erreicht. Die Unterdrückung des Rauschens ist vor allem in den Sprechpausen erkennbar. Vor allem in den tiefen Frequenzen ist die Dämpfung allerdings deutlich schwächer als im Schallmessraum. Dies lässt sich auf das angenommene diffuse Rauschfeld zurückführen, bei dem Rauschen nicht als unkorreliert betrachtet wird und somit nicht weiß ist.

Die zugehörigen Audio-Files sind auf der DVD zu finden. Die Hörproben zeigen, dass das Audiosignal im Dämpfungsfall klingt, als wäre die Quelle weit vom Hörer entfernt.

9.2.3 Subarray MVDR-Beamformer

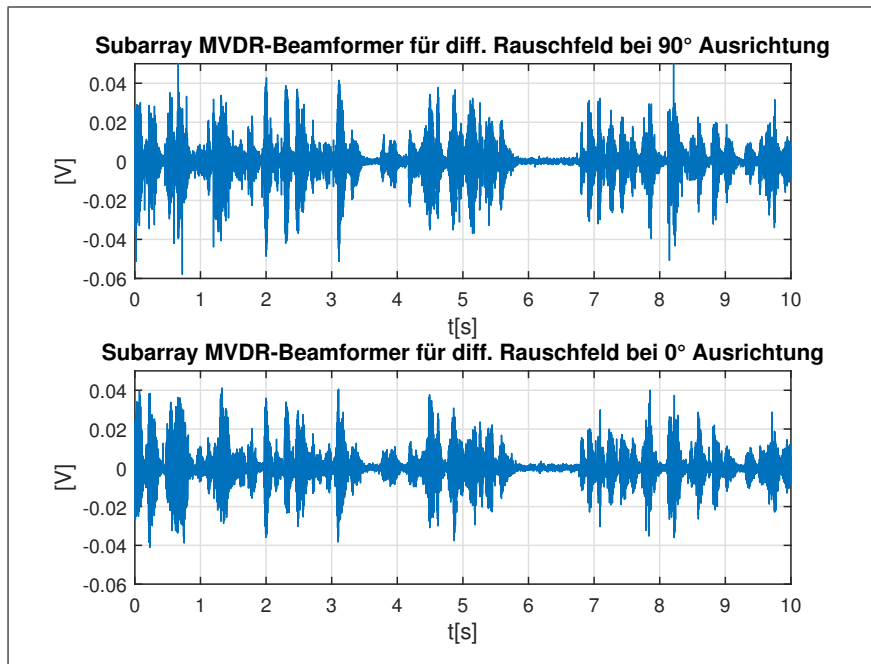


Abbildung 9.12: Ausgangssignal, Subarray MVDR-Beamformer für diffuses Rauschfeld, weibliche Sprecherin in Büroumgebung

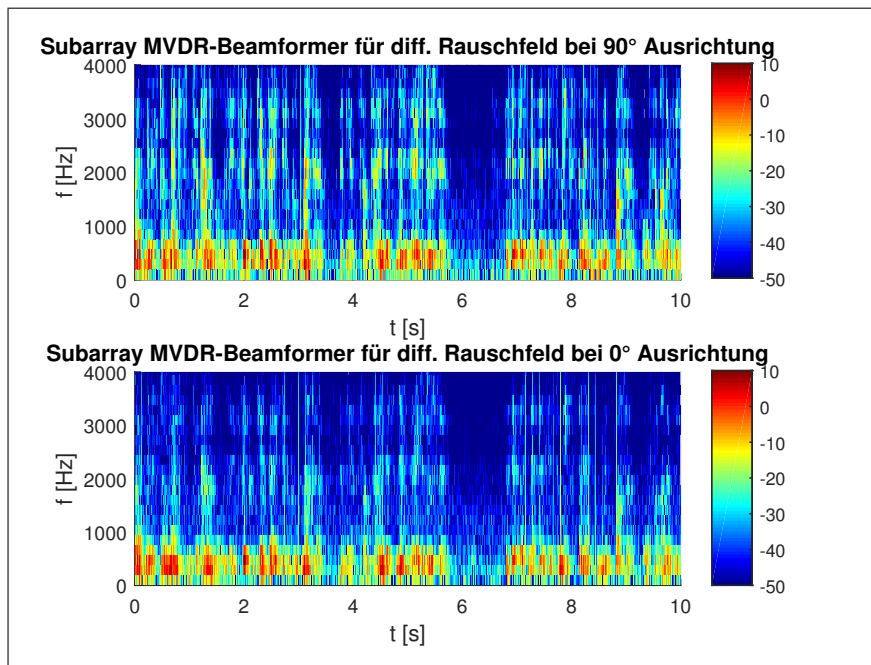


Abbildung 9.13: Spektrogramm, Ausgangssignal Subarray MVDR-Beamformer für diffuses Rauschfeld, weibliche Sprecherin in Büroumgebung

90°	0 – 1 kHz	-36,92 dB
90°	1 – 3,5 kHz	-53,22 dB
0°	0 – 1 kHz	-36,76 dB
0°	1 – 3,5 kHz	-58,67 dB

Tabelle 9.6: Durchschnittsleistung Subarray MVDR-Beamformer für diffuses Rauschfeld

Entgegen der Erwartungen, sieht das Ergebnis des MVDR-Beamformers für diffuses Rauschfeld kaum anders aus, als das des D&S-Beamformers. Die erkennbare Filterwirkung zeigt sich erneut vor allem im hochfrequenten Bereich.

9.2.4 ULA D&S-Beamformer

Da die Subarray-Beamformer nicht an die Performance aus der Messungen im Schallmessraum herankamen, wird die verbleibende Messzeit genutzt, um das reguläre ULA zu testen.

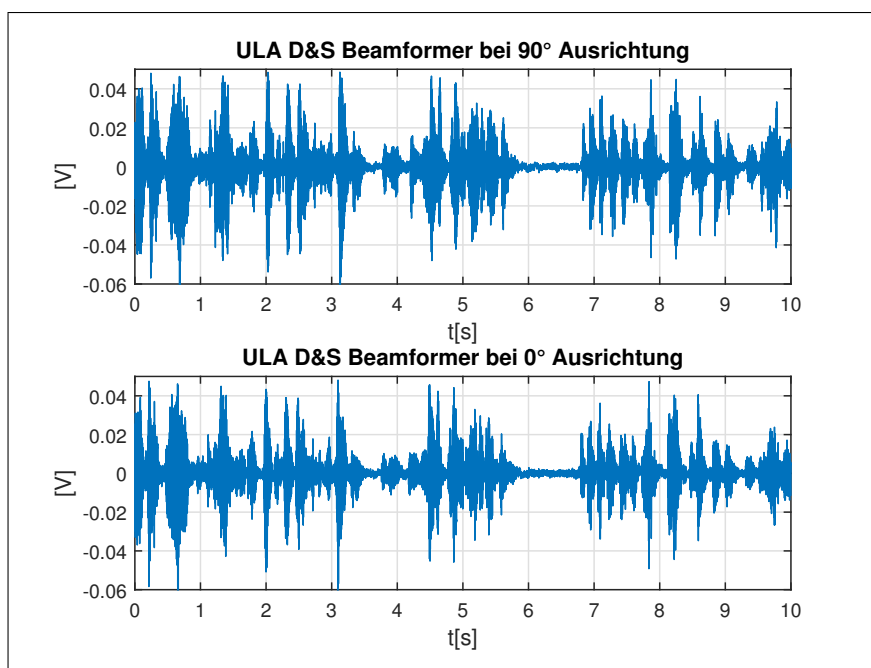


Abbildung 9.14: Ausgangssignal, D&S-Beamformer, weibliche Sprecherin in Büroumgebung

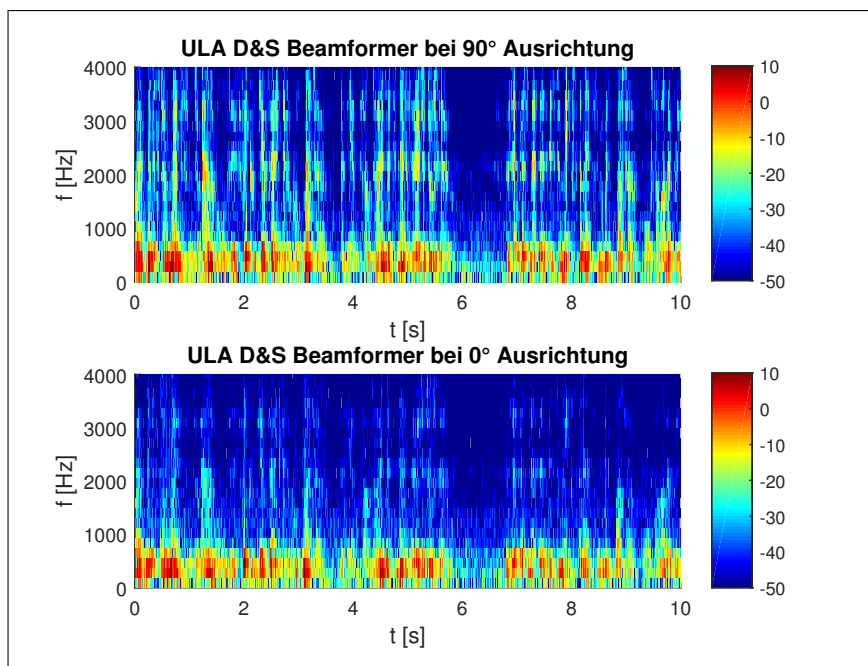


Abbildung 9.15: Spektrogramm, Ausgangssignal D&S-Beamformer, weibliche Sprecherin in Büroumgebung

90°	0 – 1 kHz	–34,19 dB
90°	1 – 3,5 kHz	–54,63 dB
0°	0 – 1 kHz	–34,92 dB
0°	1 – 3,5 kHz	–62,77 dB

Tabelle 9.7: Durchschnittsleistung ULA D&S-Beamformer

Der Vergleich zwischen Subarray und ULA D&S-Beamformer zeigt sehr deutlich die Unterschiede beider Implementierungen. Betrachtet man beide Spektrogramme bei 90° erkennt man, dass das Hintergrundrauschen des ULA wesentlich geringer ausfällt als beim Subarray. Die Dämpfungscharakteristik zeigt auch deutliche Vorteile zu Gunsten des ULA, welches zwar erst ab ca. 1000 Hz dämpft, aber dafür deutlich höhere Dämpfungswerte in den hohen Frequenzen erreicht. Die Betrachtung der Durchschnittsleistung zeigt dasselbe Bild. Die Energie im Sperrfall ist um ca. 8 dB geringer als im Empfangsfall, während das Subarray nur 4 dB erreicht. Im Fall des D&S-Beamformers lässt sich also feststellen, dass die höhere Nebenkeulendämpfung des ULA zu besseren Ergebnissen führt, als der im Gegenzug konstantere Empfangsbereichsverlauf des Subarrays. Dies liegt vor allem darin begründet, dass der Gewinn in Konstantheit der Hauptkeulenbreite zum einen durch eine schwächere Dämpfung und zum anderen durch insgesamt breitere Hauptkeule „erkauft“ wird.

9.2.5 ULA MVDR-Beamformer

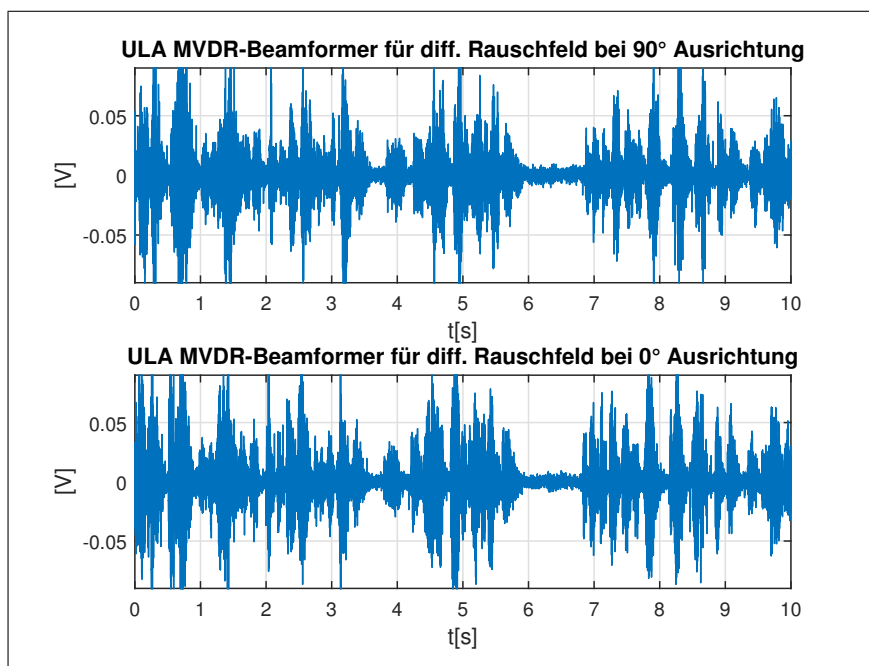


Abbildung 9.16: Ausgangssignal, MVDR-Beamformer für diffuses Rauschfeld, weibliche Sprecherin in Büroumgebung

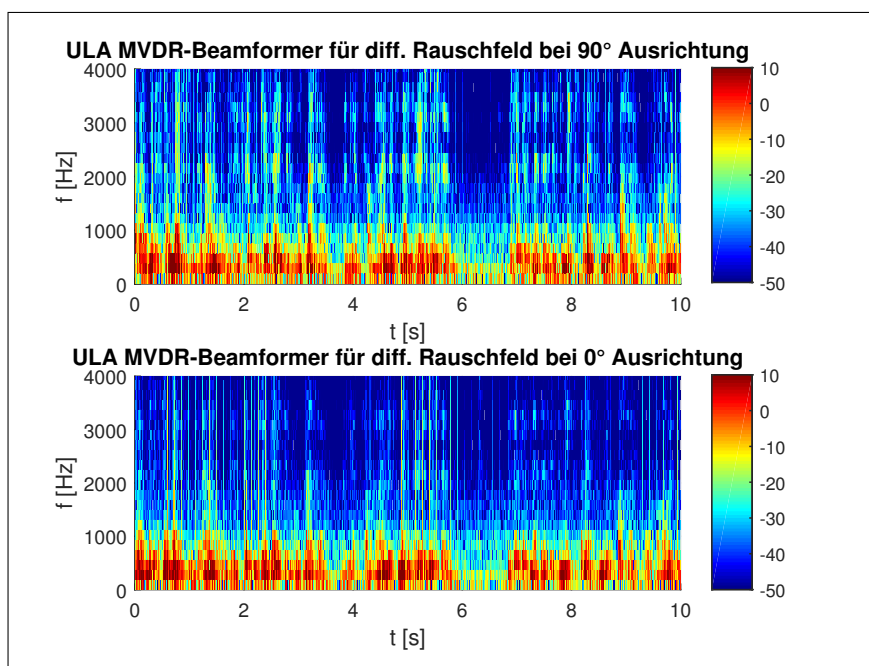


Abbildung 9.17: Spektrogramm, Ausgangssignal MVDR-Beamformer für diffuses Rauschfeld, weibliche Sprecherin in Büroumgebung

Die MVDR-Implementierung des ULA zeigt insgesamt eine höhere Signalenergie als alle anderen Implementierungen. Die erkennbare Filterwirkung ist vergleichbar zum Subarray. Anhand

90°	0 – 1 kHz	–19,32 dB
90°	1 – 3,5 kHz	–51,56 dB
0°	0 – 1 kHz	–18,84 dB
0°	1 – 3,5 kHz	–56,88 dB

Tabelle 9.8: Durchschnittsleistung ULA MVDR-Beamformer für diffuses Rauschfeld

der Spektrogramme, noch deutlicher durch Auswertung der Hörprobe, erkennt man einen relativ starken „dumpfen“ Rauschteppich. Dieser bestätigt die theoretischen Erwartungen und erklärt die höhere Signalenergie.

9.3 Bewertung

Die Messungen mit Sprachsignal haben zu unterschiedlichen Erkenntnissen geführt. Die Messungen haben gezeigt, dass im idealen Raum beide ULA-Beamformer eine in etwa gleich gute Filterwirkung erzielen. Wenn mit wenigen Reflexionen zu rechnen ist, sollte der D&S-Beamformer bzw. der MVDR-Beamformer für unkorreliertes Rauschfeld den anderen MVDR-Beamformern auf Grund des besseren Rauschunterdrückungsverhaltens vorgezogen werden. Ist weniger eine schmale Winkelselektion als ein konstanter Empfangsbereichsverlauf über der Frequenz entscheidend, bietet die Subarray-Implementierung eine gute Alternative. Bei der Verwendung eines Subarrays ist zu beachten, dass die Mikrofonanzahl pro Array deutlich reduziert ist, was insgesamt zu schlechterem Dämpfungsverhalten und höheren Nebenkeulen führt.

In reflexionsbehafteten Umgebungen, wie z.B. einem Büro, haben die Messungen gezeigt, dass bei begrenzter Arraybreite eine höhere Mikrofonanzahl einer schärferen Empfangskeule vorzuziehen ist. Während beide Arrayaufbauten die aus dem idealen Raum erwarteten Filterwirkungen zeigen, stellt sich heraus, dass mit den verwendeten Aufbauten nur Dämpfungen von 5 dB mit dem Subarray bzw. 8 dB mit dem ULA erreicht werden konnten.

Im Sinne der besseren Sprachverständlichkeit, ist aus den Messungen der D&S- bzw. der MVDR-Beamformer für unkorreliertes Rauschfeld als beste Variante hervorgegangen. Für den gewählten Anpassungsfaktor von $\mu = 0.001$ ist die Verstärkung des weißen Rauschens der MVDR-Beamformer für diffuses Rauschfeld zu hoch. Man könnte diesen Faktor zwar weiter anpassen, läuft jedoch Gefahr, die durch die diffuse Implementierung gewonnen Vorteile wieder zu verlieren.

Insgesamt kann man sagen, dass der Minimierungsansatz der MVDR-Beamformer bei guter Nachbildung des tatsächlichen Rauschfeldes gute Ergebnisse liefert. Dies hat sich vor allem im Schallmessraum gezeigt, der gut ein unkorreliertes Rauschfeld nachbildet. Bei anderen Rauschumgebungen müssen detaillierte Messungen zum tatsächlichen Rauschfeld vorgenommen

werden, da im Büro die Annahme des diffusen Rauschfeldes zu keinen guten Ergebnissen geführt hat.

10 Zusammenfassung und Fazit

Im Laufe dieser Arbeit wurden verschiedene Möglichkeiten der räumlichen Filterung von Audiosignalen durch Beamforming betrachtet. Die ersten Untersuchungen haben gezeigt, dass eine zirkulare Arraystruktur mit 5 cm Durchmesser bei 8 Mikrofonen nicht geeignet ist. Die weiteren Analysen haben ergeben, dass ein gleichverteiltes lineares Array (ULA) mit 8 Mikrofonen bei 5 cm Abstand eine nutzbare Alternative darstellt.

Literaturrecherchen haben unterschiedliche Entwurfsmethoden für Beamformer aufgezeigt. Aufbauend auf diesen Erkenntnissen wurde zunächst der klassische „Delay-and-Sum“ bzw. D&S-Beamformer evaluiert. Die Simulationen haben gezeigt, dass der normale Entwurfsansatz funktionsfähig ist, jedoch für breitbandige Signale keine konstante Charakteristik aufweist. Eine Breitbandstruktur, die eine Frequenzerlegung durch DFT verwendet, hat in diesem Fall Abhilfe geschaffen. Nachdem ein prinzipieller Beamforming-Ansatz funktionsfähig vorhanden war, wurden spezifischere Beamformer entworfen. Diese sogenannten „Minimum Variance Distortionless Response“ - oder kurz MVDR-Beamformer - basieren auf Minimierungsalgorithmen mit einer Nebenbedingung und werden für definierte (Rausch-)Umgebungen entworfen. Nachdem das Verhalten dieser Art von Beamformern simuliert und verifiziert ist, wird der Minimierungsansatz auf mehrere Nebenbedingungen erweitert. Der resultierende „Linear Constraint Minimum Variance“ oder LCMV-Beamformer bildet den aufwendigsten entworfenen Filtertyp.

Die Messungen der unterschiedlichen Beamformer in realer Umgebung fand in der reflexionsfreien Umgebung des Schallmessraums an der HAW-Hamburg statt. Hierbei konnte die Funktionalität der D&S und der MVDR-Implementierung bestätigt werden. Der LCMV-Beamformer zeigt zwar räumliche Filterwirkung, allerdings konnte die als Nebenbedingung geforderte Nullstelle nicht nachgewiesen werden. Insgesamt hat sich gezeigt, dass die Beamformer im tiefen Frequenzbereich bis ca 1000 Hz eine eher schwache Filterwirkung erzielen. Aus diesem Grund wird als Alternative zum gleichverteilten linearen Array eine logarithmische Arraystruktur aus mehreren ULA-Subarrays getestet. Durch die unterschiedlich großen Mikrofonabstände konnte gezeigt werden, dass auf Kosten der Dämpfung und der Beambreite über den gesamten Frequenzbereich ein konstanterer Filterverlauf erzielt werden kann.

Den Abschluss der Leistungsanalysen bilden Messreihen in der realen Umgebung eines normalen Büroraumes. Hier hat sich gezeigt, dass die Dämpfungsleistung der Arrays deutlich schwächer ist als im Schallmessraum. Die besten Ergebnisse in dieser Umgebung erzielten, auf Grund der höheren Mikrofonanzahl pro Array, die linear gleichverteilten Implementierungen. Die Leistung der D&S-Beamformer war in den Tests im Sinne der Sprachverständlichkeit besser als die MVDR-Varianten, da diese ein vergleichsweise schlechtes Dämpfungsverhalten für weißes Rauschen aufweisen.

Zusammenfassend lässt sich sagen, dass wenn vor allem eine gute Dämpfung abseits der Hauptempfangsrichtung gefordert ist, gleichverteilte Arrays den Subarrays vorzuziehend sind. Ist in der Anwendung vor allem eine konstante Leistung über der gesamten Frequenz gefordert, sind die logarithmischen Subarrays die bessere Wahl. Die Wahl zwischen D&S- oder MVDR-Beamformer ist von den a priori bekannten Raumeigenschaften abhängig. Kann das Rauschumfeld gut mathematisch beschrieben bzw. ausgemessen werden, ist es möglich den MVDR-Beamformer optimal anzupassen. Zudem besteht die Möglichkeit, die Verstärkung des weißen Rauschens durch einen Korrekturfaktor zu begrenzen. Ist über den Anwendungsraum nichts bekannt, empfiehlt sich eher ein D&S-Beamformer.

11 Ausblick

Beamforming für Audiosignale ist ein weites Feld und wird erst seit wenigen Jahren praktisch erforscht. Während in der Theorie bereits viele Ideen und Ansätze vorhanden sind, beginnen erst seit einigen Jahren Forscher und Firmen praktischen Nutzen hieraus zu ziehen. Im Laufe dieser Arbeit wurden die gängigsten Beamforming-Ansätze simuliert und in Hardware implementiert. In der Literatur existieren einige alternative Implementierungsansätze, die zusätzliche Erweiterungen ermöglichen. Ein von vielen Autoren genutzter Ansatz ist die Implementierung der Beamformer in der sogenannten „Generalized-Sidelobe-Canceller“-Struktur [z.B. S.34 2]:

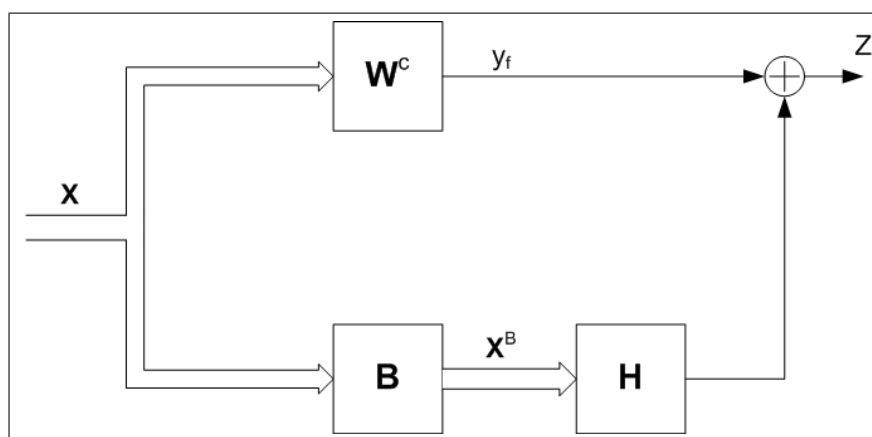


Abbildung 11.1: Beamformer in „Generalized-Sidelobe-Canceller“-Struktur

Die Struktur geht von einem zeitlich angepassten Eingangssignal \mathbf{X} aus und arbeitet mit einem fixen Beamformer \mathbf{W}^C und einer Blocking-Matrix \mathbf{B} , die das Input-Signal in einen Rauschunterraum transformiert¹. Der Vektor \mathbf{H} repräsentiert einen Optimalfilter z.B. einen Wiener Filter. Das Ergebnis ist die Subtraktion des Beamformer-Signals und des durch Blockingmatrix und Optimalfilter berechneten Rauschens.

Dieser Ansatz beinhaltet bereits einen weitem Gedanken, den Post-Filter. Da Sprachsignale für sich genommen charakteristische Eigenschaften aufweisen, wie z.B. eine wiederkehrende Sprachmelodie oder Sprachpausen, bieten sich Optimalfilter zur Nachverarbeitung der Signale an. Ein gutes Eingangsbeispiel hierfür ist die Verwendung von Wiener-Filtern. Ein gängiger Ansatz ist, dass Fehlersignal zwischen Originalsignal und Beamformer-Signal zu minimieren. Ein möglicher Wiener-Filter, der diesen Ansatz verfolgt, ist in [2, S.44] aufgeführt². Er benötigt die Rauschleistung $\phi_{V_0 V_0}$ und die Signalleistung $\phi_{S_0 S_0}$. Die Rauschleistung z.B. kann während der

¹Das bedeutet, das gewünschte Signal wird jeweils räumlich ausgefiltert also „geblockt“. Eine Möglichkeit dies zu erreichen, ist z.B. die sog. Griffith-Jim Blockingmatrix. Sie enthält pro Zeile jeweils 1 und eine -1, sodass jeweils das Wunschsignal geblockt und das Rauschen durchgelassen wird. [2, S.34]

²Die Formel lautet:
$$\mathbf{H} = \frac{\Phi_{S_0 S_0}}{\Phi_{S_0 S_0} + \Phi_{V_0 V_0}}$$

Sprachpausen gut geschätzt werden, während man mit dieser Kenntnis auch eine Abschätzung der Signalleistung treffen kann.

Die Anwendung der Subarrays ist ein weiteres neues Feld, das bisher an der HAW nicht untersucht wurde. Im Bereich der Audiolokalisierung könnte sich diese Form der Arraygeometrie als sinnvoll erweisen. Während sich in dieser Arbeit die Stärken des Subarrays nicht zeigen konnten, hat dieser Ansatz bei einer flexibleren Wahl der Arrayparameter wie Abstand der Mikrofone und Mikrofonanzahl dennoch großes Potential.

Die Stärke der Beamforming Algorithmen kommt erst dann zum Tragen, wenn verlässliche Winkelinformationen zu den Audioquellen vorliegen. Der in dieser Arbeit nicht verifizierte LCMV-Ansatz könnte vor allem bei zusätzlich bekannten Störquellen zu deutlich verbesserten Ergebnissen führen. Da für das Beamforming selbst in erster Linie nur Frequenztransformation und inverse Transformation notwendig ist, lässt es sich einfach mit Verfahren zur Lokalisierung wie MUSIC oder ESPRIT, welche diese Transformation auch benötigen, kombinieren.

Parallel zu dieser Arbeit läuft an der HAW-Hamburg eine Untersuchung zum Thema psychoakustische Analyse von Audiosignalen, die sich mit Entzerrung von Raumübertragungsfunktionen befasst. Diese Thematik bietet ebenfalls viele Kombinationsmöglichkeiten mit Beamforming und Lokalisierungsalgorithmen.

Literaturverzeichnis

- [1] K. Buckley B. Van Veen. "Beamforming: A Versatile Approach to Spatial Filtering". In: *IEEE ASSP Magazine* (1988).
- [2] M. Brandstein and D. Ward. *Microphone Arrays: Signal Processing Techniques and Applications*. Digital Signal Processing - Springer-Verlag. Springer, 2001. ISBN: 9783540419532. URL: <http://books.google.de/books?id=nND60bXSNoEC>.
- [3] DSignt. *D.Module.C6713 Technical Data Sheet*. 2010.
- [4] DSignt. *D.Module.PCM3003 Technical Data Sheet*. 2014.
- [5] T. Haulick G. Schmidt. "Signal processing for in-car communication systems". In: *Signal Processing Journal No. 86* (2006).
- [6] Jürgen Göbel. *Radartechnik - Grundlagen und Anwendungen*. VDE, 2011. ISBN: 9783800-731411.
- [7] Hagen Stübing Adam Opel GmbH. "A Demonstrator for Beamforming in C2X Communication". In: *Vehicular Technology Conference 2010* (2010). URL: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5494215.
- [8] S. Gannot J. Benesty I. Cohen. "New Insights Into the MVDR Beamformer in Room Acoustics". In: *IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, Vol. 18* (2010).
- [9] Dr. Iain McCowan. *Microphone Arrays: A Tutorial*. 2001. URL: <https://www.idiap.ch/~mccowan/arrays/tutorial.pdf>.
- [10] K. Müller. "Sprecherlokalisierung von zwei oder mehr Sprechern im Raum unter Verwendung von Mikrofonarrays". Masterthesis. HAW Hamburg, 2012.
- [11] M. Papageorgiou. *Optimierung: Statische, Dynamische, Stochastische Verfahren für die Anwendung*. Springer Berlin Heidelberg, 2012. ISBN: 9783540340133.
- [12] K. Pikora. "Korrelationsalgorithmen zur Sprecherlokalisierung mittels DSP-basiertem Echtzeitsystem und Mikrofonarray". Masterthesis. HAW Hamburg, 2011.
- [13] J. Reermann. "Echtzeit-Lokalisierung von Audioquellen mittels zirkularen Mikrofonarrays und des UCA-ESPRIT-Algorithmus". Masterthesis. HAW Hamburg, 2013.
- [14] Wainhouse Research. "Navigating the Emerging Gap in Large Conference Calls & Webcast Event Solutions". In: *Navigating the Emerging Gap in Large Conference Calls & Webcast Event Solutions* (2013). URL: <http://cinchcast.com/news/wainhouse-research-report-uncovers-emerging-gap-in-large-conference-calls-webcast-event-market/>.

-
- [15] B. Widrow. *Adaptive Signal Processing*. Pearson Education. ISBN: 9788131705322. URL: https://books.google.de/books?id=S_b4GB01WfwC.
- [16] Wikipedia. *Long Term Evolution, LTE*. 06/2015. URL: http://de.wikipedia.org/wiki/Long_Term_Evolution (visited on 06/15/2015).

Anhang

Der Anhang mit Erläuterung der Ordnerstruktur ist auf dem Hauptverzeichnis der DVD zu finden. Der Anhang enthält eine Beschreibung der mathematischen Methoden, die erstellten Matlab-Skripte und den entworfenen C-Code. Die DVD kann bei Professor Kölzer an der HAW-Hamburg oder bei Professor Schütte an der FH Westküste eingesehen werden.

[A] Lagrange Multiplier

Das Verfahren der Lagrange-Multiplikatoren wird verwendet um Minimierungsprobleme mit Nebenbedingungen zu lösen [siehe 11, S.69].

Für jede Nebenbedingung wird ein sogenannter „Lagrange Multiplikator“ (meist λ) eingeführt. Weiterhin müssen alle Nebenbedingung so umgestellt werden, dass auf einer Seite der Gleichung eine Null steht. Für den MVDR-Beamformer ergibt sich somit zunächst die folgende Ausgangssituation:

$$\min_{\mathbf{w}} \mathbf{w}^H \Phi_{XX} \mathbf{w} \quad \text{unter Bedingung} \quad \mathbf{w}^H \mathbf{d} - 1 = 0$$

Aus dem Minimierungsansatz und der Nebenbedingung wird eine Gesamtgleichung inklusive der Lagrange Multiplikatoren aufgestellt:

$$L : \quad \mathbf{w}^H \Phi_{XX} \mathbf{w} + \lambda(\mathbf{w}^H \mathbf{d} - 1)$$

Im Folgenden müssen alle partiellen Ableitungen der variablen Parameter gebildet werden. An dieser Stelle ist es ausreichend, die Ableitungen z.B. nach \mathbf{w}^H zu bilden, da Ableitungen nach \mathbf{w} dieselbe Aussage beinhalten. Anschließend wird die erhaltene Gleichung gleich Null gesetzt:

$$\frac{dL}{d\mathbf{w}^H} : \quad \Phi_{XX} \mathbf{w} + \lambda \mathbf{d} = 0$$

Hieraus resultiert die Lösung für \mathbf{w} unter Abhängigkeit von den Lagrange Multiplikatoren λ :

$$\mathbf{w} = -\lambda \Phi_{XX}^{-1} \mathbf{d}$$

Der Letzte Schritt ist die Lösung des erhaltenen Gleichungssystems der partiellen Ableitungen. Da in diesem Beispiel nur eine Ableitung existiert, kann λ durch einsetzen in die Nebenbedingung bestimmt werden:

$$\begin{aligned} \mathbf{w}^H \mathbf{d} &= 1 \\ (-\lambda \Phi_{XX}^{-1} \mathbf{d})^H \mathbf{d} &= 1 \end{aligned}$$

Unter Berücksichtigung der Rechenregel:

$$(A \cdot B)^H = B^H \cdot A^H$$

erhält man:

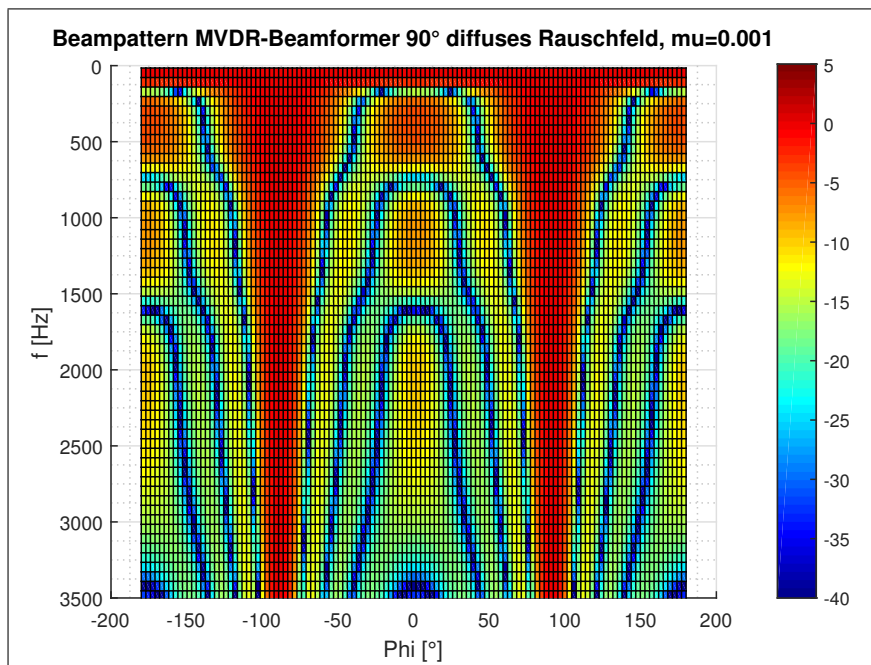
$$-\lambda(\mathbf{d}^H \Phi_{XX}^{-1} \mathbf{d}) = 1$$

$$-\lambda = \frac{1}{\mathbf{d}^H \Phi_{XX}^{-1} \mathbf{d}}$$

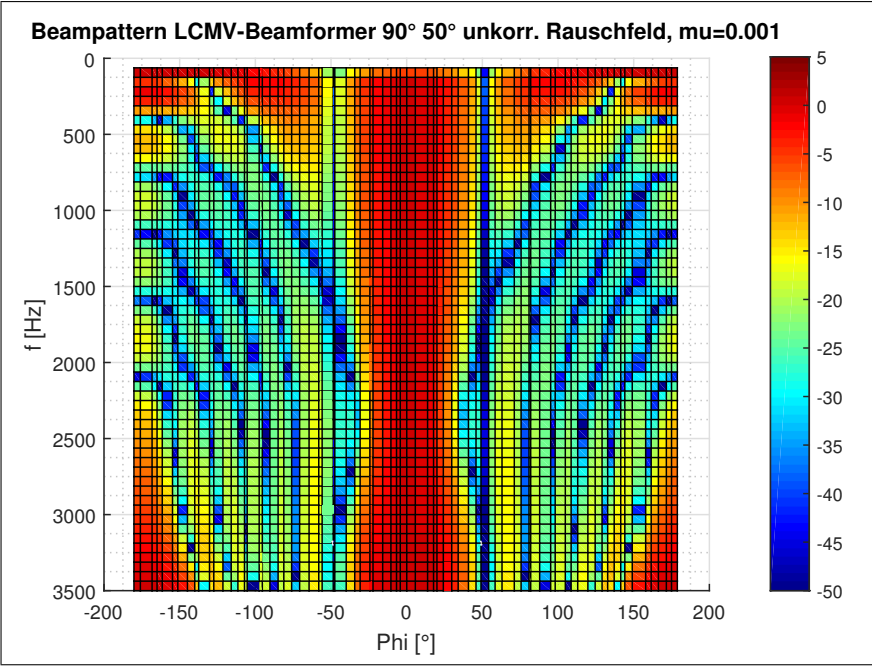
Somit ergibt sich die Lösung zu:

$$\mathbf{W}_{MVDR} = \frac{\Phi_{XX}^{-1} \mathbf{d}}{\mathbf{d}^H \Phi_{XX}^{-1} \mathbf{d}}$$

[B] Zusätzliche Abbildungen



Beampattern, 90° MVDR-Beamformers für diffuses Rauschfeld, mu=0.001



Beampattern, 0°/50° LCMV-Beamformers, mit Winkelabweichung

[C] Audiofiles

Die Hörproben der Messergebnisse sind im Ordner *Audiofiles* zu finden. Das Präfix *ULA* kennzeichnet den linearen Array-Aufbau, *SA* steht für den Subarray-Aufbau. Der mittlere Teil des Namens kennzeichnet *Delay-and-Sum*, *DaS*, oder *MVDR* für diffuses Rauschfeld, *MVDR*. Die letzten Ziffern geben den Empfangswinkel an.

[D] Schaltpläne

Die Schaltpläne sind im *Datenblaetter*-Verzeichnis zu finden.

[E] Datenblätter

Die Datenblätter zu DSP und Mikrofonarrays sind im Verzeichnis *Datenblaetter* in entsprechenden Unterordnern hinterlegt.

[F] Matlab-Code

Die Matlab-Codes befinden sich im Verzeichnis *Matlab_Code*. In den Ordnern *UCA* und *ULA* befinden sich die Skripte zur Generierung verschiedener Beampattern für definierte Parameter. Die Skripte benötigen jeweils die Funktionen, welche im Ordner *User_function* hinterlegt sind. Entsprechend muss dieser vorher in Matlab per Rechtsklick „Add to Path → Selected Folders and Subfolders“ hinzugefügt werden.

Die Generierung der Koeffizienten-Header für die Hardwareimplementierung in C wird über die Skripte im *Generate_C_header*-Ordner durchgeführt. Auch hier müssen die *User_functions* im Pfad vorhanden sein.

[G] C-Code

Die C-Codes für die Hardwareimplementierung befinden sich im *C_Code*-Ordner. Es sind für Subarray und ULA zwei separate Versionen vorhanden. Neben den im *custom_include*-Unterordner vorhandenen Koeffizientenheadern können weitere Kombinationen über die bereitgestellten Matlab Skripte erzeugt werden.

Versicherung über die Selbstständigkeit

Hiermit versichere ich, dass ich die vorliegende Arbeit im Sinne der Prüfungsordnung nach §16(5) APSO-TI-BM ohne fremde Hilfe selbstständig verfasst und nur die angegebenen Hilfsmittel benutzt habe. Wörtlich oder dem Sinn nach aus anderen Werken entnommene Stellen habe ich unter Angabe der Quellen kenntlich gemacht.

Hamburg, den 13.08.2015

Ort, Datum

Steffen Johannes Zahn

Die im Rahmen dieser Arbeit entwickelte Software kann bei Prof. Dr. Kölzer an der HAW Hamburg eingesehen werden.