

Messung von Außenohrübertragungsfunktionen
und direktionalen Raumimpulsantworten als Basis
der Auralisation realer Räume

Diplomarbeit

von

Stefan Warum

durchgeführt am

Institut für elektronische Musik und Akustik
der Universität für Musik und darstellende Kunst in Graz

Betreuer:

o.Univ.-Prof. Mag. DI Dr. Robert Höldrich

Graz, im Oktober 2003

Kurzfassung

Diese Arbeit beschreibt Verfahren zur Messung von Außenohrübertragungsfunktionen (HRTFs) und direktionalen Raumimpulsantworten. Sie bilden die Grundlage für eine dreidimensionale akustische Abbildung eines realen Raumes bei Wiedergabe über einen Kopfhörer. Nach einem Vergleich der verschiedenen zur Verfügung stehenden Impulsantwortmesstechniken wurde die logarithmische Sinus-Sweep Methode in der echtzeitfähigen, grafischen Audioprogrammiersprache Pure Data (PD) implementiert. Die Außenohrübertragungsfunktionen entstammen dem Kunstkopf "Source", der mit Hilfe eines automatisierten Drehtellers innerhalb der gesamten oberen Hemisphäre vermessen wurde. Für die Messung der direktionalen Raumimpulsantworten kann ein Beamformer in Verbindung mit einem dreidimensionalen Mikrofonarray verwendet werden. Der Delay&Sum Beamformer und der Minimum Varianz Beamformer wurden in Matlab simuliert und in Bezug auf Richtverhalten und Stabilität gegenübergestellt. Durch die Anwendung räumlicher Fensterfunktionen lässt sich eine frequenzunabhängige Richtcharakteristik erreichen.

Abstract

This thesis presents methods for measuring head-related transfer functions (HRTFs) and directional room impulse responses. They are the basis for a three-dimensional acoustic auralization of a real room in the case of sound reproduction by headphones. After comparing several different impulse response measurement techniques a method using a logarithmically swept sine as excitation signal was implemented in Pure Data (PD), a real-time graphical programming environment for computer music. The head-related transfer functions were captured from the dummy head "Source". By using an automated turntable a full data set was acquired within the upper hemisphere. For the measurement of directional room impulse responses a beamformer can be used in conjunction with a three-dimensional microphone array. The Delay&Sum Beamformer and the Minimum Variance Beamformer were simulated in Matlab and compared to each other in reference to directivity and stability. By applying spatial windowing a frequency independent behavior of the directivity is achieved.

Inhaltsverzeichnis

1	EINLEITUNG.....	1
1.1	Kapitelübersicht.....	2
2	VERGLEICH VERSCHIEDENER IMPULSANWORTMESSTECHNIKEN	3
2.1	Allgemeines.....	3
2.2	Erregersignale.....	4
2.2.1	Impulserregung.....	4
2.2.2	Rauschen.....	4
2.2.3	Pseudorauschen.....	4
2.2.4	Zeitgedehnter Impuls.....	4
2.3	Time Delay Spectrometry (TDS).....	5
2.4	Maximum Length Sequence (MLS).....	7
2.5	Sweeptechnik.....	9
2.5.1	Systemtheoretische Grundlagen.....	10
2.5.2	Generierung des Sweeps.....	13
2.5.3	Messumgebung.....	14
2.5.3.1	Hardware.....	14
2.5.3.2	Software:.....	15
3	VERMESSUNG UND OPTIMIERUNG DES KUNSTKOPFES "SOURCE"	20
3.1	Was ist ein Kunstkopf?.....	20
3.2	Räumliches Hören.....	20
3.2.1	Richtungswahrnehmung.....	20
3.2.1.1	Interaurale Zeitdifferenzen (ITD).....	21
3.2.1.2	Interaurale Pegeldifferenzen (ILD).....	22

3.2.1.3	Spektrale Cues	23
3.2.2	Distanzwahrnehmung	23
3.2.3	Außenohrübertragungsfunktionen (HRTF-Kurven):	24
3.3	Kunstkopf "Source"	26
3.3.1	Messbedingungen	26
3.3.1.1	Systemtheoretische Beschreibung	26
3.3.1.2	Simulierte Freifeldmessung	28
3.3.1.3	Lautsprecherauswahl.....	29
3.3.1.4	Lautsprecherentzerrung.....	30
3.3.1.5	Cepstralanalyse	31
3.3.2	Vermessung des "Source"	35
3.3.2.1	Freifeldübertragungsfunktion in Vorne-Richtung	36
3.3.2.2	Diffusfeldfrequenzgang	37
3.3.3	Optimierung für 0°-Richtung.....	39
3.3.3.1	Korrektur der Überhöhung um 8kHz.....	39
3.3.3.2	Korrektur des Anstiegs zwischen 300Hz und 1kHz	41
4	BEAMFORMING	44
4.1	Allgemeines	44
4.2	Grundlagen.....	44
4.3	Mikrofonarray.....	47
4.3.1	3D-Array	49
4.3.2	Zirkuläres 2D-Array.....	50
4.4	Delay&Sum Beamformer für 3D-Array.....	50
4.4.1	Theoretische Betrachtungen.....	50
4.4.2	Simulationsergebnisse.....	51
4.5	Delay&Sum Beamformer für 3D-Array mit räumlicher Fensterung der Arraygewichte	55
4.5.1	Theoretische Betrachtungen.....	55
4.5.2	Simulationsergebnisse.....	55
4.6	Minimum Varianz Beamformer	58
4.6.1	Theoretische Betrachtungen.....	58
4.6.2	Simulationsergebnisse.....	60
4.7	Delay&Sum Beamformer für zirkuläres Array mit räumlicher Fensterung der Arraygewichte	64
4.7.1	Theoretische Betrachtungen.....	64
4.7.2	Simulationsergebnisse.....	64

4.8	Delay&Sum Beamformer mit räumlicher Tiefpassfilterung der Richtcharakteristik	66
4.8.1	Theoretische Betrachtungen.....	66
4.8.2	Simulationsergebnisse.....	67
5	AUSBLICK.....	71
	LITERATURVERZEICHNIS	72
	ANHANG A	75
	ANHANG B	78
	ANHANG C	82

1 Einleitung

Die Aufführung von Sprache oder Musik ist unmittelbar mit den akustischen Eigenschaften des Raumes verbunden in dem sie dargeboten wird. Viele Verfahren in der virtuellen Akustik basieren auf der Simulation eines imaginären Raumes dessen Verhalten rein rechnerisch ermittelt wird. Um eine virtuelle Quelle in einem realen Raum abzubilden, müssen Methoden gefunden werden, um einerseits diesen Raum akustisch dreidimensional zu beschreiben und andererseits diese Information beim Hörer auch wieder spatialisiert zu reproduzieren.

Ausgangspunkt dieser Arbeit ist die spätere Auralisation des IEM-Cube (= großer Aufnahme/Abhörraum am Institut für elektronische Musik) mittels Ambisonic bei Wiedergabe über Kopfhörer. Die akustische Beschreibung des IEM-Cube erfolgt durch direktionale Raumimpulsantworten. Diese können durch ein Mikrofonarray mit einem nachfolgenden Beamformer gemessen werden. Dabei wird einer gewählten Schallquellenposition im Raum ein Set von direktionalen Raumimpulsantworten zugeordnet.

Die menschliche Richtungswahrnehmung basiert auf der Auswertung von linkem und rechtem Ohrsignal. Diese erfahren durch die Form von Oberkörper, Kopf und Ohrmuschel eine Reihe von linearen Verzerrungen, die in ihrer Gesamtheit durch die Außenohrübertragungsfunktion (HRTF) beschrieben werden. Jeder Mensch besitzt zwar eine individuelle HRTF, ist aber weitgehend in der Lage mit der generalisierten Außenohrübertragungsfunktion eines Kunstkopfes räumlich zu hören. Aufgrund seiner guten klanglichen Eigenschaften wird der Kunstkopf "Source" für die Messung der kopfbezogenen Impulsantworten verwendet.

1.1 Kapitelübersicht

Kapitel 2 bietet einen ausführlichen Überblick über die derzeit zur Verfügung stehenden Methoden zur Impulsantwortmessung. Es werden die Time Delay Spectrometry (TDS) Methode die Maximum Length Sequence (MLS) Technik und die logarithmische Sweeptechnik gegenübergestellt. Für die im Zuge dieser Arbeit verwendete Sweep-Methode wird eine geeignete Messumgebung entwickelt. Die Messung selbst sowie die Ansteuerung eines motorbetriebenen Drehtellers wird dabei in PD (Pure Data) realisiert, die Messauswertung erfolgt mit Matlab.

Kapitel 3 befasst sich mit der Vermessung und Optimierung des Kunstkopfes "Source". Einleitend werden die Grundlagen des räumlichen Hörens erörtert. Für die Vermessung des Source kommt eine simulierte Freifeldmessung zum Einsatz. Dabei wird die Messung in einem nicht reflexionsfreien Raum durchgeführt und der Freifeldanteil, also jener vor Eintreffen der ersten Reflexion ausgewertet. Diese Vermessung liefert als Ergebnis 325 binaurale Freifeldimpulsantworten für Schalleinfallrichtungen innerhalb des oberen Halbraumes. Die Übertragungsfunktion des Kunstkopfes für die Blickrichtung wird durch Applikation von Absorbermaterial für den hochfrequenten Bereich und durch ein Shelvingfilter für tiefe Frequenzen korrigiert.

Kapitel 4 behandelt den Entwurf eines Beamformers für die akustische Abtastung eines realen Raumes. Hier wird das Verhalten des Delay&Sum Beamformers und des Minimum Varianz Beamformers in Verbindung mit einem dreidimensionalen Mikrofonarray sowie einem zirkulären Array diskutiert. Grafische Darstellungen der Beamformer-Richtcharakteristika ermöglichen einen anschaulichen Vergleich der verwendeten Methoden.

2 Vergleich verschiedener Impulsantwortmesstechniken

2.1 Allgemeines

Eine der wichtigsten Eigenschaften eines Übertragungssystems ist seine Impulsantwort. Diese kann mit unterschiedlichen Verfahren gemessen werden. Ganz allgemein gelten folgende Anforderungen an die Messmethode:

- Das Anregungssignal muss absolut reproduzierbar sein.
- Der Signal-Rauschabstand der Impulsantwort muss groß genug sein um eine störungsfreie Weiterverarbeitung zu gewährleisten (> 80 dB).
- Nichtlinearitäten sollen durch die Messmethode möglichst gut unterdrückt werden.

Die verschiedenen Messverfahren unterscheiden sich durch ihre Erregersignale und den damit verbundenen Nachverarbeitungsschritten. Je nach angewandter Methode ergeben sich große Unterschiede in Bezug auf den Energiegehalt des Erregersignals und dem damit einhergehenden Signal-Stör-Abstand.

Die zur Messung von Raumimpulsantworten am besten geeigneten Methoden sollen in diesem Kapitel gegenübergestellt werden.

Methode	Anregungssignal	Nachverarbeitung
TDS ¹	Linearer Sweep	Multiplikation + Tiefpassfilterung
MLS ²	Pseudostochastisches weißes Rauschen	Kreuzkorrelation
Sweep-Technik	Logarithmischer Sinus Sweep	Entfaltung

¹ Time Delay Spectrometry, siehe Abschnitt 2.3

² Maximal Length Sequences, siehe Abschnitt 2.4

2.2 Erregersignale

Voraussetzung für ein geeignetes Erregersignal ist ein bekanntes Amplituden- und Phasenspektrum über dem gewünschten Frequenzbereich.

2.2.1 Impulserregung

Impulse sind breitbandige Signale mit hohem Crestfaktor, die bei Abstrahlung über elektroakustische Anlagen nur mit geringen Energiegehalten verzerrungsfrei wiedergegeben werden können. In der statistischen Raumakustik erfolgt eine Impulsanregung oft mittels Pistolenknall oder Funkenstrecke, die zwar einen hohen Energiegehalt besitzen, jedoch für die anschließende Signalverarbeitung nicht ausreichend reproduzierbar sind.

2.2.2 Rauschen

Rauschen kann als Impulsfolge gedeutet werden, bei der alle Frequenzen statistisch gleichmäßig verteilt auftreten. Mit Rauschen als Anregungssignal sind zwar die Parameter der statistischen Raumakustik bestimmbar, jedoch ist mit diesem stochastischen Signal die Messung von Impulsantworten auf Grund der zufälligen Amplitude und Phasenlage nicht möglich.

2.2.3 Pseudorauschen

Ziel bei der Generierung von Pseudorauschen ist die Berechnung eines Zeitsignals mit definiertem Amplitudenspektrum. Dazu werden die spektralen Anteile derart phasenverschoben, dass das daraus entstehende Zeitsignal einen stark optimierten Leistungsgehalt besitzt. Durch die bekannte Phasenlage ist dieses Erregersignal auch reproduzierbar, und somit für Impulsantwortmessungen geeignet.

2.2.4 Zeitgedehnter Impuls

Sinn und Zweck der zeitlichen Aufdehnung eines Impulses ist es, die gewünschte sehr hohe Impulsenergie über ein längeres Zeitfenster zu verteilen. Der Impuls kann als ein diskretes Spektrum genügend dicht aneinanderliegender Spektrallinien gleicher Amplitude betrachtet werden. Mit einer frequenzabhängigen Phasenverschiebung erreicht man eine Impuls-Dehnung die sich als individuelle zeitliche Verschiebung der einzelnen Spektrallinien äußert. Dieses Signal wird nun zur Anregung verwendet. Die Impulsantwort erhält man durch eine anschließende Stauchung des gemessenen Signals, die die vorangegangene Dehnung rückgängig macht.

2.3 Time Delay Spectrometry (TDS)

Dieses Verfahren erlaubt es Freifeldmessungen im schallarmen Raum zu simulieren, indem die auftretenden Reflexionen und Störgeräusche rechnerisch weitgehend ausgeschieden werden [Müller, 2001], [Stan, 2002]. Das TDS-Verfahren benutzt als Prüfsignal einen zeitlich linearen Sinussweep (d.h. $df/dt = \text{konstant}$). Dieser besitzt eine konstante Amplitude und erzeugt in Summe ein weißes Spektrum. Im Wesentlichen besteht ein TDS-System aus einem Sweep-Generator, der ein Sinus- sowie ein Kosinus-Signal steigender Frequenz erzeugt. Diese sind phasenstarr miteinander verkoppelt. Eines dieser beiden Messsignale wird dem Lautsprecher zugeführt, dessen Antwort darauf jeweils mit dem Originalsignal (z.B. Sinus) und dem um 90° phasengedrehten Anregungssignal (Kosinus) zeitverzögert multipliziert wird. Diese Zeitverzögerung muss der akustischen Laufzeit zwischen Lautsprecher und Mikrofon entsprechen um eine Lautsprecherimpulsantwort unter Freifeldbedingungen zu simulieren. Zur Übertragungsfunktion einer Reflexionsfläche gelangt man, indem als Zeitverzögerung die Laufzeit zwischen Lautsprecher, Reflexionsfläche und Mikrofon gewählt wird. Das Ausgangssignal der Multiplikatoren wird mit einem Tiefpass fixer Grenzfrequenz gefiltert und stellt dann bereits den Realteil (aus der Multiplikation mit dem Sinus) und Imaginärteil (aus der Multiplikation mit dem Kosinus) der gesuchten Übertragungsfunktion dar.

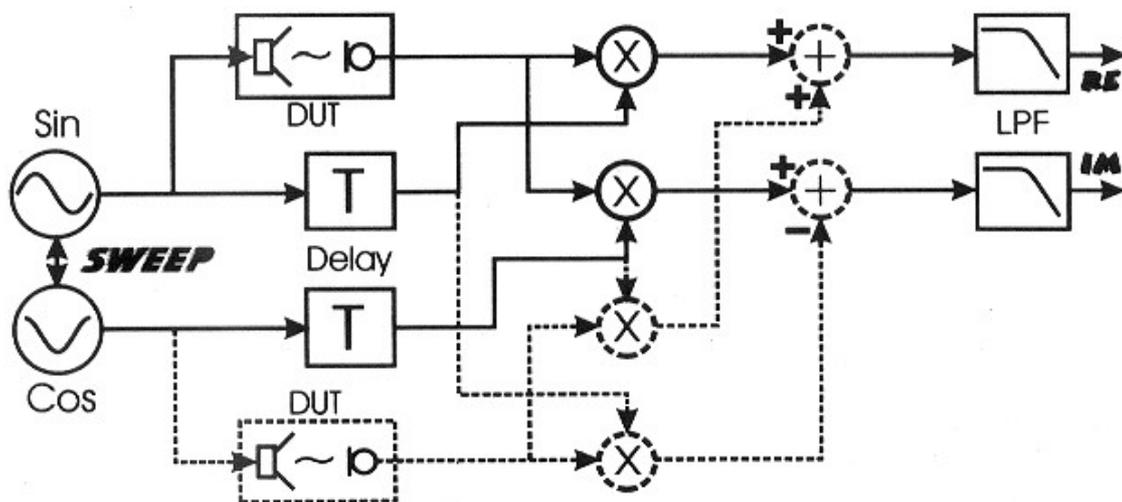


Abb. 5.1: Prinzipielle Funktionsweise eines TDS - Messsystems [Müller, 2001]

Diese Multiplikationsstufen erzeugen Summations- und Differenzfrequenzen des Eingangssignals. Die Summationsterme beider Multiplikatorausgänge müssen durch die Tiefpassfilter weggefiltert werden wohingegen die Differenzfrequenzen durchgelassen werden. Wenn die Frequenzen des vom Generator erzeugten Signals und des empfangenen

Signals nahezu übereinstimmen, dann ist die Differenzfrequenz sehr klein und wird daher vom nachgeschalteten Tiefpassfilter durchgelassen. Aus diesem Grund muss das Generatorsignal um die akustische Laufzeit von Lautsprecher zum Mikrofon verzögert der Multiplikationsstufe zugeführt werden. Praktischer ergibt sich dadurch bei Lautsprecherimpulsantworten eine gute Unterdrückung von unerwünschten Raumreflexionen, da diese ja verspätet eintreffen und somit eine niedrigere Momentanfrequenz besitzen. Sie bilden also am Multipliziererausgang eine Summationsfrequenz, die vom nachfolgenden Tiefpassfilter weggefiltert wird. Das Ergebnis der TDS-Methode ist eine Impulsantwort aus der man die Übertragungsfunktion nach Betrag und Phase erhält.

Nachteile der TDS:

- Als Anregungssignal dient ein linearer Sweep der in den meisten Messungen einen zu geringen Rauschabstand bei tiefen Frequenzen liefert. Um in diesem Frequenzbereich die Einwirkdauer des Sweeps, und somit die spektrale Leistungsdichte zu vergrößern, muss der Sweep sehr lang gemacht, oder in zwei Frequenzbereiche (z.B. < 500 Hz und > 500 Hz) aufgesplittet werden.
- An der unteren Bandgrenze, wo die Summenfrequenzen der Multiplikatoren im Bereich der Tiefpassgrenzfrequenz liegen, kommt es zu einem oszillierenden Verlauf des gemessenen Spektrums. Dagegen bietet sich ein Herabsetzen der Tiefpassgrenzfrequenz an. Zur Beibehaltung der gleichen Frequenzauflösung muss dann aber die Dauer des Sweeps proportional erhöht werden. Mit dem in Abb. 5.1 durch die strichlierten Linien dargestellten zweiten Messdurchlauf können diese Ripple-Effekte im tieffrequenten Bereich beseitigt werden.

Bei diesem gespiegelten Messaufbau wird der Lautsprecher mit einem Kosinus- anstatt mit einem Sinussignal angeregt. Der Effekt nach den Auswertoperationen ist, dass sich die Summationsterme gegenseitig auslöschen. Aus diesem Grund kann der nachgeschaltete Tiefpassfilter weggelassen werden. Um eine gesamte Raumimpulsantwort mit all ihren frühen und diffusen Reflexionen zu erhalten, muss dieser Tiefpassfilter weggelassen werden. Das bedeutete, dass man mit der TDS-Methode nur dann eine lange Raumimpulsantwort messen kann, wenn man des symmetrischen Messaufbau ohne nachgeschalteten Tiefpass verwendet.

Trotz der Beseitigung der Summationsterme mit dem gespiegelten Messdurchlauf, kann es noch immer zu Ripple-Effekten am Beginn und am Ende des Sweeps kommen. Dies sind transiente Ein- und Ausschwingvorgänge die durch das plötzlich Ein- und Ausschalten des

Sweeps entstehen. Dieses Problem kann man umgehen, indem man die Startfrequenz des Sweeps deutlich unter die interessierende Frequenz legt. Das kann einen Start des Sweep bei negativen Frequenzen erfordern, was in der Praxis ein Starten des Sweeps bei positiven Frequenzen in Richtung 0 Hz bedeutet. Bei 0 Hz angelangt steigt die Sweepfrequenz wieder an.

Vorteile der TDS:

- Durch den nachgeschalteten Tiefpassfilter werden auch die harmonischen Verzerrungen des Lautsprechers, die eine höhere Momentanfrequenz besitzen sehr gut unterdrückt.
- Die TDS-Methode mit doppelter Anregung besitzt einen großen Rauschabstand bei störungsfreier Umgebung.
- Der Crestfaktor des Sinus-Sweep beträgt nur 3 dB.
- Große Toleranz gegen Zeitinvarianz.

2.4 Maximum Length Sequence (MLS)

Bei der MLS-Technik wird der Raum mit einem pseudostochastischen, periodischen Rauschsignal angeregt, die Systemantwort gemessen und dann mit dem bekannten Erregersignal kreuzkorreliert. Als Ergebnis erhält man die periodische Impulsantwort des Raumes [Stan, 2002],[Rife, 1989].

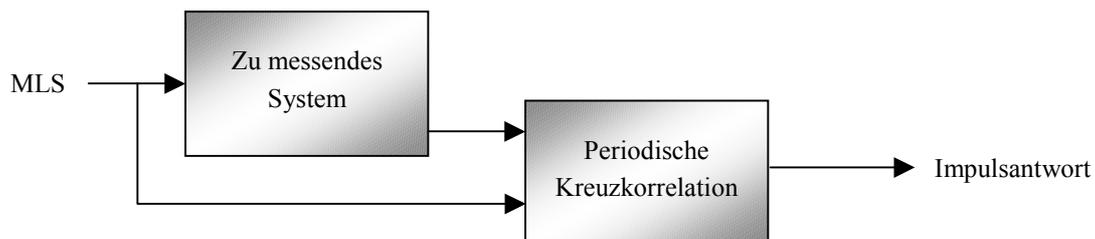


Abb. 5.2: Funktionsweise der MLS - Messung

Vorteil der MLS-Technik ist, dass nicht das Erregersignal einem Impuls gleichen muss, sondern die Autokorrelationsfunktion des Erregersignals, was auf ein weißes Spektrum hindeutet. Maximalfolgen sind periodische binäre pseudostochastische Rauschsignale, die weißem Rauschen sehr ähnlich sind, ohne dabei zufällig zu sein. Die Anzahl der Samples L einer MLS-Periode m -ter Ordnung berechnet sich aus $L = 2^m - 1$.

Die tatsächliche Periodendauer T des MLS-Signals ergibt sich aus der Beziehung

$$T = L \cdot \Delta t = \frac{L}{f_c} . \quad (5.1)$$

mit Δt ... Dauer eines logischen Zustandes

f_c ... Taktfrequenz des MLS-Signals

Aus dem Kehrwert der Periodendauer T berechnet sich die Frequenzauflösung der MLS-Messung.

Ursprünglich wurden die Werte der Maximalfolgen aus einem deterministischen Prozess mit Hilfe eines rückgekoppelten Schieberegisters gewonnen. Die Kaskadierungen der einzelnen Register wird mathematisch durch ein spezielles Polynom der Ordnung m (Anzahl der Register) bestimmt. Die aus dem Schieberegister erhaltene logische 0 und 1 Folge wird den analogen Werten $+1$ und -1 zugeordnet um eine Sequenz zu erzeugen, die symmetrisch zur Nullachse ist und somit keinen Gleichanteil besitzt. Heutzutage wird die Sequenz direkt mit einem Softwarealgorithmus generiert.

Bei der Messung wird der Pegel der MLS-Folge zwischen den einzelnen Impulsen konstant gehalten, da Einzelimpulszüge zu wenig Energie in den Raum abgeben würden. Diese Haltefunktion erster Ordnung führt im Zeitbereich zu einer Faltung des Impulszuges mit einer Rechteckfunktion, was im Frequenzbereich der Multiplikation des Spektrums mit einer Sincfunktion entspricht. Daraus resultiert ein Pegelabfall von ca. 4dB bei der halben Abtastrate, der kompensiert werden muss.

Da das Spektrum des anregenden Signals nahezu weißem Rauschen entspricht, ergibt sich aus dessen periodischer Kreuzkorrelation mit der Systemantwort die gesuchte Raumimpulsantwort. Die Korrelation kann direkt im Zeitbereich mit der schnellen Hadamard-Transformation (FHT) oder im Frequenzbereich durch konjugiert-komplexe Multiplikation durchgeführt werden.

Vorteile der MLS-Methode:

- Unempfindlichkeit gegen Störgeräusche → Aufgrund des stark unregelmäßigen Phasenverlaufs der Maximalfolge, wird jede Störkomponente in der Systemantwort, die ja nicht mit dem Anregungssignal korreliert, über die gesamte Länge der Impulsantwort gleichmäßig verteilt. Durch Mittelung mehrerer Messungen reduziert sich der Einfluss der Störgeräusche um 3dB je Verdoppelung der Messanzahl.
- Kurze Messdauer trotz hoher Frequenzauflösung.
- Hoher Energiegehalt des Messsignals, das den maximal erreichbaren Crestfaktor von 1 besitzt. Die digitalen Antialiasingfilter von Oversampling-DA-Konvertern verursachen allerdings Überschwinger an den steilen Flanken der MLS. Um Übersteuerungen des Filters zu vermeiden, ist es daher notwendig, die MLS mit 5-8dB unter Vollaussteuerung in den Konverter zu speisen.

Nachteile der MLS-Methode:

- Nichtlinearitäten und Zeitinvarianzen des Messsystems, speziell des Lautsprechers, verursachen Verzerrungsspitzen (Phantompeaks) die in der resultierenden Impulsantwort annähernd gleichmäßig verteilt sind.

2.5 Sweeptechnik

Als Messsignal wird in der Sweeptechnik ein logarithmischer Sinussweep verwendet. Aus der Dekonvolution (Entfaltung) der gemessenen Systemantwort mit dem Sweep erhält man die Raumimpulsantwort [Müller,2001], [Farina,2000].

Die MLS- und die TDS-Methode gehen von der Annahme aus, dass das zu untersuchende System perfekte Linearität und Zeitinvarianz besitzt. Da diese Bedingungen in der Praxis nicht absolut erfüllbar sind, treten bei beiden Verfahren Messfehler auf.

Verwendet man einen logarithmischen Sweep als Anregungssignal, so ist es möglich die lineare Impulsantwort und die Impulsantworten jeder einzelnen harmonischen Verzerrung zu bestimmen. Diese Messmethode ist robust gegen geringe Zeitvarianzen des Systems [Müller, 2001].

2.5.1 Systemtheoretische Grundlagen

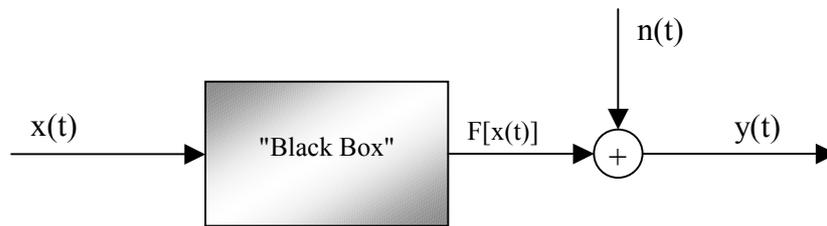


Abb. 5.3: allgemeiner Eingang/Ausgang

Betrachtet man die gesamte Messstrecke als "Blackbox", so ergibt sich das am Mikrofon empfangene Signal $y(t)$ aus der Summe der Systemantwort auf das Eingangssignal $x(t)$ und den überlagerten Störungen $n(t)$.

$$y(t) = n(t) + F[x(t)] \quad (5.2)$$

Für den Fall eines linearen und zeitinvarianten Übertragungssystems (LTI-System) entspricht die Funktion $F[x(t)]$ der Faltung des Eingangssignals $x(t)$ mit der Impulsantwort $h(t)$ des Systems.

$$y(t) = n(t) + x(t) \otimes h(t) \quad (5.3)$$

Bei raumakustischen Messungen kann man jedoch nicht von einem linearen und zeitinvarianten Übertragungssystem ausgehen. Hauptverantwortlich für die Nichtlinearitäten der Messstrecke sind die harmonischen Verzerrungen des Lautsprechers. Die Ursache für die Zeitinvarianz des Messsystems liegt einerseits in der Erwärmung der Lautsprecherspule mit zunehmender Messdauer und andererseits in den Schwankungen der Raumtemperatur und den damit verbundenen Veränderungen der Schallausbreitungsgeschwindigkeit.

Der Störanteil $n(t)$ wird durch die synchrone Mittelwertbildung mehrerer gemessener Ausgangssignale $y(t)$ reduziert. Das daraus erhaltene Signal $\hat{y}(t)$ besitzt ein um 3dB je Messanzahlverdoppelung verbessertes Signal-Rauschverhältnis.

Um $h(t)$ zu bestimmen, ist die "Entfaltung" des Ausdrucks $x(t) \otimes h(t)$ notwendig. Diese Operation ist auf zwei Varianten lösbar.

1.) Lösung im Zeitbereich:

Die Dekonvolution ist im Zeitbereich nur dann möglich, wenn ein geeignetes inverses Filter $f(t)$ gefunden werden kann, das als Faltungsergebnis mit dem Eingangssignal $x(t)$ den Diracimpuls $\delta(t)$ ergibt.

$$x(t) \otimes f(t) = \delta(t) \quad (5.4)$$

Die Impulsantwort des Systems berechnet sich aus der Faltung von der gemessenen Systemantwort $y(t)$ mit dem inversen Filter $f(t)$.

$$h(t) = y(t) \otimes f(t) \quad (5.5)$$

2.) Lösung im Frequenzbereich:

Die Entfaltung erfolgt durch die Division der Spektren $X(f)$ und $\hat{Y}(f)$. Da das einer zirkulären Dekonvolution entspricht, besteht die Gefahr von Timealiasing. Hat $x(t)$ die Länge N und $\hat{y}(t)$ die Länge M , so besitzt das Ergebnis der linearen Faltung die Länge $N+M-1$. Bevor nun die Fouriertransformation von $x(t)$ und $\hat{y}(t)$ geschieht, müssen daher beide Signale auf die Länge $N+M-1$ mit Nullen aufgefüllt werden, um Timealiasing zu vermeiden. Die Dekonvolution erfolgt nun durch Division im Frequenzbereich. Die gesuchte Impulsantwort $h(t)$ ergibt sich aus der anschließenden Rücktransformation.

$$h(t) = IFFT \left[\frac{FFT(\hat{y}(t))}{FFT(x(t))} \right] \quad (5.6)$$

Beide Lösungsmethoden verursachen eine Drehung des Sonogramms³ gegen den Uhrzeigersinn, sodass die lineare Impulsantwort genau als vertikale Linie dargestellt wird. Die Verzerrungsprodukte werden dabei nach "negativen" Zeiten verschoben und erscheinen im Fall des log. Sinus-Sweeps ebenfalls als vertikale Linien im Sonogramm.

In der Systemantwort $\hat{y}(t)$ treten bei jeder Momentanfrequenz des Anregungssignals die zugehörigen harmonischen Verzerrungen zur gleichen Zeit auf, sie haben daher die gleiche Gruppenlaufzeit. Jede harmonische Verzerrung kann als eigener log. Sweep aufgefasst werden, der zu einem früheren Zeitpunkt als das Anregungssignal gestartet wurde. Bei der Division von $\hat{Y}(f)$ durch $X(f)$ subtrahieren sich die zugehörigen Gruppenlaufzeiten, sodass alle Frequenzanteile des Anregungssignals im Zeitpunkt $t = 0$ zu einem Impuls (konstante Gruppenlaufzeit) zusammengeführt werden. Betrachtet man gleiche Frequenzen in den harmonischen Verzerrungen, so ist ersichtlich, dass diese zu früheren Zeitpunkten auftreten als die jeweilige Frequenz im Anregungssignal. Sie besitzen daher kleinere Gruppenlaufzeiten, was sich bei der Division im Frequenzbereich durch eine Überführung der harmonischen Verzerrungen in separate Impulse bei negativen Zeiten äußert (siehe Abb. 5.4 unten rechts). Somit können bei der log. Sweep-Methode die Impulsantworten für jede harmonische Verzerrung mit der zugehörigen Raumimpulsantwort bestimmt werden.

³ Das Sonogramm ist eine Darstellungsform, welche die Frequenz mit der Zeit und dem Schalldruck (bzw. Betrag der Momentan-Amplitude) verknüpft darstellt.

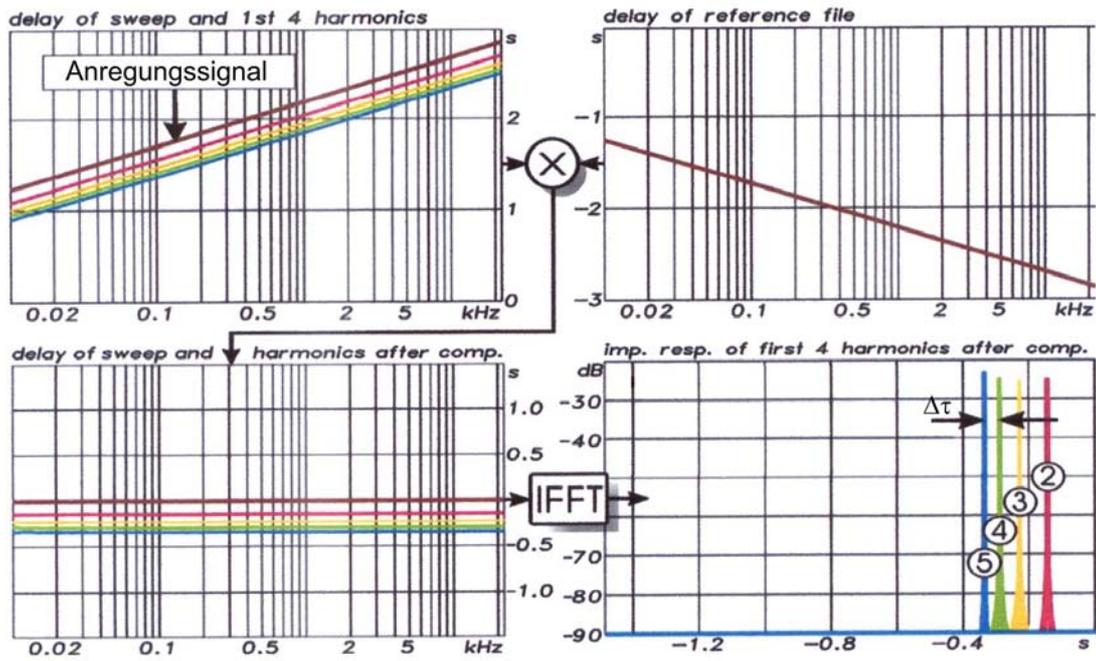


Abb. 5.4: Gruppenlaufzeit des Anregungssignals und der ersten vier harmonischen Verzerrungen (links oben), Multiplikation mit Spektrum des inversen Anregungssignals (rechts oben), entfalteter Sweep mit den Harmonischen (links unten) und Position der Impulsantworten der Harmonischen (rechts unten) [Müller, 2001]

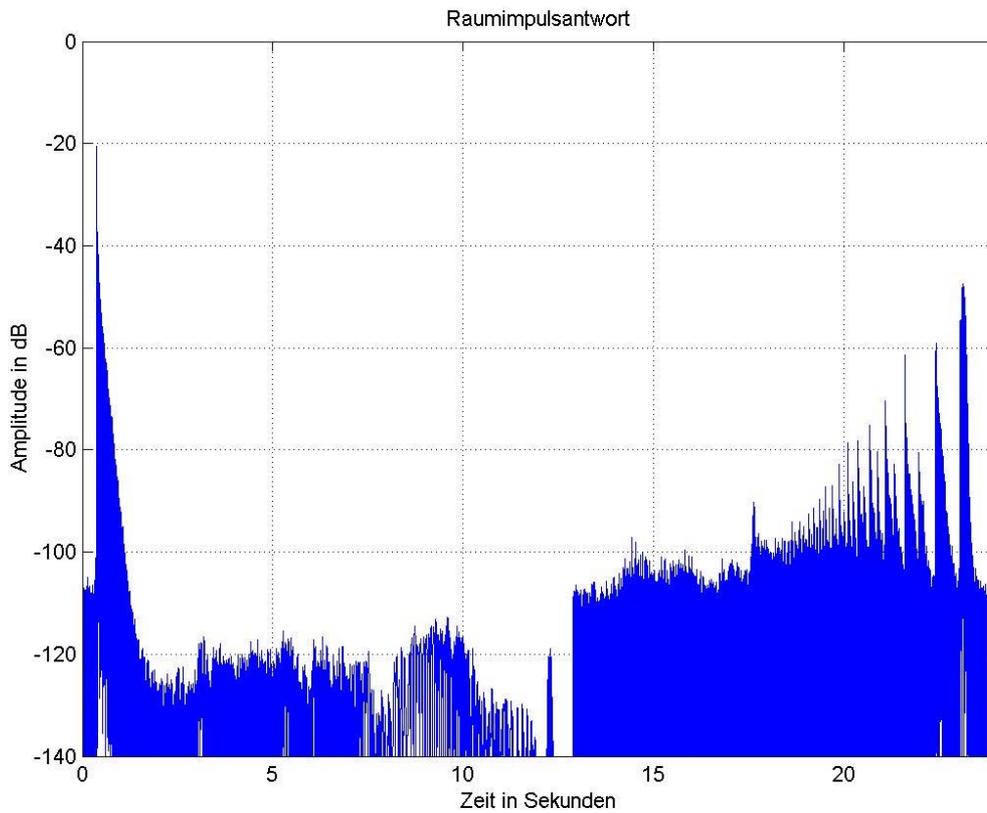


Abb. 5.5: Raumimpulsantwort mit harmonischen Verzerrungen

Wie sich aus der Abb. 5.5 erkennen lässt, werden diese in Folge der Dekonvolution im Frequenzbereich zirkulär an das Ende der gesamten Raumimpulsantwort verschoben. In der abklingenden Flanke jeder harmonischen Verzerrung befindet sich die zugehörige Raumimpulsantwort.

2.5.2 Generierung des Sweeps

Ein Sinussignal mit veränderlicher Frequenz kann mathematisch folgendermaßen beschrieben werden.

$$x(t) = \sin(f(t)) \quad (5.7)$$

Einen logarithmischen Sweep erhält man aus der Beziehung

$$x(t) = \sin \left[K \cdot \left(e^{\frac{t}{L}} - 1 \right) \right]. \quad (5.8)$$

Die Ableitung des Arguments der Sinusfunktion nach der Zeit liefert die Momentanfrequenz. Mit Hilfe der Startfrequenz ω_1 zum Zeitpunkt $t=0$ und der Stopfrequenz ω_2 zum Zeitpunkt $t=T$ lassen sich die unbekanntenen Konstanten K und L berechnen.

$$K = \frac{T \cdot \omega_1}{\ln \left(\frac{\omega_2}{\omega_1} \right)} \quad L = \frac{T}{\ln \left(\frac{\omega_2}{\omega_1} \right)} \quad T \dots \text{Sweeplänge} \quad (5.9)$$

Daraus ergibt sich die Gleichung zur Berechnung des logarithmischen Sweeps:

$$x(t) = \sin \left[\frac{\omega_1 \cdot T}{\ln \left(\frac{\omega_2}{\omega_1} \right)} \cdot \left(e^{\frac{t}{T} \ln \left(\frac{\omega_2}{\omega_1} \right)} - 1 \right) \right] \quad (5.10)$$

Durch die Ableitung des Arguments der Sinusfunktion nach der Zeit kann berechnet werden, bei welcher Zeitverschiebung Δt die obige Funktion eine Momentanfrequenz liefert, die der N-fachen der Ursprünglichen entspricht. Das Ergebnis repräsentiert die zeitliche Verschiebung der N-ten Verzerrung gegenüber der linearen Impulsantwort.

$$\Delta t = T \cdot \frac{\ln(N)}{\ln \left(\frac{\omega_2}{\omega_1} \right)} \quad (5.11)$$

$\Delta t \dots$ Abstand der harmonischen Verzerrung von der linearen Impulsantwort
 $T \dots$ Länge des Sweeps in Sekunden
 $N \dots$ Ordnung der harmonischen Verzerrung
 $\omega_1 \dots$ Startfrequenz des Sweeps
 $\omega_2 \dots$ Stopfrequenz des Sweeps

Der Zeitpunkt des Auftretens jeder einzelnen harmonischen Verzerrung ist in Bezug auf die lineare Impulsantwort als konstant anzusehen. Da Δt logarithmisch mit der Ordnung N zunimmt, verringert sich der Abstand zwischen den einzelnen harmonischen Impulsantworten mit zunehmender Ordnungszahl. Wie bereits in der Erklärung zur Abb. 5.5 erwähnt wurde befindet sich in der abklingenden Flanke jeder Harmonischen die zugehörige Raumimpulsantwort. Damit sich der abklingende Teil der Raumimpulsantwort von der Verzerrung zweiter Ordnung nicht mit der linearen Impulsantwort überlagert, muss die Sweeplänge T , bzw. die Start-/Stopfrequenz so gewählt werden, dass Δt größer als die Nachhallzeit des Messraumes ist.

2.5.3 Messumgebung

2.5.3.1 Hardware

Computersystem:

Das Messsystem basiert auf einem Laptop mit zugehöriger Audiohardware. Diese wird über einen ASIO-Treiber an das Betriebssystem angebunden, der kleine und vor allem reproduzierbare Latenzzeiten gewährleistet. Alle Messungen wurden bei einer Samplingrate von 44,1 kHz und mit 16 bit Auflösung durchgeführt. Aufgrund der hohen Qualität der DA- und AD-Konverter sowie der Messmikrofone samt Vorverstärker lässt sich eine Verfälschung der Ergebnisse durch diese Komponenten von vornherein weitgehend ausschließen.

Drehteller:

Für die Durchführung dieser Arbeit ist es in großem Umfang notwendig, Impulsantworten unter reproduzierbaren Winkeleinstellungen zu messen. Dazu wurde ein motorbetriebener Drehteller der Fa. Outline [Outline] verwendet, der um einen mechanischen Aufbau erweitert wurde, um die Vorgaben in Bezug auf Größe und Belastbarkeit zu erfüllen. Laut Datenblatt erreicht der Drehteller eine Einstellgenauigkeit von $0,25^\circ$ bei einer maximalen Winkelauflösung von $2,5^\circ$.

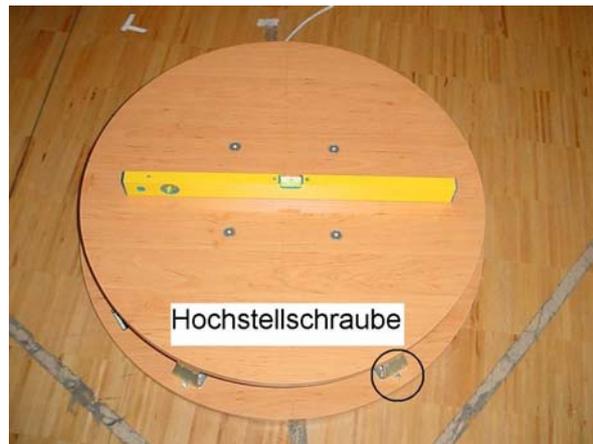


Abb. 5.6: Drehteller

Am Rand der Bodenplatte befinden sich im 60° -Abstand Hochstellschrauben zum waagrechten Einrichten des Drehtellers. Es ist zweckmäßig vorerst nur drei Schrauben innerhalb eines gleichseitigen Dreiecks zu verwenden, und die restlichen erst bei erfolgter horizontaler Justage soweit nachzudrehen, dass eine gute Kraftübertragung auf den Untergrund stattfindet.

2.5.3.2 Software:

Die praktische Durchführung der Messung wurde in Pure Data (PD) realisiert [Puckette]. PD ist eine grafische Programmierumgebung für die Audiosignalverarbeitung in Echtzeit. Die Aufgabe des Messprogramms besteht im Wesentlichen in der Wiedergabe des Testsignals, der synchronen Aufnahme der zu messenden Systemantwort und der Steuerung des Drehtellers. Die Generierung des Testsignals sowie die Auswertung der Messung wurde in MATLAB implementiert.

Wie schon oben erwähnt hat bei der Sweepmessung die Nachhallzeit des Raumes entscheidenden Einfluss auf die Wahl des Testsignals. Bei vorgegebenen Start- und Stopfrequenzen von 20 Hz und 20 kHz muss die Sweepdauer derart gewählt werden, dass sich die lineare Impulsantwort und die der ersten harmonischen Verzerrung nicht überlappen, d.h. die Verschiebung zu negativen Zeiten muss größer sein als die Länge der Raumimpulsantwort des Messraumes. Für den IEM-Cube mit einer mittleren Nachhallzeit von 0,7 s ergibt sich gemäß Gleichung (5.11) eine Sweepdauer von mindestens 7s. Außerdem muss die Aufnahmedauer ausreichend länger als die der Wiedergabe sein, um das Abklingen des Raumes nicht abzuschneiden. In diesem Fall würde sich ein Fehler bei hohen Frequenzen ergeben. Tiefe Frequenzen, die ja schon am Beginn des Messsweeps stattfinden, sind meist hinreichend abgeklungen. In MATLAB wurden Sweepfiles

verschiedener Dauer generiert, woraus je nach Länge der Raumimpulsantwort das für die Messung geeignetste ausgewählt wurde. Um transiente Störungen durch den Einschaltvorgang zu vermeiden, wurde der Beginn des Sweeps mit einem kurzen Hanningfenster eingeblendet, was grundsätzlich einen Fehler bei tiefen Frequenzen unter 25 Hz verursacht. Da in diesem Frequenzbereich aufgrund des Frequenzgangs des Messlautsprechers sowie der erreichbaren Frequenzauflösung keine validen Daten zu erwarten sind, kann der Fehler vernachlässigt werden.

PD-Patch:

In PD wurde ein Messpatch entwickelt, der alle notwendigen Parametereinstellungen über eine grafische Bedienoberfläche erlaubt. Beim Start von PD können eine Reihe von Programmeinstellungen mit Hilfe von sogenannten Flags durchgeführt werden. Zweckmäßigerweise werden diese in einem Batch-File zu einem vorkonfigurierten Startsetup zusammengefasst.

Folgende Einträge sind in der Befehlszeile unbedingt notwendig

```
allow_io 0x378 c:\Programme\pd\bin\pd -asio -path c:/.../lib -lib parallel_io
```

<code>allow_io 0x378</code>	erlaubt den Zugriff auf den Parallelport für das nachfolgende Programm
<code>c:\Programme\pd\bin\pd</code>	startet PD
<code>-asio</code>	PD verwendet den Asio-Soundkartentreiber
<code>-path c:/.../lib -lib parallel_io</code>	Pfadangabe für das PD-External zur Drehtellersteuerung

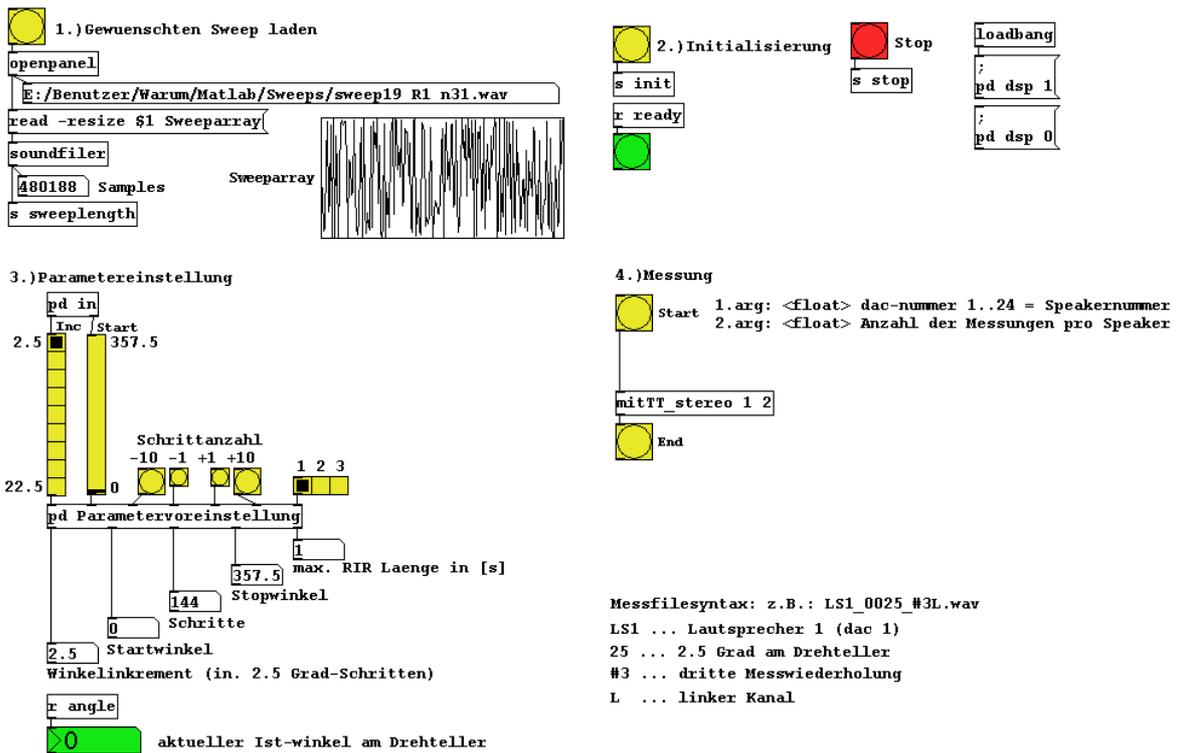


Abb. 5.7: grafische Oberfläche für die Messung in PD

Parametereingabe:

Alle für die Parametereingabe wichtigen Elemente sind farblich gekennzeichnet

1. Gewünschten Sweep laden
2. Parametereinstellung

Inc: Schrittweite des Drehtellers als Vielfache von 2,5°

Start: Startwinkel des Drehtellers

Schrittzahl:

max. RIR Länge: reserviert einen Speicherbereich, der zur Aufnahme des abklingenden Nachhalls benötigt wird.

Der Messbereich ist auf 360° beschränkt. Die Schrittzahl wird bei Überschreiten dieser Grenze automatisch auf den maximal möglichen Wert gesetzt.

3. Initialisierung

Initialisiert alle Unterprogramme mit den gewählten Einstellungen, und ist bei Abbruch einer Messung für die Rücksetzung der Variablen notwendig.

4. Messung



- a.) Für eine Einkanalmessung muss die Eingabe von "mitTT_mono" erfolgen und für eine Zweikanalmessung wird "mitTT_stereo" eingetragen.
- b.) Dieser Parameter bestimmt den verwendeten DA-Konverter Ausgangskanal.
- c.) Bestimmt die Anzahl der Messwiederholungen.

Im Zuge der Messauswertung in MATLAB werden vier Messungen gemittelt.

Dadurch verbessert sich der Signal-Rauschabstand des Ergebnisses um 6 dB.

5. Stop

Durch die Betätigung des Stop-buttons kann die Messung jederzeit abgebrochen werden. Danach ist für eine weitere Messung jedoch eine erneute Initialisierung notwendig.

Messfilesyntax:

Für die eindeutige Bezeichnung der Ergebnisfiles (Wave-Dateien) werden die momentanen Parameterwerte übernommen.

Beispiel: LS1_0025_#3L.wav
 LS1 ... Lautsprecher 1 bzw. DAC1
 0025 ... 2,5 Grad Drehtellerwinkel
 #3 ... dritte Messwiederholung
 L ... linker Kanal

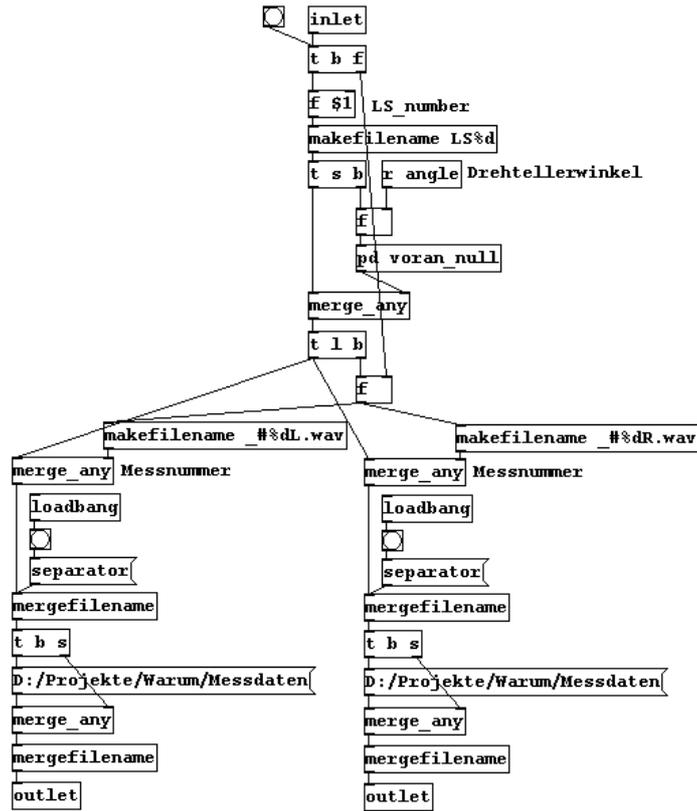


Abb. 5.8: PD-Subpatch für die Dateinamengenerierung (2-kanalig)

3 Vermessung und Optimierung des Kunstkopfes "Source"

3.1 Was ist ein Kunstkopf ?

Im Bestreben Schall so aufzuzeichnen, wie ihn das menschliche Ohr wirklich wahrnimmt, war es naheliegend ein Aufnahmesystem zu entwickeln, das versucht, den menschlichen Kopf möglichst gut nachzubilden. Die Wiedergabe dieser Aufnahmen über Kopfhörer lässt dann eine weitgehend natürliche Reproduktion der ursprünglichen räumlichen und klanglichen Situation erwarten.

Es gibt eine Reihe von Kunstköpfen am Markt, die sich in ihrer Formgebung oft erheblich unterscheiden. Um die Gründe für diesen Variantenreichtum zu verstehen, sollen zunächst kurz die wichtigsten Grundlagen des räumlichen Hörens erörtert werden.

3.2 Räumliches Hören

3.2.1 Richtungswahrnehmung

Das menschliche Gehör wertet für die Bestimmung der Schalleinfallrichtung sogenannte "auditory cues" aus, das sind in erster Linie:

- Interaurale Zeitdifferenzen (ITD, interaural time difference)
- Interaurale Pegeldifferenzen (ILD, interaural level difference)
- Spektrale Unterschiede (spectral cues)

Zweckmäßigerweise wird für die Richtungsbestimmung ein kopfbezogenes Kugelkoordinatensystem verwendet. Der Kopf des Hörers befindet sich dabei im Ursprung und die 0°-Richtung fällt sowohl für den Azimutwinkel (Horizontalebene) wie auch für den

Elevationswinkel (Medianebene) mit der Blickrichtung zusammen. Dadurch ergibt sich ein gegen Kopfbewegungen invariantes Koordinatensystem.

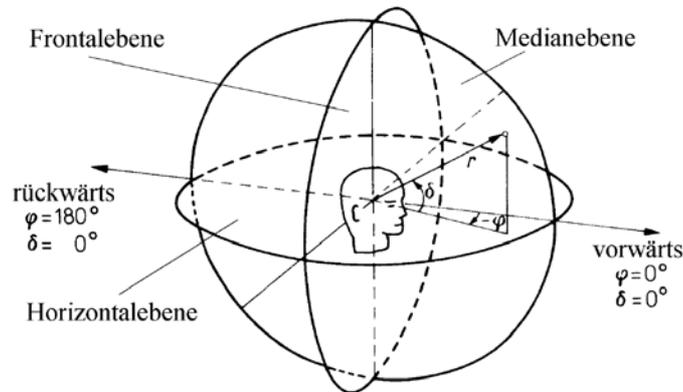


Abb. 6.1: Kopfbezogenes Koordinatensystem [Blauert, 1974]

3.2.1.1 Interaurale Zeitdifferenzen (ITD)

Bei seitlichem Schalleinfall ergeben sich aufgrund der unterschiedlichen Weglänge von der Schallquelle zum ipsilateralen⁴ bzw. kontralateralen⁵ Ohr Laufzeitdifferenzen zwischen den Ohrsignalen. Wird der Kopf vereinfacht als Kugel dargestellt, ergibt sich für ebenen Schalleinfall abhängig vom Einfallswinkel φ eine näherungsweise Laufzeitdifferenz Δt von

$$\Delta t = \frac{a}{c} \cdot (\sin \varphi + \varphi), \quad (6.1)$$

wobei a dem Kugelradius (Kopfradius) und c der Schallgeschwindigkeit entspricht.

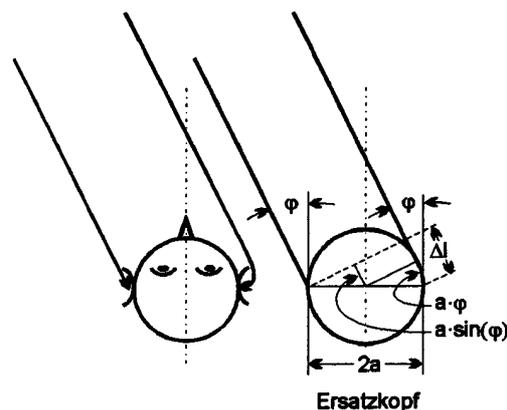


Abb. 6.2: Wegdifferenz der ebenen Schallwelle am Kugelmodell [Graber, 1997]

Eine eindeutige Zuordnung ist nur für einen begrenzten Frequenzbereich gültig, innerhalb dessen die Wegdifferenz kleiner bleibt als die halbe Wellenlänge. Für einen mittleren

⁴ ipsilaterales Ohr = dem Schallereignis zugewandtes Ohr

⁵ kontralaterales Ohr = dem Schallereignis abgewandtes Ohr

Ohrabstand von ca. 17cm erhält man somit eine Laufzeitdifferenz im Bereich von $650\mu\text{s}$, was einer oberen Grenzfrequenz von ca. 1,5kHz entspricht [Fellner, 1998]. Versuche mit gewobbelten oder getasteten höherfrequenten Sinussignalen haben allerdings gezeigt, dass das auditive System in der Lage, die ITD's auch anhand der Einhüllenden wahrzunehmen.

3.2.1.2 Interaurale Pegeldifferenzen (ILD)

Aufgrund einer Auslenkung der Schallquelle in der Horizontalebene erfährt der Schall am kontralateralen Ohr nicht nur eine Zeitverzögerung, sondern auch eine frequenzabhängige Pegelminderung. Dieser Pegelunterschied wird durch Beugungs- und Abschattungseffekte hervorgerufen. Für tiefe Frequenzen stellt der Kopf im Vergleich zur Wellenlänge ein kleines Hindernis dar. Die Schallwelle wird um den Kopf gebeugt und es sind kaum Pegelunterschiede zwischen den Ohrpositionen feststellbar. Für hohe Frequenzen findet zunehmend an der Vorderseite des Hindernisses Reflexion und an der Hinterseite Abschattung statt, was frequenzabhängig zu Differenzen von bis zu 20dB führt. Für die Richtungswahrnehmung ist der Bereich über 750Hz wirksam [Graber, 1997] Somit decken ITD's und ILD's gemeinsam den gesamten Hörbereich ab.

Für das Kugelmodell existiert eine Kegelmantelfläche (bekannt auch als "cone of confusion") auf der zwei gegenüberliegende Punkte die gleichen Laufzeit- und Pegelunterschiede hervorrufen, und somit zu Mehrdeutigkeiten führen. Beim natürlichen Hören werden daher Peilbewegungen (Drehung des Kopfes in Richtung Schallquelle) durchgeführt, die die statische Hörsituation in eine dynamische überführt, und somit weitere Informationen über das Schallfeld liefern.

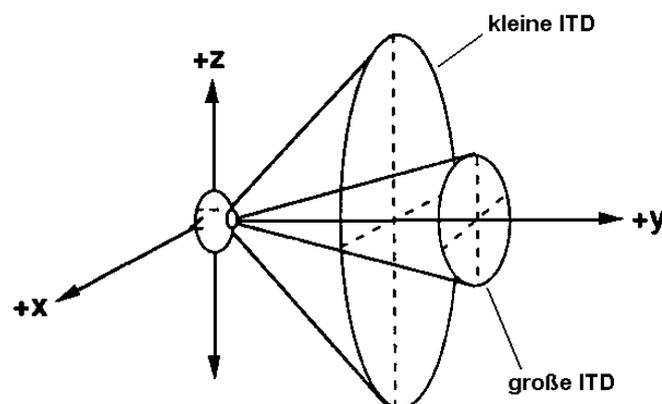


Abb. 6.3: Cone of confusion für verschiedene Laufzeitdifferenzen [Wenzel, 1993]

3.2.1.3 Spektrale Cues

Interaurale Zeit- und Pegeldifferenzen bestimmen primär die Lateralisation der Quelle. Für die Bestimmung der Einfallrichtung in der Medianebene sind diese aufgrund der Kopfsymmetrie nicht verwendbar, sodass monoaurale Merkmale dafür herangezogen werden. Das sind in erster Linie Klangfarbenverschiebungen, hervorgerufen durch die Ohrmuschel, Kopf, Schultern und Torso, die dem einfallenden Schall eine richtungsspezifische spektrale Gewichtung aufprägen. Testpersonen mit einseitig verschlossenem Ohr waren für breitbandigen Testschall fast uneingeschränkt in der Lage die korrekte Elevation festzustellen, für bandbegrenzte Signale unter 7kHz ging diese Eigenschaft auch für beidohriges Hören fast zur Gänze verloren [Blauert, 1974].

Einer Beobachtung von Blauert zur Folge lassen sich Frequenzbänder bestimmen, die direkt einem Richtungseindruck zuordenbar sind. So werden z.B. Schallereignissen mit einer Frequenz von 500Hz und 3kHz von vorne geortet, jene um 8kHz werden von oben und Signale um 1kHz von hinten lokalisiert.

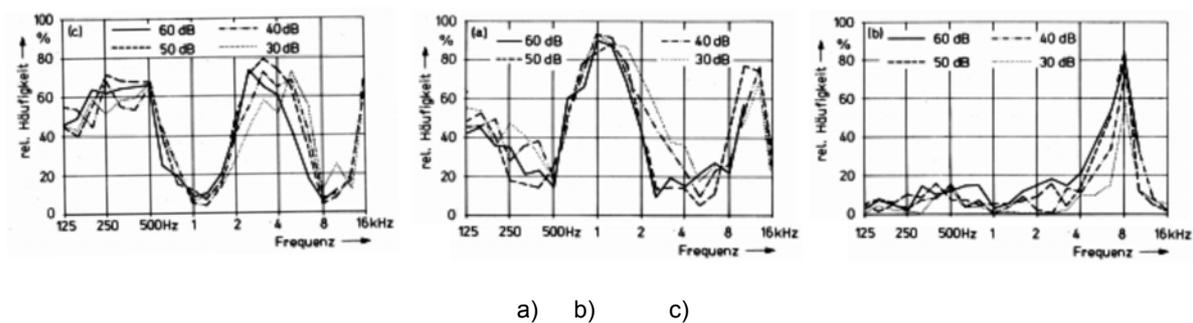


Abb. 6.4: Richtungsbestimmende Frequenzbänder [Blauert, 1974]

Von links nach rechts zeigen die Diagramme die relative Häufigkeit der Ortung von a) hinten, b) oben und c) vorne in Abhängigkeit von der Signalfrequenz.

3.2.2 Distanzwahrnehmung

Im Freifeld hängt die Distanzwahrnehmung lediglich vom Pegel am Hörort ab. Wie nicht anders zu erwarten war werden laute Schallereignisse als näher empfunden als leise. Für die Entfernungsabschätzung unter natürlichen Bedingungen stehen dem Ohr verschiedene Parameter zur Verfügung. So werden hier auch Änderungen der interauralen Pegeldifferenz bei verändertem Abstand zur Quelle ausgewertet. Für große Entfernungen (>15m) ist aufgrund der Luftdämpfung ein Abfall zu hohen Frequenzen hin hörbar. Für kleine Abstände (<3m) treten spektrale Verfärbungen aufgrund der Krümmung der eintreffenden Wellenfront am Kopf und durch den Einfluss der Ohrmuschel auf.

In geschlossenen Räumen gibt das Verhältnis von Direktschallenergie zur Energie der Reflexionen sowie deren zeitlicher Verlauf Aufschluss über Raumgröße, Raumform und den Abstand zur Quelle. Ein wichtiger Parameter ist dabei die Hallversatzzeit. Sie ist am Ort der Quelle selbst maximal, und sinkt mit steigendem Abstand.

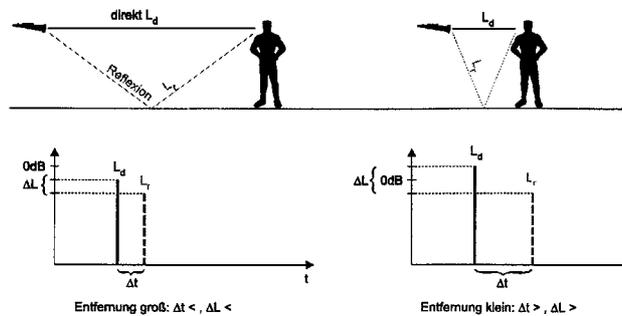


Abb. 6.5: Entfernungshören in Abhängigkeit von Hallversatzzeit und Hallpegel [Graber, 1997]

Ein Grenzfall der Entfernungswahrnehmung ist die Im-Kopf-Lokalisation (IKL). Sie beschreibt Schallereignisse, die in Entfernungen wahrgenommen werden, die kleiner sind als der Kopfradius und wird meist als unnatürlich empfunden. Am häufigsten ist die IKL bei Kopfhörerwiedergabe beobachtbar.

3.2.3 Außenohrübertragungsfunktionen (HRTF-Kurven):

Das Spektrum einer Schallquelle erfährt durch den Einfluss von Oberkörper, Schultern, Kopf und nicht zuletzt der Ohrmuschel (Pinna) eine Reihe von linearen Verzerrungen, die in ihrer Gesamtheit als kopfbezogene Außenohrübertragungsfunktionen (Head-Related Transfer Functions, HRTFs) bezeichnet werden. Die HRTF beschreibt das Verhältnis des Schalldrucks an einem Ort im Gehörgang der Versuchsperson zum Schalldruck im Koordinatenursprung, also im Kopfmittelpunkt bei Abwesenheit des Probanden.

Die binauralen HRTFs kann man sich als entfernungs- und frequenzabhängige Amplituden- und Zeitverzögerungsdifferenzen vorstellen, wobei zwischen richtungsabhängigen und nicht richtungsabhängigen Verzerrungen unterschieden werden kann. Richtungsabhängige Verzerrungen werden hauptsächlich durch die Abschattungseffekte des Kopfes, sowie durch die Richtcharakteristik der Pinna hervorgerufen, eine typische richtungsunabhängige Verzerrung stellt die Ohrkanalresonanz dar.

Besondere Bedeutung hat die Filterwirkung der Pinna für die Lokalisation in der Medianebene, da hier nur minimale interaurale Differenzen auftreten. Die komplexe Form der Ohrmuschel mit ihren unsymmetrischen Erhebungen und Kerben führt zu einer Vielzahl von Reflexionen mit Laufzeitunterschieden von bis zu $300\mu\text{s}$. Aus der Überlagerung dieser

reflektierten Anteile entstehen für verschiedene Elevationswinkel typische Einbrüche und Resonanzen im Verlauf der Außenohrübertragungsfunktion.

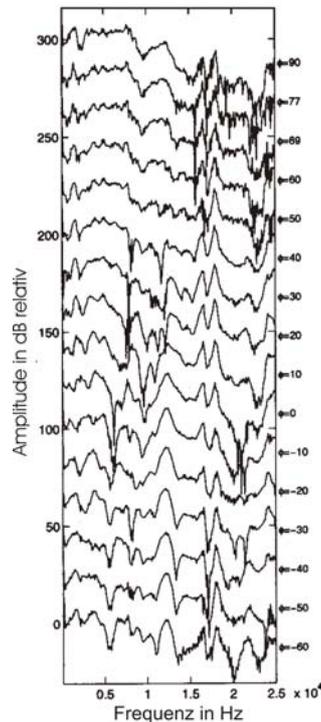


Abb. 6.6: HRTF-Kurven eines linken Ohres innerhalb der Medianebene in Abhängigkeit vom Elevationswinkel [Cheng, 1999]

Gut erkennbar sind z.B.: der Einbruch bei 7kHz, der sich mit steigendem Elevationswinkel zu höheren Frequenzen verschiebt oder die Überhöhung bei 12kHz, die zunehmend abflacht.

Nachdem HRTFs stark von der Form der Außenohren abhängen, weisen sie große intersubjektive Unterschiede auf. Dies stellt dann ein Problem dar, wenn eine Entscheidung für die Verwendung eines bestimmten Kurvensatzes für eine reale bzw. eine virtuelle binaurale Wiedergabe (vgl. [Wenzel, 1993]) gefällt werden muss. Daher greift man auf non-individuelle/generalisierte HRTFs zurück. Diese kann man durch Mittelung von HRTFs verschiedener Personen erreichen, wobei aber die Gefahr besteht, dass signifikante Minima und Maxima herausgemittelt werden. Eine andere Möglichkeit stellt die Messung von HRTFs an einem Kunstkopf dar. Beim Design eines Kunstkopfes begibt man sich unweigerlich auf eine Gratwanderung zwischen detailgetreuer Nachbildung eines realen Kopfes und der Modellierung mit vereinfachten Formen bis hin zur Kugel. Je nach Testperson sind durch stilisierte Formen teilweise besserer Ergebnisse erzielbar als mit sehr naturgetreuen Nachbildungen von Köpfen mit Pinnae, Schultern und Torso. Die am häufigsten beanstandeten Mängel bei der Wiedergabe von Kunstkopfaufnahmen sind

räumliche Verzerrungen, Vorne/Hinten - Vertauschung, die "Überelevation" und zu nahe Abbildung der Quelle bis hin zur Im-Kopf-Lokalisation. Bei Versuchen mit beweglichen Kunstköpfen, die den spontanen Kopfbewegungen des Testhörers folgen verschwand die Vorne/Hinten - Vertauschung fast zur Gänze, sogar bei der Anwendung eines Kugelflächenmikrofons [Theile, 1986]. Daraus wird klar, dass für eine erfolgreiche dreidimensionale Abbildung des Schallfeldes die dynamischen Aspekte miteinbezogen werden müssen.

3.3 Kunstkopf "Source"

Der im Weiteren verwendete Kunstkopf "Source" wurde von Franz Graf im Zuge seiner Diplomarbeit [Graf, 1999] entwickelt, und wurde mir freundlicherweise für die Dauer meiner Diplomarbeit von der AVL List GmbH zur Verfügung gestellt.



Abb. 6.7: Kunstkopf "Source"

Source hat sich in psychoakustischen Versuchsreihen als sehr gut klingendes Aufnahmesystem herausgestellt. Da dieser Kunstkopf allerdings primär als Messinstrument für die Evaluierung von Fahrzeuginnengeräuschen verwendet wird, sind objektive Parameter für die Beurteilung der Ergebnisse unerlässlich.

3.3.1 Messbedingungen

3.3.1.1 Systemtheoretische Beschreibung

Messungen von Freifeld-Impulsantworten werden im den seltensten Fällen wirklich im Freien durchgeführt, stattdessen steht dafür oft ein schalltoter Raum zur Verfügung dessen Wandverkleidung keine störenden Reflexionen verursacht, bzw. diese auf ein Minimum reduziert. Freifeldmessungen sind aber unter gewissen Einschränkungen auch in

gewöhnlichen Räumen durchführbar. Die Messsituation stellt sich dann wie in Abb. 6.8 skizziert dar.

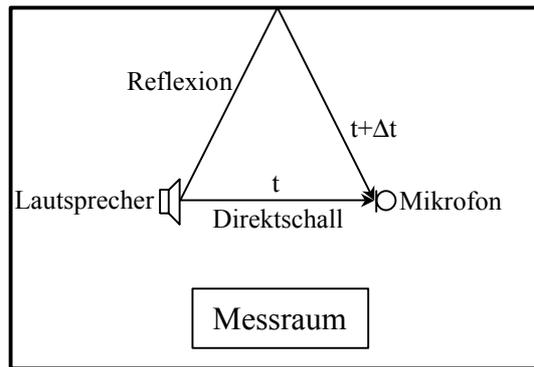


Abb. 6.8: Übliche Messsituation

Systemtheoretisch kann die Messstrecke folgendermaßen beschrieben werden:

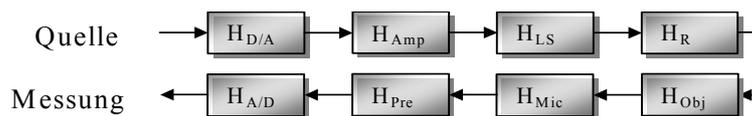


Abb. 6.9: Systemtheoretische Beschreibung der Messstrecke

Gesucht ist die Freifeld-Übertragungsfunktion des Messobjektes (H_{Obj}), in diesem Fall die des Kunstkopfes. Die Übertragungsfunktionen der D/A- und A/D-Konverter (H_{DA} , H_{AD}), der Endstufe (H_{Amp}), des Mikrofonvorverstärkers (H_{Pre}) und des Messmikrofons (H_{Mic}) können als hinreichend linear vorausgesetzt werden, dass deren Einfluss auf das Messergebnis vernachlässigbar klein ist. Der Lautsprecher (H_{LS}) und die Übertragungseigenschaften des Raumes (H_R) verursachen jedoch Messfehler, die vor der Auswertung eliminiert werden müssen. Da wir von einem LTI-System ausgehen können (auch unterstützt durch die gewählte Sweep-Messmethode) lässt sich obiges Blockschaltbild auf drei Hauptkomponenten zusammenfassen, die elektroakustische Übertragungsstrecke (H_{EAK}), den Messraum (H_R) und das zu messende Objekt.



Abb. 6.10: Messstrecke mit zusammengefassten Komponenten

3.3.1.2 Simulierte Freifeldmessung

Voraussetzung für die Freifeldmessung in reflektierender Umgebung ist, dass am Messmikrofon ein genügend großer Zeitbereich Δt zwischen dem Eintreffen des Direktschalls der Quelle und der ersten Raumreflexion vorhanden ist. Nur dieser sogenannte Freifeldanteil der Raumimpulsantwort wird für die Weiterverarbeitung der Messdaten herangezogen. Da alle Raumreflexionen erst später eintreffen ist dessen Einfluss somit aus der Messkette eliminiert.

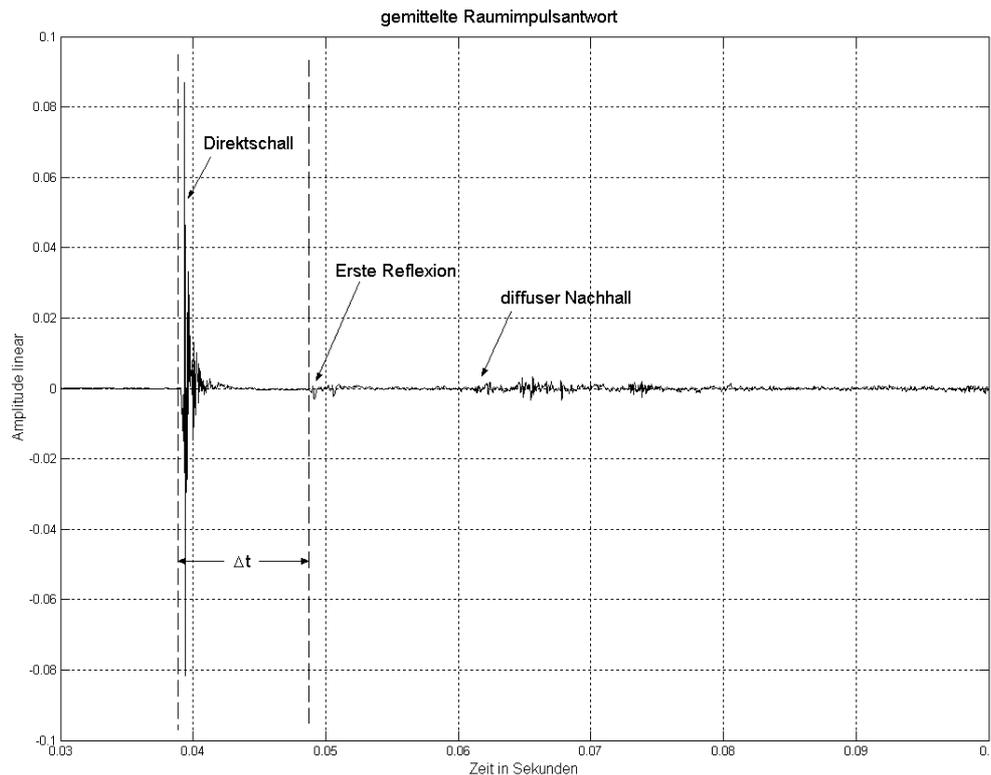


Abb. 6.11: Raumimpulsantwort

Die Länge von Δt bestimmt die untere Grenzfrequenz der Messung bzw. die erreichbare Frequenzauflösung, da für ein gültiges Ergebnis mindestens eine volle Schwingungsperiode innerhalb des Freifeldanteils Platz finden muss. Die freie Weglänge ist umso größer, je näher das Mikrofon am Lautsprecher positioniert ist, und je weiter reflektierende Flächen von dieser Anordnung entfernt sind. Da die Messung im Fernfeld des Lautsprechers stattfinden soll, ist dem Mikrofonabstand eine Untergrenze gesetzt. Die Fernfeldbedingung für den Kugelstrahler ist definiert als [Graber, 1997]

$$k \cdot r \geq 2, \quad (6.2)$$

$$\text{mit } r = \text{Abstand und } k = \frac{2 \cdot \pi}{\lambda} = \frac{2 \cdot \pi \cdot f}{c} = \text{Wellenzahl.}$$

Für den gewählten Lautsprecherabstand von 1,5m ergibt sich somit eine untere Grenzfrequenz von 72Hz, was in Anbetracht der erreichbaren Frequenzauflösung von etwa 100Hz ($\Delta t \approx 10ms$) und des Lautsprecherfrequenzganges als ausreichend erachtet werden kann.

3.3.1.3 Lautsprecherauswahl

Idealerweise sollte der Lautsprecher im gesamten zu messenden Frequenzbereich einen linearen Frequenzgang besitzen. Des Weiteren wünscht man sich von einem Messlautsprecher eine möglichst kugelförmige Abstrahlcharakteristik damit ein konstanter Frequenzgang für einen großen Winkelbereich um die Lautsprecherachse gegeben ist. Messungen, die von leicht abweichenden Messmikrofonpositionen stammen, können somit mit einer Referenzmessung entzerrt werden. Für einen ausgedehnten Frequenzgang hin zu tiefen Frequenzen sind große Lautsprechermembranen erforderlich. Je größer allerdings die Membran, und damit auch das Gehäuse im Vergleich zur Wellenlänge des wiedergegebenen Signals werden, umso gerichteter ist auch das Abstrahlverhalten. Man erreicht also immer nur einen Kompromiss zwischen den eingangs genannten Forderungen.

Auf der Suche nach einem möglichst kleinen Lautsprecher mit ausreichend breitbandigem und linearem Frequenzgang wurden Vergleichsmessungen zwischen acht Modellen durchgeführt. Die Entscheidung fiel auf den großen Tropfenlautsprecher. Details der anderen Kandidaten sind im Anhang A zusammengefasst.



Abb. 6.12: Messlautsprecher

Abb. 6.13 zeigt die Frequenzgänge des Messlautsprechers für Azimutwinkel im Bereich von 0° - 180° in 10° -Schritten. Dabei entspricht die oberste Kurve der 0° -Richtung (blau). Die weiteren Winkelabstufungen sind besonders zwischen 1kHz und 3kHz gut unterscheidbar.

Am parallelen Verlauf der Kurvenschar bis 300Hz erkennt man, dass die gewünschte Kugelcharakteristik nur für tiefe Frequenzen erreicht wird. Auffallend ist allerdings der Amplitudenoffset bei verschiedenen Drehwinkeln. Offenbar liegt das akustische Zentrum des Lautsprechers nicht auf der Drehachse, sondern bewegt sich auf einer Kreisbahn um diese. Dadurch verändert sich auch der Abstand zum Messmikrofon und führt zu Amplitudendifferenzen bei unterschiedlichen Drehwinkeln. Für hohe Frequenzen ist deutlich die zunehmende Richtwirkung erkennbar.

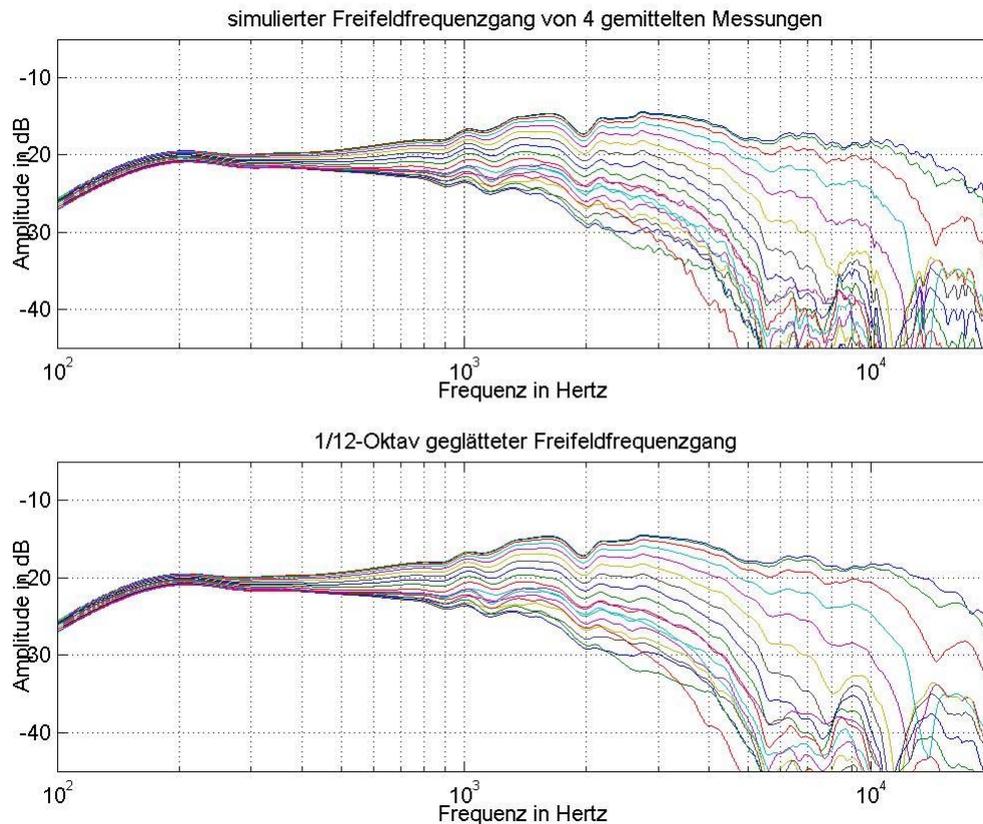


Abb. 6.13: Frequenzgänge des Messlautsprechers unter verschiedenen Azimutwinkeln

3.3.1.4 Lautsprecherentzerrung

Genaugenommen kompensiert man den Einfluss der gesamten elektroakustischen Übertragungsstrecke. Da allerdings der Lautsprecher mit Abstand den größten Fehler verursacht, wird allgemein von der Lautsprecherentzerrung gesprochen.

Der Freifeldanteil des Messergebnisses entspricht der Faltung des Messsignals $x(t)$ mit den Impulsantworten von Lautsprecher $h_{LS}(t)$ und Messobjekt $h_{Obj}(t)$.

$$y(t) = h_{LS}(t) \otimes h_{Obj}(t) \otimes x(t) \quad (6.3)$$

Direkt im Zeitbereich ist die Kompensation von $h_{LS}(t)$, nicht durchführbar, da keine zur Faltung inverse Rechenvorschrift existiert. Im Frequenzbereich ist eine Lösung durch

komplexe Division mit dem Freifeldfrequenzgang des Lautsprechers möglich. h_{LS} muss in einer eigenen Referenzmessung in Abwesenheit des Kunstkopfes ermittelt werden. Mittels Fouriertransformation lässt sich Glg.(6.3) wie folgt darstellen.

$$Y(f) = H_{LS}(f) \cdot H_{Obj}(f) \cdot X(f) \quad (6.4)$$

Entzerrung:
$$\bar{Y}(f) = \frac{Y(f)}{H_{LS}(f)} = H_{Obj}(f) \cdot X(f) \quad (6.5)$$

Nach inverser Fouriertransformation erhält man wiederum das entzerrte Zeitsignal

$$\bar{y}(t) = h_{Obj}(t) \otimes x(t). \quad (6.6)$$

Bei Anregung mit dem Dirac-Impuls $x(t) = \delta(t)$ folgt aus (6.6)

$$\bar{y}(t) = h_{Obj}(t) \otimes \delta(t) = h_{Obj}(t). \quad (6.7)$$

Voraussetzung für die Anwendbarkeit dieses Verfahrens ist, dass die Lautsprecherimpulsantwort $h_{LS}(t)$ innerhalb von Δt vollständig ausgeschwungen ist. Trifft die erste Raumreflexion noch in der Ausschwingphase des Messlautsprechers ein, werden durch das Herausschneiden (Fensterung) des Freifeldanteils der Raumimpulsantwort relevante Anteile der Lautsprecherimpulsantwort entfernt. Die Multiplikation der Raumimpulsantwort mit der Fensterfunktion im Zeitbereich entspricht im Frequenzbereich einer Faltung mit dem Fensterfrequenzgang und zieht eine Verschlechterung der Frequenzauflösung nach sich. Da der Hochtonbereich meist schnell ausgeschwungen ist, ergeben sich Fehler im Referenzfrequenzgang hauptsächlich bei den tiefen Resonanzfrequenzen.

3.3.1.5 Cepstralanalyse

Das komplexe Cepstrum⁶ einer Folge $h_{LRM}(n)$ ist definiert als die inverse Fouriertransformation des komplexen Logarithmus der Systemübertragungsfunktion.

$$h(n) \xrightarrow{DFT} H(f) \xrightarrow{\log} \log[H(f)] \xrightarrow{DFT^{-1}} h(q) \quad (6.8)$$

Die in nicht reflexionsfreier Umgebung gemessene Impulsantwort des Lautsprecher-Raum-Mikrofon Systems ergibt sich aus der Faltung der Einzelimpulsantworten.

$$h_{LRM}(n) = h_{LS}(n) \otimes h_R(n) \otimes h_{Mic}(n) \quad (6.9)$$

$$H_{LRM}(f) = H_{LS}(f) \cdot H_R(f) \cdot H_{Mic}(f) \quad (6.10)$$

⁶ Die Bezeichnung "Cepstrum" wurde 1963 von Bogert et al eingeführt. Sie entstand aus dem Umdrehen der ersten Silbe des Wortes "Spectrum". Entsprechend wird das Cepstrum über der "Quefrenz", abgeleitet aus "Frequenz" aufgetragen.

$$\begin{aligned} \log[H_{LRM}(f)] &= \log[H_{LS}(f) \cdot H_R(f) \cdot H_{Mic}(f)] = \\ &= \log[H_{LS}(f)] + \log[H_R(f)] + \log[H_{Mic}(f)] \end{aligned} \quad (6.11)$$

$$h_{LRM}(q) = h_{LS}(q) + h_R(q) + h_{Mic}(q) \quad (6.12)$$

Durch die Logarithmierung von $H_{LRM}(f)$ kann die multiplikative Verknüpfung von $H_{LS}(f)$, $H_R(f)$ und $H_{Mic}(f)$ in eine additive Verknüpfung übergeführt werden. Dementsprechend setzt sich auch das komplexe Cepstrum aus der Addition der Teilcepstra zusammen. Der logarithmierte Frequenzgang kann bei der inversen DFT zu Aliasingproblemen führen. Abhilfe schafft vorheriges Zero-padding der Ausgangsfolge, wodurch eine höhere Abtastung des Frequenzganges erreicht wird.

Zwei wichtige Eigenschaften werden nun zur Trennung von Lautsprecher- und Raumimpulsantwort ausgenutzt:

- a) Signale mit genügend flachem Frequenzgang klingen im Cepstralbereich schneller ab als im Zeitbereich.
- b) Zeitverzögerte minimalphasige Anteile der ursprünglichen Impulsantwort sind auch im Cepstralbereich um den gleichen Wert verschoben.

Ad a):

Lautsprecher und Mikrofone besitzen für gewöhnlich flache Frequenzgänge innerhalb der Übertragungsbandbreite. Für tiefe Frequenzen kann diese Forderung allerdings nicht mehr erfüllt werden, da Lautsprecher gleichspannungsentkoppelt sind und somit Hochpasscharakteristik aufweisen. Der Bass-Rolloff und die damit einhergehenden Nichtlinearitäten im Phasenverlauf führen zu einer Aufspreizung des Cepstrums zu höheren Queffrenzen. Damit besteht die Gefahr, dass sich wie im Zeitbereich Teile des Lautsprechercepstrums mit dem der ersten Reflexionen überlappen.

Ad b):

Ist $h_R(n)$ minimalphasig und besitzt die Form

$$h_R(n) = \delta(n - n_0) + \sum_{i=1}^{\infty} h_{Ri}(n - n_i), \quad (6.13)$$

dann gilt

$$\begin{aligned} h_R(q) &= 0, & q < q_1 - q_0 \\ h_R(q) &\neq 0, & q \geq q_1 - q_0 \end{aligned}, \text{ wobei } q_1 = n_1 \text{ und } q_0 = n_0. \quad (6.14)$$

Die ersten Raumreflexionen erfüllen diese Bedingung ausreichend gut, und liefern somit auch im Cepstrum erst ab Quefrenzen $\geq q_1$ einen Beitrag zum Kurvenverlauf. Für spät eintreffende Reflexionen ist die Minimalphasigkeit nicht mehr gegeben, und es kommt wieder zu Überlappungen des Lautsprechercepstrums mit dem der Raumreflexionen. Trotzdem liefert eine Fensterung im Cepstralbereich meist bessere Ergebnisse als direkt im Zeitbereich.

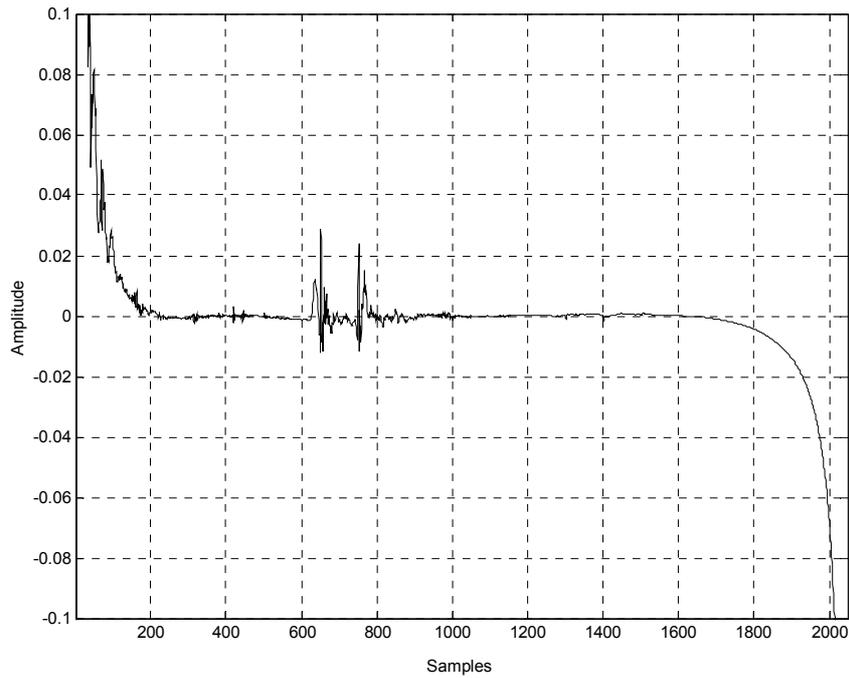


Abb. 6.14: komplexes Cepstrum $h_R(q)$

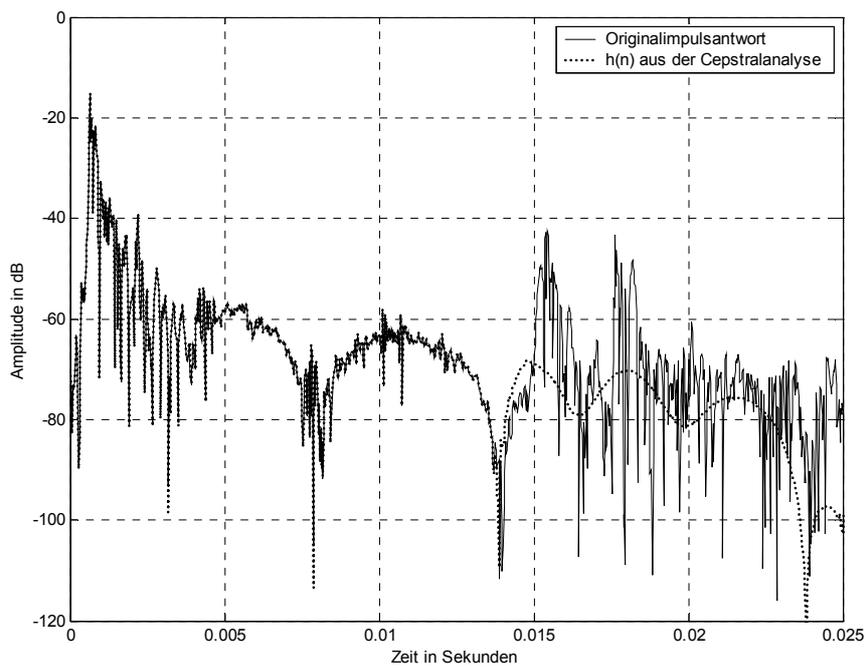


Abb. 6.15: Ursprüngliche und im Cepstralbereich gefensterter Impulsantwort

Nach Fensterung mit $w(q)$ und Rücktransformation erhält man eine von den ersten Reflexionen befreite Lautsprecherimpulsantwort.

$$h_{LRM}(q) \xrightarrow{w(q)} h_{LM}(q) \xrightarrow{DFT} \log[H_{LM}(f)] \xrightarrow{\exp} H_{LM}(f) \xrightarrow{DFT^{-1}} h_{LM}(n) \quad (6.15)$$

Für die Ermittlung des zugehörigen Lautsprecher-Referenzfrequenzganges steht nun ein größeres reflexionsfreies Zeitfenster zur Verfügung. Der Kurvenverlauf ist zu tiefen Frequenzen hin ausgeweitet und es ist ein schärferer Amplitudenabfall erkennbar. [Bauman, 1984], [Bauman, 1985], [Niedrist, 1993], [Salamouris, 1992]

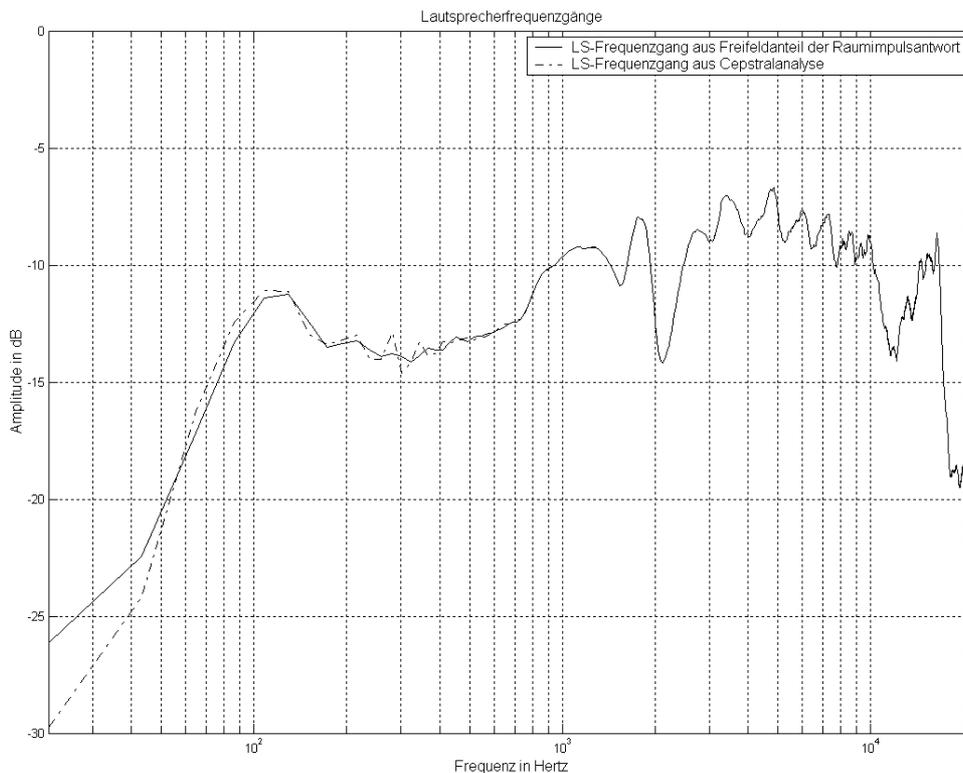


Abb. 6.16: Referenzfrequenzgang aus Direktsignal und nach cepstraler Bearbeitung

3.3.2 Vermessung des "Source"

An Source wurden insgesamt 325 binaurale Freifeldimpulsantworten für Schalleinfallrichtungen innerhalb des oberen Halbraumes gemessen. Bezogen auf den Kunstkopf liegen die Positionen des Messlautsprechers somit auf einer Halbkugel mit 1,5m Radius. Die Anordnung der Messpunkte erfolgt in horizontalen Schnittebenen dieser fiktiven Halbkugel, dadurch ist nur für veränderte Elevationswinkel eine (händische) Neupositionierung der Schallquelle notwendig. Das Weiterdrehen in Azimutrichtung erfolgt mit Hilfe der Drehtellerautomatisierung. Die erreichbare Winkelauflösung ist durch den Drehteller auf 2,5° begrenzt, was aber im Bereich des "minimum audible angle" (Blauert verwendet dafür den Begriff Lokalisationsunschärfe) in Blickrichtung liegt, und somit als ausreichend klein angesehen werden kann. Für steigende Elevationen sinkt das Auflösungsvermögen des menschlichen Gehörs und es verkleinern sich zunehmend die

Radien der Schnittkreise. Daher ist auch eine geringere Anzahl an Messpunkten ausreichend.

Der Messpunktraster wurde wie folgt festgelegt:

Elevationswinkel	Azimet-Winkelauflösung	Anzahl der Messpunkte
0° (Horizontalebene)	2,5°	144
15°	5°	72
30°	5°	72
50°	15°	24
75°	30°	12
90°	-	1
Summe		325

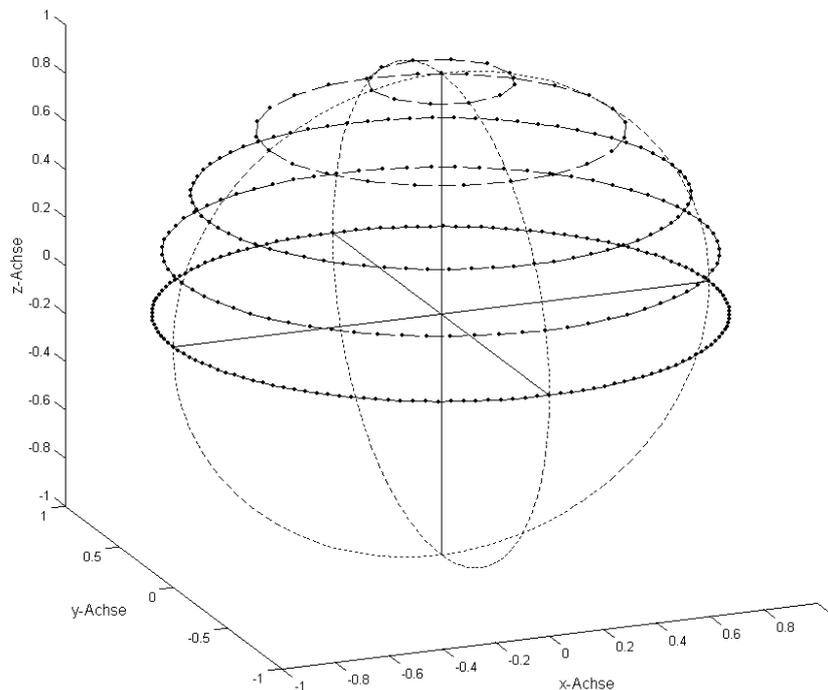


Abb. 6.17: Messpunktraster für die Vermessung von Source

3.3.2.1 Freifeldübertragungsfunktion in Vorne-Richtung

Es ist für einen Kunstkopf naheliegend als Referenz die Vorne-Richtung (Azimetwinkel = Elevationswinkel = 0°) zu wählen. Die erreichte Zeitauflösung beträgt $\Delta t \cong 9ms$, bei einem Lautsprecherabstand von 1,5m. Für die 1/12-Oktav-Glättung kommt ein Moving-Average

Filter zur Anwendung. Die Messung bestätigt weitgehend die Simulationsergebnisse aus der Entwicklungsphase des Source.

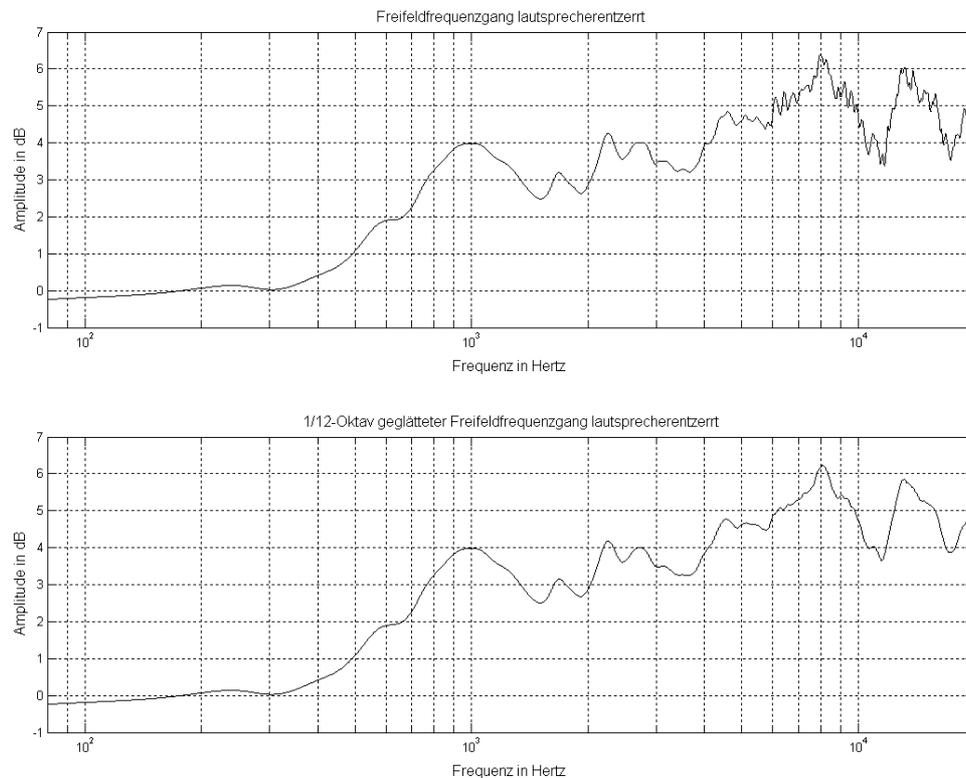


Abb. 6.18: Freifeldübertragungsfunktion des Source

3.3.2.2 Diffusfeldfrequenzgang

Bei der Schallausbreitung in Räumen bildet sich neben dem durch die Schallquelle abgestrahlten Schall (Direktschallfeld) durch Reflexionen an den Wänden, der Decke, dem Boden und den Einrichtungsgegenständen ein diffuses Schallfeld aus. Der diffuse Schall ist im Raum gleichmäßig verteilt. Nahe der Schallquelle, überwiegt der Direktschallanteil, der mit der Entfernung von der Schallquelle rasch abnimmt. In einiger Entfernung von der Schallquelle überwiegt der diffuse Schallanteil. Die Entfernung von der Schallquelle, in der beide Anteile, d.h. Direktschall und Diffusschall, gleich groß sind, wird als Hallradius bezeichnet. Die Schallenergiedichte im Direktschallfeld w_{dir} hängt von der Schallleistung und der Entfernung von der Schallquelle ab. Im Gegensatz hierzu ist die Schallenergiedichte im diffusen Schallfeld w_{diff} von der Entfernung unabhängig. Sie wird von der Schallleistung und der äquivalenten Schallabsorptionsfläche bestimmt.

Der Diffusfeldfrequenzgang ergibt sich aus der Summation der Gesamtenergie über alle Schalleinfallrichtungen. Jedem Messpunkt ist eine komplexe Freifeldübertragungsfunktion

und ein Flächenanteil auf einer gedachten Kugeloberfläche zugeordnet. Durch Bildung der Betragsquadrate der Frequenzgänge erhält man deren Energieinhalte. Diese werden mit dem zugehörigen Flächenanteil gewichtet, aufsummiert und auf die Kugelgesamtoberfläche normiert. Die Quadratwurzel dieser Summe ergibt den Diffusfeldfrequenzgang.

$$H_{Diff}(f) = \sqrt{\frac{\sum_i |H_i(f)|^2 \cdot A_i}{A_{ges}}} \quad (6.16)$$

wobei

$$A_i = \int_{\varphi_1}^{\varphi_2} \int_{\vartheta_1}^{\vartheta_2} r \cdot d\varphi \cdot \cos \vartheta \cdot r \cdot d\vartheta \quad (6.17)$$

- H_{Diff} ... Diffusfeldfrequenzgang
- H_i ... Übertragungsfunktion für eine Richtung
- A_i ... Flächenanteil für eine Richtung
- A_{ges} ... Kugeloberfläche $4 \cdot \pi \cdot r^2$
- r ... Kugelradius
- φ ... Azimutwinkel
- ϑ ... Elevationswinkel

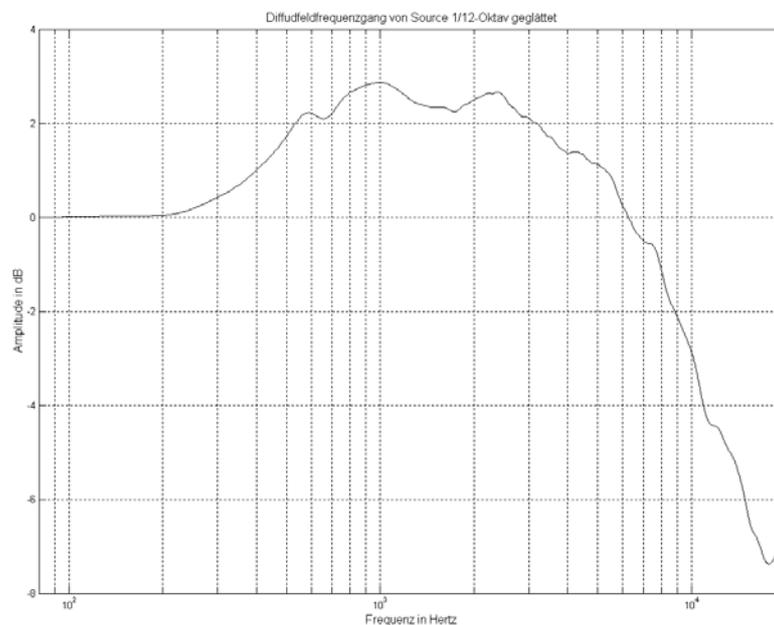


Abb. 6.19: Diffusfeldfrequenzgang

Betrachtet man nur einen nach vorne orientierten Kugelausschnitt von $45 \times 45^\circ$ erhält man einen mittleren Frequenzgang für die Blickrichtung.

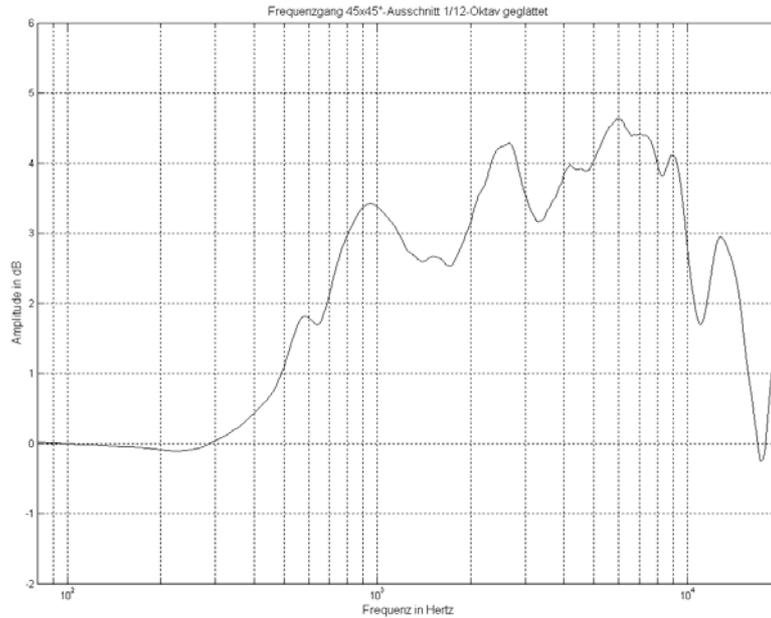


Abb. 6.20: mittlerer Frequenzgang in Blickrichtung

3.3.3 Optimierung für 0°-Richtung

Der Kunstkopf "Source" wird hauptsächlich in der akustischen Fahrzeugmesstechnik eingesetzt. Für ein Messgerät wird natürlich ein möglichst lineares Übertragungsverhalten für die Referenzrichtung angestrebt. In Abb. 6.18 lassen sich zwei Hauptproblemzonen in der Freifeld-Übertragungsfunktion erkennen, die Überhöhung um 8kHz und der Anstieg zwischen 300Hz und 1kHz. Die Optimierung des Frequenzganges soll soweit möglich mittels konstruktiver Methoden erreicht werden.

3.3.3.1 Korrektur der Überhöhung um 8kHz

Es wurde versucht durch Anbringen von Absorbermaterialien auf den Seitenflächen des Source eine Bedämpfung des Frequenzbereiches über 4kHz zu erreichen. Hier ist die Wellenlänge noch klein genug, dass schon mit praktikabel dünnen Schichten eine Linearisierung des Übertragungsverhaltens erwartet werden kann. Bei senkrechtem Schalleinfall auf einen Absorber vor einem schallharten Trägermaterial muss die Schichtdicke größer sein als $\lambda/4$. Hier nimmt die Schallschnelle und somit die Molekülbewegung in der Luft den größten Wert an und die Energie wird aufgrund von Reibungsverlusten im Material am effektivsten in Wärme umgewandelt. Durch den schiefen Schalleinfall auf die Seitenflächen des Kunstkopfes (19°) vergrößert sich die

effektive Weglänge im Absorber, was sich zusätzlich positiv auf die notwendige Materialstärke auswirkt.

Folgende Materialien wurden getestet:

1. Schaumstoffabsorber mittel 4,11 und 20mm
2. Schaumstoffabsorber weich 5,11 und 20mm (gelb)
3. Korkplatte 4mm
4. Filzteppich 4mm
5. Gummibelag 3 und 6mm
6. Schaumstoffabsorber fest 3mm



Abb. 6.21: Absorbermaterialien

Ergebnisse:

Die Messkurven aller Materialien sind im Anhang B abgebildet.

- ad 1. verursacht im Hochtonbereich Dämpfungen $>20\text{dB}$. Die Filterwirkung ist geprägt von einem dominierenden Einbruch, dieser setzt erwartungsgemäß für dickere Schichten bei tieferen Frequenzen ein. Die 4mm-Variante erreicht zwar die gewünschte Korrektur zwischen 4kHz und 8kHz, beeinträchtigt allerdings darüber liegende Frequenzbereiche so stark, dass insgesamt von keiner Verbesserung gesprochen werden kann.
- ad 2. Wirkung ähnlich wie 1. aber mit noch schmalbandigerem und tieferem Einbruch \Rightarrow unbrauchbar.
- ad 3. Die Korkplatte erreicht eine gleichmäßige Bedämpfung ab 3kHz, ohne starke Unregelmäßigkeiten im weiteren Verlauf zu verursachen. Der resultierende

Frequenzgang bietet eine deutliche Verbesserung der Linearität zu hohen Frequenzen hin.

- ad 4. verhält sich ähnlich wie der 4mm Schaumstoffabsorber (mittel)
- ad 5. Der Gummibelag lässt kaum Veränderungen im Übertragungsverhalten des Source erkennen. Ein Zusammenhang mit der Materialstärke ist auch nicht nachvollziehbar.
- ad 6. verhält sich ähnlich wie 1. Der Frequenzgangeinbruch liegt wenig störend am oberen Ende des Hörbereichs, die Absorptionswirkung ist aber im gewünschten Frequenzbereich zu schwach ausgeprägt.

Das beste Ergebnis liefert somit die 4mm Korkplatte.

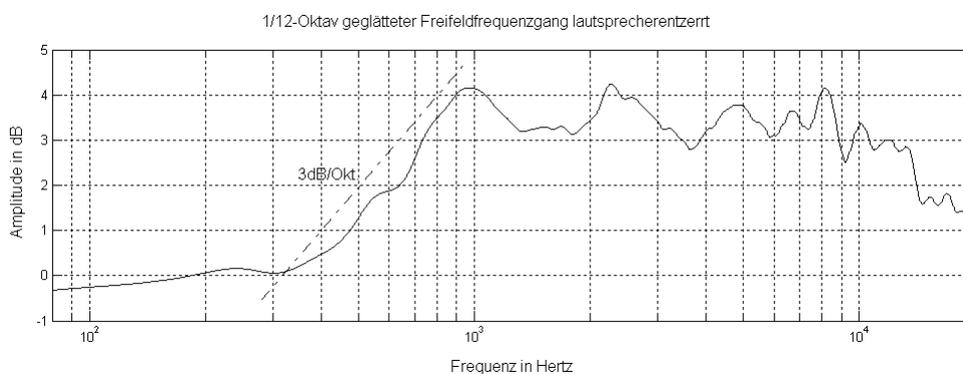


Abb. 6.22: Freifeldfrequenzgang Source mit 4mm Korkplatte

3.3.3.2 Korrektur des Anstiegs zwischen 300Hz und 1kHz

Der Frequenzbereich unter 1kHz mit Wellenlängen ab etwa 34cm (für Luftschall) bleibt durch die Applikation von Absorbermaterialien völlig unbeeinflusst. Mittels konstruktiver Maßnahmen ist dieses Problem unter grundsätzlicher Beibehaltung der Formgebung des Kunstkopfes nicht lösbar. Somit bleibt nur noch die Möglichkeit einer elektronischen Entzerrung. Der Frequenzgang in Abb. 6.22 verläuft bis 300Hz linear und steigt dann bis 1kHz mit etwa 3dB/Oktave an. Für die Korrektur bietet sich ein Tiefshelvingfilter an. Dieses Filter erlaubt das Anheben bzw. Absenken des gesamten Frequenzbereiches unterhalb einer bestimmten Grenzfrequenz. Diese Übertragungsfunktion erreicht man durch Parallelschaltung eines Tiefpassfilters mit der Gleichspannungsverstärkung V_0 und eines Systems mit der Übertragungsfunktion $H(z)=1$.

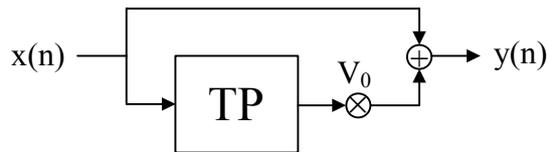


Abb. 6.23: Filterstruktur eines Tiefen-Shelvingfilters

Theoretisch besitzt das Shelvingfilter 1.Ordnung eine Flankensteilheit von 6dB/Oktave. Dieser Wert wird allerdings nur für hohe Verstärkungen V_0 und in einiger Entfernung von der Grenzfrequenz f_g erreicht. Die Grenzfrequenz ist definiert als der -3 dB-Punkt des Tiefpasszweiges. Bei der hier erforderlichen Anhebung von nur 3-4dB befindet man sich stets in den flachen Übergangsbereichen und erreicht eine Flankensteilheit von nur knapp 1,4dB. Für die Realisierung der Entzerrung kommt daher ein Shelvingfilter 2.Ordnung zum Einsatz [Zölzer, 1997].

Die Übertragungsfunktion hat folgende Form:

$$H(z) = \frac{b_0 + b_1 \cdot z^{-1} + b_2 \cdot z^{-2}}{1 + a_1 \cdot z^{-1} + a_2 \cdot z^{-2}} \text{ mit den Koeffizienten}$$

$$a_1 = \frac{2 \cdot (K^2 - 1)}{1 + \sqrt{2} \cdot K + K^2}$$

$$a_2 = \frac{1 - \sqrt{2} \cdot K + K^2}{1 + \sqrt{2} \cdot K + K^2}$$

$$b_0 = \frac{1 + \sqrt{2} \cdot V_0 \cdot K + V_0 \cdot K^2}{1 + \sqrt{2} \cdot K + K^2}$$

$$b_1 = \frac{2 \cdot (V_0 \cdot K^2 - 1)}{1 + \sqrt{2} \cdot K + K^2}$$

$$b_2 = \frac{1 - \sqrt{2} \cdot V_0 \cdot K + V_0 \cdot K^2}{1 + \sqrt{2} \cdot K + K^2}, \text{ wobei}$$

$$K = \tan\left(\frac{\omega_g \cdot T}{2}\right).$$

Der lineare Verstärkungsfaktor V_0 berechnet sich aus dem logarithmischen G mit $V_0 = 10^{\frac{G}{20}}$. Gute Ergebnisse liefert ein Filter mit Verstärkung $G=3,5$ dB bei einer Grenzfrequenz von $f_g=450$ Hz.

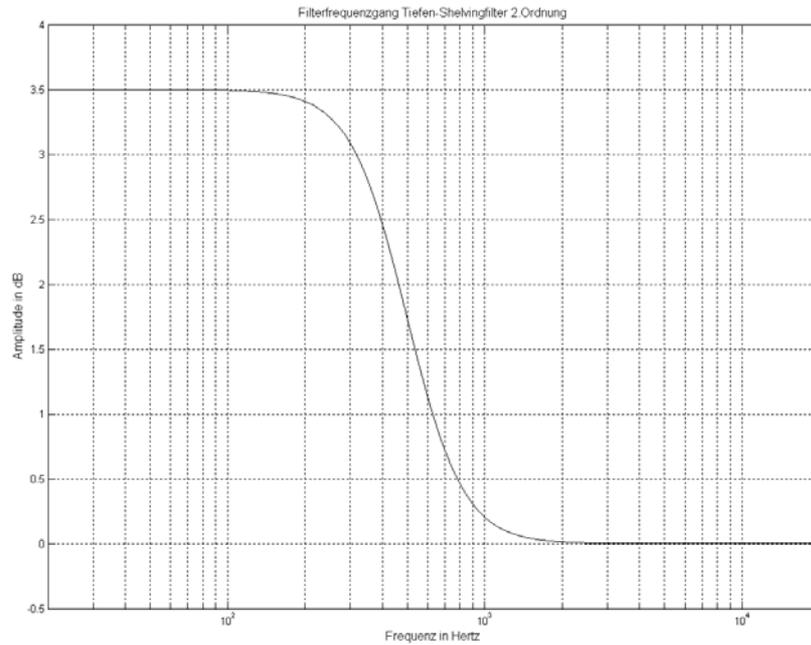


Abb. 6.24: Tiefen-Shelvingfilter mit $G=3,5\text{dB}$ und $f_g=450\text{Hz}$

Durch die Korrekturmaßnahmen erreicht man eine gute Linearität der Übertragungsfunktion für die Vorne-Richtung. Der Kurvenverlauf bleibt für den gesamten relevanten Frequenzbereich innerhalb eines $\pm 1\text{dB}$ -Schlauches.

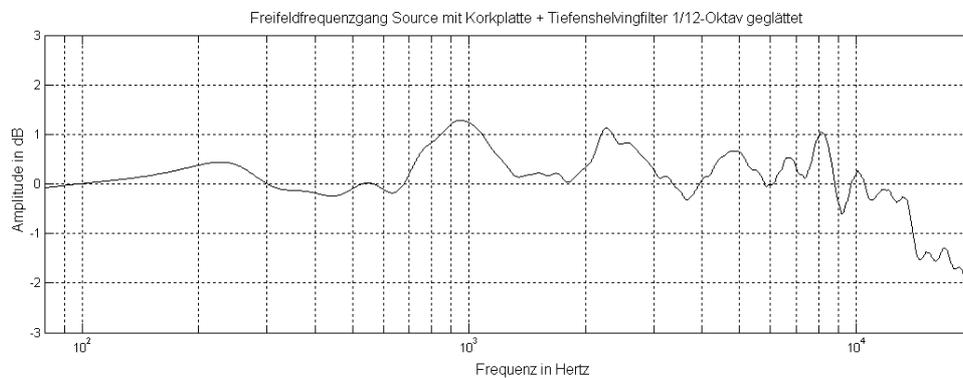


Abb. 6.25: Optimierte Freifeldübertragungsfunktion von Source

4 Beamforming

4.1 Allgemeines

Ein Beamformer ist ein räumliches Filter, das in Verbindung mit einem Sensorarray die Trennung von Signalen anhand deren unterschiedlicher Einfallrichtungen bewirkt. Die Einzelsignale werden derart verarbeitet, dass die Empfindlichkeit der Gesamtanordnung eine Richtungsabhängigkeit aufweist. In dieser Arbeit wird ein Beamformer in Verbindung mit einem Mikrofonarray verwendet, um direktionale Raumimpulsantworten zu messen. Diese sind gemeinsam mit den HRTF-Kurven des Kunstkopfes "Source" die Grundlage für eine spätere Auralisation des IEM-Cube (= großer Aufnahme/Abhörraum am Institut für elektronische Musik).

4.2 Grundlagen

Ein Mikrofonarray besteht aus N räumlich angeordneten Einzelmikrofonen. Trifft nun das Signal einer Schallquelle $s(t)$ auf dieses Array ergibt sich abhängig von der Schalleinfallrichtung für jeden Sensorpunkt ein individuell zeitverzögertes Eingangssignal $x_i(t)$, wobei $x = 1 \dots N$.

Die Schalleinfallrichtung wird durch den Azimutwinkel Φ und den Elevationswinkel Θ innerhalb eines Kugelkoordinatensystems angegeben dessen Ursprung mit dem des Arrays übereinstimmt.

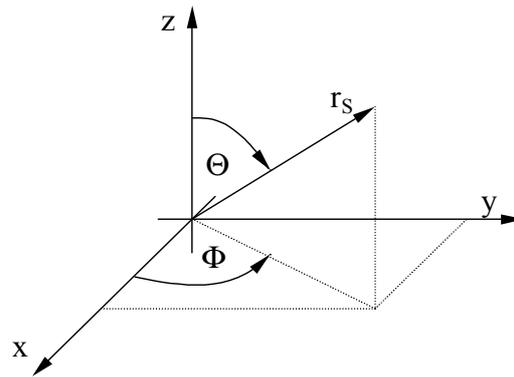


Abb. 7.1: Richtungsbestimmung im Kugelkoordinatensystem

Die Ausformung der Richtcharakteristik des Arrays erfolgt für schmalbandige Eingangssignale durch Gewichtung der Mikrofonsignale mit den komplexen Gewichtungsfaktoren w_i und anschließender Addition. Für breitbandige Signale benötigt man FIR-Filter⁷, wobei die Übertragungsfunktion des Filters in jedem Frequenzpunkt mit dem jeweiligen Schmalbandfall übereinstimmt.

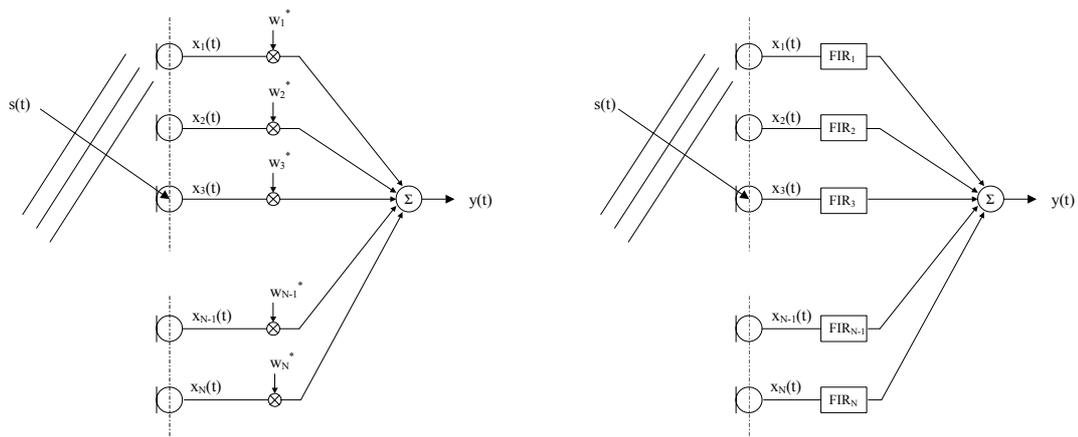


Abb. 7.2: Beamformerstruktur für schmal- und breitbandige Eingangssignale

⁷ FIR-Filter: *finite impulse response filter*, Filter mit endlicher Impulsantwort

Für alle weiteren Betrachtungen wird angenommen, dass sich das Array im Fernfeld⁸ der Quelle befindet, dadurch sind folgende Vereinfachungen möglich:

- Die Schallausbreitung erfolgt in ebenen Wellen, somit ist die Schalleinfallrichtung für alle Mikrofonpositionen gleich und vom Schallquellenabstand unabhängig. Der neue räumliche und somit auch zeitliche Bezugspunkt für die Ermittlung der Wegdifferenzen ist der Arrayursprung.
- Die Ausbreitungsdämpfung ist für alle Mikrofonpositionen gleich, daher treten an den Sensoren keine Pegeldifferenzen auf.

Für schmalbandige Quellsignale ist aufgrund dieser Fernfeldnäherung der Laufzeitunterschied zwischen den Mikrofonsignalen und dem Arrayursprung als reine Phasenverschiebung

$$\Delta\varphi = -\frac{\omega}{c} \cdot \vec{r}_i \cdot \vec{e}_S \quad (7.1)$$

darstellbar. Die Projektion des Mikrofonaufpunktvektors auf die Schalleinfallrichtung $\vec{r}_i \cdot \vec{e}_S$ (das Innenprodukt der beiden Vektoren) ergibt dabei die Wegdifferenz des Schalls zwischen dem jeweiligen Mikrofon und dem Ursprung.

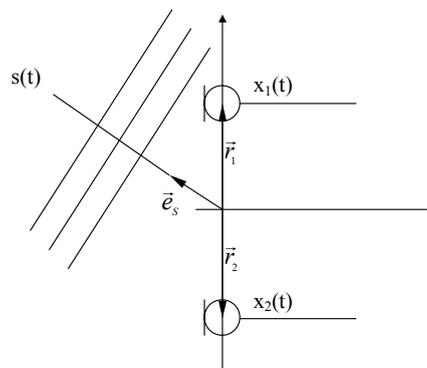


Abb. 7.3: Phasenverschiebung der Mikrofonsignale

Ein Sensorsignal lautet somit:

$$x_i(t) = s(t) \cdot \Gamma(\vec{e}_S) \cdot e^{j\frac{\omega}{c} \vec{r}_i \cdot \vec{e}_S}, \quad (7.2)$$

⁸ Ist man weit von einer Schallquelle entfernt, so nimmt der Betrag der Schallschnelle genauso wie der Schalldruck mit $1/r$ ab, und lässt sich abgesehen davon wie eine ebene Welle betrachten. Nahe an der Schallquelle nimmt die Schnelle aber nicht mit $1/r$, sondern mit $1/r^2$ ab. Man unterscheidet je nach Verhalten der Schnelle zwischen Nahfeld (Abnahme mit $1/r^2$) und Fernfeld (Abnahme mit $1/r$). Der Betrag der Schallschnelle ist sowohl von der Entfernung, als auch von der abgestrahlten Wellenlänge abhängig. Den Übergang zwischen Nahfeld und Fernfeld findet man ungefähr bei $r = \lambda/2\pi$. Der Punkt des Übergangs ist also frequenzabhängig.

wobei $\Gamma(\vec{e}_s)$ die Richtcharakteristik des Mikrofons angibt. Bei Verwendung von Mikrofonen mit Kugelcharakteristik gilt daher für alle Schalleinfallrichtungen $\Gamma(\vec{e}_s)=1$.

Für die kompakte Darstellung aller Eingangssignale bietet sich die Vektornotation wie folgt an:

$$\mathbf{x}=\mathbf{s}(t)\cdot\mathbf{d} \quad (7.3)$$

mit dem Eingangsdatenvektor

$$\mathbf{x}=[x_1(t),x_2(t),x_3(t),\dots,x_N(t)]$$

und dem Steuervektor

$$\mathbf{d}(\Theta,\Phi,\omega)=\left[e^{j\frac{\omega}{c}\vec{r}_1\cdot\vec{e}_s},e^{j\frac{\omega}{c}\vec{r}_2\cdot\vec{e}_s},e^{j\frac{\omega}{c}\vec{r}_3\cdot\vec{e}_s},\dots,e^{j\frac{\omega}{c}\vec{r}_N\cdot\vec{e}_s}\right]^T. \quad (7.4)$$

Für das Ausgangssignal des Beamformers gilt somit⁹:

$$y=\mathbf{w}^H\mathbf{x} \quad (7.5)$$

4.3 Mikrofonarray

Die Wahl der Mikrofonpositionen beeinflusst stark erzielbaren Richtcharakteristika. Insbesondere die wirksame Bandbreite lässt sich direkt mit der Arraygeometrie in Verbindung bringen.

Die obere Grenzfrequenz ist durch die kleinste Sensordistanz des Arrays bestimmt. Durch die Arraymikrofone erfolgt eine räumliche Abtastung des Schallfeldes. In Analogie zum Samplingtheorem im Zeitbereich, wo das Abtastintervall kleiner als die Hälfte der kleinsten auftretenden Periodendauer sein muss, gilt bei der räumlichen Abtastung, dass die wirksame Sensordistanz (= Wegdifferenz in Schalleinfallrichtung) kleiner sein muss als die halbe Wellenlänge des Eingangssignals.

$$\Delta x \leq \frac{c}{2 \cdot f_{\max} \cdot \sin \alpha_{\max}} \quad (7.6)$$

Δx ... Sensorabstand

c ... Schallgeschwindigkeit

f_{\max} ... größte Signalfrequenz

α_{\max} ... größter Schalleinfallswinkel

⁹ Das hochgestellte H steht für hermitesch (=konjugiert komplex und transponiert)

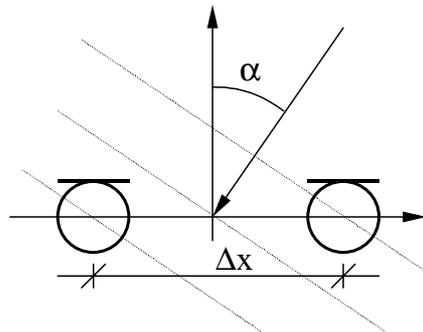


Abb. 7.4: räumliche Abtastung des Schallfeldes

Ist Beziehung (7.6) nicht erfüllt, kommt es zu räumlichem Aliasing. In der Richtcharakteristik bilden sich Nebenkeulen aus und eine eindeutige Richtungszuordnung hochfrequenter Signalanteile ist nicht mehr möglich [van der Wal, 1996].

Um bei tiefen Frequenzen noch eine ausgeprägte Richtwirkung zu erzielen, muss die Gesamtausdehnung des Arrays (Apertur) möglichst groß sein. Verteilt man die Mikrofonpositionen über diese Länge in gleichmäßigen Abständen, ergibt sich eine stark frequenzabhängige Richtcharakteristik, wobei die Hauptkeule mit zunehmender Frequenz immer enger wird. Eine gleichmäßige Keulenform erhält man, indem man die Mikrofonsignale derart frequenz- und positionsabhängig gewichtet (Filterung), dass sich die Apertur des Arrays direkt proportional zur Wellenlänge verändert. Somit werden bei tiefen Signalfrequenzen auch weiter vom Arrayursprung entfernte Mikrofone vom Beamformer miteinbezogen wohingegen bei hohen Signalfrequenzen nur jene in der Nähe des Ursprungs berücksichtigt werden. Erreicht wird dieses Verhalten durch eine räumliche Fensterung, wobei die Fensterbreite mit zunehmender Signalfrequenz immer schmaler wird. Die Filterverläufe in den einzelnen Mikrofonkanälen ergeben sich punktweise für jede Frequenz aus der jeweiligen Fensterfunktion. Um Sprünge in der Übertragungsfunktion bei Erreichen der nächsten Mikrofonposition zu vermeiden, verwendet man Fensterfunktionen mit glatten Verläufen wie z.B.: das Hanning-Fenster oder das Gauß-Fenster. Da weiter vom Ursprung entfernte Mikrofone nur für zunehmend tiefere Signalanteile verwendet werden, ist es ohne Aliasingfehler zu begehen möglich, den Mikrofonabstand zu vergrößern. Man erreicht durch diesen Zusammenhang bei gleichbleibender Mikrofonanzahl und einer logarithmischen Verteilung eine größere Arrayapertur.

Mikrofonarrays können in ein-, zwei- oder dreidimensionalen Anordnungen aufgebaut werden. Eindimensionale Linienarrays mit omnidirektionalen Mikrofonen besitzen um ihre Achse rotationssymmetrische Richtcharakteristiken, für zweidimensionale Arrays ist die Richtwirkung dementsprechend symmetrisch zur Arrayebene. Um eine eindeutige

Zuordnung des Beamformersignals zur Einfallrichtung im Raum zu erreichen, muss ein dreidimensionales Mikrofonarray verwendet werden.

4.3.1 3D-Array

Das in dieser Arbeit verwendete dreidimensionale Array entstammt einer vorhergegangenen Diplomarbeit [Krejci, 2001]. Im Zuge dieser Arbeit wurden alle Raumimpulsantworten der 24 Lautsprecher des IEM-Cube zu den insgesamt 53 Arraymikrofonen gemessen. Diese sind in sieben Linienarrays angeordnet. Drei davon liegen auf den Achsen eines kartesischen Koordinatenkreuzes, die weiteren vier in den Raumdiagonalen. Alle Mikrofonpositionen befinden sich innerhalb eines Quaders mit den Abmessungen 130x130x28cm. Der minimale Mikrofonabstand beträgt in horizontaler Richtung 8cm in vertikaler Richtung 6cm. Die obere Grenzfrequenz des Arrays liegt somit bei etwa 2,1kHz. Die Abstände vergrößern sich mit zunehmender Entfernung vom Arrayursprung. Eine Tabelle mit den genauen Koordinaten findet sich in Anhang C.

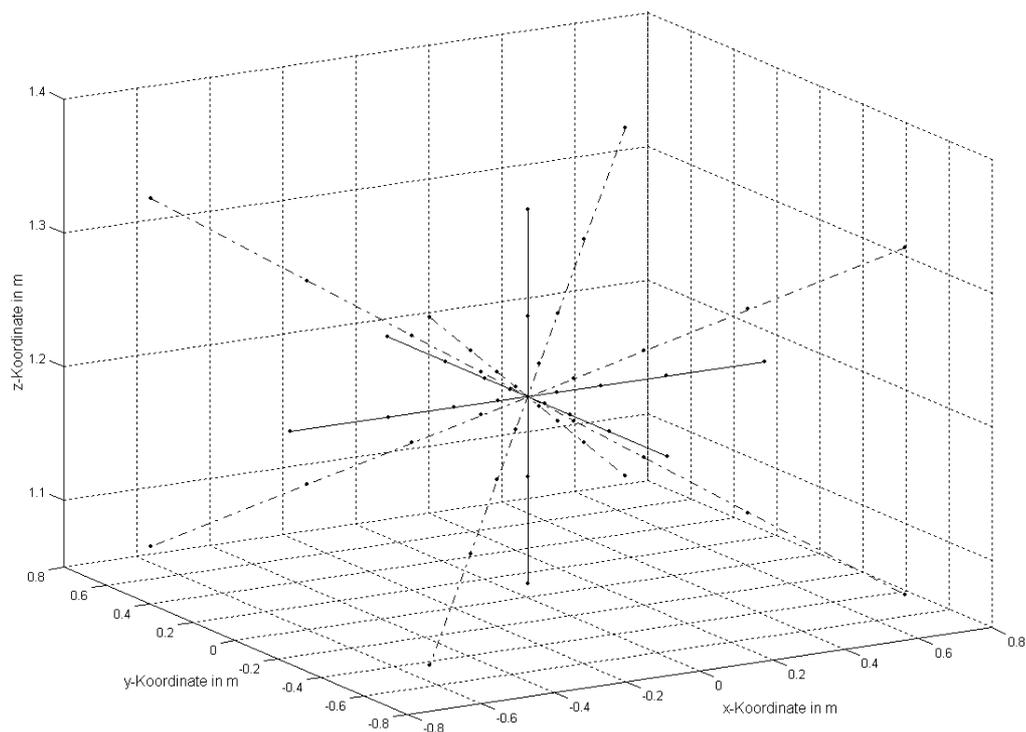


Abb. 7.5: dreidimensionales Mikrofonarray

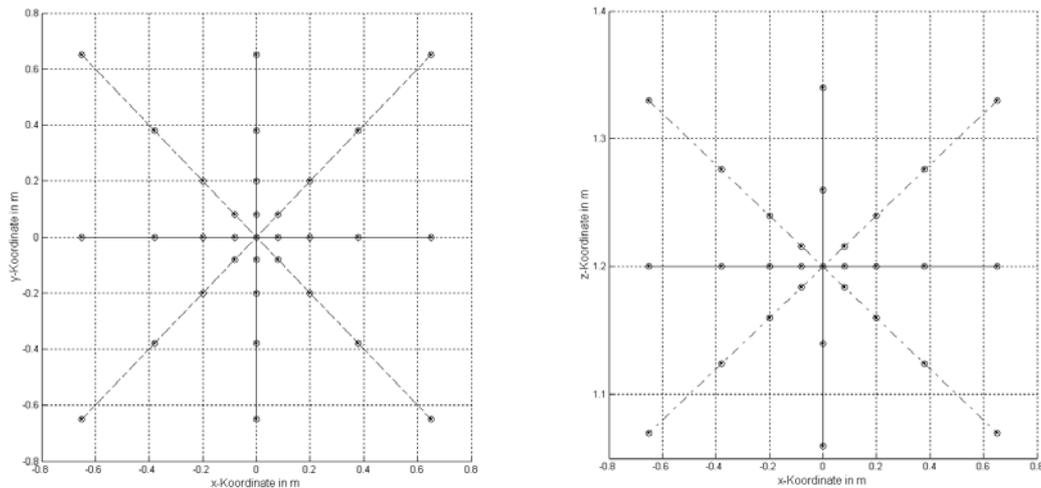


Abb. 7.6: Grund- und Aufriss des Mikrofonarrays

4.3.2 Zirkuläres 2D-Array

Aufgrund der Ergebnisse mit dem vorgegebenen Array wurde entschieden, eine sogenannte isotrope Arraygeometrie zu verwenden [Baysal, 2003], die eine gleichbleibende Richtcharakteristik für unterschiedliche Schalleinfallrichtungen bietet. Die einfachste isotrope Struktur ist ein ebenes kreisrundes Array mit gleichmäßig auf dem Umfang verteilten Mikrofonpositionen. Untersucht wird eine Variante mit 144 Sensorpunkten auf einem Kreis mit 0,5m Radius. Das ergibt einen minimalen Sensorabstand von etwa 22mm bei einer Apertur von 1m und damit eine obere Grenzfrequenz von fast 8kHz.

4.4 Delay&Sum Beamformer für 3D-Array

Die einfachste Möglichkeit zu Erzielung der Richtwirkung stellt der Delay&Sum Beamformer dar. Wie der Name schon vermuten lässt, werden die Mikrofon-signale zuerst entsprechend verzögert (Delay) und anschließend aufsummiert (Sum). Die Delayzeiten werden so gewählt, dass sich die Mikrofon-signale für eine bestimmte Schalleinfallrichtung phasengleich überlagern (konstruktive Interferenz), dadurch ergibt sich eine Hauptrichtung mit maximaler Energieaufnahme.

4.4.1 Theoretische Betrachtungen

Die Sensorgewichte sind reine Phasenterme. Für die Normierung des Ausgangssignals werden diese um einen Verstärkungsfaktor a erweitert.

$$w_i^* = a \cdot e^{-j\frac{\omega}{c}\vec{r}_i \cdot \vec{r}_S} \quad (7.7)$$

In Hauptrichtung (Φ_0, Θ_0) des Beamformers ergibt sich eine verzerrungsfreie Abbildung der Schallquelle, wenn für eine Gesamtverstärkung des Arrays von 1 gilt:

$$y(\Phi_0, \Theta_0) = \mathbf{w}^H \mathbf{x} = \mathbf{w}^H s(t) \cdot \mathbf{d}(\Phi_0, \Theta_0) = s(t) \quad (7.8)$$

daraus folgt:

$$\mathbf{w}^H \cdot \mathbf{d}(\Phi_0, \Theta_0) = a \cdot \begin{bmatrix} e^{-j\frac{\omega}{c}\vec{r}_1 \cdot \vec{r}_S} & e^{-j\frac{\omega}{c}\vec{r}_2 \cdot \vec{r}_S} & \dots & e^{-j\frac{\omega}{c}\vec{r}_N \cdot \vec{r}_S} \end{bmatrix} \begin{bmatrix} e^{j\frac{\omega}{c}\vec{r}_1 \cdot \vec{r}_S} \\ e^{j\frac{\omega}{c}\vec{r}_2 \cdot \vec{r}_S} \\ \vdots \\ e^{j\frac{\omega}{c}\vec{r}_N \cdot \vec{r}_S} \end{bmatrix} = a \cdot N = 1 \quad (7.9)$$

und daraus der Gewichtsvektor

$$\mathbf{w}_{DS} = \frac{1}{N} \mathbf{d}(\Phi_0, \Theta_0) \quad (7.10)$$

wobei N die Sensoranzahl angibt.

Der Delay&Sum Beamformer bietet maximalen Signal-Störabstand (SNR) für unkorrelierte Störungen, wie etwa das Eigenrauschen der elektronischen Komponenten. Der SNR steigt proportional mit der Anzahl der verwendeten Mikrofone. Für gerichtete und/oder korrelierte Störungen, wie sie in realen Räumen durch die Reflexionen auftreten, erreicht der Delay&Sum Beamformer deutlich schlechtere Werte.

4.4.2 Simulationsergebnisse

Als Resultat der Matlab-Simulation erhält man einen dreidimensionalen Plot der Richtcharakteristik des Beamformers. Bei der Simulation werden die Amplitudenwerte mit einem Winkelinkrement von $2,5^\circ$ in Azimut- wie auch Elevationsrichtung berechnet. Dies stellt einen guten Kompromiss zwischen Darstellungsgenauigkeit und Berechnungsaufwand dar. Die dreidimensionale Darstellungsform gibt zwar einen raschen und anschaulichen Überblick über das tendenzielle Verhalten der zu untersuchenden Anordnung. Für die Datenanalyse, insbesondere das Ablesen von Pegel und Breite der Haupt- und Nebenkeulen, sind allerdings zweidimensionale, achsenparallele Schnitte besser geeignet.

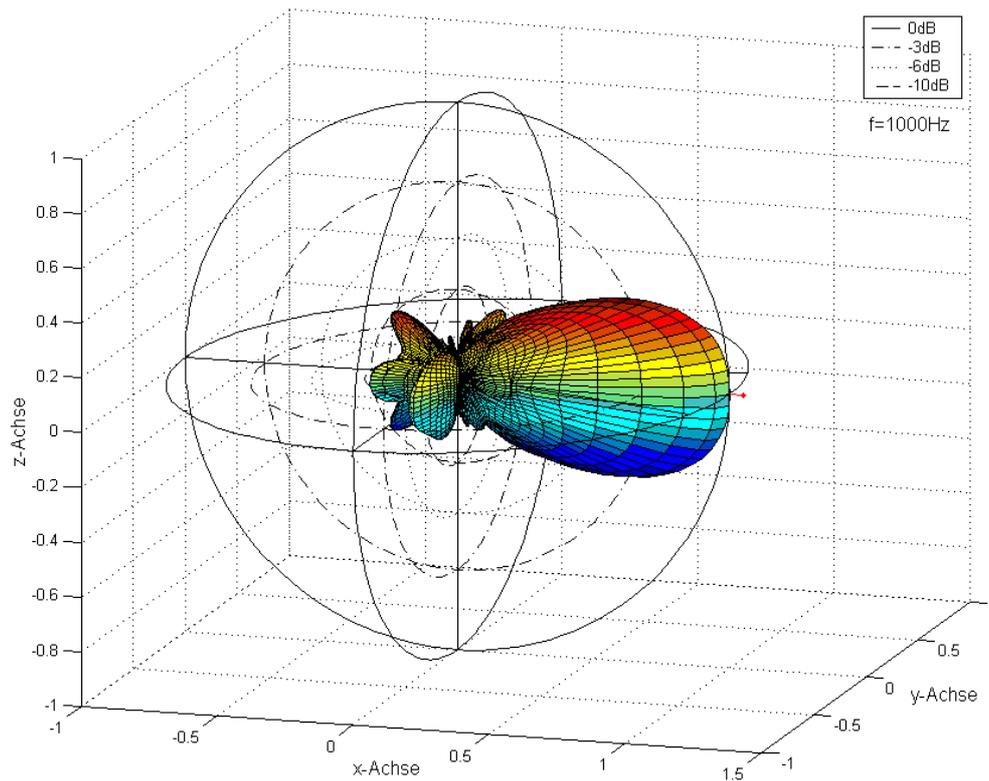
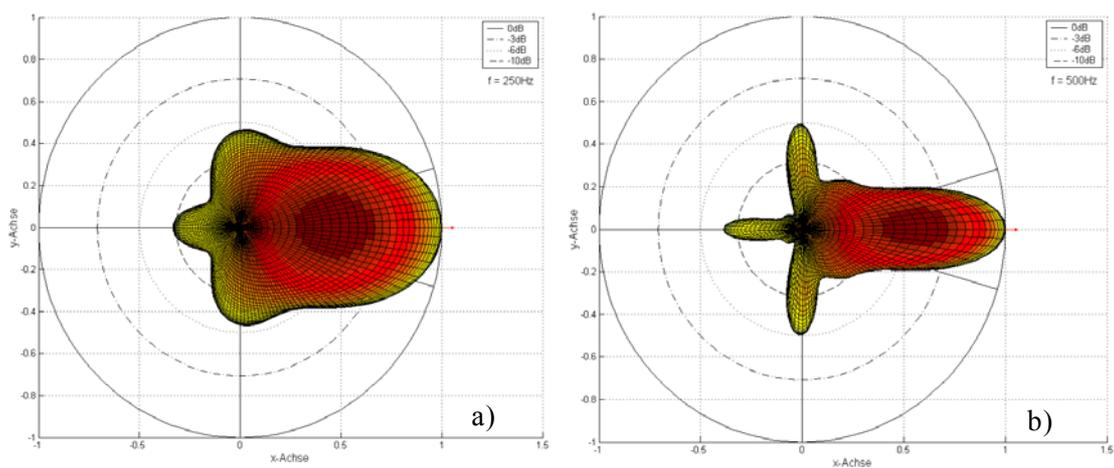


Abb. 7.7: Dreidimensionaler Beamplot, 3D-Mikrofonarray [vgl. Krejci, 2001]

Die folgenden Abbildungen stellen die Richtcharakteristiken bei Ausrichtung in der Horizontalebene (x -Achse) und für 90° Elevation (z -Achse) für das 3D-Array bei verschiedenen repräsentativen Frequenzen innerhalb der verwendbaren Bandbreite des Mikrofonarrays nach [Krejci, 2001] dar.



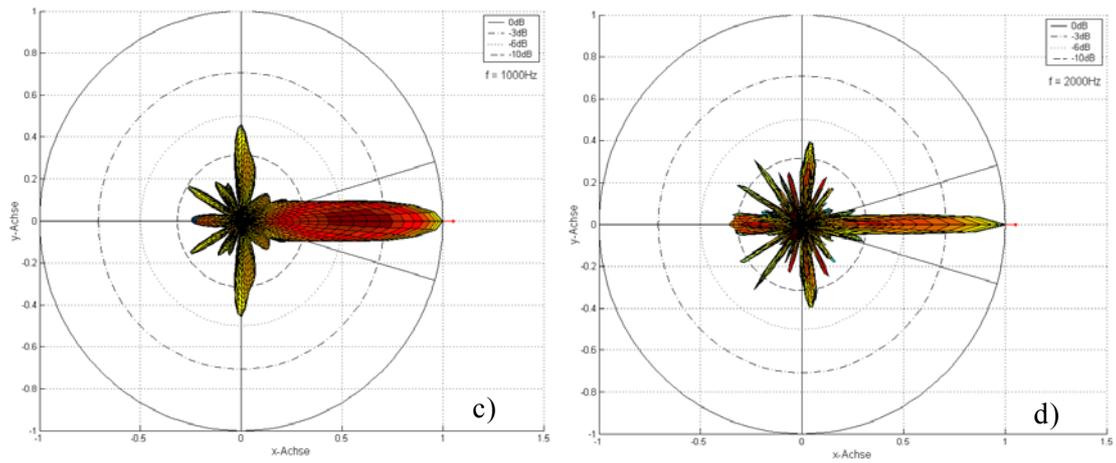


Abb. 7.8: Beamplots des Delay&Sum Beamformers bei horizontaler Haupttrichtung für a) 250Hz, b) 500Hz, c) 1kHz und d) 2kHz

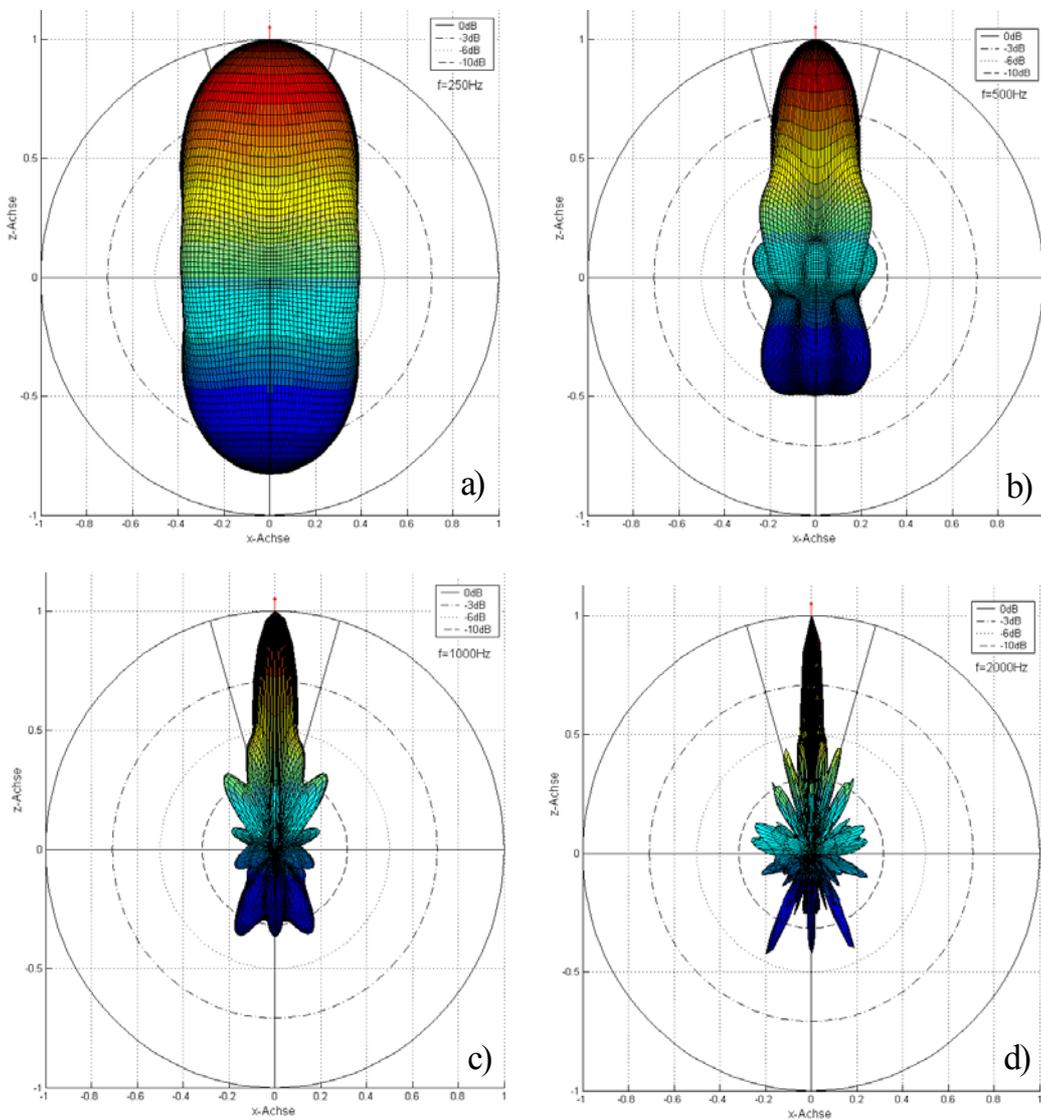


Abb. 7.9: Beamplots des Delay&Sum Beamformers bei 90° Elevation der Haupttrichtung für a) 250Hz, b) 500Hz, c) 1kHz und d) 2kHz

Die Hauptkeule wird mit steigender Frequenz immer schmaler. Damit sinkt im homogenen diffusen Schallfeld auch die aufgenommene Schallenergie. Dieses nichtlineare Übertragungsverhalten macht den Delay&Sum Beamformer für die angestrebte Messung breitbandiger direktonaler Raumimpulsantworten unbrauchbar.

Anhand der Beamplots aus Abb. 7.9 lassen sich gut die Zusammenhänge zwischen Nebenkeulendämpfung und Arraygeometrie erkennen. In vertikaler Richtung, in der das Array nur 1/5 der horizontalen Ausdehnung besitzt, sind Nebenkeulen viel stärker ausgeprägt als in der Horizontalebene. Besonders bei tiefen Frequenzen macht sich die kleine Apertur in z-Richtung negativ bemerkbar. (vgl. Abb. 7.9a mit Abb. 7.8a)

Alle Schallanteile die in Richtung der Nebenkeule eintreffen werden auf die Hauptrichtung abgebildet, das führt zu einer räumlichen Verschmierung des abgetasteten Schallfeldes. Die Nebenkeulendämpfung muss daher so groß sein, dass bei Überlagerung der richtungsbezogenen Beamformersignale bei der späteren Auralisation des Raumes eine genügend großer Pegeldifferenz zwischen einem korrekt positionierten Richtungssignal und seiner falsch positionierten Kopien bestehen bleibt. Dieser Wert wird in der Literatur in Anlehnung an den Signal-Rauschabstand oft als Signal-Interferenzabstand bezeichnet. Folgende Abbildung zeigt die Lokalisation der Phantomschallquelle im Schallfeld zweier Lautsprecher in Abhängigkeit von deren Pegeldifferenz.

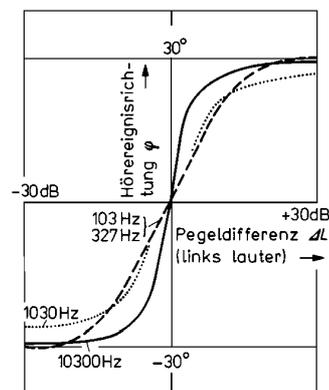


Abb. 7.10: Höreignisrichtung bei Lautsprecheraufstellung unter $\pm 30^\circ$ um die Medianebene für kohärente terzbreite Tonimpulse vgl. [Blauert, 1974]

Man erkennt, dass zumindest ein Pegelunterschied von 10dB erreicht werden sollte, um die Fehler in der Richtungsabbildung in akzeptablen Grenzen zu halten. Umgelegt auf den Beamplot einer einzelnen Richtung sind eher Werte $>15\text{dB}$ für die Nebenkeulendämpfung anzustreben.

Der Delay&Sum Beamformer ist sehr robust gegen Ungenauigkeiten in der Mikrofonpositionierung. Bei der Vermessung des 3D-Arrays wurden die Mikrofone händisch eingerichtet, daher muss davon ausgegangen werden, dass die realen von den idealen Messpositionen abweichen. In der Simulation wurde ein von den idealen Positionen abgeleiteter Steuervektor auf ein Array angewendet, dessen Mikrofonkoordinaten für jede Raumrichtung zufällig verteilt im Bereich von $\pm 1\text{cm}$ um die Sollwerte schwanken. An den Beampatterns in Abb. 7.8 und Abb. 7.9 ist erkennbar, dass diese Abweichungen zwar leichte Unsymmetrien hervorrufen, aber die Stabilität des Delay&Sum Beamformers nicht negativ beeinträchtigen.

4.5 Delay&Sum Beamformer für 3D-Array mit räumlicher Fensterung der Arraygewichte

4.5.1 Theoretische Betrachtungen

Im Hinblick auf die weitere Verwendbarkeit der direktionalen Raumimpulsantworten in einem Ambisonic-System 5. Ordnung sollte die Hauptkeule möglichst frequenzunabhängig einen Winkelbereich von $\frac{360^\circ}{2 \cdot \text{Ambisonicordnung} + 1} = 32.7^\circ$ (Ambisonic-Winkel) abdecken.

Dieser ist durch die beiden Linien links und rechts der Hauptrichtung in Abb. 7.8 und Abb. 7.9 jeweils markiert. Um bei der Überlagerung der Impulsantworten einen linearen Leistungsverlauf in den Überlappungsbereichen zu gewährleisten, muss die Amplitude der Hauptkeule im oben definierten Ambisonic-Winkel gegenüber der Hauptrichtung auf -3dB abgesunken sein.

4.5.2 Simulationsergebnisse

Um der Verengung der Hauptkeule entgegenzuwirken wurde ein Gaußfenster implementiert das die Arrayapertur frequenzabhängig verkleinert. Dabei werden die Mikrofonsignale der weiter vom Ursprung entfernten Sensoren stärker bedämpft als die von nähergelegenen, wobei die Fensterbreite mit steigender Frequenz abnimmt (vgl. Abschnitt 4.3).

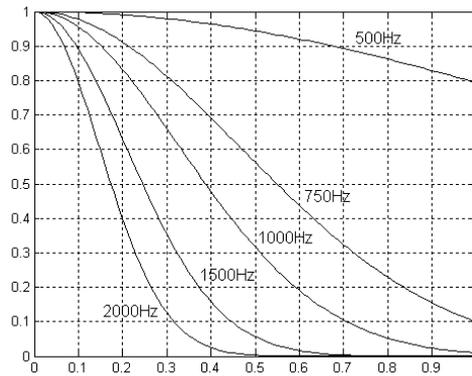


Abb. 7.11: Fensterfunktionen für verschiedene Signalfrequenzen.

Die x-Achse in Abb. 7.11 ist normiert auf den Abstand des am weitesten vom Ursprung entfernten Sensorpunktes.

Für die Darstellung der Beamplots wurden die gleichen Parameter verwendet, wie in Abb. 7.9.

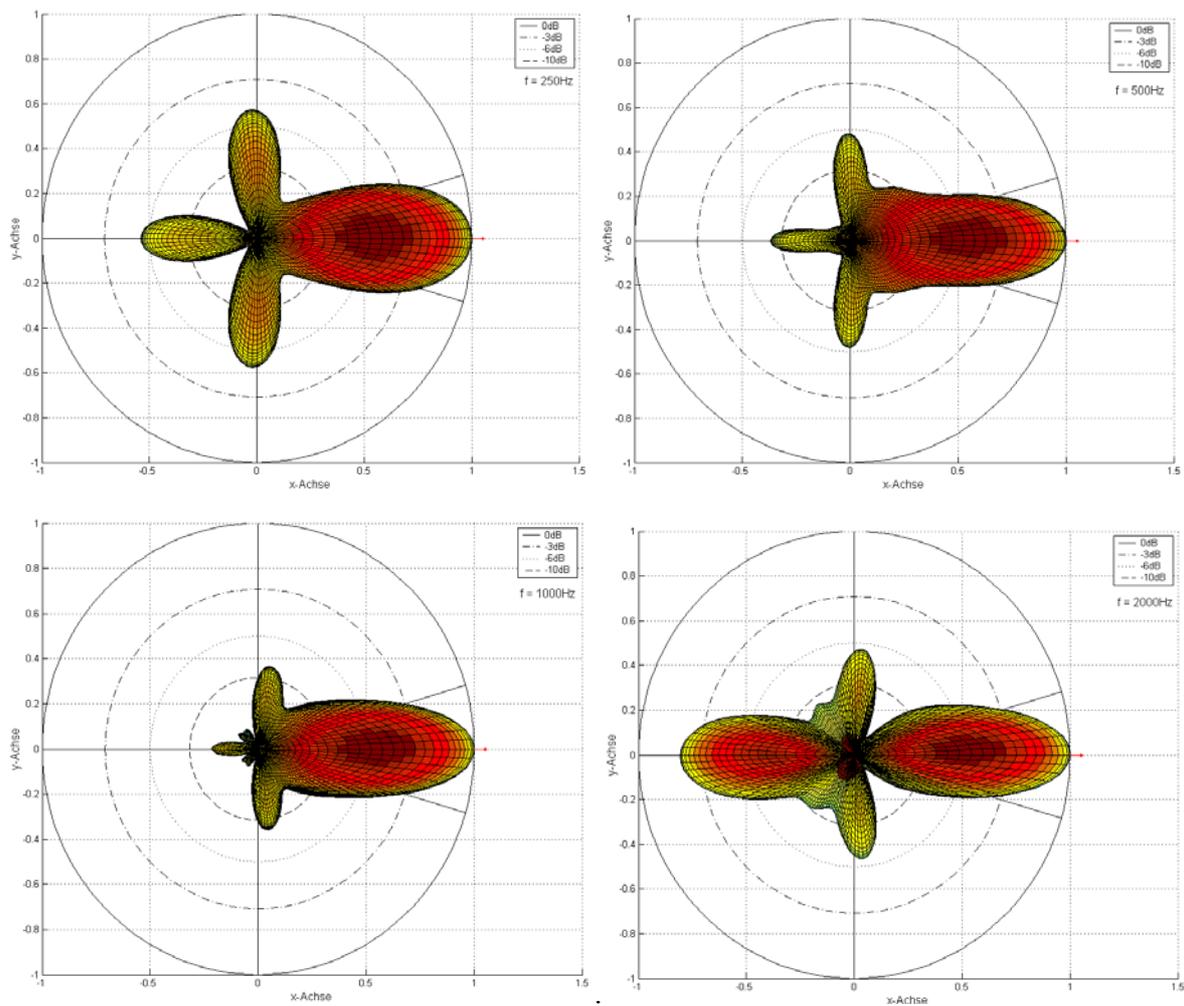


Abb. 7.12: Beamplots des Delay&Sum Beamformers mit Gaußgewichtung bei horizontaler Hauptrichtung für 250Hz, 500Hz, 1kHz und 2kHz.

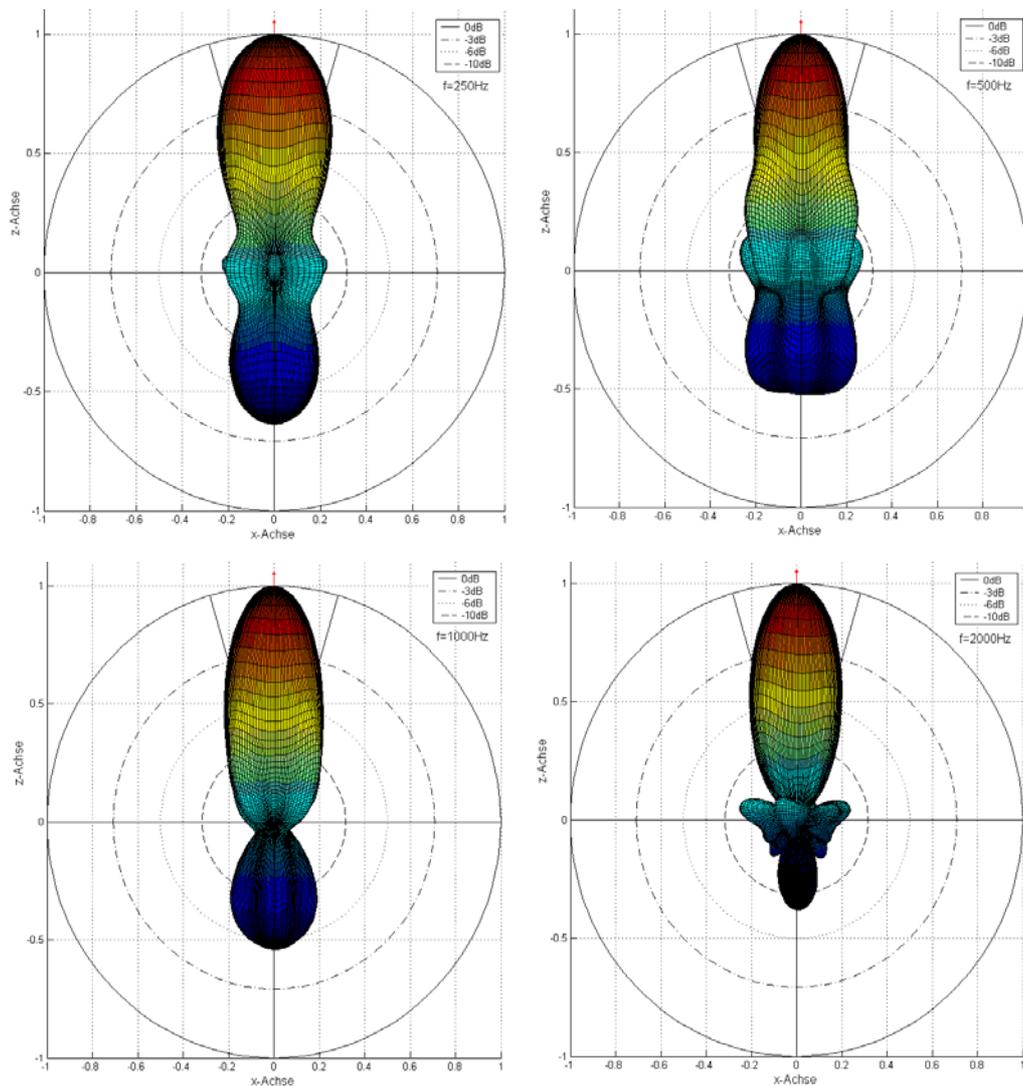


Abb. 7.13: Beamplots des Delay&Sum Beamformers mit Gaußgewichtung bei 90° Elevation der Hauptrichtung für 250Hz, 500Hz, 1kHz und 2kHz.

Die Hauptkeulenbreite bleibt wie gefordert im betrachteten Frequenzbereich annähernd konstant. Die Nebenkeulendämpfung erreicht allerdings auch nicht die Vorgabe von -15dB . Jenseits der oberen Grenzfrequenz ist eine Kontrolle der Nebenkeulen durch die auftretenden Mehrdeutigkeiten nicht mehr möglich.

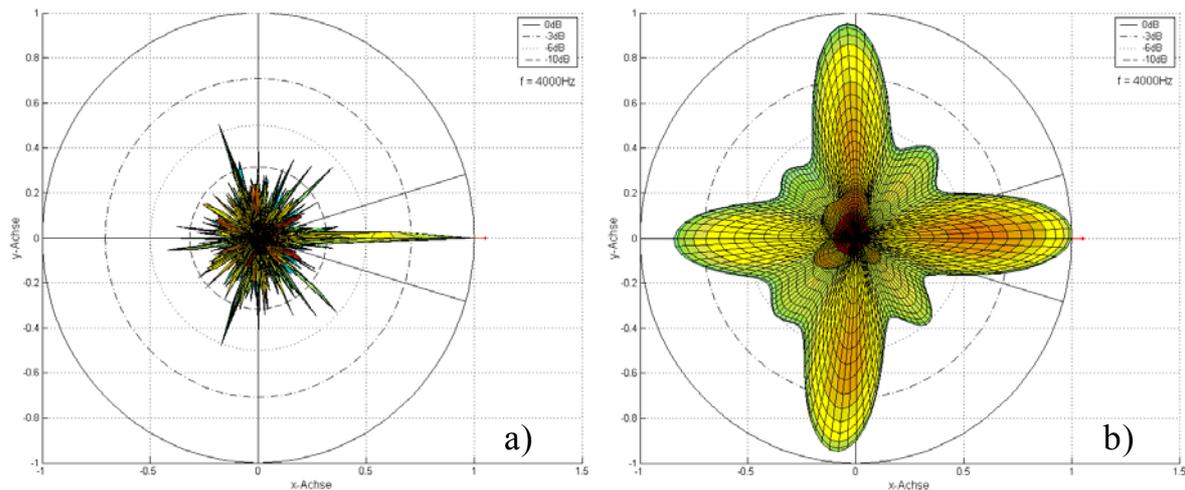


Abb. 7.14: Beamplots zufolge räumlichen Aliasings des Delay&Sum Beamformers ohne a) und mit b) Gaußgewichtung bei horizontaler Hauptrichtung für 4kHz

4.6 Minimum Varianz Beamformer

Der Minimum Varianz Beamformer optimiert den Signal-Rauschabstand des Arrays indem er die Gesamtleistung unter der Nebenbedingung einer unverzerrten Wiedergabe aus der Wunschrichtung minimiert.

4.6.1 Theoretische Betrachtungen

Das Ausgangssignal des Beamformers ist laut Gleichung (7.5)

$$y = \mathbf{w}^H \mathbf{x}$$

Die Varianz bzw. der Erwartungswert für die Ausgangsleistung des Beamformers berechnet sich zu [Van Veen, 1988]

$$E\{|y|^2\} = E\{|\mathbf{w}^H \mathbf{x}|^2\} = E\{(\mathbf{w}^H \mathbf{x})(\mathbf{w}^H \mathbf{x})^H\} = \mathbf{w}^H E\{\mathbf{x}\mathbf{x}^H\} \mathbf{w} = \mathbf{w}^H \mathbf{R}_{xx} \mathbf{w} \quad (7.11)$$

Unter der Voraussetzung quasistationärer Eingangsdaten ist die Korrelationsmatrix \mathbf{R}_{xx} zeitinvariant. In der praktischen Anwendung müsste das Eingangssignal jeweils zeitlich gefenstert und die entsprechenden Korrelationsmatrizen bestimmt werden. Der Entwurf des Beamformers wird anhand einer definierten Korrelationsmatrix hergeleitet.

Die Übertragungsfunktion des Beamformers für einen bestimmten Schalleinfallswinkel lautet

$$\mathbf{w}^H \mathbf{d}(\Phi, \Theta) = g \quad (7.12)$$

Man kann nun eine Vorgabe für die gewählte Richtung treffen, indem man den Faktor g festlegt, und die Optimierung unter der Nebenbedingung durchführt, dass Gleichung (7.13) erfüllt bleibt.

$$\min_{\mathbf{w}} \{ \mathbf{w}^H \mathbf{R}_{xx} \mathbf{w} \}, \quad \mathbf{d}(\Phi_0, \Theta_0)^H \mathbf{w} = g \quad (7.13)$$

Mit Hilfe der Lagrange'schen Multiplikatoren ist die Lösung von Gleichung (7.13) möglich.

$$\mathbf{w}_{opt} = g \cdot \frac{\mathbf{R}_{xx}^{-1} \mathbf{d}(\Phi_0, \Theta_0)}{\mathbf{d}^H(\Phi_0, \Theta_0) \mathbf{R}_{xx}^{-1} \mathbf{d}(\Phi_0, \Theta_0)} \quad (7.14)$$

Um eine unverzerrte Wiedergabe für die Hauptrichtung (Φ_0, Θ_0) zu erreichen, setzt man für $g = 1$ ein. Die Anzahl der Nebenbedingungen L ist nicht auf eine beschränkt, sondern kann abhängig von der Sensoranzahl N des Arrays auf mehrere ausgeweitet werden. Jede Nebenbedingung belegt einen der N Freiheitsgrade im Gewichtsvektor \mathbf{w} , der dann für die Minimierung der Varianz nicht mehr zur Verfügung steht. Mit den zusätzlichen Vorgaben lassen sich z.B.: Nullstellen in der Richtung bekannter Störungen erzeugen oder die Ausbildung der Hauptkeule durch Stützstellen in die gewünschte Form zwingen.

Man fasst L linear unabhängige Vektoren \mathbf{d}_i für die Vorgabenrichtungen zu einer Matrix $\mathbf{C} = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_L]$ zusammen und erhält die verallgemeinerte Form von (7.13)

$$\min_{\mathbf{w}} \{ \mathbf{w}^H \mathbf{R}_{xx} \mathbf{w} \}, \quad \mathbf{C}^H \mathbf{w} = \mathbf{g} \quad (7.15)$$

Der Vektor \mathbf{g} enthält die Vorgabenwerte (Nebenbedingungen) für jede Spalte von \mathbf{C} , wobei die Hauptrichtung inkludiert ist. Der optimale Gewichtsvektor lautet:

$$\mathbf{w}_{opt} = \mathbf{R}_{xx}^{-1} \mathbf{C} (\mathbf{C}^H \mathbf{R}_{xx}^{-1} \mathbf{C})^{-1} \mathbf{g} \quad (7.16)$$

Im Allgemeinen ist die Korrelationsmatrix \mathbf{R}_{xx} nicht bekannt. Ein Lösungsansatz ergibt sich durch Aufspaltung der Korrelationsmatrix in Wunsch- und Störkomponente [Cox, 1987].

$$\mathbf{R}_{xx} = \sigma_s^2 \mathbf{d}(\Phi_0, \Theta_0) \mathbf{d}(\Phi_0, \Theta_0)^H + \sigma_n^2 \mathbf{Q} \quad (7.17)$$

σ_s ... Signalleistung

σ_n ... Störleistung

\mathbf{Q} ... Kreuzkorrelationsmatrix des Störanteils

Die Optimierung unter Einbeziehung der Nebenbedingung $\mathbf{d}(\Phi_0, \Theta_0)^H \mathbf{w} = 1$ ergibt

$$\min_{\mathbf{w}} \{ \mathbf{w}^H \mathbf{R}_{xx} \mathbf{w} \} = \sigma_s^2 + \sigma_n^2 \min_{\mathbf{w}} \{ \mathbf{w}^H \mathbf{Q} \mathbf{w} \} \quad (7.18)$$

und daraus die optimalen Gewichte

$$\mathbf{w}_{opt} = \frac{\mathbf{Q}^{-1} \mathbf{d}(\Phi_0, \Theta_0)}{\mathbf{d}(\Phi_0, \Theta_0)^H \mathbf{Q}^{-1} \mathbf{d}(\Phi_0, \Theta_0)} \quad (7.19)$$

Für sphärisch isotrop verteilte Störungen und unter der Voraussetzung, dass nur Mikrofone mit Kugelcharakteristik zum Einsatz kommen, lässt sich \mathbf{Q} analytisch berechnen.

$$[\mathbf{Q}]_{i,k} = \frac{1}{4\pi} \int_{\Omega} e^{j\frac{\omega}{c}(\vec{r}_k - \vec{r}_i) \cdot \vec{e}_{\Omega}} d\Omega = \frac{\sin\left(\frac{\omega}{c} r_{ik}\right)}{\frac{\omega}{c} r_{ik}} \quad (7.20)$$

\vec{r}_i, \vec{r}_k ... Ortsvektoren der Mikrofonpositionen

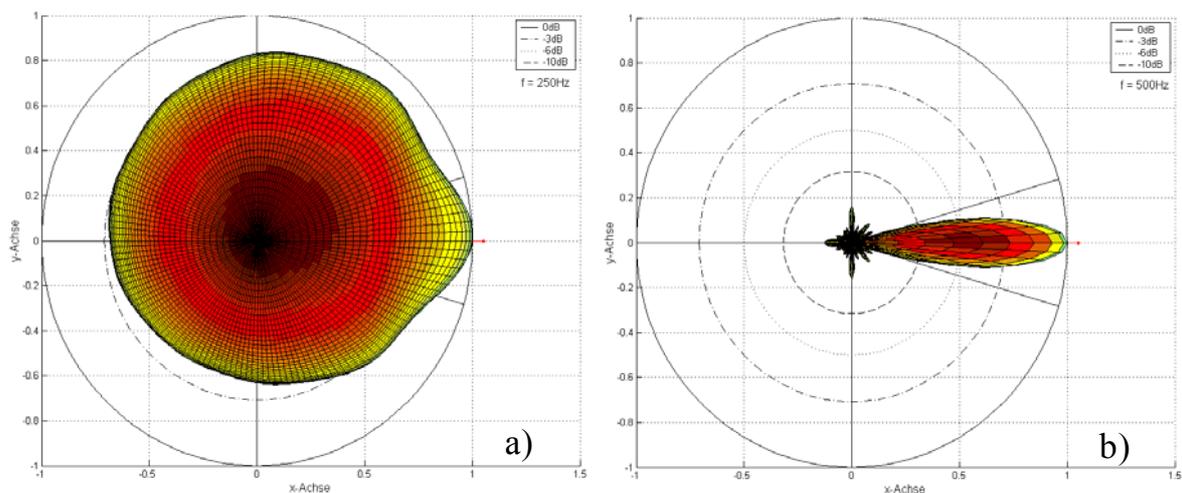
r_{ik} ... Abstand zwischen zwei Mikrofonpositionen

\vec{r}_{Ω} ... Richtungsvektor zum Flächenelement $d\Omega$

4.6.2 Simulationsergebnisse

Unter idealen Bedingungen, wie sie in der Simulation vorherrschen liefert der Minimum Varianz Beamformer deutlich bessere Ergebnisse als ein Delay&Sum Beamformer. Aufgrund der hohen Empfindlichkeit gegenüber Abweichungen von diesem Idealzustand ist eine praktische Umsetzung des Minimum Varianz Beamformers in dieser Form nicht durchführbar. Die Ergebnisse sind daher eher theoretischer Natur, bringen aber trotzdem nützliche Erkenntnisse für einen späteren robusten Entwurf.

Verwendet man keine weiteren Nebenbedingungen, als die minimal notwendige für die Hauptrichtung, zeigt der Minimum Varianz Beamformer ein ähnlich frequenzabhängiges Verhalten wie schon zuvor der Delay&Sum Beamformer. Auch die Regeln betreffend die obere Grenzfrequenz sowie der Zusammenhang zwischen Arrayapertur und erreichbarer Richtwirkung bei tiefen Frequenzen bleiben erhalten.



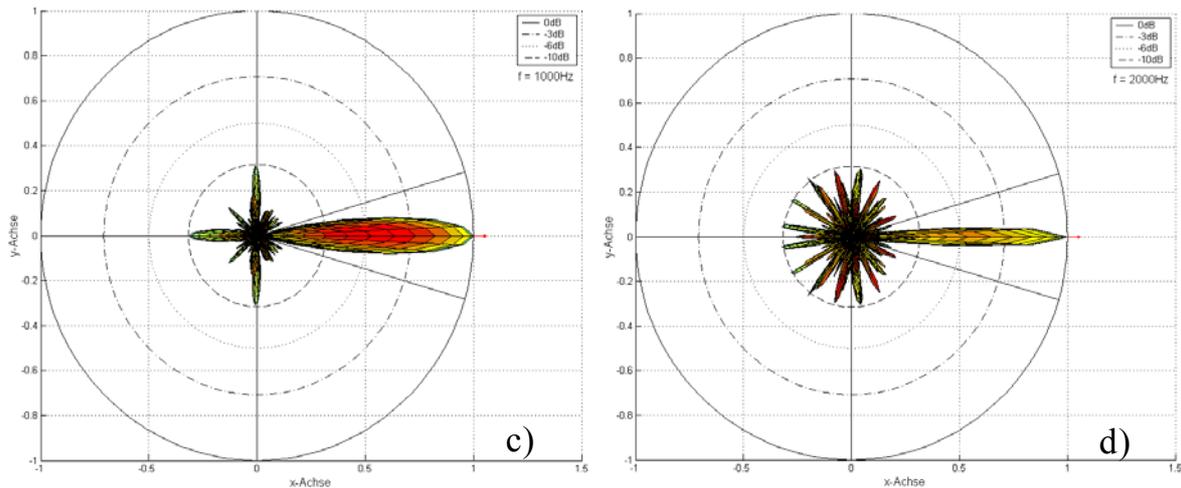


Abb. 7.15: Beamplots des Minimum Varianz Beamformers ohne weitere Nebenbedingungen für a) 250, b) 500, c) 1000 und d) 2000 Hz.

Die unsymmetrische Ausformung der Richtcharakteristik in Abb. 7.15.a) entsteht durch Rundungsfehler bei der Matrixinversion von \mathbf{Q} . Für tiefe Frequenzen nehmen dicht beieinander liegende Mikrofone nahezu die gleiche Schallinformation auf. Das führt zu einer fast singulären Matrix \mathbf{Q} , deren Inversion an die Grenzen der Zahlenauflösung der Simulationssoftware stößt. Für den Frequenzbereich unter 450Hz (Mindestmikrofonabstand ohne räumliches Aliasing $\Delta x_{\min} \ll \sim 38\text{cm}$) wurde die Sensoranzahl von ursprünglich 53 auf 27 Mikrofone reduziert, indem jede zweite Mikrofonposition vom Ursprung aus gezählt weggelassen wurde. Die Matrixinversion ist dadurch ohne Fehler durchführbar. Das "Umschalten" der Arraygeometrie bei 450Hz verursacht eine Sprungstelle im Verlauf der Sensorgewichte und somit auch in der Richtcharakteristik über der Frequenz. Diese scharfe Kante kann bei der Approximation der FIR-Filter in den Mikrofonkanälen zu Problemen führen.

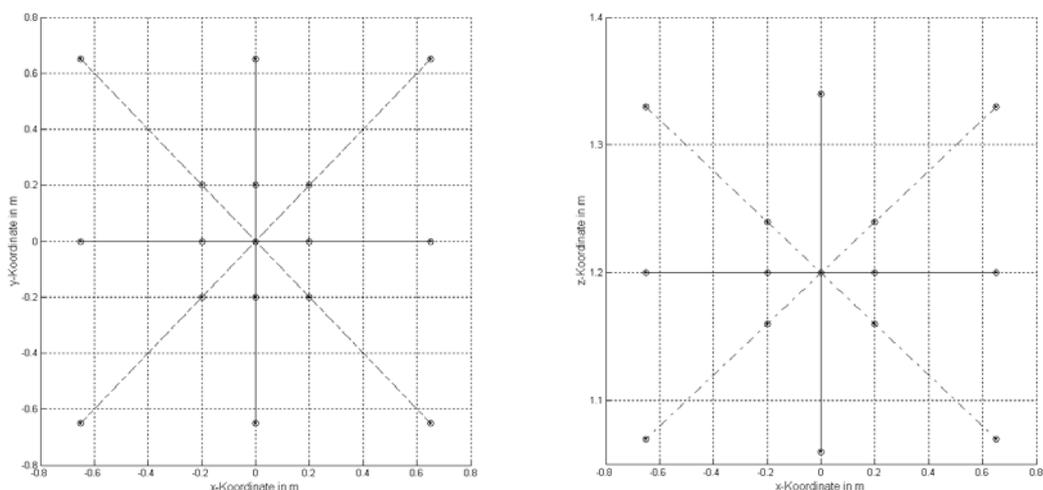


Abb. 7.16: Array mit verringerter Sensorzahl für tiefe Frequenzen (vergleiche dazu Abb. 7.6)

Um die Breite der Hauptkeule konstant zu halten, werden Stützstellen durch weitere Nebenbedingungen festgelegt. Die Amplitudenwerte für die Vorgaben besitzen einen kosinusförmigen Verlauf, wobei die Nulldurchgänge ($\pm\pi/2$) mit den angrenzenden Hauptkeulen zusammenfallen. Der Leistungsverlauf bleibt daher bei der Überlagerung der richtungsbezogenen Beamformersignale gemäß

$$\cos^2(\alpha) + \cos^2\left(\alpha - \frac{\pi}{2}\right) = \cos^2(\alpha) + \sin^2(\alpha) = 1 \quad (7.21)$$

im gesamten Überlappungsbereich konstant.

Für die folgenden Abbildungen wurden vier zusätzliche Stützstellen verwendet, zwei für die -3dB -Punkte und zwei nochmals in deren Winkelhalbierenden mit der Hauptrichtung.

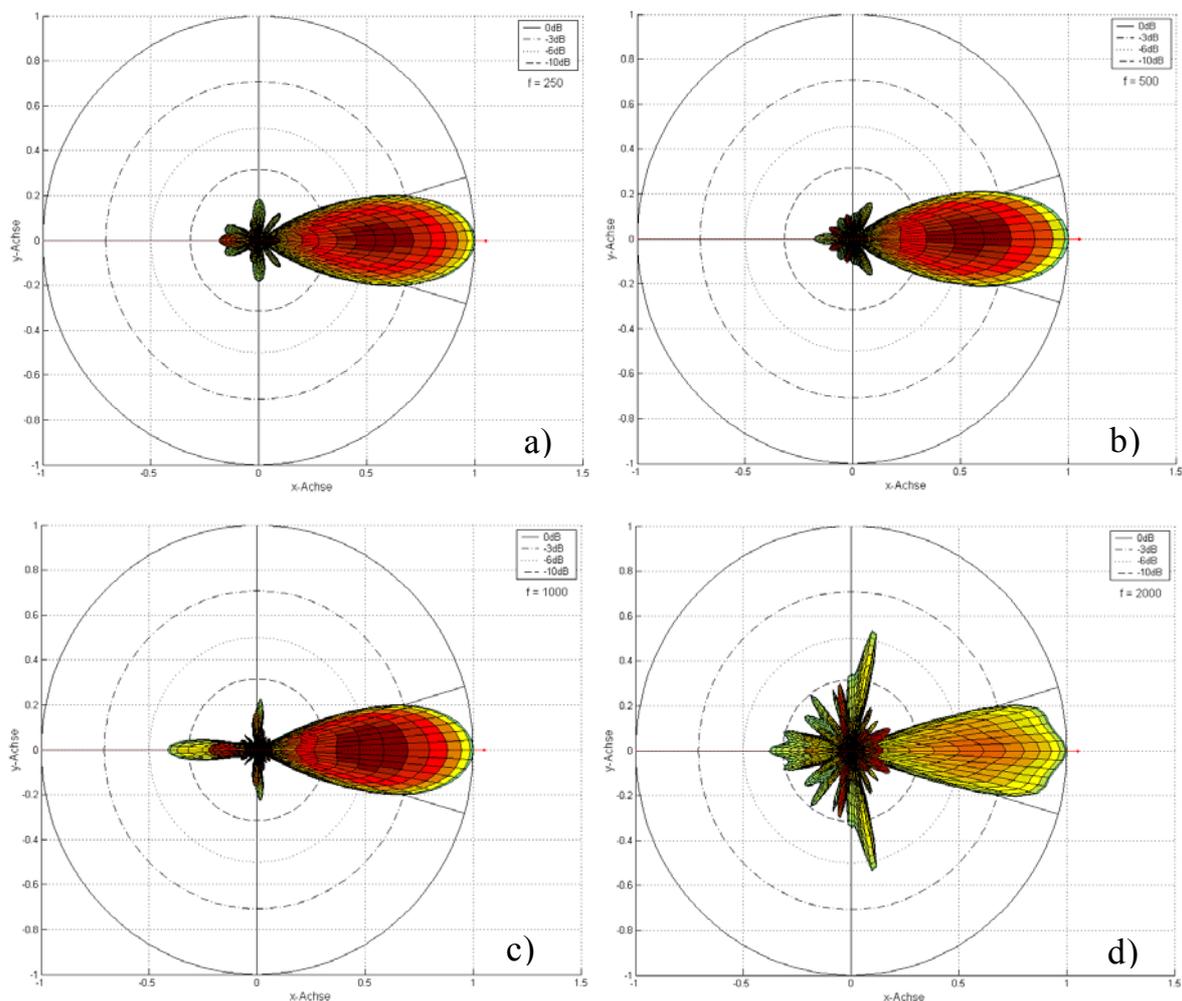


Abb. 7.17: Minimum Varianz Beamformer mit 4 Stützstellen in der Hauptkeule für a) 250Hz, b) 500Hz, c) 1kHz und d) 2kHz.

Bei 2kHz erkennt man bereits, dass die Hauptkeule nicht mehr gleichmäßig rund verläuft sondern zwischen den Stützstellen leicht einbricht. Für hohe Frequenzen ist daher eine größere Anzahl von Vorgaben erforderlich, um die gewünschte Form beizubehalten.

Der Minimum Varianz Beamformer ist sehr empfindlich gegen Ungenauigkeiten in der Mikrofonpositionierung. Durch die große Dynamik der Sensorgewichte von über 50dB haben bereits Abweichungen im Millimeterbereich eine völlig unbrauchbare Richtcharakteristik zur Folge. Folgende Abbildungen zeigen die Beträge der komplexen Gewichtungsfaktoren für 1kHz sowie einen damit berechneten Beamplot bei zufälligen Sensorpositionierfehlern von ± 1 mm. Die Hauptrichtung war auf einen Azimutwinkel von 0° eingestellt, also entlang der x-Achse (vgl. dazu Abb. 7.17c).

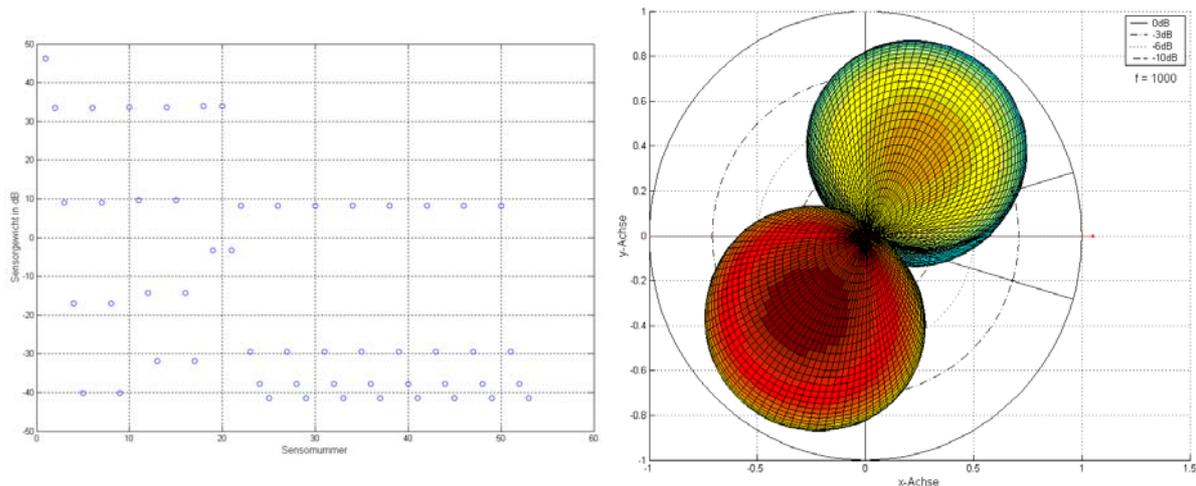


Abb. 7.18: Sensorgewichte für den Minimum Varianz Beamformer bei 1kHz (links) und Beamplot für Positionierfehler von ± 1 mm (rechts).

Erwähnenswert ist, dass eine Arraygeometrie, bei der die Sensorpositionen nicht symmetrisch angeordnet sind bessere Ergebnisse erzielt als eine exakt symmetrische Variante (Verhinderung von Eigenmoden). Dazu wurden alle Sensorpositionen zufällig um ± 1 cm variiert und die Simulation mit diesem adaptierten Array nochmals durchgeführt. Für weite Frequenzbereiche ergibt sich eine Verbesserung der Nebenkeulendämpfung von mehr als 5dB. Die Arraygewichte bleiben auch in einem um ca. 20dB kleineren Wertebereich. Dieses Verhalten gilt nicht nur für die 0° -Richtung, die ja auf einer Arraysymmetrieebene liegt, sondern ist auch für alle anderen Schalleinfallswinkel gegeben.

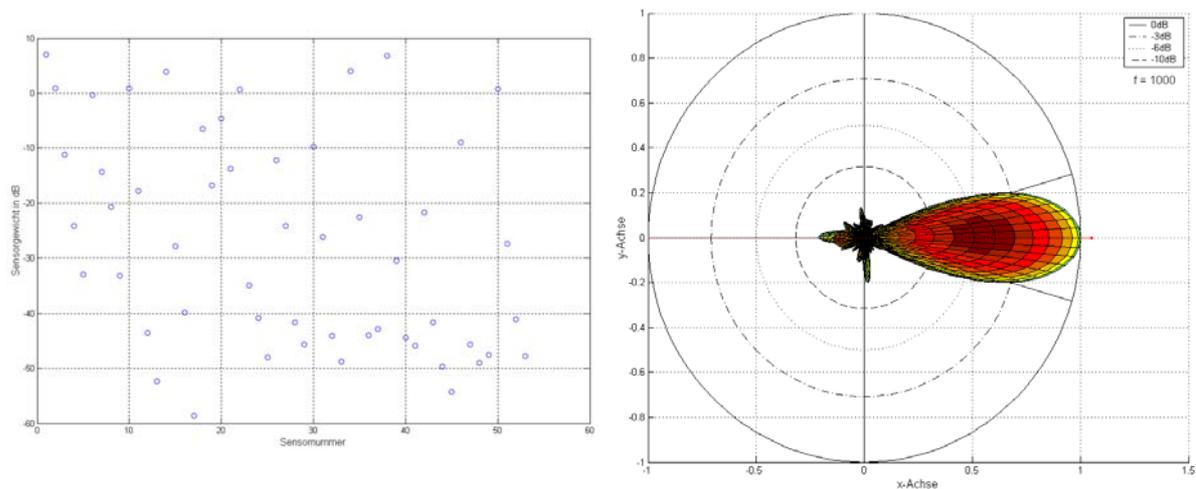


Abb. 7.19: Sensorgewichte und Beamplot bei leichten Abweichungen in der Sensorsymmetrie (vergleiche dazu Abb. 7.17c)

4.7 Delay&Sum Beamformer für zirkuläres Array mit räumlicher Fensterung der Arraygewichte

Die Verwendung des 3D-Arrays führt zu Ergebnissen, welche die gestellten Anforderungen nicht erfüllen (Delay&Sum Beamformer) bzw. für eine praktische Umsetzung zu instabil sind (Minimum Varianz Beamformer). Um eine realisierbare Lösung zu erhalten, wurde für alle weiteren Untersuchungen das in 4.3.2 beschriebene zirkuläre Array verwendet.

4.7.1 Theoretische Betrachtungen

Wie zu erwarten war, ist die Richtcharakteristik des reinen Delay&Sum Beamformers stark frequenzabhängig. Eine Fensterfunktion in Abhängigkeit von der Entfernung des Mikrofons zum Ursprung wie sie in Abschnitt 4.5 verwendet wird, ist durch die punktsymmetrische Anordnung der Mikrofone allerdings nicht möglich. Stattdessen erreicht man eine frequenzunabhängige Hauptkeulenbreite durch eine zukzessive Einschränkung der wirksamen Sensoranzahl mit steigender Signalfrequenz.

4.7.2 Simulationsergebnisse

Als Gewichtsfunktion kommt wiederum ein Gaussfenster zum Einsatz, dessen Breite mit steigender Frequenz abnimmt. Man kann sich dieses Gaussfenster so über dem Kreisarray aufgetragen vorstellen, dass die Fenstersymmetrieachse über jenem Sensor liegt, der in die Hauptrichtung des Beamformers weist. Entfernt man sich nun entlang dem Kreisumfang von

der Hauptrichtung, folgen die Gewichtungsfaktoren für die "darunterliegenden" Sensoren dem Fensterverlauf und werden dementsprechend immer kleiner.

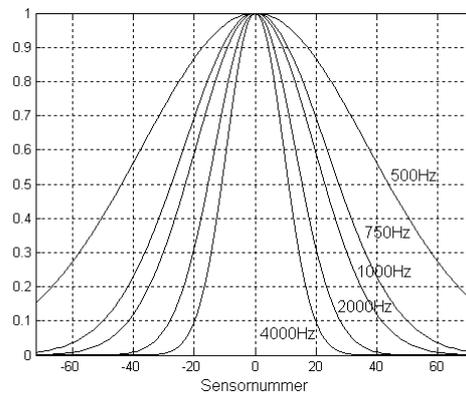
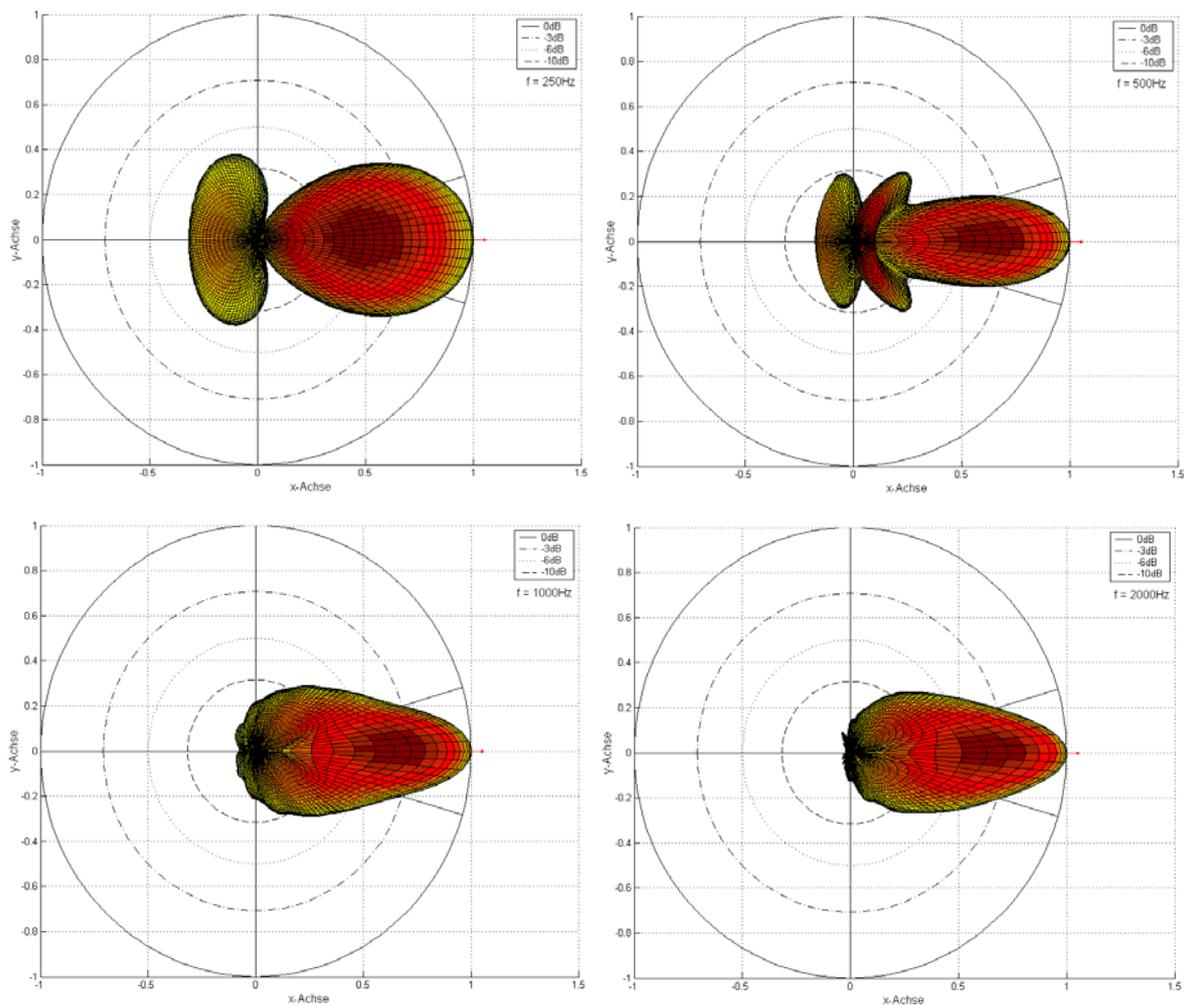


Abb. 7.20: Fensterfunktionen für verschiedene Signalfrequenzen

Diese Methode liefert brauchbare Ergebnisse im Bereich von 500Hz-7kHz.



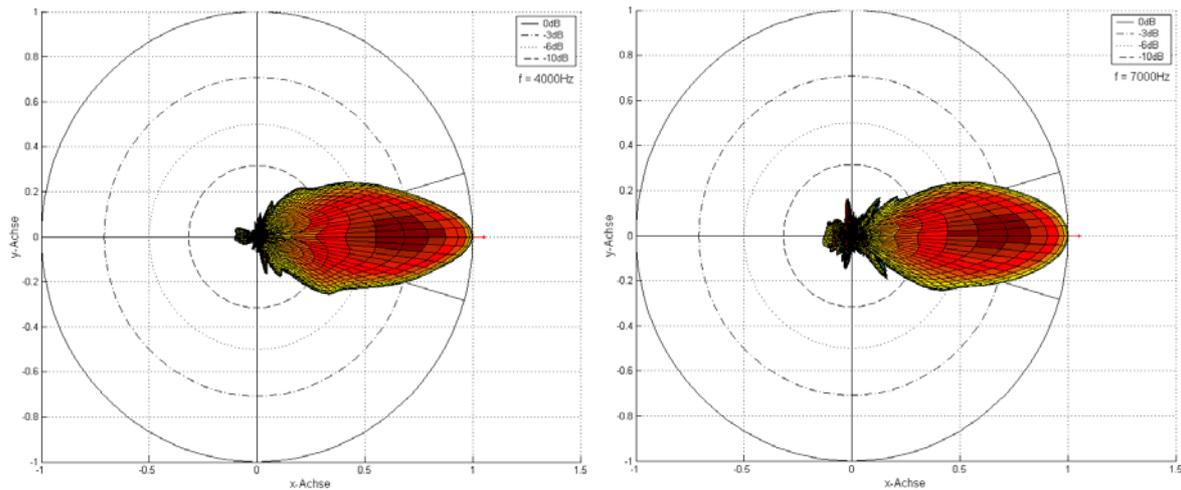


Abb. 7.21: Beamplots des zirkulären Arrays mit räumlicher Gaussgewichtung der Mikrofonsignale in der Horizontalebene für 250Hz, 500Hz, 1kHz, 2kHz, 4kHz und 7kHz

Besonders bei hohen Frequenzen bleibt die gewünschte Richtcharakteristik über weite Bereiche konstant. Die geringen Abweichungen von der Symmetrie ergeben sich aus der in die Berechnung miteinbezogenen Positionierengenauigkeit der Arraymikrofone um $\pm 0.5\text{cm}$ vom idealen Wert, was auch für die gute Robustheit dieses Beamformers spricht.

4.8 Delay&Sum Beamformer mit räumlicher Tiefpassfilterung der Richtcharakteristik

4.8.1 Theoretische Betrachtungen

Eine Möglichkeit die Hauptkeulenbreite zu steuern besteht in der räumlichen Filterung der Richtcharakteristik. Dazu wurde ein FIR-Tiefpassfilter implementiert und wie nachfolgend beschrieben in horizontaler Richtung auf die dreidimensionale Richtcharakteristik angewendet.

Für die Darstellung der Beamplots existiert eine Matrix, die jeder Kombination aus Azimut- und Elevationswinkel einen Amplitudenwert zuordnet. Bei der gewählten Winkelauflösung von 2.5° besitzt diese eine Dimension von 144×72 Einträgen.

$$A_{i,j} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdot & a_{1,72} \\ a_{2,2} & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ a_{144,1} & \cdot & \cdot & a_{144,72} \end{bmatrix}$$

Jede Spalte beschreibt die äquidistant abgetastete Richtcharakteristik für einen Elevationswinkel. Die Filterung erfolgt nun durch spaltenweise zirkuläre Faltung der Folge der Spaltenwerte mit der Fensterfunktion.

Für eine bessere Übersichtlichkeit der mathematischen Beschreibung wird hier nur der zweidimensionale Fall innerhalb der Horizontalebene betrachtet.

Für eine (beliebig gewählte) Schalleinfallrichtung Φ_0 gilt:

$$\underbrace{\begin{bmatrix} \bar{y}(\Phi_1) \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \bar{y}(\Phi_N) \end{bmatrix}}_{\text{Outputvektor}} = \underbrace{\begin{bmatrix} u_0 & u_1 & u_2 & \dots & u_2 & u_1 \\ u_1 & u_0 & u_1 & u_2 & \dots & u_2 \\ u_2 & u_1 & u_0 & u_1 & u_2 & \dots \\ \dots & u_2 & u_1 & u_0 & u_1 & u_2 \\ u_2 & \dots & u_2 & u_1 & u_0 & u_1 \\ u_1 & u_2 & \dots & u_2 & u_1 & u_0 \end{bmatrix}}_{\text{Mittelungsmatrix } u} \cdot \underbrace{\begin{bmatrix} w(\Phi_1)_1 & w(\Phi_1)_2 & \cdot & \cdot & \cdot & w(\Phi_1)_N \\ w(\Phi_2)_1 & w(\Phi_2)_2 & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ w(\Phi_N)_1 & \cdot & \cdot & \cdot & \cdot & w(\Phi_N)_N \end{bmatrix}}_{\text{Gewichtsmatrix } w} \cdot \underbrace{\begin{bmatrix} d_1(\Phi_0) \\ d_2(\Phi_0) \\ \cdot \\ \cdot \\ \cdot \\ d_N(\Phi_0) \end{bmatrix}}_{\mathbf{d}(\Phi_0)} \cdot s(t) \quad (7.22)$$

Die Zeilen der Gewichtsmatrix \mathbf{w} entsprechen den Sensorgewichten für jeweils eine Hauptrichtung $\Phi_{1..N}$. $\mathbf{d}(\Phi_0)$ ist der Steuervektor für die angenommene Schalleinfallrichtung. $\mathbf{w} \cdot \mathbf{d}$ ergibt einen Spaltenvektor, der dem Verlauf des Beamplots folgt. Dabei führt eine Änderung der Schalleinfallrichtung bei konstanten Arraygewichten zum gleichen Ergebnis, wie die Änderung der Arrayausrichtung um den gleichen Wert bei gleichbleibender Schalleinfallrichtung. Durch die Mittelungsmatrix \mathbf{u} erfährt die Richtcharakteristik eine räumliche Tiefpassfilterung. Das hat eine Aufweitung der Hauptkeule auf einen weitgehend frequenzunabhängigen Öffnungswinkel zur Folge und es kommt zu einer Quellenaufweitung bzw. Quellenverschmierung, d.h. je breiter die Hauptkeule wird, umso schlechter ist auch das räumliche Auflösungsvermögen des Beamformers. Daher ist es ausreichend den Raum in größeren Winkelinkrementen abzutasten (Downsampling) indem man nur jeden n -ten Wert von \bar{y} für eine Raumaualisation verwendet. Die Überlagerung aller Beamplots sollte dabei einen möglichst linearen Leistungsverlauf ergeben.

4.8.2 Simulationsergebnisse

Für die Glättung der Richtcharakteristik konnte nach einigen Iterationsschritten eine gut funktionierende Fensterfunktion gefunden werden.

$$w(n) = 2 \cdot e^{\frac{-n^2}{275}} - e^{\frac{-n^2}{650}}, \text{ mit } n = -50 \dots 50 \quad (7.23)$$

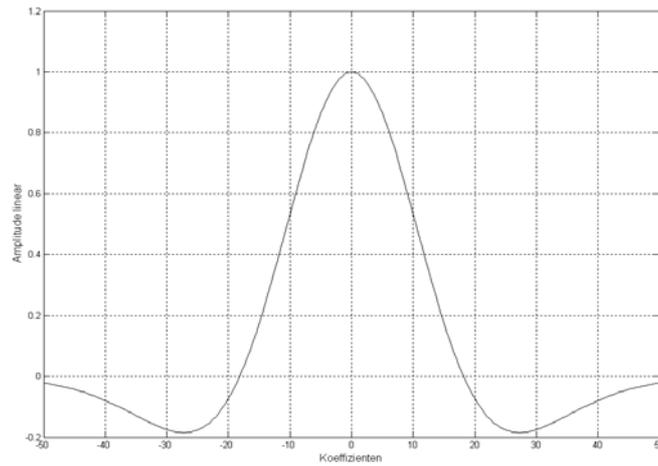
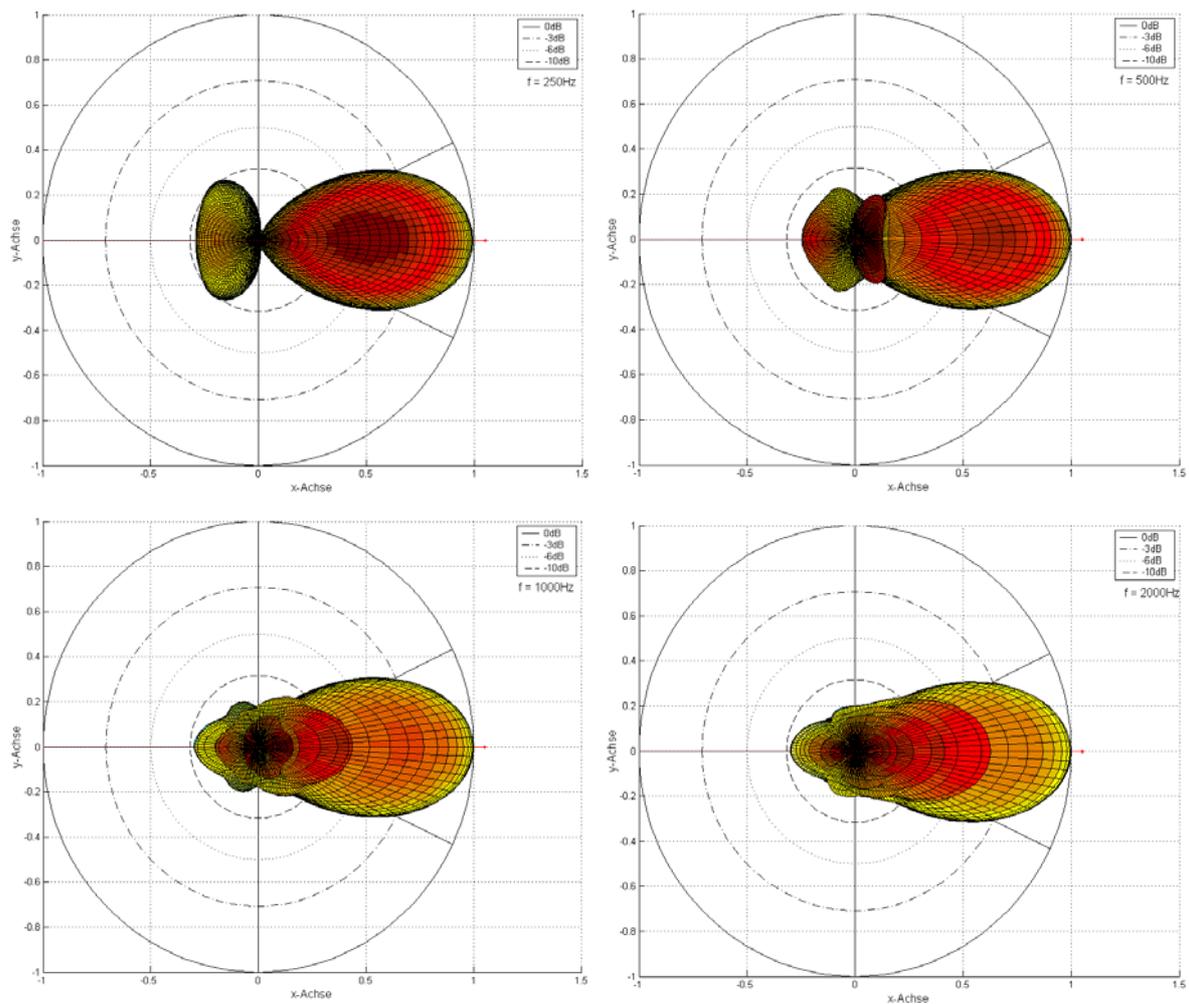


Abb. 7.22: Fensterfunktion für die räumliche Tiefpassfilterung

Die Filterlänge beträgt somit 101 Punkte.

Nach der räumlichen Filterung erhält man in der Horizontalebene eine gleichbleibende Hauptkeulbreite bis etwa 7kHz bei einer Nebenkeulendämpfung $>10\text{dB}$.



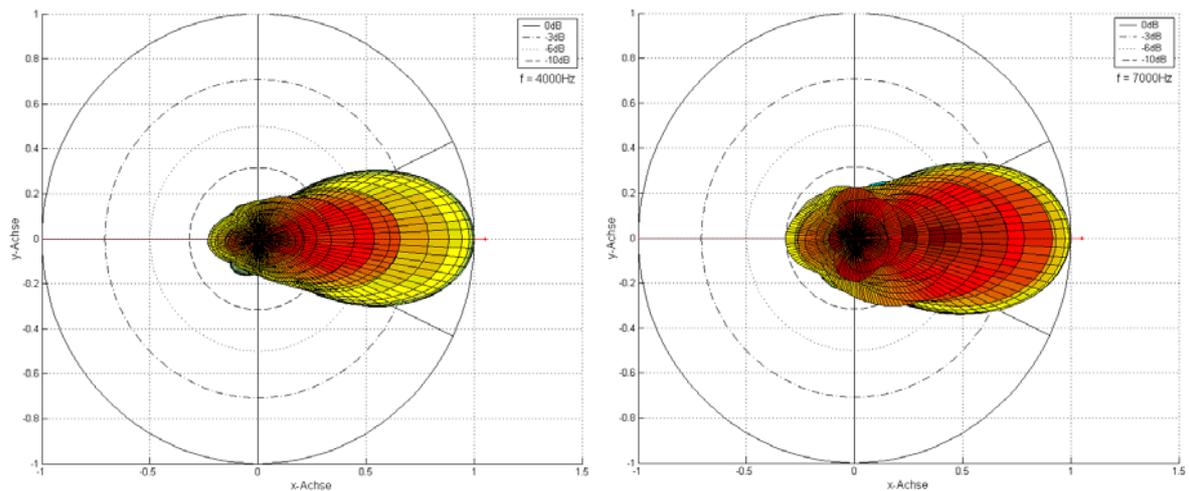


Abb. 7.23: Beamplots des zirkulären Arrays mit räumlicher Tiefpassfilterung in der Horizontalebene für 250Hz, 500Hz, 1kHz, 2kHz, 4kHz und 7.5kHz

In vertikaler Richtung kommt es ebenfalls zu einer Aufweitung der Hauptkeule und zusätzlich zu einer drastischen Verschlechterung der Nebenkeulendämpfung, die den Mindestansprüchen für die Messung gerichteter Impulsantworten nicht mehr genügt.

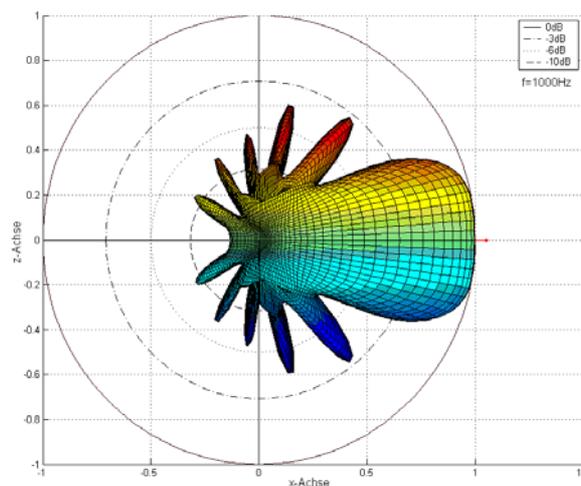


Abb. 7.24: Beamplots des zirkulären Arrays mit räumlicher Tiefpassfilterung in vertikaler Richtung für 1kHz

Durch den zweidimensionalen Aufbau des Arrays ist die Richtwirkung symmetrisch um die Sensorebene und daher für elevierte Hauptrichtungen zweideutig. Die Verbesserungsmöglichkeiten durch Ausdehnung der Geometrie um weitere parallel versetzte Kreisarrays wurde für den Schmalbandfall bei 1kHz überprüft. Akzeptable Ergebnisse brachte ein Array mit vier im Abstand von 16cm übereinander liegenden Mikrofonebenen. Dieser Wert ist rein empirisch für 1kHz ermittelt. Gleichbleibende

Richtcharakteristika sind wiederum nur durch eine frequenzabhängige Apertur in z-Richtung erreichbar.

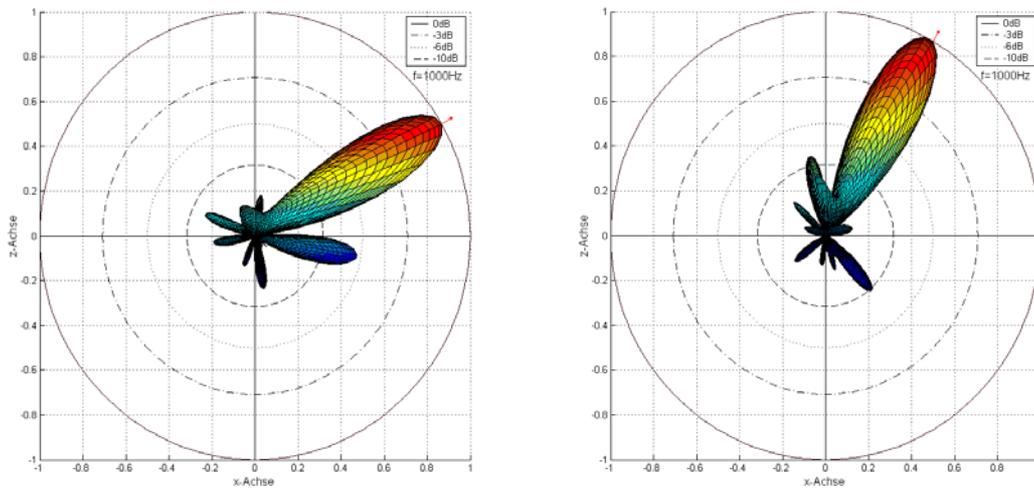


Abb. 7.25: Beamplots eines zylindrischen Arrays mit vier Sensorebenen und räumlicher Tiefpassfilterung in der Horizontalebene für 1kHz und 30° bzw. 60° Elevation

Da für die breitbandige Verwendbarkeit dieser Geometrie eine viel höhere Anzahl von parallelen Kreisarrays notwendig wäre, bleibt zu prüfen, ob eine weniger aufwändige Anordnung gefunden werden kann, die vergleichbare Ergebnisse liefert

5 Ausblick

Eine Weiterentwicklung des Beamformers ist in erster Linie in zwei Bereichen sinnvoll:

1. Durch Optimierung der Arraygeometrie können die Grenzen des verarbeitbaren Frequenzbereiches ausgedehnt sowie der Nachteil der Mehrdeutigkeit eines planaren Arrays beseitigt werden. Denkbar ist die Verwendung einer dreidimensionalen zirkulären Struktur, die sowohl durch die Vergrößerung der maximalen Apertur wie auch durch die Verkleinerung des minimalen Sensorabstandes deutliche Verbesserungen gegenüber den in dieser Arbeit verwendeten Mikrofonarrays bietet. Die Messbarkeit der Raumimpulsantworten ist durch den Einsatz des Drehtellers dabei weiterhin weitgehend automatisierbar.
2. Durch Beamforming-Algorithmen die eine ausreichende Robustheit des Minimum Varianz Beamformers gegen Abweichungen von den idealen Bedingungen (Schallquellen- und Sensorpositionierung, begrenzter Signal/Rauschabstand des Messsystems, Störquellen im Aufnahmeraum) gewährleisten, ist eine Verbesserung der Direktivität bei gleichbleibender Arraygeometrie zu erwarten.

Die in Matlab entwickelten Simulationsprogramme können durch eine grafische Oberfläche in der Bedienbarkeit stark vereinfacht werden.

Literaturverzeichnis

[Bauman, 1984]

Bauman P. D., Lipshitz S. P., Scott T. C., Vanderkooy J.: "Cepstral Techniques for Transducer Measurement", 76th AES Convention, New York 1984, Preprint No. 2172

[Bauman, 1985]

Bauman P. D., Lipshitz S. P., Vanderkooy J.: "Cepstral Techniques for Transducer Measurement, Part 2", 79th AES Convention, New York 1984, Preprint No. 2302

[Baysal, 2003]

Baysal Ü., Moses R. L.: "On the Geometry of Isotropic Arrays" IEEE Trans. Signal Processing, vol. 51, No. 6, Juni 2003

[Blauert, 1974]

Blauert J.: "Räumliches Hören", S. Hirzel Verlag Stuttgart, 1974

[Cheng,1999]

Cheng C., Wakefield G. H.: "Introduction to Head-Related Transfer Functions (HRTFs): Representations of HRTFs in Time, Frequency and Space" 107th AES Convention, New York 1999, Preprint No. 5026

[Cox, 1987]

Cox, H., Zeskind, R., Owen, M.: "Robust adaptive beamforming", IEEE Trans. ASSP Vol. ASSP-35 No. 10, Oktober 1987

[Farina, 2000]

Farina A.: „Simultaneous measurements of impulse response and distortion with a swept.sine technique“, J. Audio Eng. Soc. (Abstracts), vol. 48, p. 350, April 2000

[Fellner, 1998]

Fellner M.: "Recheneffiziente Modellierung von Außenohrübertragungsfunktionen"
Diplomarbeit an der TU Graz, 1998

[Graber, 1997]

Graber G.: "Tontechnik und interdisziplinäres Sinnen – Eine Grundlegende Fragestellung",
Habilitationsschrift eingereicht an der TU Graz, 1997

[Graf, 1999]

Graf F.: "Entwicklung eines Aufnahmesystems für psychoakustische Analysen",
Diplomarbeit an der TU Graz, März 1999

[Krejci, 2001]

Krejci P. M.: "Entwicklung eines Mehrkanaligen Systems für eine virtuelle Raumakustik",
Diplomarbeit an der TU Graz, Oktober 2001

[Müller, 2001]

Müller S. and Massarani P.: "Transfer-Function Measurement with Sweeps". J. Audio. Eng. Soc., Vol. 49, No. 6, p.443-471, Juni 2001.

[Niedrist, 1993]

Niedrist, G.: "Echo Suppression for Loudspeaker-Mikrophone Systems Measurements". J. Audio Eng. Soc., Vol. 41, p.143-153, No. 3, März 1993.

[Outline]

www.outline.it, Technische Daten zum verwendeten Drehteller ET2-ST2

[Puckette]

Puckette M., pd (pure data), open source real time computer music software,
http://cra.ucsd.edu/~msp/Pd_documentation/

[Salamouris, 1992]

Salamouris S. I.: "A Method for Extending the Measured Response of a Loudspeaker towards the Low Frequencies, with the use of Cepstral Analysis". Presented at the 92nd Convention of the Audio Engineering Society, J. Audio Eng. Soc., Preprint No. 3233, März 1992

[Stan, 2002]

Stan G. B., Embrechts J. J. and Archambeau D.: „Comparison of Different Impulse Response Measurements Techniques“. J. Audio. Eng. Soc., vol. 50, No. 4, April 2002

[Theile, 1986]

Theile G.: "Das Kugelflächenmikrofon", Bericht zur 14. Tonmeistertagung S.277-293, 1986

[van der Wal, 1996]

van der Wal M., Start E. W., de Vries D.: "Design of Logarithmically Spaced Constant-Directivity Transducer Arrays", J. Audio. Eng. Soc., vol. 44, No. 6, Juni 1996

[Van Veen, 1988]

Van Veen B. D., Buckley K. M.: "Beamforming: A Versatile Approach to Spatial Filtering", IEEE ASSP Magazine, April 1988

[Wenzel, 1993]

Wenzel E. M., Arruda M., Kistler D. J., Wightman F. L.: "Localization using nonindividualized head-related transfer functions", J. Acoust. Soc. Am., 94:111-123, 1993

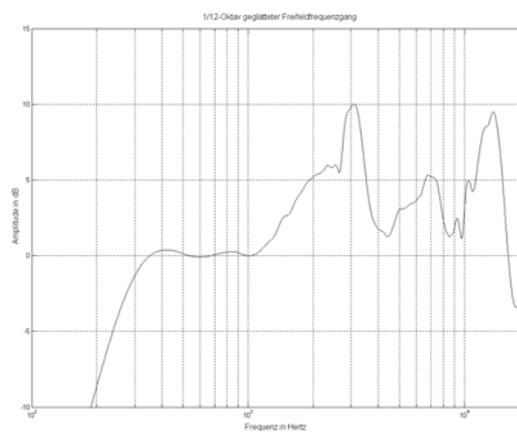
[Zölzer, 1997]

Zölzer U.: "Digitale Audiosignalverarbeitung", B.G. Teubner Verlag Stuttgart, 1997

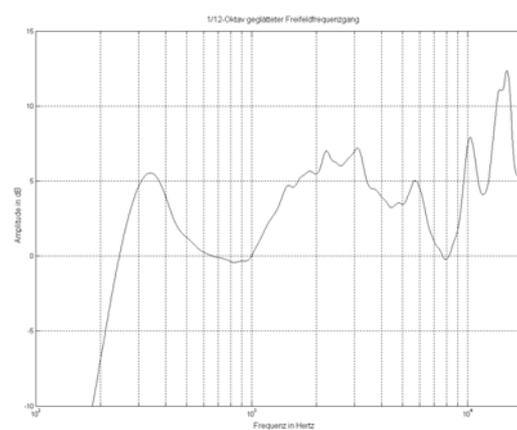
Anhang A

Ergebnisse der Vergleichsmessungen für die Auswahl des Messlautsprechers:

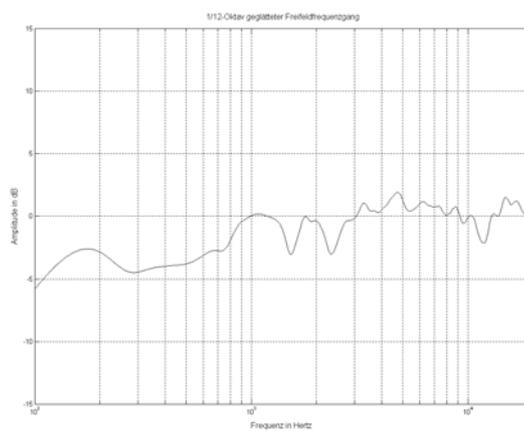
Bose Acoustimass 6 (Satellit des Surround-Systems):



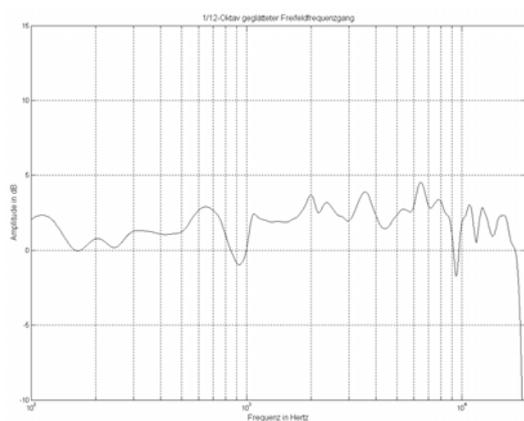
JBL 301SAT:



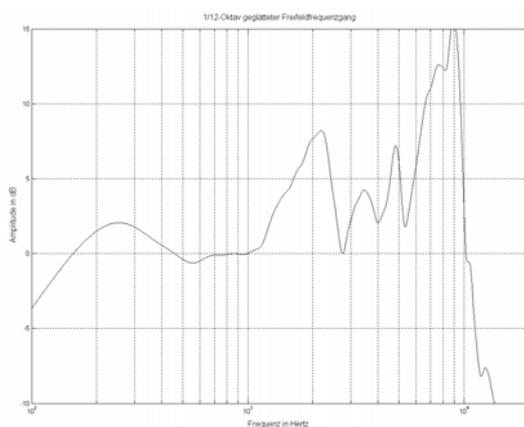
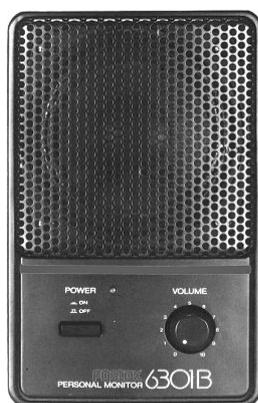
Yamaha MSP5:



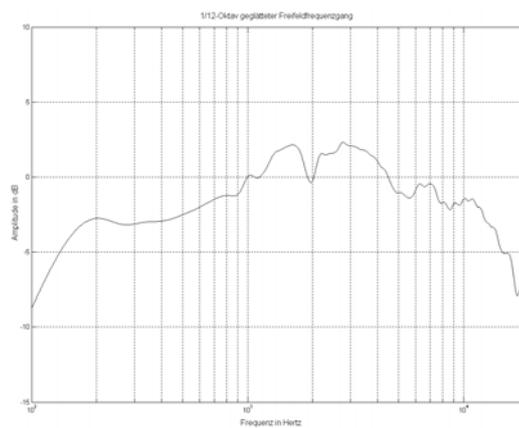
Tannoy System600A:



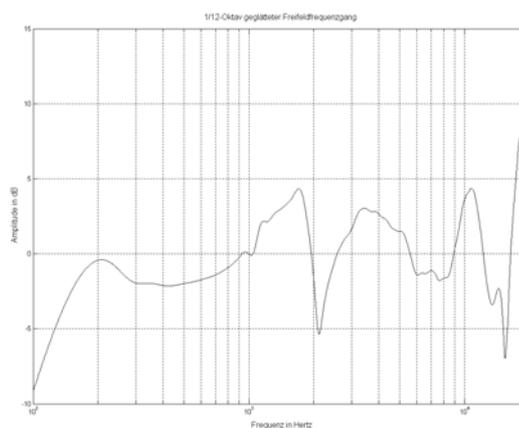
Fostex 6301B:



Großer Tropfenlautsprecher:



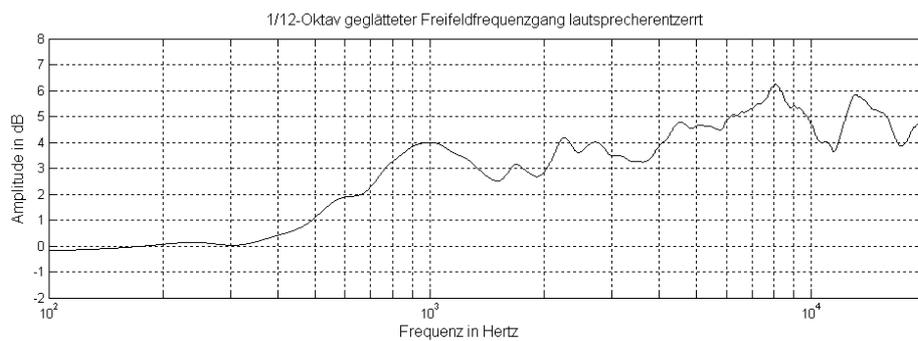
Kleiner Tropfenlautsprecher:



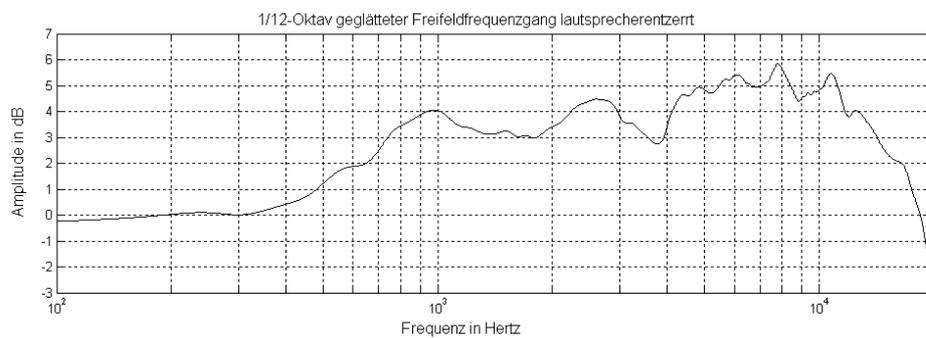
Anhang B

Folgende Abbildungen zeigen die Ergebnisse der Vergleichsmessungen für den Kunstkopf "Source" mit verschiedenen Absorbermaterialien zur Optimierung des Frequenzgangs.

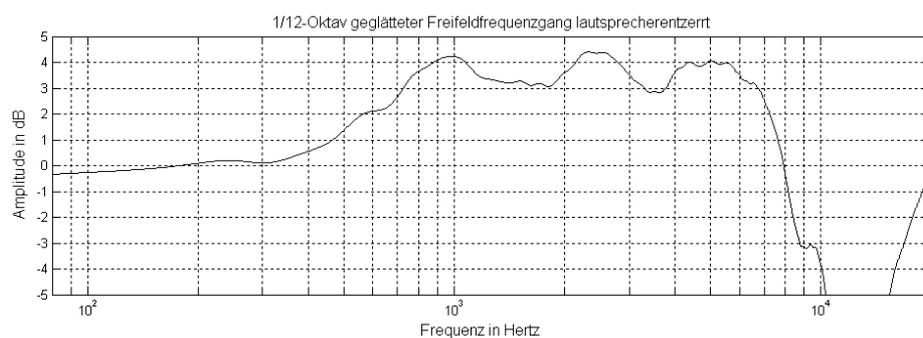
Freifeldfrequenzgang von Source ohne Absorber:



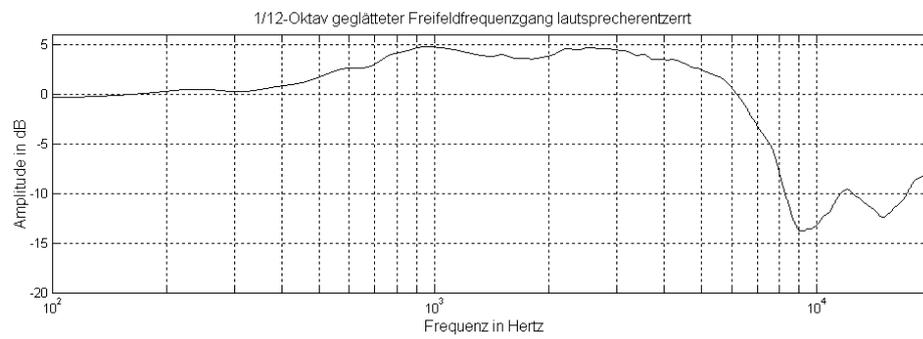
mit 3mm Schaumstoffabsorber (feste Konsistenz):



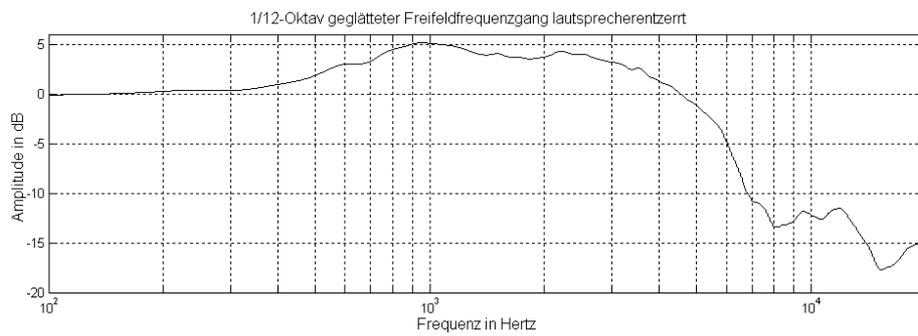
mit 4mm Schaumstoffabsorber (mittelweich):



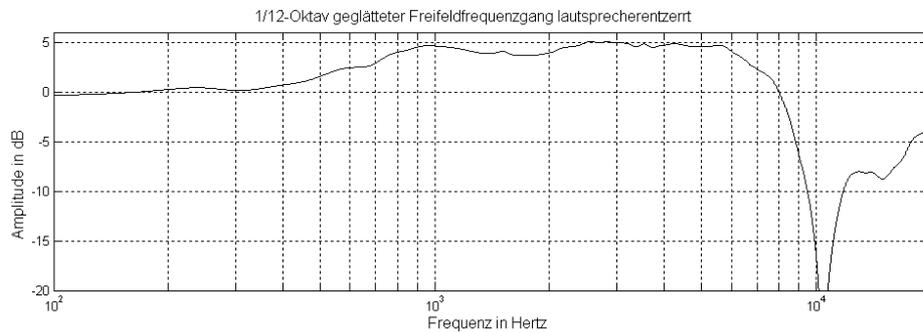
mit 11mm Schaumstoffabsorber (mittelweich):



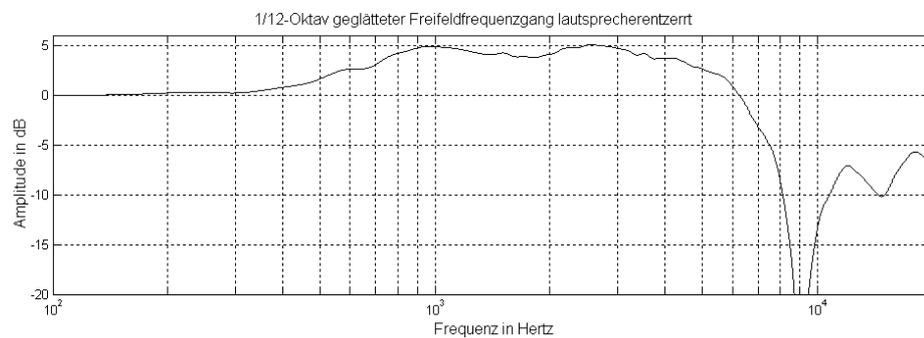
mit 20mm Schaumstoffabsorber (mittelweich):



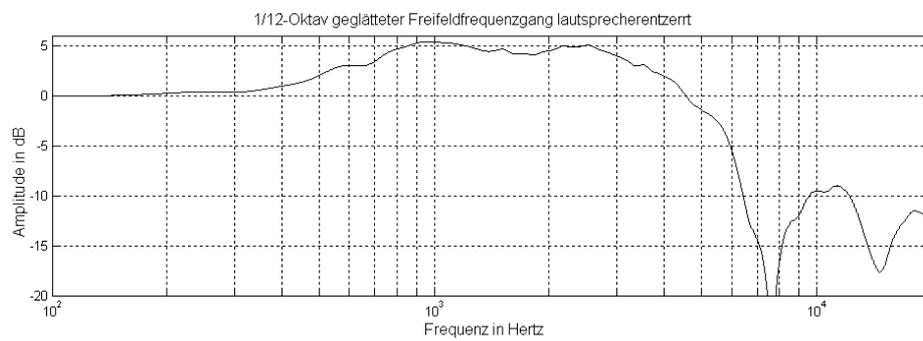
mit 5mm Schaumstoffabsorber (weich):



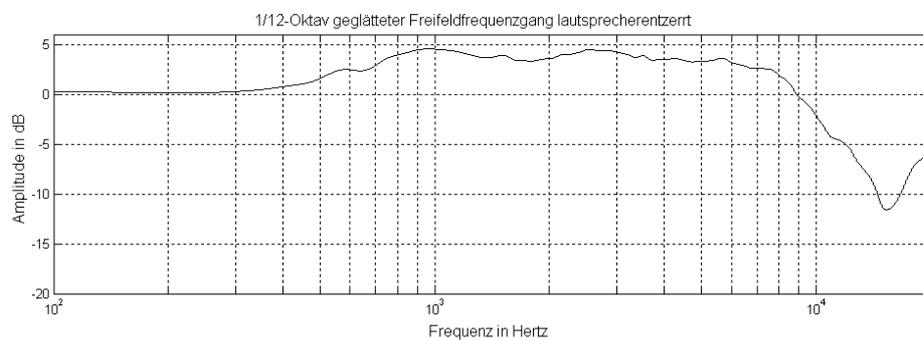
mit 11mm Schaumstoffabsorber (weich):



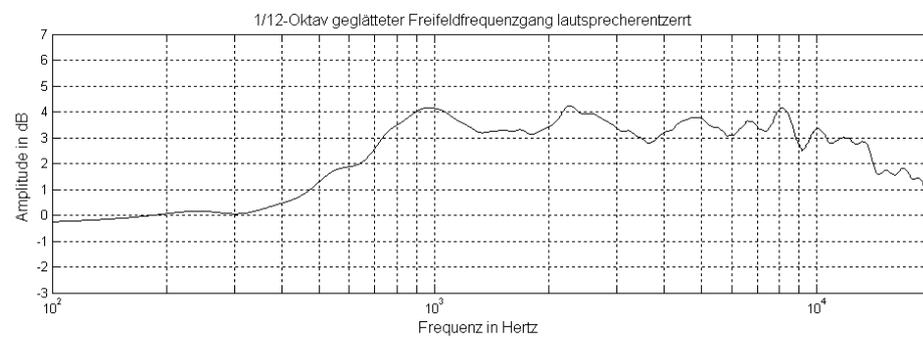
mit 20mm Schaumstoffabsorber (weich):



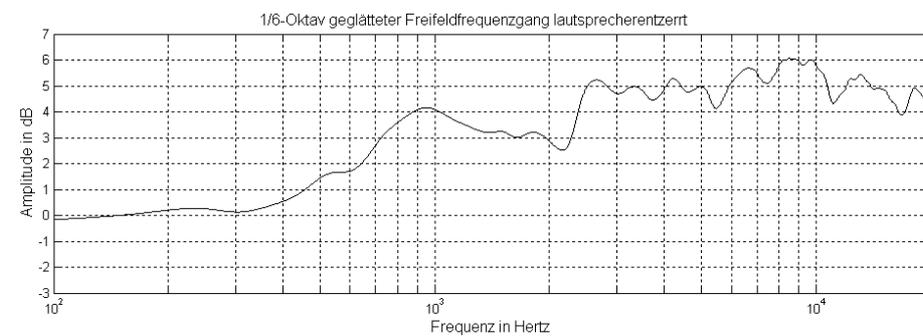
mit 4mm Filzteppich:



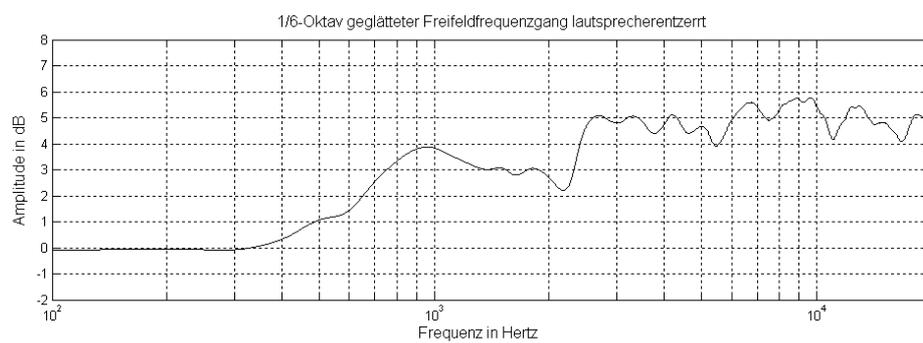
mit 5mm Korkplatte:



mit 3mm Gummibelag:



mit 6mm Gummibelag:



Anhang C

Sensorpositionen des dreidimensionalen Mikrofonarrays:

Mic-Pos.-Nr.	x	y	z
1	0,000	0,000	1,200
2	0,080	0,000	1,200
3	0,200	0,000	1,200
4	0,380	0,000	1,200
5	0,650	0,000	1,200
6	-0,080	0,000	1,200
7	-0,200	0,000	1,200
8	-0,380	0,000	1,200
9	-0,650	0,000	1,200
10	0,000	0,080	1,200
11	0,000	0,200	1,200
12	0,000	0,380	1,200
13	0,000	0,650	1,200
14	0,000	-0,080	1,200
15	0,000	-0,200	1,200
16	0,000	-0,380	1,200
17	0,000	-0,650	1,200
18	0,000	0,000	1,260
19	0,000	0,000	1,340
20	0,000	0,000	1,140
21	0,000	0,000	1,060

22	0,080	0,080	1,216
23	0,200	0,200	1,240
24	0,380	0,380	1,276
25	0,650	0,650	1,330
26	-0,080	-0,080	1,184
27	-0,200	-0,200	1,160
28	-0,380	-0,380	1,124
29	-0,650	-0,650	1,070
30	-0,080	-0,080	1,216
31	-0,200	-0,200	1,240
32	-0,380	-0,380	1,276
33	-0,650	-0,650	1,330
34	0,080	0,080	1,184
35	0,200	0,200	1,160
36	0,380	0,380	1,124
37	0,650	0,650	1,070
38	-0,080	0,080	1,216
39	-0,200	0,200	1,240
40	-0,380	0,380	1,276
41	-0,650	0,650	1,330
42	0,080	-0,080	1,184
43	0,200	-0,200	1,160
44	0,380	-0,380	1,124
45	0,650	-0,650	1,070
46	0,080	-0,080	1,216
47	0,200	-0,200	1,240
48	0,380	-0,380	1,276
49	0,650	-0,650	1,330
50	-0,080	0,080	1,184
51	-0,200	0,200	1,160
52	-0,380	0,380	1,124
53	-0,650	0,650	1,070