

**Untersuchung verschiedener Verfahren zur
messtechnischen Bestimmung und
Nachbildung der Akustik insbesondere von
Fahrzeugaudiosystemen**

Dissertation

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften

(Dr.-Ing.)

der Technischen Fakultät

der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Markus Christoph

Kiel 2013

1. Berichterstatter: Prof. Dr.-Ing. Gerhard Schmidt
2. Berichterstatter: Prof. Dr.-Ing. Henning Puder
Datum der mündlichen Prüfung: 16. September 2013

Kiel, im Mai 2013

Markus Christoph

Dipl.-Ing. (FH) Markus Christoph

Eidesstattliche Erklärung:

Hiermit erkläre ich an Eides statt, dass ich die vorgelegte Dissertation mit dem Titel „Untersuchung verschiedener Verfahren zur messtechnischen Bestimmung und Nachbildung der Akustik insbesondere von Fahrzeugaudiosystemen“ selbständig und ohne unerlaubte Hilfe angefertigt habe und dass ich die Arbeit noch keinem anderen Fachbereich bzw. noch keiner anderen Fakultät vorgelegt habe.

Kiel, den 02.05.2013

Unterschrift

Hiermit erkläre ich, dass gegen mich kein strafrechtliches Ermittlungsverfahren schwebt.

Kiel, den 02.05.2013

Unterschrift

Hiermit erkläre ich, dass die Dissertation nach den Regeln guter wissenschaftlicher Praxis (Standard wissenschaftlichen Arbeitens nach den Empfehlungen der DFG) abgefasst wurde.

Kiel, den 02.05.2013

Unterschrift

Abstract

Bei der Verfolgung des ursprünglichen Ziels der Dissertation ein lautsprecherbasiertes System zur akustischen Dokumentation zu finden wurde schnell erkannt, dass bereits einige etablierte Verfahren zur Decodierung, d.h. für die Rekonstruktion von zuvor, meist künstlich codierter Wellenfelder, wie etwa die Wellenfeldsynthese oder das Higher-Order-Ambisonic (HOA)-Verfahren, existieren. Dagegen stellte sich im Laufe der Arbeit heraus, dass im Bereich der Codierung, also der messtechnischen Bestimmung von Wellenfeldern, noch ein höherer Forschungsbedarf existiert. Zudem bestand von Beginn an das Interesse vorrangig Fahrzeugaudiosysteme zu auralisieren. Aus diesen Gründen entwickelte sich die Untersuchung unterschiedlicher Verfahren zur messtechnischen Bestimmung der Akustik, insbesondere für Audiosysteme in Kraftfahrzeugen, zum Schwerpunkt dieser Arbeit.

Nach grundlegender Untersuchung unterschiedlichster Verfahren zur messtechnischen Bestimmung der Akustik, kristallisierte sich die Verwendung zirkularer bzw. sphärischer Mikrophonarrays, wie sie im HOA-Verfahren Verwendung finden, als erfolgversprechendste Methode heraus. Unter Verwendung der HOA-Theorie wurden dabei vorrangig sphärische, d.h. kugelförmige Mikrophonarrays zur dreidimensionalen, messtechnischen Bestimmung von Wellenfeldern näher untersucht. Dabei konnten Wege zur Verbesserung bestehender Messsysteme durch Modifikationen in ihrem Design gefunden werden.

Messungen bestätigten dabei die Wirksamkeit der neuen Designs. So konnte etwa die wirksame Bandbreite durch ein neuartiges Kugelmikrophonarray durch eine einfache, rein mechanische Modifikation nahezu verdoppelt werden, ohne dabei die Anzahl der Sensoren zu erhöhen. Aber auch Veränderung im Filterdesign wurden untersucht, was zu einem neuartigen Algorithmus führte, welcher in der Lage ist Filter zu erzeugen, sodass das Mikrophonarray einerseits eine gewisse, minimale *White Noise Gain*-Schwelle nicht unterschreitet und zudem, unabhängig von der betrachteten Frequenz, stets einen gleichen, maximalen Pegel in der Hauptempfangsrichtung aufweist. Schließlich wurden noch weitere Konzepte vorgestellt wie sich die Wirkung der Mikrophonarrays für die Codierung durch einfache, praxistaugliche Modifikationen weiter steigern lässt.

Neben der messtechnischen Bestimmung der Akustik wurden zudem noch Verfahren zur Nachbildung bzw. zur Verbesserung der Akustik untersucht, wobei hierzu die Inverse-Filter-Theorie näher betrachtet wurde. Basierend auf Aufnahmen in einem Fahrzeug wurde dieses Prinzip untersucht. Der Fokus wurde dabei auf den Filterentwurf gelegt, wobei

eines der Ziele war, diese so zu entwerfen, dass sie ein möglichst geringes *pre-ringing* aufweisen. Hintergrund hierbei ist, dass klassische Entwurfsverfahren Filter erzeugen, welche unter einem zum Teil starken Vorläuten leiden, was sich akustisch als störend auswirkt und folglich weitestgehend vermieden werden sollte. Der Tribut, der dabei gezahlt werden musste war, die zeitliche Komprimierung der resultierenden, d.h. mit den inversen Filtern gewichteten Raumimpulsantworten, aufzugeben. Dies stellte sich letztendlich jedoch als wesentlich akzeptabler heraus als das Vorhandensein des zuvor genannten *pre-ringing*.

Danksagung und Widmung

Danksagung

Es mir eine Ehre als erstes meinem Doktorvater Prof. Dr.-Ing. Gerhard Schmidt dank-sagen zu dürfen. Nur durch seinen unermüdlichen Einsatz, ich denke hier nur an die zahlreichen Sitzungen im Promotionsausschuss der Uni Kiel, in denen sich Herr Prof. Dr.-Ing. Gerhard Schmidt für meine Belange eingesetzt hat sowie sein unerschütterliches Vertrauen in meine Person und meine Fähigkeiten, wurde eine Durchführung dieser Arbeit überhaupt erst ermöglicht. Dabei erwies sich Herr Schmidt stets als kompetenter, aufgeschlossener und verständnisvoller Mentor, an dem man sich zu jeder Zeit mit Belangen jeglicher Art wenden konnte. Durch ihn kam ich auch mit anderen Doktoranden wie etwa Herrn Dipl.-Phys. Christian Lüke oder Herrn Dipl.-Ing. Jochen Withopf in Kontakt, die ich trotz der geringen Zeit, die wir gemeinsam an der Uni Kiel verbrachten, kennen und schätzen gelernt habe. Auch Frau Silvia Schuchard wird mir stets als sympathische und hilfsbereite Sekretärin von Herrn Schmidt in guter Erinnerung bleiben, ebenso wie Herr Dipl.-Ing. Torge Rabsch ohne dessen Hilfe, manches Problem mit dem Netzwerk bzw. mit dem Internet nicht zu beheben gewesen wäre.

Dank gilt auch Herrn Prof. Dr.-Ing. Henning Puder, der sich freundlicherweise spontan dazu bereit erklärt hat, die Zweitkorrektur dieser Arbeit zu übernehmen, wie auch Herrn Prof. Dr.-Ing. Ulrich Heute, der speziell für mich eine Leistungsbewertung im Studienfach *Advanced Digital Signal Processing* anfertigte und abnahm.

Auch der Firma Harman/Becker Automotive Systems GmbH in Straubing, in welcher ich während der Erstellung dieser Arbeit unentwegt in der Abteilung *Center of Competence - Amplifier (CoC-Amp)* unter der Leitung von Herrn Dipl.-Ing. (FH), Dipl.-Phys. (Univ.) Gerhard Pfaffinger, genauer gesagt in der Gruppe *Basic Research*, welche von Herrn Dipl.-Ing. (FH) Franz Lorenz geleitet wird, als Ingenieur tätig war, gilt mein Dank. Ohne diesen „festen Anker“, welcher einerseits für eine finanzielle Absicherung, andererseits aber auch für viel Freude und Abwechslung sorgte sowie der Gewissheit, sich in einer fortbildungsfreundlichen, der Forschung wohlgesonnen Umgebung zu wägen, wäre eine erfolgreiche Durchführung dieser Arbeit nicht denkbar gewesen. Hier gilt insbesondere Herrn Gerhard Pfaffinger als Leiter der Abteilung *CoC-Amp* mein besonderer Dank für die von seiner Seite stets vorhandene Bereitschaft, meine Belange auch gegenüber der

Firma Harman/Becker intern zu unterstützen und meist auch durchzusetzen, was sich etwa durch die Übernahme von Reisekosten an die Uni Kiel oder zu diversen Konferenzen äußerte.

Selbstverständlich darf das private Umfeld nicht außer Acht gelassen werden. Hierbei möchte ich meiner Familie, insbesondere jedoch meiner langjährigen Lebensgefährtin Frau Dipl. Kauffrau (Univ.) Carmen Haas sowie ihrer Familie für das Interesse an dieser Arbeit sowie das Verständnis für die vielen Stunden in denen ich mit Abwesenheit glänzte und der aktiven Unterstützung dieser Arbeit durch ein kritisches Korrekturlesen, insbesondere der Kommasetzung, danken. Bedanken möchte ich mich auch bei Herrn BSc, Dip Law (GDL) Andrew Darby für das Korrekturlesen sämtlicher in Englisch verfassten Beiträge, die während der Erstellung dieser Arbeit entstanden. Auch Herrn Dr.-Ing. Richard Huber, den ich schon seit meiner Kindheit als Freund bezeichnen darf, möchte ich für all die guten Ratschläge sowie für die Übernahme der „Erstkorrektur“ und der Beseitigung der ersten, groben Rechtschreibfehler meinen Dank aussprechen - ich weiß, dass dies mit an Sicherheit grenzender Wahrscheinlichkeit, vor allem für einen fachfremden wie ihm, nicht leicht war.

Widmung

Widmen möchte ich diese Arbeit meiner, bereits in der Anfangsphase dieser Arbeit, leider viel zu früh verstorbenen Mutter Frau Marianne Christoph. Ohne ihre Hartnäckigkeit, bei meinen Geschwistern und mir von klein auf stets auf ein gewissenhaftes, schulisches Arbeiten zu achten, ohne dabei unmögliches von uns zu verlangen, wäre mein Leben wohl anders verlaufen und an eine akademische Laufbahn nicht zu denken gewesen. Auch die Art, wie sie persönliche Interessen stets hinter den Interessen ihrer Kinder gestellt hat um uns Dinge, wie etwa mir ein Studium der Elektrotechnik zu ermöglichen, erfüllen mich nach wie vor mit höchstem Respekt und tiefer Schuld. Somit war es für mich ein persönliches Anliegen, meine Mutter, wenn auch posthum, durch diese Arbeit mit Stolz zu erfüllen. Das Zustandekommen dieser Arbeit beweist: Deine Mühen waren nicht vergebens - Danke.

Inhaltsverzeichnis

Abkürzungen und Notation	xii
1 Einleitung	1
1.1 Motivation	1
1.2 Gliederung der Arbeit	2
2 Inverse Filter	5
2.1 Optimale Lautsprecheraufstellung	9
2.2 Modifizierungen beim Entwurf der inversen Filter	10
2.3 Regularisierung	11
2.3.1 Wahl des Regularisierungsparameters β	13
2.4 Berechnung der inversen Filter mit Hilfe der adjungierten Matrix	14
2.5 Beispiel und praktische Modifikationen	17
2.5.1 Ursprüngliche Situation - Situation ohne Filterung	19
2.5.2 Situation bei alleiniger Verwendung der Filter zur Kompensation des Übersprechens	22
2.5.3 Situation bei Verwendung der vollständigen, inversen Filter	22
2.5.4 Praktische Modifikationen beim Entwurf des inversen Filters $G(j\omega)$	27
2.6 Probleme und Ausblick	31
2.7 Zusammenfassung	34
3 Ambisonic 1. Ordnung	37
3.1 Ambisonic Codier/Decodier-Beispiel	41
3.1.1 Decodervariationen	50
3.2 Zusammenfassung	51
4 Ambisonic höherer Ordnung	53
4.1 Zerlegung des Klangfeldes	54

4.2	Codierung und Decodierung	62
4.2.1	Der Codierprozess von ebenen Wellen	62
4.2.2	Der Codierprozess von Kugelwellen	64
4.2.3	Der Decodierprozess	68
4.2.4	Lautsprecheraufstellung und Decodermatrix	70
4.2.5	Verallgemeinerung der Decodervariationen	72
4.2.6	Nahfeldeffekt bei der Decodierung	87
4.3	Zusammenfassung	92
5	HOA-Codierung natürlicher Klangfelder mit Hilfe von Mikrophon-	
	arrays	93
5.1	Allgemeine Fourier-Bessel-Reihe	93
5.2	Mikrophonarrays	95
5.2.1	Kugelförmige Mikrophonarrays	97
5.3	Untersuchung unterschiedlicher sphärischer Mikrophonarrays	100
5.3.1	Offene Kugel mit omnidirektionalen Mikrophonen	101
5.3.2	Offene Kugel mit gerichteten Mikrophonen 1. Ordnung	101
5.3.3	Geschlossene Kugel mit omnidirektionalen Mikrophonen	104
5.3.4	Geschlossene Kugel mit Kardioid-Mikrophonen 1. Ordnung	110
5.3.5	Zusammenfassung	115
5.4	Praktischer Entwurf eines geschlossenen, kugelförmigen Mikrophonarrays zur Messung von HOA-Komponenten	117
5.4.1	Codierung eines natürlichen Klangfeldes	119
5.4.2	Optimale Wahl des Regularisierungsparameters ϵ	123
5.5	Beschreibung des Optimierungsalgorithmus	127
5.5.1	Berechnung des <i>white noise gain</i> ($WNG_{dB}(\omega)$)	127
5.5.2	Berechnung des Array-Gewinns ($G_{dB}(\omega)$)	129
5.5.3	Optimierungsalgorithmus	130
5.5.4	Simulationsergebnisse bei Anwendung unterschiedlicher <i>equalizing-</i> Filter	133
5.6	Praxisrelevante Faktoren die beim Bau eines Kugelmikrophonarrays beach- tet werden sollten	142
5.6.1	Positionierung der Mikrophone auf der Kugeloberfläche	142
5.6.2	Festlegung des Kugelradius	143
5.6.3	Möglichkeiten zur Erhöhung der räumlichen Abtastfrequenz	145

5.6.4	Empfindlichkeit bezüglich der Exemplarstreuungen der Sensoren . . .	150
5.7	Zusammenfassung	153
6	Verifikation unterschiedlicher, sphärischer Mikrophonarrays in der Praxis	155
6.1	Vergleich eines geschlossenen, kugelförmigen Mikrophonarrays mit und ohne räumlichen Tiefpassfilter	156
6.1.1	Messergebnisse der geschlossenen, kugelförmigen Mikrophonarrays mit und ohne räumlichen Tiefpassfilter	157
6.2	Untersuchung eines offenen, kugelförmigen Mikrophonarrays geringer Größe	160
6.2.1	Beispielhafter Entwurf eines realisierbaren, robusten, offenen Kugelmikrophonarrays	162
6.3	Zusammenfassung	179
7	Zusammenfassung und Ausblick	181
7.1	Zusammenfassung	181
7.2	Ausblick	187
A	Anhang	195
A.1	Differentielle Mikrophone	195
A.1.1	Berechnung der Verzögerungszeiten T_i	197
A.1.2	Berechnung der Arrayparameter differentieller Mikrophone	198
A.1.3	Modelle nach Gerzon zur Vorhersage der subjektiv empfundenen Horizontallokalisation	201
A.2	Matrizen	206
A.2.1	Berechnung der inversen Matrix mit Hilfe der adjungierten Matrix	206
A.2.2	Hermitesche Matrix	207
A.3	Besselfunktionen	209
A.3.1	Sphärische Besselfunktionen	211
A.4	Entwurf eines modalen <i>beamformer</i>	215
A.4.1	Praktische Anwendungen modaler <i>beamformer</i>	220
A.4.2	Rotation des modalen <i>beamformer</i>	223
A.5	Reguläre und quasi-reguläre Körper	231
A.5.1	Der Tetraeder	233
A.5.2	Der Hexaeder	237
A.5.3	Der Oktaeder	239

A.5.4	Der Dodekaeder	241
A.5.5	Der Ikosaeder	243
A.5.6	Der Buckyball	245
A.5.7	Der verwendete Ikosaederstumpf	247
Literaturverzeichnis		251

Abkürzungen

A/D	Analog/Digital
AES	Audio Engineering Society
ASINC	Angular SINC-Funktion
ASW	Apparent Source Width oder Auditory Source Width
BMW	Bayerische Motoren Werke
CSINC	Circular SINC-Funktion
dB	Dezibel
DI	Directivity Index
2D	Zweidimensional
3D	Dreidimensional
EDPL	Electro Dynamic Planar Loudspeaker
FBR	Front to Back Ratio
FIR	Finite Impulse Response
FFT	Fast Fourier-Transformation
GPU	Graphical Processing Unit
HOA	Higher Order Ambisonic
HRTF	Head Related Transfer Function
IEEE	Institute of Electrical and Electronics Engineers
IFFT	Inverse Fast Fourier-Transformation
IIR	Infinite Impulse Response
ILD	Interaural Level Difference
ITD	Interaural Time Difference
LMS	Least Mean Square
LS	Least Square
MELMS	Multiple Error Least Mean Square
NLMS	Normalized Least Mean Square
OSD	Optimal Source Distribution
PC	Personal Computer
RIR	Room Impulse Response
SI-Funktion	Sinus Cardinalis-Funktion
SNR	Signal to Noise Ratio

SVD	Singular Value Decomposition
TP	Tiefpass
VST	Virtual Studio Technology
WFS	Wave Field Synthesis
WNG	White Noise Gain

Notation

Konventionen und Operatoren

\mathbf{I}	Einheitsmatrix
\mathbf{E}	Fehlermatrix
$\mathbf{A}_{N \times M}$	Matrix \mathbf{A} mit N Zeilen und M Spalten
$a_{N,M}$	Matrizelement der N . Zeile und M . Spalte
\mathbf{A}^T	Transponierte der Matrix \mathbf{A}
\mathbf{A}^*	Komplex Konjugierte der Matrix \mathbf{A}
\mathbf{x}	Vektor \mathbf{x}
\bar{x}	Mittelwert bzw. geglätteter Wert der Größe x
\mathbf{A}^{-1}	Inverse der Matrix \mathbf{A}
$\ \mathbf{x}\ $	2-Norm des Arguments \mathbf{x}
$\ \mathbf{x}\ _p$	p-Norm des Arguments \mathbf{x}
$f'(x)$	Ableitung der Funktion $f(x)$
$\frac{\partial f}{\partial x_i}$	Partielle Ableitung der Funktion $f(x_1, \dots, x_n)$, welche von n Variablen (x_1, \dots, x_n) abhängt nach der Variablen x_i . Alternativ kann dies auch mit $\frac{df}{dx_i}$ bezeichnet werden
\mathbf{A}^+	Pseudoinverse der Matrix \mathbf{A}
\mathbf{A}^H	Komplex konjugierte Transponierte der Matrix \mathbf{A} ($\mathbf{A}^H = (\mathbf{A}^*)^T$)
$\text{Rg}(\mathbf{A})$	Rang der Matrix \mathbf{A}
$\text{diag}(\mathbf{x})$	Diagonalmatrix, welche eindeutig durch den Vektor \mathbf{x} definiert wird
$\text{roots}(f(x))$	Nullstellen der Funktion $f(x)$
$\det(\mathbf{A})$	Determinante der Matrix \mathbf{A}
$\text{adj}(\mathbf{A})$	Adjungierte der Matrix \mathbf{A}
\mathbf{A}_{ij}	Untermatrix von \mathbf{A} , die sich ergibt, wenn aus \mathbf{A} die i . Zeile und j . Spalte entfernt wird
$\det(\mathbf{A}_{ij})$	Minor von \mathbf{A}
$\lceil \dots \rceil$	Aufrunden zum nächst größeren Integerwert
$\lfloor \dots \rfloor$	Abrunden auf den nächsten Ganzzahlwert
$\Re\{X(j\omega)\}$	Realteil des Signals $X(j\omega)$
$\Im\{X(j\omega)\}$	Imaginärteil des Signals $X(j\omega)$

Formelzeichen

a	Radius des Kugelmikrophons in [m]
$B_{m,n}^\sigma$	Ambisonickomponente m . Ordnung, n . Grades mit hochgestellten Index $\sigma \in [-1, +1]$
\mathbf{B}	Vektor, welcher sämtliche $N = (2M + 1)$ (2D), bzw. $N = (M + 1)^2$ (3D), Ambisonicsignale bzw. Ambisonickanäle beinhaltet
c	Schallgeschwindigkeit in $\left[\frac{\text{m}}{\text{s}}\right]$
\mathbf{C}	$N \times L$ <i>Re-Encoding</i> Matrix (Inverse Decodermatrix)
d_{Head}	Kopfdurchmesser in [m]
d_{Mic}	Mikrophonabstand in [m]
\mathbf{D}	$N \times L$ Decodermatrix
$d(\Theta_0, \varphi_0, \omega)$	Ausgangssignal des <i>beamformer</i> , welcher in die Richtung (Θ_0, φ_0) zeigt
\mathbf{e}	Schallenergievektor
$EQ_m(ka)$	Radiale <i>equalizing</i> -Funktionen mit $0 \leq m \leq M$ zur Messung der Ambisonickomponenten bis zur M . Ordnung
$e_M(kr)$	<i>Truncation error</i> = Fehler, der bei der Berechnung eines Schallfeldes entsteht, welche mit Hilfe der Fourier-Bessel-Reihe durchgeführt wird. Der Index M bezeichnet dabei die maximale Ordnung, ab welcher die Fourier-Bessel-Reihe abgebrochen wird.
f_A	Abtastfrequenz in [Hz]
f	Frequenz in [Hz]
f_c	Eckfrequenz eines parametrischen Filters in [Hz]
f_{c_i}	Grenzfrequenz des i . <i>equalizing</i> -Filters für differentielle Mikrophone in [Hz]
$F_m(kr)$	Nahfeldmodellierungsfunktion
$F_m^{\left(\frac{r}{c}\right)}(j\omega)$	Nahfeldmodellierungsfunktion, wobei der hochgestellte Index $\left(\frac{r}{c}\right)$ der Latenzzeit in [s] entspricht, die sich aus der endlichen Schallgeschwindigkeit c in $\left[\frac{\text{m}}{\text{s}}\right]$ und dem Abstand vom Koordinatenursprung r in [m] ergibt
$g_k(\mathbf{r} \mid \mathbf{r}_{\text{SRC}})$	Greensche Funktion, welche die Übertragung zwischen der Quelle (\mathbf{r}_{SRC}) und dem Beobachtungspunkt (\mathbf{r}) beschreibt
G_i	Verstärkungsfaktor des i . Lautsprechers
$G(\omega)$	Array-Gewinn
$H_m^{(1)}(z)$	Hankelfunktion 1. Art, m . Ordnung des Arguments z (Besselfunktion 3. Art, m . Ordnung des Arguments z für austretende (zylindrische) Wellen)
$h_m^{(1)}(z)$	Sphärische Hankelfunktion 1. Art, m . Ordnung des Arguments z (sphärische Besselfunktion 3. Art, m . Ordnung des Arguments z für austretende (zylindrische) Wellen)

$H_m^{(2)}(z)$	Hankelfunktion 2. Art, m . Ordnung des Arguments z (Besselfunktion 3. Art, m . Ordnung des Arguments z für eintretende (zylindrische) Wellen)
$h_m^{(2)}(z)$	Sphärische Hankelfunktion 2. Art, m . Ordnung des Arguments z (sphärische Besselfunktion 3. Art, m . Ordnung des Arguments z für eintretende (zylindrische) Wellen)
$H_q(\Theta_q, \varphi_q, \omega)$	Rauschsignal des q . Mikrophons eines <i>beamformer</i> , verursacht durch dessen Eigenrauschen
j	Imaginäre Einheit von komplexen Zahlen ($j = \sqrt{-1}$)
$J_m(z)$	Besselfunktion 1. Art, m . Ordnung des Arguments z
$j_m(z)$	Sphärische Besselfunktion 1. Art, m . Ordnung des Arguments z
k	Wellenzahl ($k = \frac{2\pi f}{c} = \frac{\omega}{c}$)
\mathbf{k}	Wellenvektor
$K(\omega)$	Suszeptibilität
K^{2D}	Anzahl der Ambisonickomponenten $B_{m,n}^\sigma$ für die Rekonstruktion eines zweidimensionalen Wellenfeldes bis zur M . Ordnung wobei gilt: $K^{2D} = (2M + 1)$
K^{3D}	Anzahl der Ambisonickomponenten $B_{m,n}^\sigma$ für die Rekonstruktion eines dreidimensionalen Wellenfeldes bis zur M . Ordnung wobei gilt: $K^{3D} = (M + 1)^2$
L	Anzahl der Lautsprecher
l	Quellenabmessung (Größe der Quelle) in [m]
M	Anzahl der Mikrophone
M	Maximale Ordnung eines HOA-Systems
N	Anzahl der Ambisonicsignale bzw. der Ambisonickanäle
p	Index zur Kennzeichnung der Polyederflächen
$p(\mathbf{r}, t)$	Schalldruck am Ort \mathbf{r} zum Zeitpunkt t
$P_n(z)$	Legendrepolynom n . Ordnung
$P_{m,n}(z)$	Legendrefunktion n . Ordnung, m . Grades
$\tilde{P}_{m,n}(z)$	Schmidt-semi-normalisierte Version der Legendrefunktion
$P_{M+1}^{-1}(0)$	Größte Nullstelle eines Legendrepolynoms M . Ordnung innerhalb eines Wertebereichs von ± 1
$p_a(\Theta_q, \varphi_q)$	Schalldruck auf einer Kugel mit Radius a an der Position (Θ_q, φ_q)
Q	Anzahl der Mikrophone im Array
q	Richtfaktor (<i>directivity factor</i>) $(0, \dots, 2)$
$q_{\mathbf{v}\text{Opt}}$	Optimaler Richtfaktor des Schallschnellektors \mathbf{v}
$q_{\mathbf{e}\text{Opt}}$	Optimaler Richtfaktor des Schallenergievektors \mathbf{e}
$r_{\mathbf{v}}(q)$	Länge des Vektors \mathbf{v} in Abhängigkeit des Richtfaktors q
$r_{\mathbf{e}}(q)$	Länge des Vektors \mathbf{e} in Abhängigkeit des Richtfaktors q
\mathbf{r}	Ortsvektor zum Beobachtungspunkt

\mathbf{r}_S	Ortsvektor zum Lautsprecher S
\mathbf{r}_{SRC}	Ortsvektor zur Schallquelle (S ource = SRC)
r_S	Radius, d.h. Abstand vom Koordinatenursprung zum Lautsprechers S in [m]
r_{SRC}	Radius, d.h. Abstand vom Koordinatenursprung zur Schallquelle in [m]
$\sphericalangle(\mathbf{a} - \mathbf{b})$	Winkeldifferenz zwischen den Vektoren \mathbf{a} und \mathbf{b}
$\mathbf{a} \cdot \mathbf{b}$	Skalarprodukt der Vektoren \mathbf{a} und \mathbf{b} ($\mathbf{a} \cdot \mathbf{b} = \mathbf{a} \mathbf{b} \cos(\sphericalangle(\mathbf{a}, \mathbf{b}))$)
S_k	Signal der k . virtuellen Quelle
T_i	i . Verzögerungselement in [s]
$T_m(ka)$	Regularisierungsfunktionen mit $0 \leq m \leq M$ zur Beschränkung der M <i>equalizing</i> -Filter $EQ_m(ka)$
T_{60}	Nachhallzeit = Zeit in [s] die vergeht, bis der Schalldruck, ausgehend von seinem Maximum, um 60 [dB] gefallen ist.
\mathbf{v}	Schallschnellevektor
V	Volumen in [m ³]
$W_i(\omega)$	i . <i>equalizing</i> -Tiefpassfilter für differentielle Mikrophone
$W_m(ka)$	Radiale Funktionen, mit $0 \leq m \leq M$ zur Messung der Ambisonickomponenten bis zur M . Ordnung
$X(j\omega)$	Fouriertransformierte eines Zeitsignal $x(t)$
y_l	Ansteuersignal des l . Lautsprechers im Zeitbereich
$Y_l(j\omega)$	Ansteuersignal des l . Lautsprechers im Spektralbereich
\mathbf{Y}	Vektor, welcher alle L Lautsprecher Ansteuersignale beinhaltet
$Y_{m,n}^\sigma(\Theta, \varphi)$	Winkelfunktion bzw. sphärischen Harmonische m . Ordnung, n . Grades mit hochgestellten Index $\sigma \in [-1, +1]$ zur Unterscheidung des imaginären und reellen Anteils
$Y_{m,n}^\sigma(\Theta, \varphi)^{(2D)}$	Sphärischen Harmonische für den zweidimensionalen (2D) Fall
$Y_{m,n}^\sigma(\Theta, \varphi)^{(3D)}$	Sphärischen Harmonische für den dreidimensionalen (3D) Fall
$Y_m(z)$	Neumannfunktion m . Ordnung des Arguments z (Besselfunktion 2. Art, m . Ordnung des Arguments z)
$y_m(z)$	Sphärische Neumannfunktion m . Ordnung des Arguments z (sphärische Besselfunktion 2. Art, m . Ordnung des Arguments z)
Z	Akustische (Feld/Kenn)Impedanz in $[\frac{\text{N}\cdot\text{s}}{\text{m}^3}]$ ($Z = \rho c$)
α	Richtcharakteristikparameter für den gilt: $0 \leq \alpha \leq 1$
α	Glättungskoeffizient
β	Regularisierungsparameter
γ	Winkeldifferenz zwischen zwei Vektoren
γ_l	Differenz des Winkels zwischen dem Vektor zum Lautsprecher l ($\mathbf{r}_{S_l} = (r_{S_l}, \Theta_{S_l}, \varphi_{S_l})$), mit $l \in [1, \dots, L]$ und dem zur Schallquelle i
γ_{Mic}	Winkeldifferenz zwischen zwei benachbarten Mikrophenen in [rad]

$\delta_{i,j}$	Kronecker-Delta-Funktion
ϵ	Regularisierungsparameter
κ	Skalierungsfaktor $\kappa > 1$ mit dem der Radius der soliden Kugel a gewichtet werden muss, so dass gilt: $r_{\text{Mic}} = \kappa a$
λ	Wellenlänge in [m]
λ_n	n . Eigenwert einer Matrix
ρ	Dichte des Mediums (Luft) in $\left[\frac{\text{kg}}{\text{m}^3}\right]$
τ_d	Verzögerungszeit in [s]
Θ	Horizontaler Einfallswinkel gerichteter Schallwellen
$\tilde{\Theta}$	Horizontale Empfangsrichtung von Mikrofonarrays
$\tilde{\Theta}$	Empfundene Horizontallokalisation
Θ_0	Hauptempfangsrichtung in [Grad]
$\Theta_{-\infty}$	Nullstelle der Richtcharakteristik eines Mikrofonarrays in [Grad]
Θ_S	Azimut des Lautsprechers S
Θ_{SRC}	Azimut der Schallquelle (S ource = SRC)
Θ_q	Azimut des q . Mikrophons des Kugelmikrofonarrays
φ	Vertikaler Einfallswinkel gerichteter Schallwellen
$\hat{\varphi}_l$	Vertikaler Aufstellwinkel des l . Lautsprechers
φ_{SRC}	Elevation der Schallquelle (S ource = SRC)
φ_S	Elevation des Lautsprechers S
φ_q	Elevation des q . Mikrophons des Kugelmikrofonarrays
$\Psi(\omega, \Theta)$	Richtdiagramm eines Mikrofonarrays
ω	Kreisfrequenz in $\left[\frac{1}{\text{s}}\right]$ ($\omega = 2\pi f$)
ω_k	Diskrete Kreisfrequenz in $\left[\frac{1}{\text{s}}\right]$ ($\omega_k = \frac{2\pi k}{N} f_A$) an der Frequenzstelle k , bei einer FFT-Länge von N und einer Abtastfrequenz von f_A ([Hz])
Δ	Laplace-Operator
∇	Nabla-Operator
\mathbb{R}	Menge der reellen Zahlen
\mathbb{R}_0^+	Menge der positiven, reellen Zahlen inklusive der Null
\mathbb{R}^+	Menge der positiven, reellen Zahlen ohne Null
\mathbb{Z}	Menge der ganzen Zahlen
\mathbb{C}	Menge der komplexen Zahlen
\mathbb{N}	Menge der natürlichen Zahlen

Kapitel 1

Einleitung

1.1 Motivation

Diese Arbeit wurde mit dem Ziel begonnen, ein Verfahren zu finden, mit welchem sich die Akustik einer bestimmten Position im Raum „dokumentieren“, also speichern und zu einem späteren Zeitpunkt und an einem anderen Ort möglichst originalgetreu reproduzieren lässt. Dieser Vorgang, der auch als *Auralisation* bezeichnet wird, lässt sich bereits recht gut mit Hilfe von kopfhörerbasierten Algorithmen, wie z.B. dem BRS¹-Verfahren, welches bereits in [HKP⁺99] vorgestellt wurde, verwirklichen. Kopfhörerbasierte Auralisationsverfahren liefern nach wie vor die mit Abstand besten akustischen Ergebnisse und bieten sich daher für Vergleiche, z.B. mit den in dieser Arbeit behandelten lautsprecherbasierten Verfahren, als Referenz an, wobei in diesem Zusammenhang explizit die Methode aus [Gri09] erwähnt werden sollte. Zwar sind lautsprecherbasierte Verfahren schwieriger zu beherrschen, bieten aber gegenüber kopfhörerbasierten Varianten einige Vorteile, wie z.B. eine realistische Wiedergabe tieffrequenter Schallanteile (Bass-Performance) oder der Möglichkeit eine Auralisation gleichzeitig für mehrere Personen zu gewähren (großer räumlicher Wirkungsbereich), um nur einige zu nennen. Im Rahmen dieser Arbeit wurden dabei folgende Verfahren näher untersucht:

- das inverse Filter-Verfahren,
- das Ambisonic-Verfahren 1. Ordnung sowie
- das Ambisonic-Verfahren höherer Ordnung (HOA²).

Aufgrund dessen, dass in dieser Arbeit das HOA-Verfahren genau untersucht wird und der Tatsache, dass sich dieses theoretisch in das Verfahren der Wellenfeldsynthese (WFS³)

¹BRS = engl. *Binaural Room Scanning*

²HOA = engl. *Higher Order Ambisonic*

³WFS = engl. *Wave Field Synthesis*

umrechnen lässt (siehe z.B. [DMN03, Pol05b]), wurde hier das WFS-Verfahren nicht weiter behandelt. Die Erklärung hierfür liegt in dem Faktum, dass die Wellenfeldsynthese ihre Vorteile bei der Reproduktion von Wellenfeldern in einem größeren Hörbereich hat, wohingegen das HOA-Verfahren ihre Stärke speziell bei der Wiedergabe an einem bestimmten Punkt im Raum, dem sogenannten *sweet spot* hat, an dem dann in einem wesentlich größeren Spektralbereich die Nachbildung des gewünschten Wellenfeldes gelingt. Da letzteres in wesentlich besserer Übereinstimmung mit dem hier gesteckten Vorgaben, eine lautsprecherbasierte Auralisation speziell für Fahrzeuginnenräume zu entwickeln, ist, wobei die Positionen im Raum bekanntermaßen durch die Sitzpositionen vorgegeben und somit bekannt sind, wurde die Entscheidung getroffen sich in dieser Arbeit ausschließlich auf das HOA-Verfahren und nicht auf das WFS-Verfahren zu konzentrieren, was den Hauptteil dieser Arbeit darstellt.

1.2 Gliederung der Arbeit

In Kapitel 2 wird auf die inverse Filter-Theorie eingegangen, wobei in unserem Fall der Schwerpunkt nicht in der Verwendung von adaptiven Systemen, wie etwa in [Gar00] liegt, sondern in der Realisierung der transauralen Wiedergabe, wie aus [CB89] bekannt. Zunächst wird das Prinzip der transauralen Wiedergabe am einfachsten Beispiel, d.h. am Beispiel des transaural-Stereo aufgezeigt. Anschließend werden Optimierungsmaßnahmen wie z.B. eine optimale Lautsprecheraufstellung (OSD⁴) diskutiert, bevor auf Besonderheiten bei der Berechnung der inversen Filtermatrix (adjungierte Matrix) sowie auf die Notwendigkeit einer Regularisierung beim Filterentwurf eingegangen wird. Ein Beispiel verdeutlicht dabei die Wirkungsweise der in dieser Arbeit neu entwickelten Regularisierung. Zudem eröffnet es Einblicke über praktische Grenzen dieses Verfahrens.

Kapitel 3 dient als Einstieg in den Schwerpunkt dieser Arbeit, welches die Vorstellung des Ambisonic-Verfahrens umfasst. Dabei wird das Ambisonic-Verfahren am einfachsten Beispiel, d.h. an einem System 1. Ordnung vorgestellt. Begonnen wird dabei mit der Interpretation der B-Format-Signale (W, X, Y, Z), welche die Grundlage für die messtechnische Bestimmung, d.h. für die Decodierung eines beliebigen Wellenfeldes, darstellen. Dabei erfasst das omnidirektionale Mikrofon des Klangfeldmikrophons die 0. sphärische Harmonische (W -Signal), wohingegen die drei verbleibenden, jeweils senkrecht zueinander stehenden, bidirektionalen Mikrofone des Klangfeldmikrophons die drei sphärischen Harmonischen 1. Ordnung, d.h. die X -, Y - und Z -Signale aufnehmen. Dies kann auch so interpretiert werden, dass das W -Signal den Schalldruck am Messort liefert, wohingegen die drei verbleibenden X -, Y - und Z -Signale die Schallschnelle (gerichtete Größe) in allen drei Raumrichtungen aufzeichnen, aus denen sich dann die dreidimensionale Einfallrichtung des Schalls bestimmen lässt. Mit diesem Wissen ist es dann ein Leichtes, auch virtuelle Signale aus einer beliebigen, dreidimensionalen Raumrichtung zum Abhör-

⁴OSD = engl. *Optimal Source Distribution*

ort einfallen zu lassen, was nichts anderes als eine Codierung des Wellenfeldes darstellt. Im anschließenden Codier/Decodier-Beispiel wird unter anderem verdeutlicht, dass sich die, bei der Codierung generierten Gewichtungsfaktoren für die Lautsprecher im Zielraum, was auch als *panning* bezeichnet wird, auch sehr anschaulich als Richtdiagramme darstellen lassen und somit auf plastische Weise aufzeigen, wie gut sich virtuelle Quellen mit dem vorhandenen System räumlich zuordnen lassen. Schließlich werden unterschiedliche Optimierungsverfahren (Länge des Schallschnellektors \mathbf{v} bzw. Länge des Schallenergievektors \mathbf{e}) vorgestellt und Methoden, wie zwischen unterschiedlichen Optimierungsverfahren über der Frequenz mit Hilfe z.B. von *shelving*-Filter umgeschaltet werden kann, beschrieben. Mit einer Diskussion bezüglich der Wirkungsweise von Fensterfunktionen, wie z.B. das Hann-Fenster, auf die *panning*-Funktion schließt dieses Kapitel.

In Kapitel 4 wird nun das im vorherigen Kapitel aufgezeigte Ambisonic-Prinzip verallgemeinert, d.h. auch für Ordnungen von $M > 1$ aufgezeigt. Die Decodierung, d.h. die Erzeugung eines zuvor künstlich oder durch Messung codierten Wellenfeldes, ist dabei auch für höhere Ordnungen bereits gut verstanden. So existieren unterschiedliche Decodervarianten etwa zur Optimierung der Lokalisation im *sweet spot*, d.h. der interauralen Laufzeitdifferenz (ITD⁵), was durch den sogenannten Basic- oder Max r_v -Decoder geschieht, für die Maximierung des Schallenergievektors \mathbf{e} (Max r_e -Decoder) bzw. zur Vergrößerung der Fläche im Wiedergabesystems innerhalb dessen die Lokalisation erhalten bleibt (In-Phase-Decoder). Auch eine Erzeugung von innenliegenden Quellen, d.h. sich innerhalb des Abhörsystems (z.B. Lautsprecherring) befindlichen Quellen und somit auch von unterschiedlichen Abständen zur Quelle (Bühnenabstand), ist mit Einführung der *Nahfeldkompensation* möglich. Diese ist in dem aus [Dan03] bekannten, relativ neuen Codierverfahren namens HOA-NFC⁶ bereits intrinsisch berücksichtigt. Ferner können bei Verwendung von regulären Lautsprecheranordnungen Wellenfelder sowohl zwei- als auch dreidimensional reproduziert werden.

Anders gestaltet sich die Situation bei der Codierung natürlicher Wellenfelder, was in Kapitel 5 behandelt wird. Vor Forcierung des Ambisonicverfahrens waren auf diesem Gebiet so gut wie keine Methoden bekannt, mit deren Hilfe man ein beliebiges Wellenfeld messtechnisch erfassen und anschließend codieren konnte. Erst das Ambisonicverfahren, welches auf einer der Fourieranalyse bzw. Fouriersynthese nicht unähnlichen Zusammensetzung des Wellenfeldes anhand von sphärischen Harmonischen, die hier als Basisfunktionen agierten, basiert, konnten auch in dem Bereich der Codierung von Wellenfelder erhebliche Fortschritte erzielt werden. Mit Hilfe der Ambisonictheorie ist es nun möglich, Mikrophonarrays zu entwickeln, welche die Aufgabe der Codierung wesentlich besser erfüllen als vorhergehende Verfahren. Da dieser Bereich noch relativ neu ist, ist es nicht verwunderlich, dass hier noch viel Forschungs- und Verbesserungsbedarf herrscht. Erst wenn man ein Wellenfeld in ausreichender Genauigkeit messen, d.h. codieren kann, ist es

⁵ITD = engl. *Interaural Time Delay*

⁶HOA-NFC = engl. *Higher Order Ambisonic - Near Field Compensated*

möglich, dieses auch (weitestgehend unverfälscht) zu reproduzieren! Aus diesem Grund ist diesem Bereich, d.h. der messtechnischen Erfassung (Codierung) von Wellenfeldern mit Hilfe von Mikrophonarrays der Schwerpunkt dieser Arbeit gewidmet. Ein Teil, der in diesem Kapitel behandelten Verbesserungen, widmet sich dem Entwurf von Regularisierungsfiltren, welche aufgrund physikalischer Grenzen z.B. der Sensoren (Eigenrauschen der Mikrophone) bzw. der beim Bau des Mikrophonarrays verwendeten Mechanik (Abweichungen der Sensorpositionen) berücksichtigt werden müssen. Zudem werden in diesem Kapitel unterschiedliche Ansätze zur Erhöhung der räumlichen Abtastfrequenz bzw. der Realisierung einer räumlichen Tiefpassfilterung diskutiert, was als Grundlage zukünftiger Arbeiten betrachtet werden kann.

Im Kapitel 6 werden schließlich zwei unterschiedliche, sphärische Mikrophonarrays in der Praxis untersucht. Eines der dabei behandelten Probleme betrifft eine neuartige, rein mechanische Form der Umsetzung einer räumlichen Tiefpassfilterung zur Erhöhung der räumlichen Abtastfrequenz, welche auf einem geschlossenen Kugelmikrophon basiert. Ein weiteres Problem befasst sich mit einer einfach skalierbaren, aufwandsgünstigen Realisierung eines offenen Kugelmikrophonarrays. Erst mit einer erfolgreichen, messtechnisch basierten Methode zur Codierung eines beliebigen, natürlichen Wellenfeldes an einem bestimmten Ort im Quellraum, kann die Voraussetzung für eine erfolgreiche Auralisation geschaffen werden und somit das anfängliche Ziel dieser Arbeit erreicht werden.

Im Kapitel 7 werden die Ergebnisse dieser Arbeit zusammengefasst und Wege aufgezeigt, welche Möglichkeiten der Verbesserung bzw. Verfeinerung noch existieren, die in dieser Arbeit nicht behandelt wurden.

Im Kapitel A befinden sich Erläuterungen, die für das Verständnis der Arbeit wichtig sind, jedoch nicht unmittelbar zu den Grundlagen gezählt werden können. Hier wird etwa auf die Berechnung differentieller Mikrophone 1. Ordnung sowie auf die Berechnung diverser Kennwerte von *beamformer* bzw. Mikrophonarrays, wie etwa der Suszeptibilität eingegangen; Grundlagen für die Berechnung der in dieser Arbeit verwendeten Matrizen dargeboten; Besselfunktionen, insbesondere sphärische Besselfunktionen vorgestellt; einen Überblick über den Entwurf modaler *beamformer* gegeben und schließlich Eigenschaften regulärer und quasi-regulärer Körper, wie etwa die von platonischen Körpern, erörtert.

Kapitel 2

Inverse Filter

Die Aufgabe eines *inversen Filter* $C(j\omega)$ ist es, eine beliebige Übertragungsfunktion $H(j\omega)$ so zu entzerren, dass dessen Einfluss neutralisiert wird, d.h.

$$C(j\omega) H(j\omega) = e^{-j\omega\tau_d}, \text{ wobei}$$

$$\tau_d \geq 0, \text{ mit}$$

$$\tau_d = \text{Verzögerungszeit (engl. } \mathbf{modelling\ delay} \text{) in [s], die mindestens die akustische Wegstrecke von der Quelle (Lautsprecher) zur Senke (Mikrofon) abdecken muss.} \quad (2.1)$$

In unserem Fall werden inverse Filter zur *Raumentzerrung*, was auch als *equalizing* bezeichnet wird, verwendet. Theoretisch kann dies an beliebig vielen Punkten im Raum geschehen, sofern es sich zumindest um ein *bestimmtes* bzw. um ein *überbestimmtes* System handelt. Dies bedeutet, dass sich im Zielraum mehr Lautsprecher L (Quellen) als Mikrophone M (Senken), also Punkte an denen die Entzerrung erreicht werden soll, befinden müssen, d.h. $L \geq M$.

Abbildung 2.1 veranschaulicht dieses Prinzip in Anlehnung an die Erkenntnisse aus [Gar00], wobei hier *adaptive Filter*¹, z.B. in Form des MELMS² oder eines anderen Adaptionsverfahrens, verwendet werden, um die Koeffizienten der inversen Filtermatrix $\mathbf{C}(j\omega)$ iterativ zu bestimmen.

Ist man lediglich an einer binauralen Reproduktion, d.h. an einer Entzerrung an den Orten der beiden Ohren interessiert, spricht man meist nicht mehr von *inversen Filtern*, sondern von CTC³-Filter. Diese Filter sollen einerseits das Übersprechen vom linken Kanal auf das rechte Ohr und vom rechten Kanal auf das linke Ohr unterdrücken. Andererseits besteht ihre Aufgabe aber auch darin, die Übertragung vom linken Kanal zum linken Ohr

¹Ein *adaptives Filter* bezeichnet in der digitalen Signalverarbeitung ein digitales Filter welches seine Übertragungsfunktion selbständig, auf sich ändernde äußere Gegebenheiten, wie z.B. eine sich ändernde Raumimpulsantwort (RIR), einstellen kann.

²MELMS = engl. *Multiple Error Least Mean Square*

³CTC = engl. *Crosstalk Cancellation*

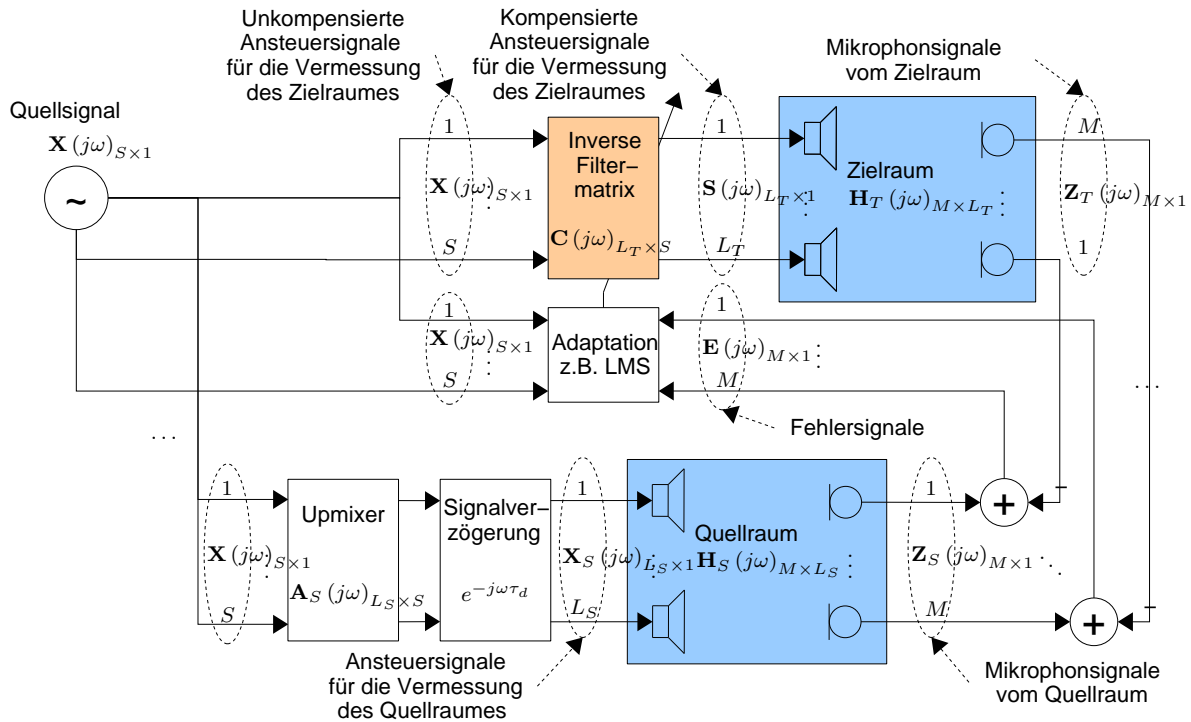


Abbildung 2.1: Darstellung des Prinzips der adaptiven, inversen Filterung.

sowie vom rechten Kanal zum rechten Ohr zu entzerren, was bereits in Gleichung 2.1 beschrieben wurde.

Aus [CB89] ist bekannt, dass man eine *transaurale Wiedergabe*⁴ mit zwei oder mehr Lautsprechern, deren Ansteuersignale zuvor mit Hilfe eines entsprechend entworfenen Filternetzwerks $C(j\omega)$, wobei man hier, wie bereits erwähnt, nun von CTC-Filter spricht, vorverzerrt wurden, realisieren kann.

Abbildung 2.2 verdeutlicht das Prinzip der transauralen Wiedergabe am *transaural-Stereo*⁵.

Bei genauer Betrachtung von Abbildung 2.2 erkennt man, dass sich die Ohrsignale $Z_L(j\omega)$ und $Z_R(j\omega)$ wie folgt aus den Eingangssignalen $X_L(j\omega)$ und $X_R(j\omega)$ ergeben:

⁴Der Begriff *transaurale Wiedergabe* bezeichnet ein Verfahren bei dem versucht wird, Eingangssignale, die über zwei oder mehr Lautsprecher in einen beliebigen Raum abgespielt werden, mit Hilfe eines vorgeschalteten Filternetzwerks so zu verändern, dass an den Ohren binaurale Signale ankommen, also Signale, die „perfekte“ Kopfhörer an den Ohren erzeugen würden.

⁵*Transaural-Stereo* bezeichnet eine transaurale Wiedergabe bei der ausschließlich zwei Lautsprecher verwendet werden.

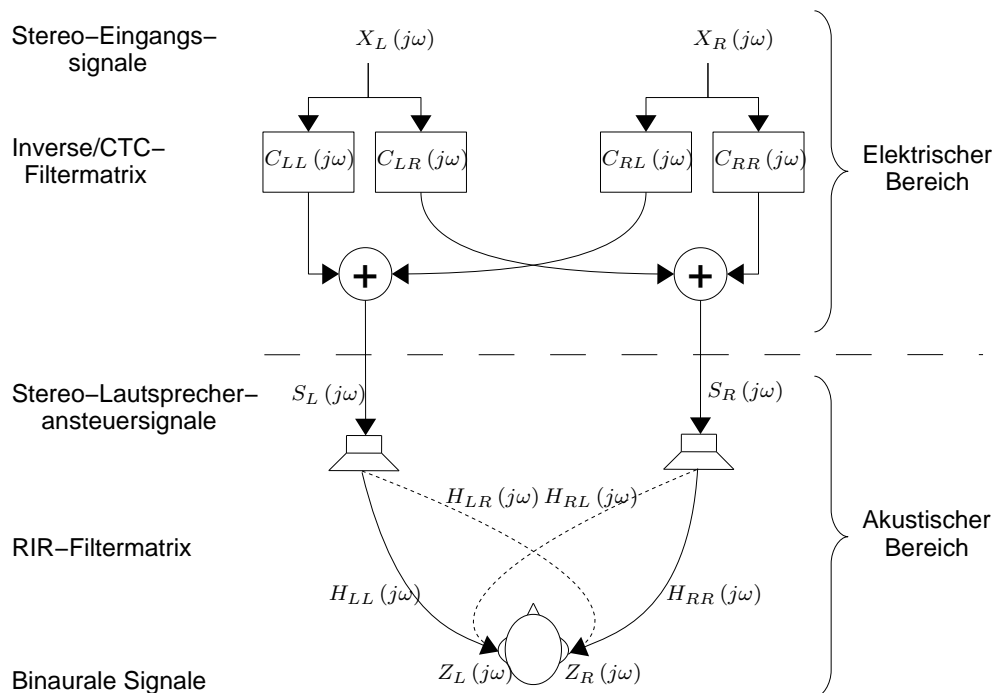


Abbildung 2.2: Darstellung des Prinzips der Übersprechdämpfung zur transauralen Wiedergabe eines Stereosignals.

(a) Berechnung der Ohrsignale:

$$Z_L(j\omega) = S_L(j\omega) H_{LL}(j\omega) + S_R(j\omega) H_{RL}(j\omega),$$

$$Z_R(j\omega) = S_L(j\omega) H_{LR}(j\omega) + S_R(j\omega) H_{RR}(j\omega), \text{ oder als Matrix:}$$

$$\begin{pmatrix} Z_L(j\omega) \\ Z_R(j\omega) \end{pmatrix} = \begin{pmatrix} H_{LL}(j\omega) & H_{RL}(j\omega) \\ H_{LR}(j\omega) & H_{RR}(j\omega) \end{pmatrix} \begin{pmatrix} S_L(j\omega) \\ S_R(j\omega) \end{pmatrix}, \text{ bzw. in kompakter Form:}$$

$$\mathbf{Z}(j\omega) = \mathbf{H}(j\omega) \mathbf{S}(j\omega), \text{ mit}$$

$\mathbf{Z}(j\omega) = 2 \times 1$ -Matrix, welche die beiden Ohrsignale $Z_L(j\omega)$ und $Z_R(j\omega)$ enthält,

$\mathbf{H}(j\omega) = 2 \times 2$ -Matrix, welche die Übertragungsfunktionen der RIRs⁶ von den beiden Lautsprecheransteuersignalen $S_L(j\omega)$ und $S_R(j\omega)$ zu den beiden Ohrsignalen $Z_L(j\omega)$ und $Z_R(j\omega)$ enthält. (2.2)

⁶RIR = engl. **R**oom **I**mpulse **R**esponse

(b) Berechnung der Lautsprecheransteuersignale:

$$\begin{aligned}
S_L(j\omega) &= X_L(j\omega) C_{LL}(j\omega) + X_R(j\omega) C_{RL}(j\omega), \\
S_R(j\omega) &= X_L(j\omega) C_{LR}(j\omega) + X_R(j\omega) C_{RR}(j\omega), \text{ oder als Matrix:} \\
\begin{pmatrix} S_L(j\omega) \\ S_R(j\omega) \end{pmatrix} &= \begin{pmatrix} C_{LL}(j\omega) & C_{RL}(j\omega) \\ C_{LR}(j\omega) & C_{RR}(j\omega) \end{pmatrix} \begin{pmatrix} X_L(j\omega) \\ X_R(j\omega) \end{pmatrix}, \text{ bzw. in kompakter Form:} \\
\mathbf{S}(j\omega) &= \mathbf{C}(j\omega) \mathbf{X}(j\omega), \text{ mit} \\
\mathbf{X}(j\omega) &= 2 \times 1\text{-Matrix, welche die beiden Eingangssignale } X_L(j\omega) \text{ und} \\
&\quad X_R(j\omega) \text{ enthalt,} \\
\mathbf{C}(j\omega) &= 2 \times 2\text{-Matrix, welche die Ubertragungsfunktionen der CTC-} \\
&\quad \text{Filter enthalt, mit denen die beiden Eingangssignale } X_L(j\omega) \\
&\quad \text{und } X_R(j\omega) \text{ gefiltert werden mussen, bevor sie zu den} \\
&\quad \text{Lautsprecheransteuersignalen } S_L(j\omega) \text{ und } S_R(j\omega) \text{ werden.} \quad (2.3)
\end{aligned}$$

Setzt man nun die Ergebnisse der Gleichung 2.3 in Gleichung 2.2 ein, folgt:

$$\mathbf{Z}(j\omega) = \mathbf{H}(j\omega) \mathbf{C}(j\omega) \mathbf{X}(j\omega). \quad (2.4)$$

Eine korrekte, binaurale Reproduktion der Eingangssignale wird erzielt, wenn gilt:

$$\mathbf{Z}(j\omega) = \mathbf{X}(j\omega) e^{-j\omega\tau_d}. \quad (2.5)$$

Somit lassen sich die CTC-Filter $\mathbf{C}(j\omega)$ direkt aus Gleichung 2.4 und 2.5 wie folgt bestimmen:

$$\begin{aligned}
\mathbf{C}(j\omega) &= \mathbf{H}^{-1}(j\omega) e^{-j\omega\tau_d}, \text{ mit} \\
\mathbf{H}^{-1}(j\omega) &= \text{Inverse der RIR-Matrix } \mathbf{H}(j\omega). \quad (2.6)
\end{aligned}$$

Allerdings kann $\mathbf{H}(j\omega)$ in Gleichung 2.6 nur dann invertiert werden, wenn es sich um eine *quadratische Matrix* handelt. Dies ist der Fall, wenn, wie im obigen Beispiel, lediglich zwei Lautsprecher fur die Ubersprechkompensation verwendet werden, was jedoch nicht zwingend erforderlich ist. Werden mehr als zwei Lautsprecher fur die Ubersprechkompensation verwendet, ist $\mathbf{H}(j\omega)$ nicht mehr symmetrisch und kann somit nicht mehr direkt invertiert werden. In diesem Fall kann dessen Inverse und somit auch die benotigten inversen Filter nur noch approximativ mit Hilfe der sogenannten *Pseudoinversen*, auf die genauer in Kapitel 4.2.4 eingegangen wird, berechnet werden.

Durch Verwendung der Pseudoinversen lasst sich die Berechnung der fur die Realisierung der Ubersprechdampfung benotigten inversen bzw. CTC-Filter wie folgt verallge-

meinern:

$$\begin{aligned} \mathbf{C}(j\omega) &= \mathbf{H}^+(j\omega) e^{-j\omega\tau_d}, \text{ mit} \\ \mathbf{H}^+(j\omega) &= \text{Pseudoinverse der RIR-Matrix } \mathbf{H}(j\omega), \text{ wobei gilt:} \\ \mathbf{H}^+(j\omega) &= (\mathbf{H}^H(j\omega) \mathbf{H}(j\omega))^{-1} \mathbf{H}^H(j\omega). \end{aligned} \quad (2.7)$$

2.1 Optimale Lautsprecheraufstellung

Untersuchungen aus [TTN01] ergaben, dass die Qualität der CTC-Methode sowohl von der Menge als auch von der Art, wie die verwendeten Lautsprecher im Raum verteilt sind, abhängt.

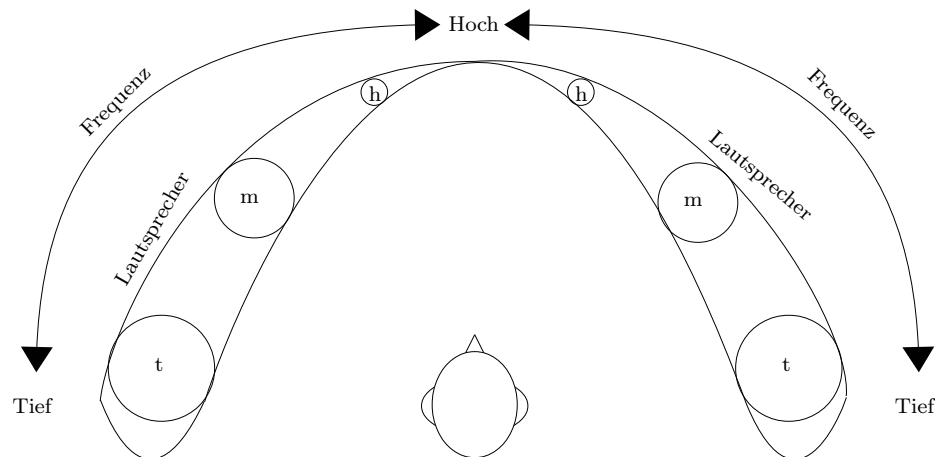


Abbildung 2.3: Darstellung der optimalen, frequenzabhängigen Lautsprecherverteilung (h =Hochtöner, m =Mitteltöner, t =Tieftöner) für die transaurale Wiedergabe eines Stereosignals.

In [TTN01] wurde festgestellt, dass über der Frequenz eine maximale Übersprechdämpfung erreicht werden kann, wenn eine Lautsprecherverteilung verwendet wird, wie sie schematisch in Abbildung 2.3 dargestellt ist.

Dabei spricht man von einer *optimalen Lautsprecherverteilung* (OSD⁷), wenn die für die Übersprechkompensation verwendeten, symmetrisch zur Medianebene angeordneten Lautsprecherpaare mit steigender Frequenz einen immer größer werdenden, horizontalen Aufstellwinkel aufweisen. Als Ergebnis einer solchen OSD erhält man, wie bereits erwähnt, eine maximierte Übersprechdämpfung, die zudem mit relativ einfach zu realisierenden CTC-Filter erreicht werden kann, da diese, bedingt durch die OSD, lediglich eine geringe Dynamik aufweisen. Zudem vergrößert sich durch die OSD der *sweet spot*, also der Raum, in dem die Übersprechdämpfung wirkt.

⁷OSD = engl. **O**ptimal **S**ource **D**istribution

Aus [PR10, MFV11] ist ferner bekannt, dass sich der *sweet spot* weiter steigern lässt, wenn die optimal verteilten Lautsprecher zusätzlich in der Höhe, also in einem Elevationswinkel von $\varphi > 0^\circ$, angebracht werden.

Eine weitere Erkenntnis aus [TTN01] ist, dass für die Übersprechkompensation tiefer Frequenzen von ca. $f < 400$ [Hz] bei normaler Kopfgröße sehr hohe Energien benötigt werden, um überhaupt eine CTC-Wirkung zu erzielen, die sich zudem auf einen sehr kleinen, örtlichen Bereich beschränkt. Glücklicherweise ist unsere Fähigkeit zur Lokalisierung in diesem Bereich, aufgrund der kleinen ILD⁸, welche sich wegen der großen Wellenlängen ergibt und der sich erst aufbauenden ITD⁹, ziemlich schlecht, weshalb im Prinzip hier ohnehin keine CTC benötigt wird.

2.2 Modifizierungen beim Entwurf der inversen Filter

Theoretisch werden mit Hilfe der bislang vorgestellten Verfahren zur Bestimmung der inversen Filter, gegeben z.B. durch Gleichung 2.6 bzw. verallgemeinert durch Gleichung 2.7, ideale Ergebnisse erzeugt, die so in der Praxis jedoch nicht haltbar sind.

Hintergrund dabei ist, dass die inversen Filter meist über eine viel zu hohe Dynamik verfügen, was sich damit erklärt, dass die RIR-Filtermatrix $\mathbf{H}(j\omega)$ Übertragungsfunktionen enthält, die sehr starke, spektrale Einbrüche aufweisen.

Diese meist schmalbandigen, spektralen Einbrüche innerhalb der Raumimpulsantworten resultieren aus der Überlagerung des Direktschalls mit den vom Raum abhängigen Reflexionen. Dabei kommt es bei einer gleichphasigen Überlagerung zu einer Erhöhung des Pegels und bei einer gegenphasigen Superposition zu einer Verringerung des Pegels, die bis zur vollkommenen Auslöschung reichen kann. Da diese Überlagerungen jedoch mit beliebiger Phasenlage auftreten können, führt dies am jeweiligen Messort zu einer Veränderung des dort vorherrschenden Pegels, was gemeinhin als *Kammfiltereffekt* bezeichnet wird.

Ein Filter, welches diese schmalbandigen Einbrüche, die, wie erwähnt sehr tief sein können, kompensieren soll, muss zwangsweise sehr hohe, schmalbandige Anhebungen, d.h. eine sehr hohe Dynamik, aufweisen. Einerseits können solch hohen Leistungen, die durchaus in den Bereich von mehreren 100 [dB] liegen können, in einem praktischen System nicht realisiert werden, da es natürliche Grenzen, z.B. vom Netzteil, von den Lautsprechern usw. gibt, die eine Umsetzung unmöglich machen. Andererseits werden hohe, schmalbandige Anhebungen akustisch als störend empfunden, was in diesem Zusammenhang jedoch vorrangig in der Tatsache begründet ist, dass solche Filter ein ausgeprägtes *preringing*¹⁰ aufweisen. Dieses hier als *preringing* bezeichnete Phänomen resultiert beim Entwurf inverser Filter aus der Tatsache, dass deren Impulsantwort einen *akausalen* Teil aufweist,

⁸Das Akronym ILD steht für *Interaural Level Difference* und bezeichnet die interaurale Pegeldifferenz.

⁹Das Akronym ITD steht für *Interaural Time Difference* und bezeichnet die interaurale Zeitdifferenz.

¹⁰Der Begriff *preringing* (Vorläuten), was in der Akustik zuweilen auch als *temporal diffusion* bezeichnet wird, beschreibt akustisch wahrnehmbare Geräusche, die *vor* dem eigentlichen Nutzsignal eintreffen.

d.h. vor Erscheinen des Hauptimpulses bereits Filterkoeffizienten mit Betragswerten > 0 vorhanden sind.

Da theoretisch perfekte inverse Filter, die etwa nach Gleichung 2.6 oder Gleichung 2.7 entworfen wurden, zwangsweise unter solchen Problemen leiden, ist es notwendig den Entwurf so zu modifizieren, dass die resultierenden Filter eine höhere Praxistauglichkeit erreichen.

2.3 Regularisierung

Aus [KNHOB98] ist eine Methode bekannt, welche eine *Regularisierung* im Spektralbereich vollzieht. Wie wir später noch sehen werden, ist eine der Aufgaben dieser Regularisierung, die Energie, mit welcher die L_T Lautsprecher im Zielraum angesteuert werden, auf einem maximalen, durch den sogenannten *Regularisierungsparameter* β festgelegten, Wert zu begrenzen. Somit lassen sich die Längen der inversen Filter $\mathbf{C}(j\omega)$ reduzieren, wodurch sich auch das *preringing* verringert.

Basierend auf dem in Abbildung 2.1 dargestellten Signalfloss lässt sich das inverse Filterprinzip mathematisch wie folgt zusammenfassen:

1. Bestimmung der M Mikrophonsignale im Quellraum $\mathbf{Z}_S(j\omega)_{M \times 1}$, welche als Referenz dienen:

$$\begin{aligned}
 \mathbf{Z}_S(j\omega)_{M \times 1} &= e^{-j\omega\tau_d} \mathbf{H}_S(j\omega)_{M \times L_S} \mathbf{X}_S(j\omega)_{L_S \times 1}, \text{ wobei} \\
 \mathbf{X}_S(j\omega)_{L_S \times 1} &= \mathbf{A}_S(j\omega)_{L_S \times S} \mathbf{X}(j\omega)_{S \times 1}, \text{ mit} \\
 \mathbf{X}(j\omega)_{S \times 1} &= \text{Vektor, welcher die } S \text{ Eingangssignale enthält,} \\
 \mathbf{A}_S(j\omega)_{L_S \times S} &= \text{Upmixing-Matrix, welche die } S \text{ Eingangssignale auf die } L_S \\
 &\quad \text{Lautsprecher des Quellraumes verteilt,} \\
 \mathbf{X}_S(j\omega)_{L_S \times 1} &= \text{Vektor, welcher die Ansteuersignale der } L_S \text{ Lautsprecher des} \\
 &\quad \text{Quellraumes beinhaltet,} \\
 \mathbf{H}_S(j\omega)_{M \times L_S} &= \text{Matrix, welche die Übertragungsfunktionen des Quellraumes} \\
 &\quad \text{zwischen den } L_S \text{ Lautsprechern und den } M \text{ dort} \\
 &\quad \text{angebrachten Mikrofonen enthält,} \\
 \tau_d &= \text{Modelling delay zur Erzeugung kausaler, inverser Filter,} \\
 \mathbf{Z}_S(j\omega)_{M \times 1} &= \text{Vektor, welcher die Signale der } M, \text{ sich im Quellraum} \\
 &\quad \text{befindlichen Mikrophone enthält.} \tag{2.8}
 \end{aligned}$$

2. Bestimmung der M Mikrophonsignale im Zielraum $\mathbf{Z}_T(j\omega)_{M \times 1}$, welche kontinuierlich gemessen werden können:

$$\begin{aligned}
\mathbf{Z}_T(j\omega)_{M \times 1} &= \mathbf{H}_T(j\omega)_{M \times L_T} \mathbf{S}(j\omega)_{L_T \times 1}, \text{ wobei} \\
\mathbf{S}(j\omega)_{L_T \times 1} &= \mathbf{C}(j\omega)_{L_T \times S} \mathbf{X}(j\omega)_{S \times 1}, \text{ mit} \\
\mathbf{C}(j\omega)_{L_T \times S} &= \text{Matrix, welche die Übertragungsfunktionen der } L_T \times S, \\
&\quad \text{inversen Filter enthält,} \\
\mathbf{S}(j\omega)_{L_S \times 1} &= \text{Vektor, welcher die Ansteuersignale der } L_T \text{ Lautsprecher} \\
&\quad \text{des Zielraumes beinhaltet (Ausgangssignale der inversen} \\
&\quad \text{Filter Matrix } \mathbf{C}(j\omega)_{L_T \times S}), \\
\mathbf{H}_T(j\omega)_{M \times L_T} &= \text{Matrix, welche die Übertragungsfunktionen des Zielraumes} \\
&\quad \text{zwischen den } L_T \text{ Lautsprechern und den } M \text{ dort} \\
&\quad \text{angebrachten Mikrofonen enthält,} \\
\mathbf{Z}_T(j\omega)_{M \times 1} &= \text{Vektor, welcher die Signale der } M, \text{ sich im Zielraum} \\
&\quad \text{befindlichen Mikrophone enthält.} \tag{2.9}
\end{aligned}$$

3. Bestimmung der M Fehlersignale $\mathbf{E}(j\omega)_{M \times 1}$:

$$\begin{aligned}
\mathbf{E}(j\omega)_{M \times 1} &= \mathbf{Z}_S(j\omega)_{M \times 1} - \mathbf{Z}_T(j\omega)_{M \times 1}, \text{ mit} \\
\mathbf{E}(j\omega)_{M \times 1} &= \text{Vektor der die } M \text{ Fehlersignale beinhaltet, welche Auskunft} \\
&\quad \text{über die Qualität der inversen Filter } \mathbf{C}(j\omega)_{L_T \times S} \text{ geben.} \tag{2.10}
\end{aligned}$$

Basierend auf den M Fehlersignalen $\mathbf{E}(j\omega)$ sowie auf den Ansteuersignalen der L_T Lautsprecher des Zielraumes $\mathbf{S}(j\omega)$ lässt sich gemäß [KNHOB98] eine Kostenfunktion $J(j\omega)$ wie folgt aufstellen:

$$\begin{aligned}
J(j\omega) &= \mathbf{E}^H(j\omega) \mathbf{E}(j\omega) + \beta \mathbf{S}^H(j\omega) \mathbf{S}(j\omega), \text{ mit} \\
J(j\omega) &= \text{Kostenfunktion der inversen Filterberechnung,} \\
\beta &= \text{Regularisierungsparameter,} \tag{2.11}
\end{aligned}$$

die es beim Entwurf der inversen Filter $\mathbf{C}(j\omega)$ zu minimieren gilt.

Wie man aus Gleichung 2.11 erkennen kann, ist $J(j\omega)$ rein reell, obwohl die beiden Vektoren $\mathbf{E}(j\omega)$ und $\mathbf{S}(j\omega)$ komplexe Signale beinhalten können.

Der erste Teil der Kostenfunktion $J(j\omega)$ aus Gleichung 2.11 stellt ein Maß für die Gesamtleistung der Fehlersignale $\mathbf{E}(j\omega)$ bei der (diskreten) Kreisfrequenz $\omega_{(k)}$ dar. Der Regularisierungsparameter β , der aus einem Bereich von $\beta = [0, \dots, \infty]$ stammt, gibt hingegen an, wie stark der zweite Teil in die Kostenfunktion einfließt. Dieser beinhaltet die Leistung der Ausgangssignale der inversen Filtermatrix $\mathbf{S}(j\omega)$, die mit Hilfe von β

auf ein gewünschtes, maximales Maß zu begrenzen sind. Je nachdem wie β gewählt wird, setzt sich die Kostenfunktion lediglich aus den Fehlersignalen $\mathbf{E}(j\omega)$ ($\beta = 0$) oder aus den den Ausgangssignalen der inversen Filtermatrix $\mathbf{S}(j\omega)$ ($\beta = \infty$) zusammen und wird somit entsprechend minimiert.

Nach [KNHOB98] ergeben sich die regularisierten, inversen Filter im Spektralbereich für jede (diskrete) Frequenz $\omega_{(k)}$ wie folgt:

$$\begin{aligned} \mathbf{C}(j\omega)_{L_T \times L_S} &= (\mathbf{H}_T^H(j\omega)_{M \times L_T} \mathbf{H}_T(j\omega)_{M \times L_T} + \beta \mathbf{I}_{L_S \times L_S})^{-1} \\ &\quad \mathbf{H}_T^H(j\omega)_{M \times L_T} \mathbf{H}_S(j\omega)_{M \times L_S} \mathbf{A}_S(j\omega)_{L_S \times S} e^{-j\omega\tau_d}, \text{ mit} \\ \mathbf{I}_{L_S \times L_S} &= \text{Einheitsmatrix der Größe } L_S \times L_S. \end{aligned} \quad (2.12)$$

Wird keine Regularisierung angewendet ($\beta = 0$), geht Gleichung 2.12 mit $\mathbf{H}_S(j\omega)_{M \times L_S} \mathbf{A}_S(j\omega)_{L_S \times S} = \mathbf{I}_{M \times S}$, d.h. bei Verwendung eines „neutralen“ Quellraumes, in die, aus Gleichung 2.7 bekannte Form über.

2.3.1 Wahl des Regularisierungsparameters β

Aus [FU99] ist bekannt, dass β über der Frequenz in Abhängigkeit von den Übertragungseigenschaften des bzw. der verwendeten Lautsprecher gewählt werden sollte.

In dem normalerweise vom Hersteller angegebenen Arbeitsbereich des Lautsprechers, in dem dieser mehr oder weniger linear arbeitet, sollte demnach ein kleiner Regularisierungsparameter β verwendet werden, wohingegen ober und unterhalb dieses Bereiches eine stärkere Regularisierung anzuwenden ist.

Dies ist auch leicht nachzuvollziehen, da es physikalisch bedingt ist, dass Lautsprecher außerhalb ihres Arbeitsbereiches keinen hohen Schalldruck liefern können. Um dies beim Entwurf der inversen Filter zu berücksichtigen, gilt es hier vorrangig die Ausgangsleistung mit Hilfe von β so zu reduzieren, dass die Lautsprecher nicht übersteuern und somit kein höheres Maß an nichtlinearen Verzerrungen erzeugen. Da es sich bei elektrodynamischen Lautsprechern, welche nach wie vor die mit Abstand am weitest verbreitete Form darstellen, physikalisch betrachtet um Systeme 2. Ordnung handelt, reicht es gemäß [CS08] aus, wenn β mit der Flankensteilheit eines Filter 2. Ordnung hin zu hohen und tiefen Frequenzen ansteigt. Eine Abweichung von dieser Regel besteht jedoch, wenn „Bassreflexboxen“ verwendet werden, da es sich dabei um Systeme 4. Ordnung handelt, was im Verlauf von β durch entsprechend steilflankigere Übergänge berücksichtigt werden muss.

Abbildung 2.4 zeigt einen typischen Verlauf des Regularisierungsparameters über der Frequenz, wie etwa aus [CCFS10] bekannt. Wie man sieht wurde zusätzlich für hohe und tiefe Frequenzen eine maximale Schranke für β eingeführt, wobei der Autor dieser Arbeit der Meinung ist, dass auf eine solche, aus den oben aus [CS08] bekannten Gründen, durchaus verzichtet werden kann, was in [CCFS10] so nicht explizit erwähnt wurde.

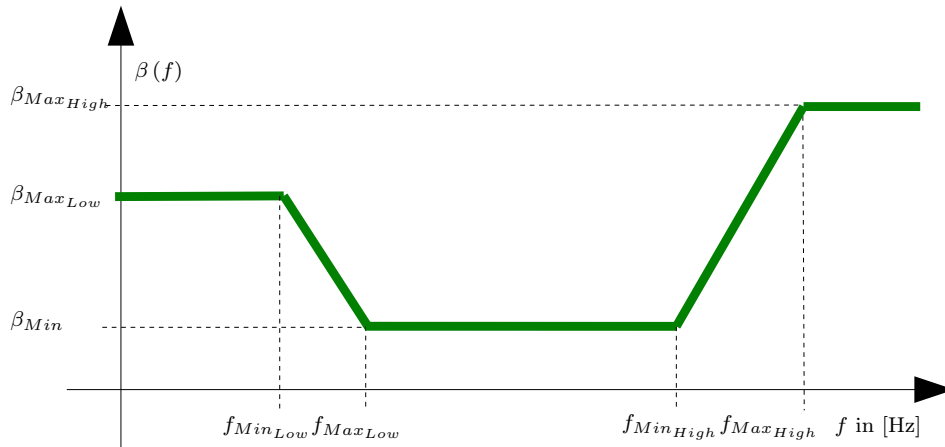


Abbildung 2.4: Darstellung eines typischen Verlaufes für den Regularisierungsparameter β über der Frequenz.

2.4 Berechnung der inversen Filter mit Hilfe der adjungierten Matrix

Gemäß des *laplaceschen Entwicklungssatzes*, der detailliert im Anhang A.2.1 behandelt wird, kann eine inverse Matrix \mathbf{A}^{-1} im Allgemeinen auch wie folgt bestimmt werden:

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \text{adj}(\mathbf{A}). \quad (2.13)$$

Angewendet auf das, aus Abbildung 2.2 bekannte, *transaural-Stereo-Prinzip*, bedeutet dies, dass sich die hierfür nötigen inversen Filter $\mathbf{C}(j\omega)$ auch wie folgt darstellen lassen:

$$\begin{aligned} \mathbf{C}(j\omega) &= \frac{1}{\det(\mathbf{H}(j\omega))} \text{adj}(\mathbf{H}(j\omega)) \\ &= \frac{1}{\det(\mathbf{H}(j\omega))} \begin{pmatrix} H_{RR}(j\omega) & -H_{RL}(j\omega) \\ -H_{LR}(j\omega) & H_{LL}(j\omega) \end{pmatrix}, \text{ mit} \\ \det(\mathbf{H}(j\omega)) &= H_{LL}(j\omega) H_{RR}(j\omega) - H_{LR}(j\omega) H_{RL}(j\omega), \text{ oder ausgeschrieben} \\ C_{LL}(j\omega) &= \frac{1}{\det(\mathbf{H}(j\omega))} H_{RR}(j\omega), \\ C_{LR}(j\omega) &= -\frac{1}{\det(\mathbf{H}(j\omega))} H_{LR}(j\omega), \\ C_{RL}(j\omega) &= -\frac{1}{\det(\mathbf{H}(j\omega))} H_{RL}(j\omega), \\ C_{RR}(j\omega) &= \frac{1}{\det(\mathbf{H}(j\omega))} H_{LL}(j\omega). \end{aligned} \quad (2.14)$$

Dieses Prinzip, welches bereits aus [FU97, FU98, FU99, FBA00, Sof02] bekannt ist, zeichnet sich, wie Gleichung 2.14 belegt, durch eine strikte Trennung der Übersprechdämpfung und des Kompensationsfilters $\frac{1}{\det(\mathbf{H}(j\omega))}$ aus. Anders ausgedrückt bedeutet dies, dass bei Vernachlässigung des Kompensationsfilters, also für $\frac{1}{\det(\mathbf{H}(j\omega))} = 1$, es zwar zu einer spektralen Verfärbung am linken und rechten Ohr kommt, jedoch davon, d.h. bei Verwendung der Adjungierten der RIR-Matrix als inverse Filtermatrix, also für $\mathbf{C}(j\omega) = \text{adj}(\mathbf{H}(j\omega))$, die Übersprechkompensation nicht gestört wird.

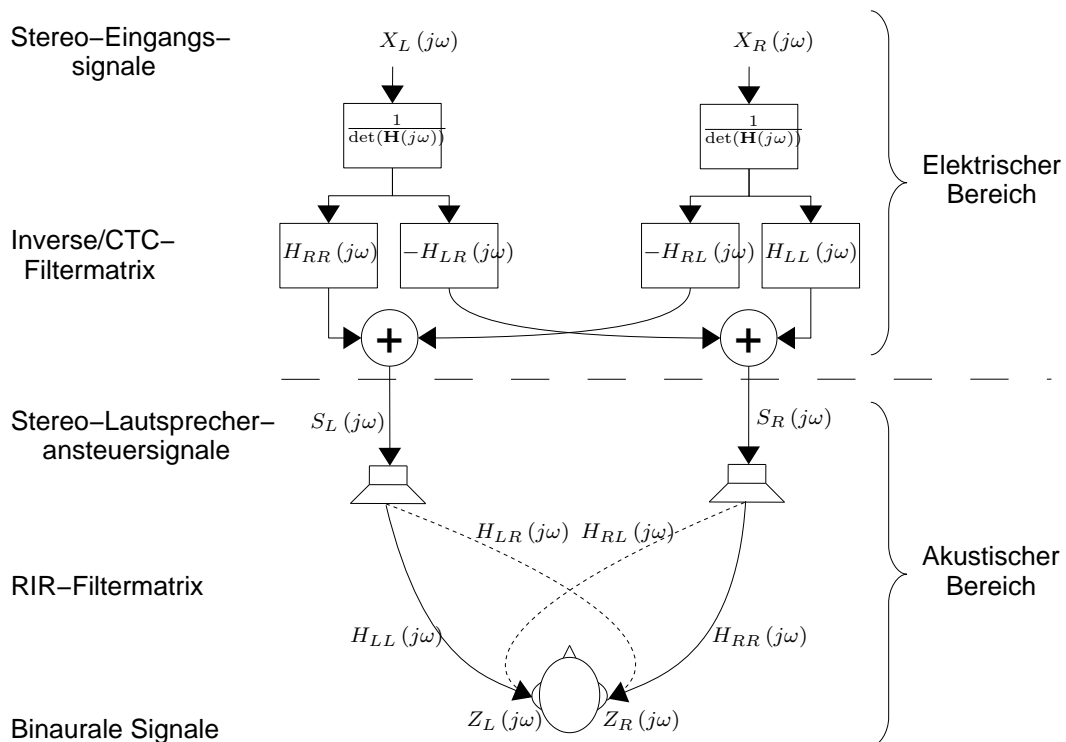


Abbildung 2.5: Darstellung des vereinfachten Prinzips der inversen Filterung zur transauralen Wiedergabe eines Stereosignals.

In [Kim03] wurde gezeigt, dass diese Form der Erzeugung der benötigten inversen Filter $\mathbf{C}(j\omega)$ einer Bestimmung der *Optimalfilter*, die auch als *Wiener-Filter* bezeichnet werden, im Zeitbereich gleichkommt. Diese Vorgehensweise, die je nach Anzahl der Lautsprecher bzw. der Punkte im Raum, an denen diese „optimalen“ Verhältnisse erreicht werden sollen, ziemlich rechenintensiv werden kann, was vor allem auf die Berechnung der Determinante $\det(\mathbf{H}(j\omega))$ sowie der adjungierten Matrix $\text{adj}(\mathbf{H}(j\omega))$ zutrifft, hat den Vorteil, dass sich dadurch *kausale* Filter erzeugen lassen, die, wie bereits erwähnt, kein *preringing* aufweisen. Solange man, wie in dem hier behandelten Fall, von stationären Verhältnissen ausgehen kann, d.h. solange es dem Hörer nicht gestattet ist, sich frei im Raum zu bewegen und man dementsprechend weiß, wo er sich im Raum befindet, genügt es, die Berechnung der inversen Filter nur ein einziges Mal durchzuführen, weshalb in diesem Fall die oben erwähnte Rechenintensität kein Problem darstellt. Anders

verhält es sich, wenn diese Methode, wie in [MFV11] in einem *virtuellen Raum*, in dem sich ein Hörer frei bewegen darf, angewendet werden soll. Dann müssen die inversen Filter dynamisch, quasi in *Echtzeit*, erzeugt werden, weshalb deren Entwurf schnell gehen muss. Eine Lösung für dieses Problem bietet die in [MFV11] erwähnte *spektrale Faktorisierung*, mit deren Hilfe sich kausale *Wiener-Filter* direkt im Spektralbereich erzeugen lassen, was wesentlich weniger Rechenzeit benötigt als die in Gleichung 2.13 aufgezeigte Methode. Da diese jedoch die gleichen Ergebnisse wie die Optimalfilter, welche gemäß Gleichung 2.13 im Zeitbereich entworfen wurden, liefert, wird an dieser Stelle auf eine nähere Erläuterung verzichtet. Weiterführende Information dazu findet der interessierte Leser in [Vai93, Ell01]. Wie Abbildung 2.5 verdeutlicht, lassen sich einerseits, die für die Übersprechdämpfung nötigen Filter ohne jeglichen Rechenaufwand bestimmen, da hierzu nur die bereits vorhandenen RIR entsprechend umsortiert und für die Kreuzpfade phaseninvertiert angewendet werden müssen. Andererseits zeichnen sich die RIR-Filter $\mathbf{H}(j\omega)$ durch ein strikt kausales Verhalten aus, wodurch es zu keinem *preringing* kommt. Einzig und allein die Bestimmung des Kompensationsfilter $\frac{1}{\det(\mathbf{H}(j\omega))}$ kann zu einem *preringing* führen, wobei dessen Einfluss z.B. durch Verwendung der in Abschnitt 2.3 aufgezeigten *Regularisierung* reduziert werden kann. Gänzlich vermeiden lässt sich das *preringing*, indem man die Phase von $\frac{1}{\det(\mathbf{H}(j\omega))}$ ignoriert und lediglich dessen *minimalphasigen Anteil* invertiert, indem man das Zähler- und Nennerpolynom vertauscht. Der minimalphasige Anteil einer beliebigen, gemischtphasigen Impulsantwort lässt sich dabei z.B. mit Hilfe der

aus [NA79] bekannten Methode auf recht einfache Weise im *Cepstralbereich*¹¹ bestimmen:

$$\begin{aligned}
 \mathbf{h}_{\text{Min}\varphi} &= \Re\{IFFT\{\exp\{FFT\{\text{diag}(\mathbf{w}) \mathbf{h}_{\text{ReCep}}\}\}\}\}, \text{ wobei} \\
 \mathbf{h}_{\text{ReCep}} &= \Re\{IFFT\{\ln\{|FFT\{\mathbf{h}\}|\}\}\}, \text{ und} \\
 \mathbf{w} &= \begin{cases} \begin{bmatrix} 1, \underbrace{2, 2, \dots, 2}_{\lceil \frac{N}{2} \rceil - 1}, \underbrace{1, 0, 0, \dots, 0}_{\lceil \frac{N}{2} \rceil - 1} \end{bmatrix}^T, & \text{wenn } N \text{ gerade,} \\ \begin{bmatrix} 1, \underbrace{2, 2, \dots, 2}_{\lceil \frac{N}{2} \rceil - 1}, \underbrace{0, 0, \dots, 0}_{\lceil \frac{N}{2} \rceil - 1} \end{bmatrix}^T, & \text{wenn } N \text{ ungerade} \end{cases}, \text{ wobei} \\
 \mathbf{h} &= \text{Spaltenvektor, welcher die } N \text{ Filterkoeffizienten der ursprünglichen,} \\
 &\quad \text{gemischtphasigen Impulsantwort enthält,} \\
 \mathbf{h}_{\text{ReCep}} &= \text{Spaltenvektor, welcher die } N \text{ Werte des reellen Cepstrums der} \\
 &\quad \text{ursprünglichen, gemischtphasigen Impulsantwort enthält,} \\
 \mathbf{w} &= \text{Fensterfunktion der Länge } N, \text{ mit welcher das reelle Cepstrum der} \\
 &\quad \text{gemischtphasigen Impulsantwort } \mathbf{h}_{\text{ReCep}} \text{ gewichtet werden muss,} \\
 \mathbf{h}_{\text{Min}\varphi} &= \text{Spaltenvektor, welcher die } N \text{ Filterkoeffizienten der minimalphasigen,} \\
 &\quad \text{Variante der ursprünglichen, gemischtphasigen Impulsantwort } \mathbf{h} \\
 &\quad \text{enthält,} \\
 \lceil \cdot \rceil &= \text{Aufrunden.} \tag{2.15}
 \end{aligned}$$

2.5 Beispiel und praktische Modifikationen

Nachfolgend wird anhand eines Beispiels das in Kapitel 2.4 vorgestellte Verfahren in der Praxis getestet. Dazu wurden binaurale Messungen an der vorderen linken Position eines BMW¹² E61 (5er BMW Touring), dessen Lautsprecheranordnung aus Abbildung 2.6 ersichtlich ist, durchgeführt.

Ergänzend ist zu der in Abbildung 2.6 aufgezeigten Anordnung der Lautsprecher im Fahrzeuginnenraum zu sagen, dass die Hoch- und Mitteltöner, welche sich auf der linken und rechten Seite sowohl vorne als auch hinten befinden, über einen Kondensator miteinander passiv gekoppelt waren. Dadurch beläuft sich die Anzahl der voneinander unabhängigen Kanäle auf 9, die sich wie folgt angeben lassen:

¹¹Unter dem Begriff *Cepstrum* versteht man das Spektrum des logarithmierten (natürlicher Logarithmus) Frequenzspektrums (Fouriertransformierte) eines Signals. Die Bezeichnung Cepstrum leitet sich aus dem englischen Begriff *spectrum* ab, bei dem die ersten vier Buchstaben in der Reihenfolge umgekehrt wurden.

¹²BMW = Bayerische Motoren Werke

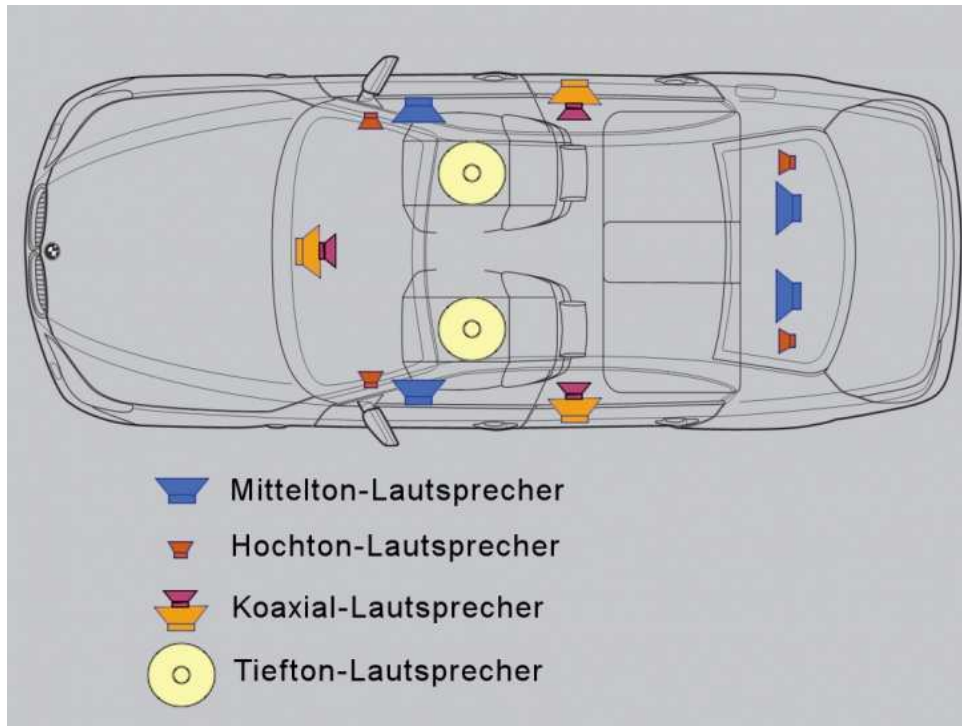


Abbildung 2.6: Typische Lautsprecheranordnung bei BMW Fahrzeugen. [Bmw07]

- FL: Engl. **Front Left** (Vorne links, passive Kopplung zwischen einem Hoch- und Mitteltonlautsprecher),
- FR: Engl. **Front Right** (Vorne rechts, passive Kopplung zwischen einem Hoch- und Mitteltonlautsprecher),
- SL: Engl. **Side Left** (Seite links, Koaxiallautsprecher¹³),
- SR: Engl. **Side Right** (Seite rechts, Koaxiallautsprecher),
- RL: Engl. **Rear Left** (Hinten links, passive Kopplung zwischen einem Hoch- und Mitteltonlautsprecher),
- RR: Engl. **Rear Right** (Hinten rechts, passive Kopplung zwischen einem Hoch- und Mitteltonlautsprecher),
- WL: Engl. **Woofer Left** (Tieftonlautsprecher, welcher sich unter dem linken Vordersitz befindet),
- WR: Engl. **Woofer Right** (Tieftonlautsprecher, welcher sich unter dem rechten Vordersitz befindet),
- C: **Centerlautsprecher** (Koaxiallautsprecher).

Für den Test wurden lediglich die beiden Kanäle FL und FR ausgewählt um das Prinzip des transaural-Stereo zu verifizieren.

2.5.1 Ursprüngliche Situation - Situation ohne Filterung

Die in Abbildung 2.7 dargestellten Messungen verdeutlichen die ursprüngliche Situation, d.h. die Situation die vorherrscht bevor Kompensationsfilter $\mathbf{C}(j\omega)$, wie etwa in Abbildung 2.2 aufgezeigt, angewendet wurden.

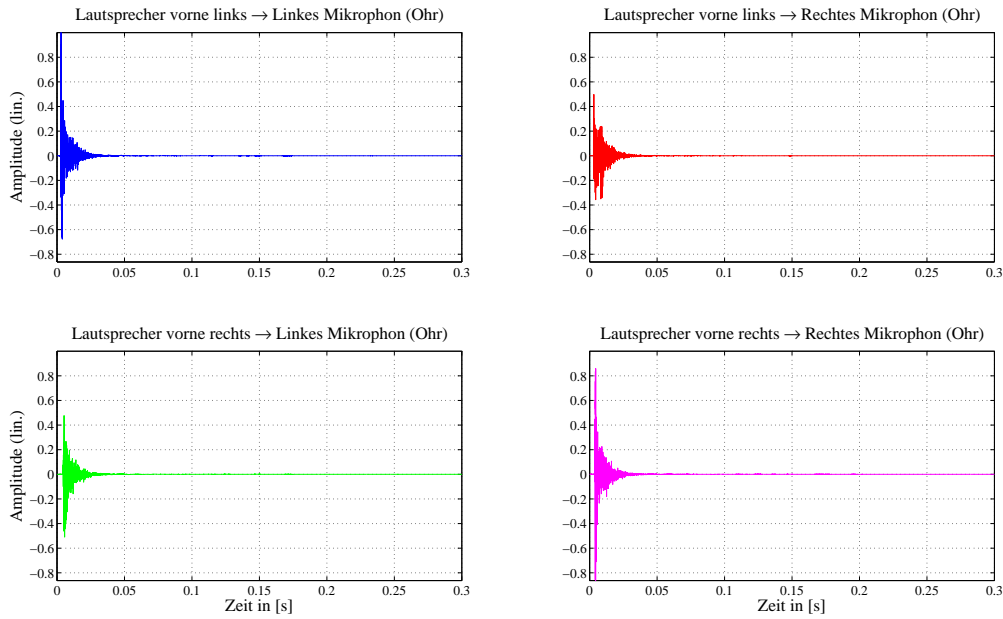
Dabei folgen die obigen, sowie alle nachfolgenden Darstellungen dieses Beispiels stets dem gleichen Schema, d.h. die oberen Teilbilder verdeutlichen, welche Signale an beiden Ohren gemessen werden, wenn der linke (FL) und die beiden unteren Teilbilder, wenn der rechte Lautsprecher (FR) spielt, wobei die beiden linken Teilbilder die Situation am linken und die beiden rechten Teilbilder die gemessenen Signale am rechten Ohr kennzeichnen.

Die *Impulsantworten* geben dabei Auskunft über *zeitliche Aspekte*, zu denen unter anderem auch das bereits mehrfach erwähnte *preringing* sowie die sich durch die Filterung veränderte *Nachhallzeit* zählt. Dagegen informieren die *Betragsfrequenzgänge* über *spektrale Aspekte*, zu denen unter anderem Veränderungen in der *Klangfarbe* sowie die Höhe der erreichten *Übersprechdämpfung* zählen.

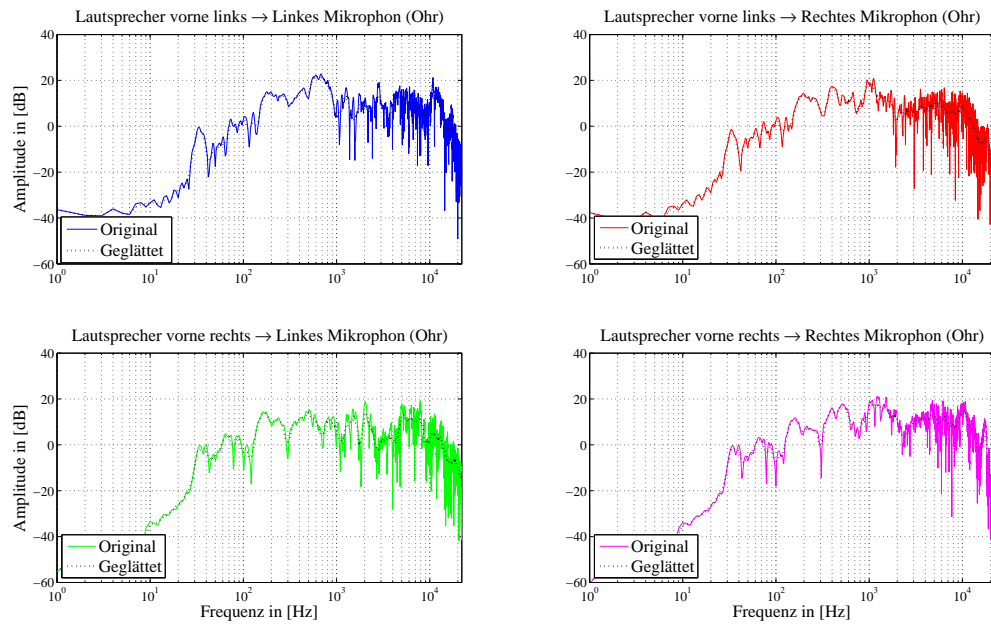
Zur besseren Deutung der Betragsfrequenzgänge wurden neben den unveränderten Daten noch geglättete Varianten berechnet und über die Originalverläufe gezeichnet.

Unter Berücksichtigung der Frequenzauflösung unseres Gehörs wurde dabei eine *nicht-*

¹³Der Begriff *Koaxiallautsprecher*, der in der Praxis des öfteren auch nur kurz als *Koax* bezeichnet wird, steht für eine Anordnung von zwei Lautsprechern (*2-Wege-System*), einem Mittel- und einem Hochtöner, die sich auf einer Achse befinden, wobei der Hochtöner über dem Zentrum des Mitteltöners angebracht ist.



(a) Impulsantworten



(b) Betragsfrequenzgänge

Abbildung 2.7: Darstellung der Impulsantworten und der zugehörigen Betragsfrequenzgänge zwischen den beiden verwendeten Lautsprechern FL und FR und den beiden Mikrofonen der binauralen Messungen (linkes und rechtes Ohrsignal) ohne Filterung.

lineare Glättung verwendet, die sich mathematisch wie folgt beschreiben lässt:

$$\begin{aligned} \bar{A}(j\omega_n) &= \frac{1}{\min\{N-1, \lceil n\alpha - \frac{1}{2} \rceil\} - \max\{0, \lceil \frac{n}{\alpha} - \frac{1}{2} \rceil\}} \sum_{k=\max\{0, \lceil \frac{n}{\alpha} - \frac{1}{2} \rceil\}}^{\min\{N-1, \lceil n\alpha - \frac{1}{2} \rceil\}} |A(j\omega_k)|, \text{ mit} \\ n &= [0, \dots, N-1] \text{ (Diskreter Frequenzindex des geglätteten Wertes),} \\ \lceil x \rceil &= \text{Aufrunden von } x, \\ \left\lfloor x - \frac{1}{2} \right\rfloor &= \text{Runden von } x, \\ \alpha &= \text{Glättungskoeffizient (z.B. } \frac{\text{Oktave}}{3} \text{ - Glättung } \Rightarrow \alpha = 2^{\frac{1}{3}}), \\ \bar{A}(j\omega_k) &= \text{Geglätteter Wert von } A(j\omega_k), \\ \omega_k &= \frac{2\pi k}{N} f_A = \text{Kreisfrequenz in } \left[\frac{1}{s} \right], \\ f_A &= \text{Abtastfrequenz in [Hz],} \\ k &= \text{Diskreter Frequenzindex ([Bin]) des ungeglätteten Wertes, wobei} \\ &\quad k \in [0, \dots, N-1], \\ N &= \text{Länge der FFT.} \end{aligned} \tag{2.16}$$

Mit Hilfe der in Gleichung 2.16 beschriebenen, nichtlinearen Glättung lassen sich schmalbandige „Ausreißer“ im Betragsfrequenzgang auf eine natürliche, d.h. dem menschlichen Gehör nachempfundene, Art automatisch unterdrücken. Bei tiefen Frequenzen, bei dem unser Gehör bekanntermaßen eine hohe spektrale Auflösung besitzt, findet nur eine geringe Glättung statt, wohingegen im oberen Spektralbereich, in welchem die Frequenzauflösung unseres Gehörs gering ist, eine hohe Glättung verwendet wird. In dem hier behandelten Beispiel wurde durchwegs mit einem Glättungskoeffizienten von $\alpha = 2^{\frac{1}{9}}$, was einer $\frac{\text{Terz}}{3}$ - bzw. einer $\frac{\text{Oktave}}{9}$ -Glättung entspricht, gearbeitet. Wie man unschwer am oberen rechten sowie am unteren linken Teilbild erkennen kann, nimmt im ursprünglichen Fall, also ohne Anwendung der CTC-Filter, das jeweils kontralaterale Ohr ein Signal in nahezu der gleichen Stärke nur mit einer etwas anders gestalteten spektralen Zusammensetzung auf als das ipsilaterale Ohr, welche im oberen linken sowie im unteren rechten Teilbild dargestellt ist. Bei genauer Betrachtung erkennt man zudem, dass das ipsilaterale Signal des linken Ohrs (oberes linkes Teilbild) etwas stärker ist als das des rechten Ohrs (unteres rechtes Teilbild), was, wie Abbildung 2.6 veranschaulicht, von der asymmetrischen Sitzposition (vorne links) herrührt. Zudem kann aus den Betragsfrequenzgängen entnommen werden, dass die verwendeten Kanäle (FL und FR) ab einer Frequenz von $f_c \approx 150$ [Hz] nahezu den gesamten akustischen Frequenzbereich des menschlichen Gehörs, also bis $f \approx 20$ [kHz], abdecken, was mit der hier verwendeten, passiven Kopplung eines Hoch- und Mitteltöners erklärt werden kann.

2.5.2 Situation bei alleiniger Verwendung der Filter zur Kompensation des Übersprechens

Werden lediglich die CTC-Filter ohne Berücksichtigung des Kompensationsfilters $\frac{1}{\det(\mathbf{H}(j\omega))}$ verwendet, stellen sich die in Abbildung 2.8 aufgezeigten Ergebnisse ein.

Dabei wurden die CTC-Filter $\mathbf{C}(j\omega) = \text{adj}(\mathbf{H}(j\omega))$ nicht in der vollen Länge der gemessenen Raumimpulsantworten $\mathbf{H}(j\omega)$, die mit einer Länge von $t = 300$ [ms] $\hat{=}$ 13230 [Samples] @ $f_A = 44100$ [Hz] bestimmt wurden, umgesetzt, sondern auf eine praxistaugliche Länge von $t \approx 21.53$ [ms] $\hat{=}$ 2048 [Samples] @ $f_A = 44100$ [Hz] gekürzt.

Die maximale Länge der RIR von $t = 300$ [ms] kam dabei nicht von ungefähr, sondern begründet sich in der Tatsache, dass selbst im tieffrequenten Bereich, also bei Messungen, bei denen ein Tieftöner (z.B. WL oder WR) als Anregungsquelle verwendet wurde, es in Kraftfahrzeuginnenräumen zu einem Abfall von mindestens 60 [dB] kommt, folglich die T_{60} ¹⁴-Zeit von Autoinnenräumen üblicherweise nicht größer als $t = 300$ [ms] ist.

Die Länge der CTC-Filter von $t \approx 21.53$ [ms] $\hat{=}$ 2048 [Samples] @ $f_A = 44100$ [Hz] hingegen hat ihren Ursprung vorrangig in der damit erreichbaren Dämpfung, die wie z.B. ein Vergleich der oberen beiden Teilbilder von Abbildung 2.8(b) verdeutlicht, bei ≈ 40 [dB] liegt, was mehr als ausreichend sein dürfte. Andererseits sollte der Wert aber auch durch eine Zweierpotenz darstellbar sein ($2048 = 2^{11}$), damit sich das Filter einfach, z.B. mit Hilfe der schnellen Faltung, im Spektralbereich realisieren lässt. Dazu wird bekanntermaßen meist eine FFT verwendet, welche mindestens diese Anforderung (bei Verwendung einer *Radix-2*-Variante) an die Länge stellt.

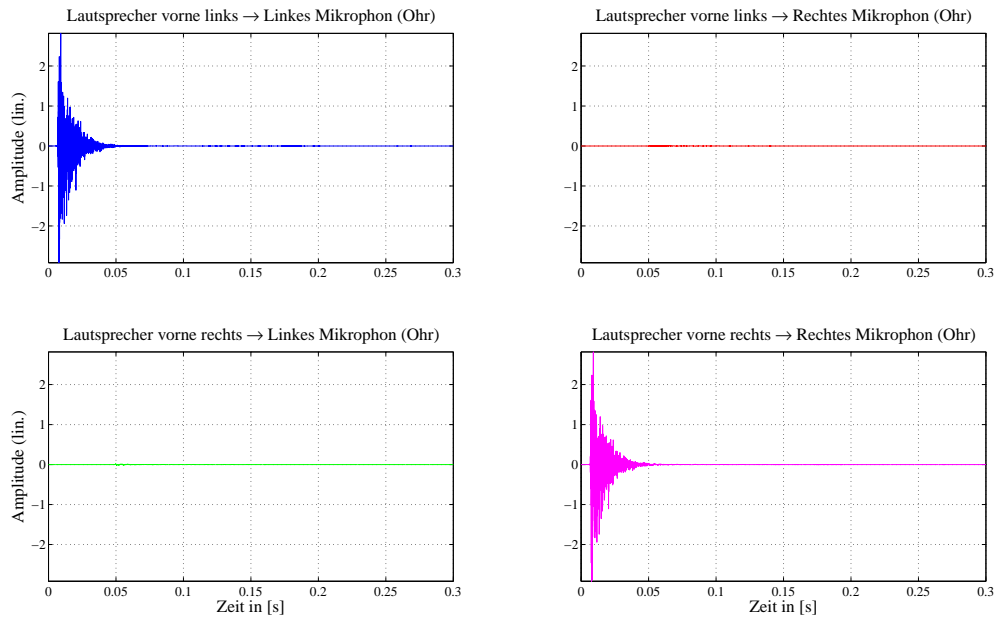
Ein Vergleich der ipsilateralen Betragsfrequenzgänge (oberes linkes und unteres rechtes Teilbild) der Abbildungen 2.7(b) und 2.8(b) zeigt, dass es nun bei Verwendung der CTC-Filter $\mathbf{C}(j\omega) = \text{adj}(\mathbf{H}(j\omega))$ zu einer Veränderung der Klangfarbe kommt, was am Deutlichsten anhand des spektralen Einbruchs bei $f \approx 2.5$ [kHz] zu erkennen ist.

Ferner zeigt ein Vergleich der ipsilateralen Impulsantworten, also der oberen linken und unteren rechten Teilbilder der Abbildungen 2.7(a) und 2.8(a), dass es nach Anwendung der CTC-Filter $\mathbf{C}(j\omega) = \text{adj}(\mathbf{H}(j\omega))$ zu einer Vergrößerung der Nachhallzeit kommt, was jedoch in Anbetracht ihrer Form (exponentieller Abfall) sowie ihrer gesamten Dauer ($t < 5$ [ms]) akustisch kaum auffallen dürfte.

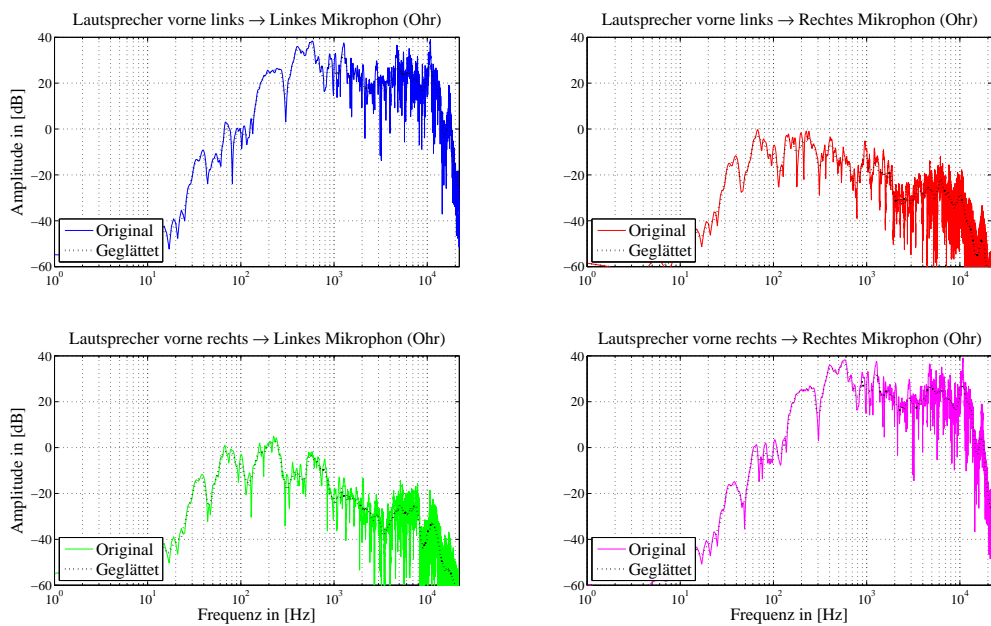
2.5.3 Situation bei Verwendung der vollständigen, inversen Filter

Werden die im vorherigen Abschnitt 2.5.2 verwendeten CTC-Filter $\mathbf{C}(j\omega) = \text{adj}(\mathbf{H}(j\omega))$ nun noch um das (ideale) Kompensationsfilter $\frac{1}{\det(\mathbf{H}(j\omega))}$, welches nachfolgend mit $G(j\omega)$ bezeichnet wird, erweitert, lassen sich die dort formulierten Probleme, nämlich die klang-

¹⁴Die T_{60} -Zeit beschreibt die Zeit, die vergeht bis der Schalldruckpegel, ausgehend von seinem Maximum, um 60 [dB] gefallen ist.



(a) Impulsantworten



(b) Betragsfrequenzgänge

Abbildung 2.8: Darstellung der Impulsantworten und der zugehörigen Betragsfrequenzgänge zwischen den beiden verwendeten Lautsprechern FL und FR und den beiden Mikrofonen der binauralen Messungen (linkes und rechtes Ohrsignal), bei alleiniger Verwendung der Filter zur Kompensation des Übersprechens ($\mathbf{C}(j\omega) = \text{adj}(\mathbf{H}(j\omega))$).

lichen Verfärbungen sowie die Verlängerung der Nachhallzeit theoretisch perfekt korrigieren.

Aus Kapitel 2.3 und 2.4 ist bekannt, dass sich $G(j\omega) = \frac{1}{\det(\mathbf{H}(j\omega))}$ im Regelfall aufgrund der nicht-minimalphasigen Natur von $\det(\mathbf{H}(j\omega))$ nicht vollständig, sondern sich, z.B. mit Hilfe der in Kapitel 2.3 vorgestellten Methode, lediglich approximativ bestimmen lässt.

Wendet man diese Methode auf die Invertierung nur eines Filters an, reduziert sich Gleichung 2.12 zu:

$$\begin{aligned} G(j\omega_k) &= \frac{\det(\mathbf{H}(j\omega_k))^*}{(\det(\mathbf{H}(j\omega_k))^* \det(\mathbf{H}(j\omega_k)) + \beta(j\omega_k))}, \text{ mit} \\ \det(\mathbf{H}(j\omega_k)) &= H_{LL}(j\omega_k) H_{RR}(j\omega_k) - H_{LR}(j\omega_k) H_{RL}(j\omega_k), \\ k &= [0, \dots, N-1] \text{ (Diskreter Frequenzindex),} \\ N &= \text{Länge der FFT,} \\ \beta(j\omega_k) &= \text{Regularisierungsparameter bei der diskreten Frequenz } k. \end{aligned} \quad (2.17)$$

Wie bereits erörtert, weisen die beiden verwendeten Kanäle FL und FR lediglich unterhalb von $f \approx 150$ [Hz] einen starken Pegelabfall auf, was bei der Wahl des Regularisierungsparameters β entsprechend berücksichtigt werden muss. Diese verdeutlicht der in Abbildung 2.9 dargestellte Verlauf des für das Beispiel verwendeten Regularisierungsparameters.

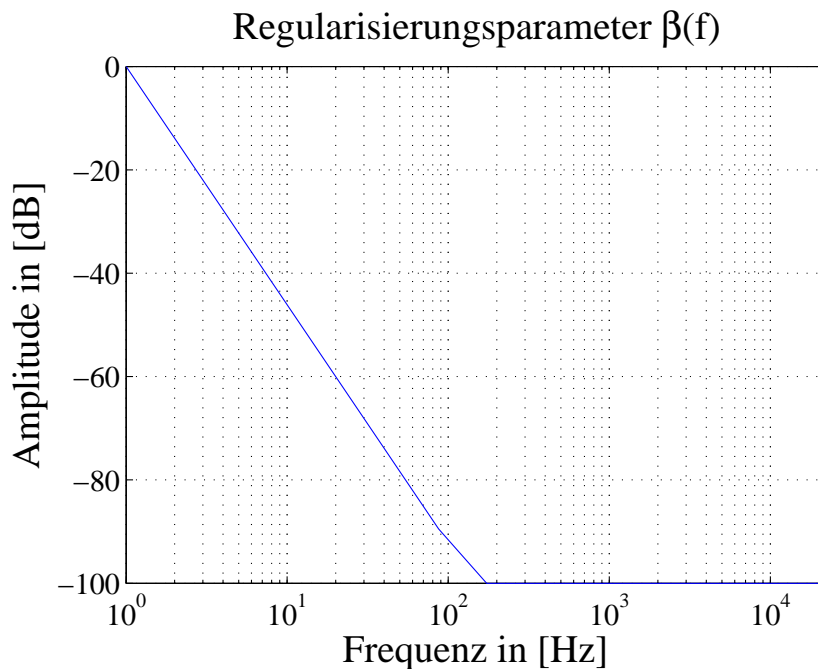


Abbildung 2.9: Darstellung des Verlaufes, des für das Beispiel verwendeten Regularisierungsparameters β über der Frequenz.

Bei einer gewählten Länge des inversen Filters von 512 [Samples] @ $f_A = 44100$ [Hz]

ergibt sich mit Gleichung 2.17 folgende in Abbildung 2.10 dargestellte Impulsantwort des inversen Filters $G(j\omega)$.

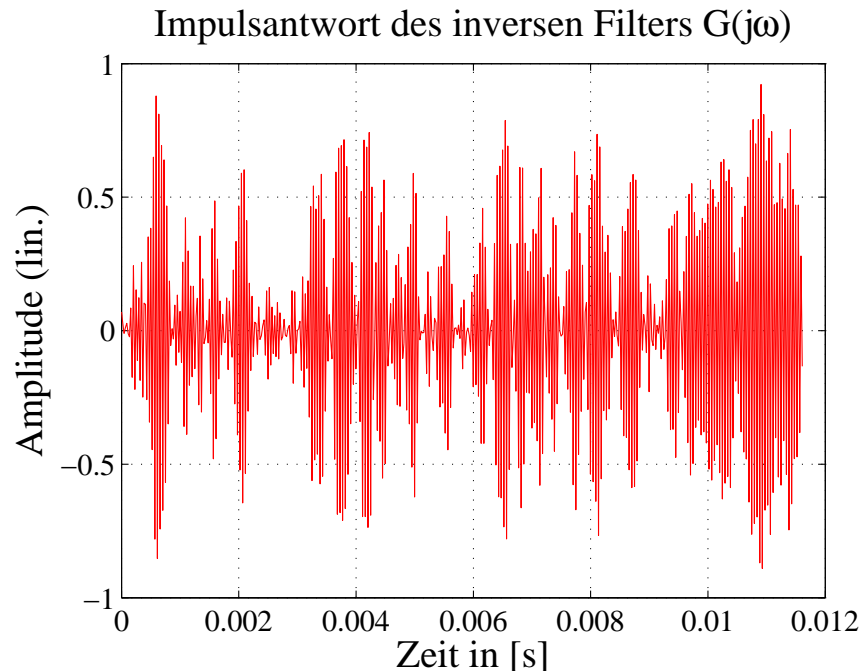


Abbildung 2.10: Darstellung der Impulsantwort des inversen Filters $G(j\omega)$, welches nach Gleichung 2.17, unter Berücksichtigung des Regularisierungsparameters $\beta(j\omega)$ aus Abbildung 2.9, berechnet wurde.

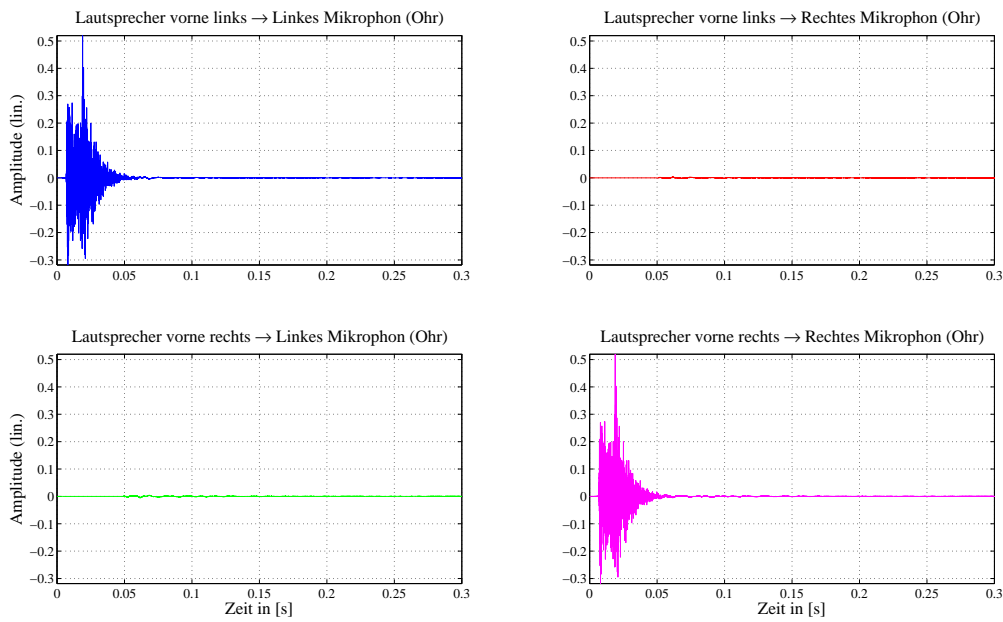
Wie man aus Abbildung 2.10 sieht, weist die Impulsantwort des inversen Filter $G(j\omega)$ eine Form auf, die keinen exponentiellen Abfall mehr zeigt, also nichts mit einer minimalphasigen Impulsantwort gemein hat. Dies führt je nach Länge des inversen Filter zu akustisch wahrnehmbaren Störungen, wobei sich diese in dem hier behandelten Beispiel, d.h. bei einer Filterlänge von 512 [Samples] @ $f_A = 44100$ [Hz], noch im Rahmen halten dürften.

Hintergrund dieser außergewöhnlichen Form ist, dass mit der in Gleichung 2.17 beschriebenen Entwurfsmethode zwar eine Regularisierung im Spektralbereich stattfindet, welche die Ausgangsleistung in bestimmten Frequenzbereichen in Abhängigkeit von β auf ein Maximum beschränkt, jedoch die Methode an sich nach wie vor versucht, das optimale inverse Filter nach *Betrag und Phase* zu approximieren.

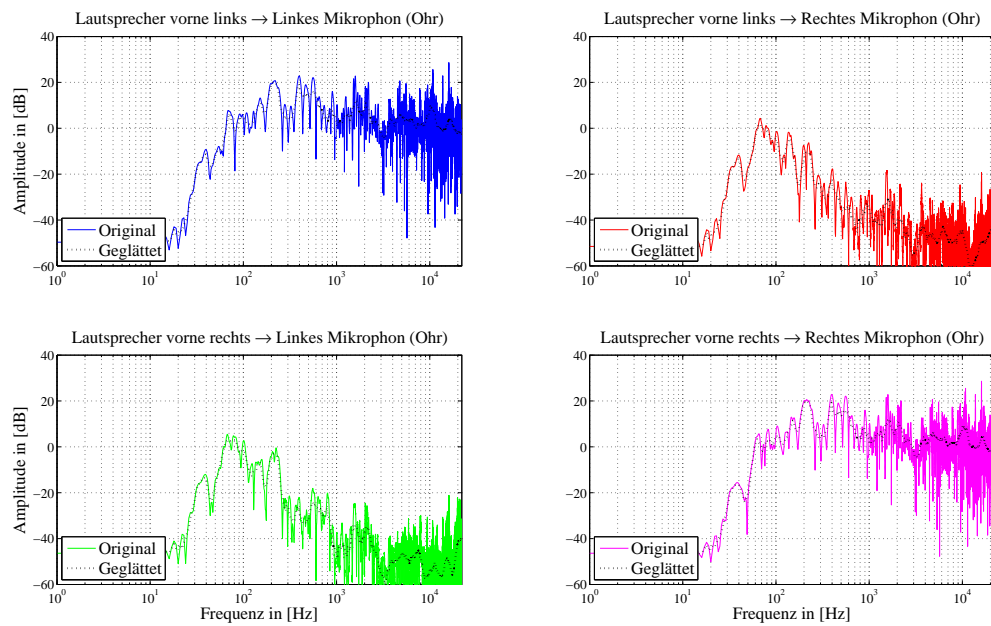
Bildlich gesprochen ist es die Aufgabe des Betrags, die klanglichen Verfärbungen und die der Phase, die Nachhallzeit zu minimieren, so dass in Summe eine Kronecker-Delta-Funktion resultiert.

Wendet man das inverse Filter aus Abbildung 2.10 an, resultieren die in Abbildung 2.11 dargestellten Ergebnisse.

Dabei verdeutlichen die ipsilateralen Darstellungen aus Abbildung 2.11, dass einerseits nach wie vor eine sehr gute Übersprechdämpfung erreicht wird. Andererseits zeigen die



(a) Impulsantworten



(b) Betragsfrequenzgänge

Abbildung 2.11: Darstellung der Impulsantworten und der zugehörigen Betragsfrequenzgänge zwischen den beiden verwendeten Lautsprechern FL und FR und der beiden Mikrophone der binauralen Messungen (linkes und rechtes Ohrsignal), bei Verwendung des inversen Filters $G(j\omega)$ aus Abbildung 2.10.

kontralateralen Darstellungen, dass zum einen das in Kapitel 2.5.2 erwähnte Problem der Veränderung der Klangfarbe zwar nicht gelöst aber zumindest gemildert werden konnte, zum anderen sich jedoch dafür die Problematik der Verlängerung der Nachhallzeiten im Vergleich zu Abbildung 2.8(a) sogar noch verschärft hat, wobei hierfür die Ursache in der unzureichenden Länge des inversen Filters $G(j\omega)$ liegt.

Somit besteht folgender Konflikt:

- Für die Reduzierung der Nachhallzeit muss das inverse Filter $G(j\omega)$ eine ausreichend große Länge besitzen.
- Ein nach Gleichung 2.17 entworfenes inverses Filter $G(j\omega)$ weist eine nicht-minimalphasige Impulsantwort auf, die mit steigender Länge zu größeren akustisch wahrnehmbaren Artefakten führt.

Da sich dieser Konflikt in der Praxis nicht lösen lässt, wurde ein praxistauglicher Kompromiss erarbeitet, welcher nachfolgend vorgestellt wird.

2.5.4 Praktische Modifikationen beim Entwurf des inversen Filters $G(j\omega)$

Ausgangspunkt des in diesem Abschnitt beschriebenen Entwurfs des inversen Filters $G(j\omega)$ ist der Verzicht einer komplexen Invertierung, genauer gesagt wird die Invertierung der Phase von $\det(\mathbf{H}(j\omega))$ aufgegeben und stattdessen lediglich dessen minimalphasiger Anteil, der sich bekanntermaßen perfekt invertieren lässt, für die Berechnung verwendet, was bereits in Kapitel 2.4 angedeutet wurde.

Allerdings darf die in Kapitel 2.4 beschriebene Methode zur Berechnung des minimalphasigen, inversen Filters (siehe Gleichung 2.15) so nicht direkt verwendet werden, da ansonsten aufgrund schmalbandiger, hoher Anhebungen, wieder die Problematik des „Klingelns“ latent wäre. Dies lässt sich jedoch wieder mit Hilfe einer *Regularisierung* beheben, die wie folgt zusammengefasst werden kann:

1. Verwendung einer aus Gleichung 2.16 bekannten nicht-linearen Glättung, mit welcher sich in physiologisch korrekter Art, schmalbandige spektrale „Ausreißer“ vermeiden lassen.
2. Verwendung eines frequenzabhängigen Regularisierungsparameters β , der dafür sorgt, dass das inverse Filter $G(j\omega)$ keine zu hohen Anhebungen aufweist.

Der Algorithmus zur Bestimmung des in Abbildung 2.12 dargestellten inversen Filters lässt sich wie folgt beschreiben:

1. Berechnung der Übertragungsfunktion der Determinante $\det(\mathbf{H}(j\omega_k))$ nach Gleichung 2.14 bzw. Gleichung 2.17.

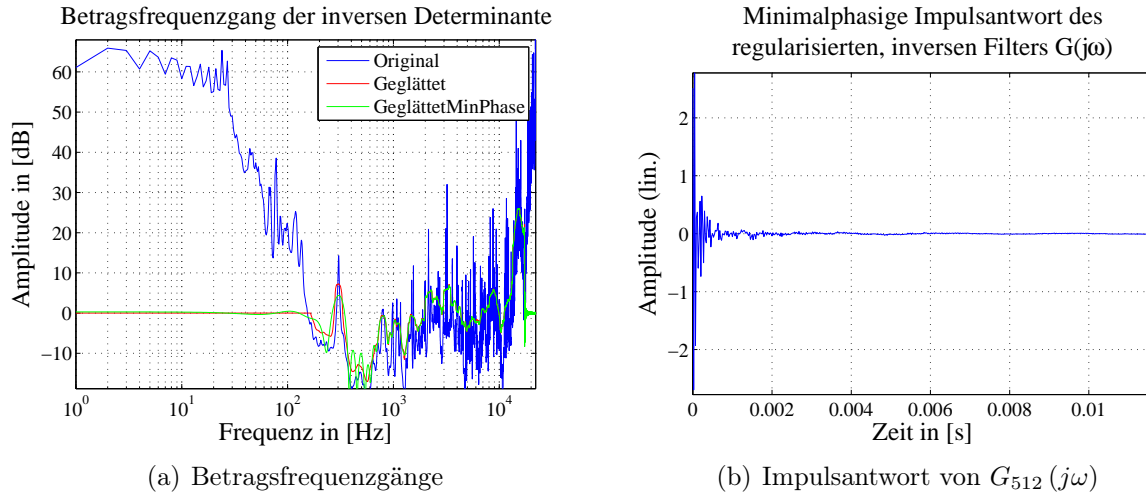


Abbildung 2.12: Darstellung der Betragsfrequenzgänge des minimalphasigen, inversen Filters $G(j\omega)$ vor (Blau) und nach Anwendung der Regularisierung (Rot) sowie einer Umsetzung mit Hilfe eines FIR-Filters der Länge 512 [Samples] (Grün), dessen Impulsantwort aus dem rechten Teilbild ebenfalls ersichtlich ist.

2. Konvertierung von $\det(\mathbf{H}(j\omega_k))$ in den Zeitbereich $h_{\det}[n] = IFFT\{\det(\mathbf{H}(j\omega_k))\}$.
3. Berechnung von $h_{\det_{\text{Min}\varphi}}[n]$, d.h. der minimalphasigen Variante von $h_{\det}[n]$, mit Hilfe der Gleichung 2.15.
4. Berechnung der Impulsantwort des inversen Filters $h_{\text{Inv}}[n]$ durch Vertauschung des Zähler- und Nennerpolynoms, d.h. durch Verwendung der Filterkoeffizienten $h_{\det_{\text{Min}\varphi}}[n]$ als Nennerpolynom.
5. Konvertierung von $h_{\text{Inv}}[n]$ in den Spektralbereich $H_{\text{Inv}}(j\omega_k) = FFT\{h_{\text{Inv}}[n]\}$.
6. Nichtlineare Glättung von $H_{\text{Inv}}(j\omega_k) = FFT\{h_{\text{Inv}}[n]\}$ nach Gleichung 2.16 zur Erzeugung von $\overline{H}_{\text{Inv}}(j\omega_k)$.
7. Skalierung von $\overline{H}_{\text{Inv}}(j\omega_k)$, so dass diese bei $k = k_{\text{Max}} = \frac{f_c}{f_A} N$ einen Wert von 0 [dB] erreicht.
8. Regularisierung:
 - (a) Beschränkung von $\overline{H}_{\text{Inv}}(j\omega_k)$ für $k = [0, \dots, k_{\text{Max}}]$ auf 0 [dB].
 - (b) Suchen des Wertes k_{Min} , d.h. des größten k -Werts, für den gilt: $H_{\text{Inv}}(j\omega_k) < 0$ [dB].
 - (c) Beschränkung von $\overline{H}_{\text{Inv}}(j\omega_k)$ für $k = [k_{\text{Min}}, \dots, N - 1]$ auf 0 [dB].

(d) → Regularisierungsparameter $\beta(j\omega_k)$:

$$\beta(j\omega_k) = \begin{cases} 0 & \text{für } k = [0, \dots, k_{\text{Max}}] \\ 0 & \text{für } k = [k_{\text{Min}}, \dots, N - 1] \\ 1 & \text{sonst.} \end{cases} \quad (2.18)$$

9. Berechnung eines *linearphasigen* FIR-Filters $h_{\text{InvLin}\varphi}[n]$ aus der nichtlinear geglätteten, regularisierten Übertragungsfunktion $\overline{H}_{\text{Inv}}(j\omega_k)$, z.B. nach dem aus [Hes93] bekannten *Frequenzabtastverfahren*¹⁵, welches für den Entwurf eines Filters des Typs 1 & 2 wie folgt angegeben werden kann:

(a) Berechnung des Betragsfrequenzgangs:

$$\begin{aligned} |H_{\text{InvLin}\varphi}(0)| &= \overline{H}_{\text{Inv}}(0), \\ |H_{\text{InvLin}\varphi}(j\omega_n)| &= \overline{H}_{\text{Inv}}(j\omega_k), \text{ für } n = [1, \dots, R - 1], \text{ und} \\ &k = \left\lceil n \left(\frac{N - 1}{R - 1} \right) - \frac{1}{2} \right\rceil, \text{ wobei} \\ N &= \text{Länge von } \overline{H}_{\text{Inv}}(j\omega) (\hat{=} \text{Länge der 1. FFT}), \\ R &= \text{Länge des linearphasigen FIR-Filters } (\hat{=} \text{Länge} \\ &\text{der 2. FFT bzw. halbe Länge von } h_{\text{InvMin}\varphi}) \end{aligned} \quad (2.19)$$

(b) Berechnung des Phasengangs:

$$\begin{aligned} \angle H_{\text{InvLin}\varphi}(j\omega_n) &= \begin{cases} -\left(\frac{R-2}{R-1}\right) \pi n, & \text{für } n = [0, \dots, \lceil (\frac{R}{2} - 1) - \frac{1}{2} \rceil] \\ \left(\frac{R-2}{R-1}\right) \pi ((R-1) - n), & \text{für } n = [\lceil (\frac{R}{2} - 1) + \frac{1}{2} \rceil, \dots, R - 1] \end{cases} \\ \angle H_{\text{InvLin}\varphi}(j\omega) &= \text{Phasengang von } H_{\text{InvLin}\varphi}(j\omega) \text{ (linearer Verlauf)} \end{aligned} \quad (2.20)$$

(c) Berechnung der Impulsantwort:

$$\begin{aligned} h_{\text{InvLin}\varphi}[n] &= \Re\{IFFT\{|H_{\text{InvLin}\varphi}(j\omega_n)| e^{j\angle H_{\text{InvLin}\varphi}(j\omega)}\}\}, \text{ mit} \\ n &= [0, \dots, R - 1]. \end{aligned} \quad (2.21)$$

10. Berechnung von $h_{\text{InvMin}\varphi}[n]$, d.h. der minimalphasigen Variante von $h_{\text{InvLin}\varphi}[n]$, mit Hilfe der Gleichung 2.15.
11. Reduzierung der Länge von $h_{\text{InvMin}\varphi}[n]$ auf ein gewünschtes Maß, z.B. auf 512 [Samples] @ $f_A = 44100$ [Hz].

Ferner lässt sich das so entworfene Filter auch leicht, z.B. mit Hilfe des Entwurfsverfahrens nach *Steiglitz-McBride*, welches in *MATLAB*[®] direkt mit dem Befehl

¹⁵Frequenzabtastverfahren = engl. *Frequency sampling method*

`[NumCoefs,DenCoefs] = stmcb(h,FirOrd,IirOrd)` angewendet werden kann, als IIR-Filter beliebiger Länge approximieren, wobei h für die Impulsantwort des Filters steht, das es zu approximieren gilt (z.B. $h = h_{\text{InvMin}\varphi}$), `FirOrd` die Ordnung des Zählerpolynoms und `IirOrd` die des Nennerpolynoms bezeichnet. Als Ergebnis erhält man die Koeffizienten des Zählerpolynoms `NumCoefs` (FIR-Filterkoeffizienten) sowie die des Nennerpolynoms `DenCoefs` (IIR-Filterkoeffizienten).

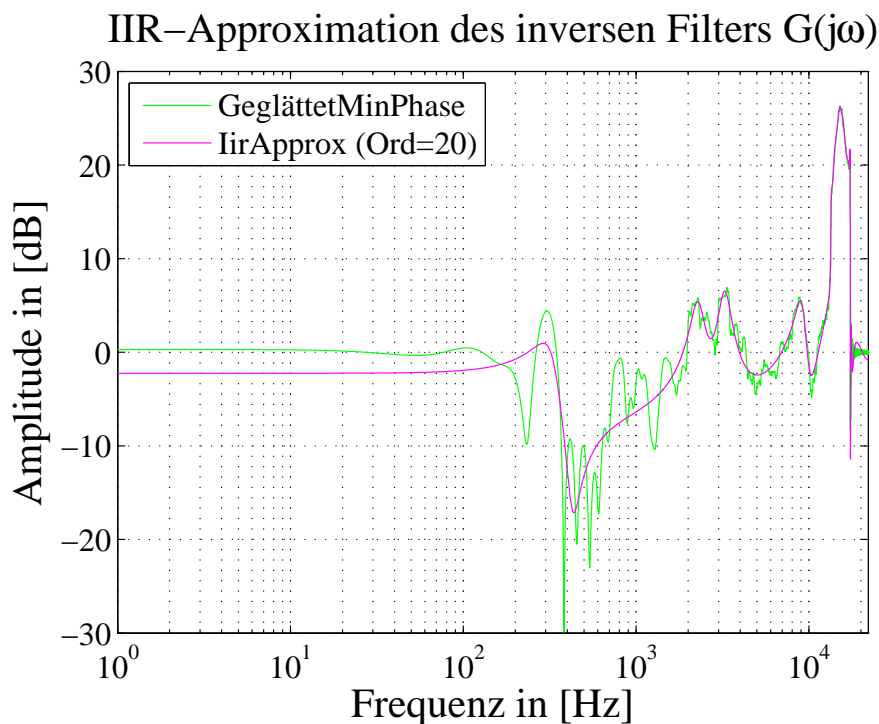


Abbildung 2.13: Approximation des inversen Filters $G(j\omega)$ mit einem IIR-Filter der Ordnung: `FirOrd=IirOrd = 20`.

Abbildung 2.13 zeigt das Ergebnis einer solchen IIR-Approximation, bei Verwendung eines Zähler- und Nennerpolynoms der Länge `FirOrd=IirOrd = 20`.

Wie man sieht, kann damit das gewünschte Filter approximativ angenähert werden, wobei auffällt, dass die Approximation hin zu hohen Frequenzen besser wird.

Aus diesem Grund sind andere Algorithmen, wie etwa die in [KP06] oder [Ban10] vorgestellten Verfahren, welche eine spektrale Gewichtung bzw. eine gezielte Platzierung der Polstellen des IIR-Filters, für die Annäherung einer beliebigen, meist minimalphasigen Impulsantwort zulassen, besser geeignet als die hier verwendete Methode nach Steiglitz-McBride. Diese liefert zwar meist brauchbare Ergebnisse wie in diesem Beispiel, lässt jedoch so gut wie keine Eingriffsmöglichkeit in die Approximation zu und ist somit sehr unflexibel. Egal welches Verfahren man auch wählt, das IIR-Filter sollte aufgrund des robusteren Verhaltens sowie der geringeren *Quantisierungsfehler* stets mit Hilfe von Filtern 2. Ordnung, sogenannten *biquadratischen Filtern*, die auch unter dem Begriff *Biquads*

bekannt sind, realisiert werden. In dem hier behandelten Fall macht eine Umsetzung von $h_{\text{InvMin}\varphi}[n]$ als IIR-Filter jedoch ohnehin keinen Sinn, da dieses am einfachsten und kostengünstigsten mit den Filtern zur Kompensation des Übersprechens $\text{adj}(\mathbf{H}(j\omega))$ durch Faltung integriert werden kann.

Wendet man das, in Abbildung 2.12 dargestellte inverse Filter $G(j\omega)$ mit einer Länge von 512 [Samples] @ $f_A = 44100$ [Hz] an, resultieren daraus die in Abbildung 2.14 dargestellten Ergebnisse.

Wie man unschwer aus Abbildung 2.14 erkennen kann, bleibt die Übersprechdämpfung von $G(j\omega)$ abermals unbeeinflusst, d.h. es wird nach wie vor eine Dämpfung von ≈ 40 [dB] erreicht.

Ferner weisen nun die ipsilateralen Impulsantworten aus Abbildung 2.14(a) einen strikt minimalphasigen Charakter auf, welcher zu keinen akustisch wahrnehmbaren Artefakten führt, obwohl diese im Vergleich zu den aus Abbildung 2.7(a) dargestellten Verhältnissen ohne Filterung eine etwas höhere Nachhallzeit aufweisen.

Dafür ist aus Abbildung 2.14(b) unschwer zu erkennen, dass nun die ipsilateralen Betragsfrequenzgänge keine klanglichen Verfärbungen mehr aufweisen zumindest keine, die störend wahrnehmbar wären.

Somit stellt die in diesem Abschnitt vorgestellte Methode eine wesentliche Verbesserung gegenüber den bereits bekannten zuvor behandelten Methoden dar.

Wie bereits erwähnt, kann die in Kapitel 2.4 beschriebene Methode auch auf mehrere Personen, wobei dann auch mehrere Lautsprecher benötigt werden, erweitert werden. Solange es sich dabei um ein bestimmtes System handelt, d.h. die Anzahl der Lautsprecher L gleich der Anzahl der Mikrophone M ist ($M = L$), kann das Verfahren aus Gleichung 2.13 auch direkt verwendet werden. Allerdings kommt es dadurch zu einer schlechter konditionierten Matrix, was wiederum zu einer wesentlich schlechteren Übersprechdämpfung führt, was meist durch eine Verlängerung der inversen Filter $\mathbf{C}(j\omega)$ wieder ausgeglichen werden kann.

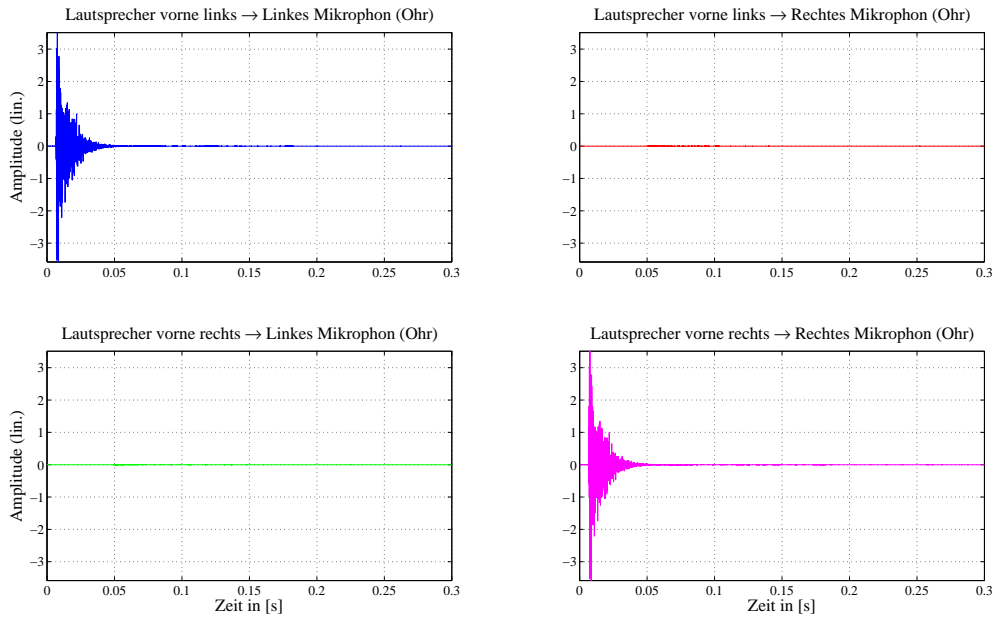
Zudem müssen die Ausgangspegel reduziert werden, damit es zu keinen Hochläufen aufgrund eventuell instabiler CTC-Filtern kommt.

Aus diesen Gründen sollte idealerweise stets ein *transaural-Stereo* System verwendet werden, welches optimalerweise über mehrere, gemäß Abbildung 2.3 angeordnete Lautsprecher in unterschiedlichen Spektralbereichen betrieben wird.

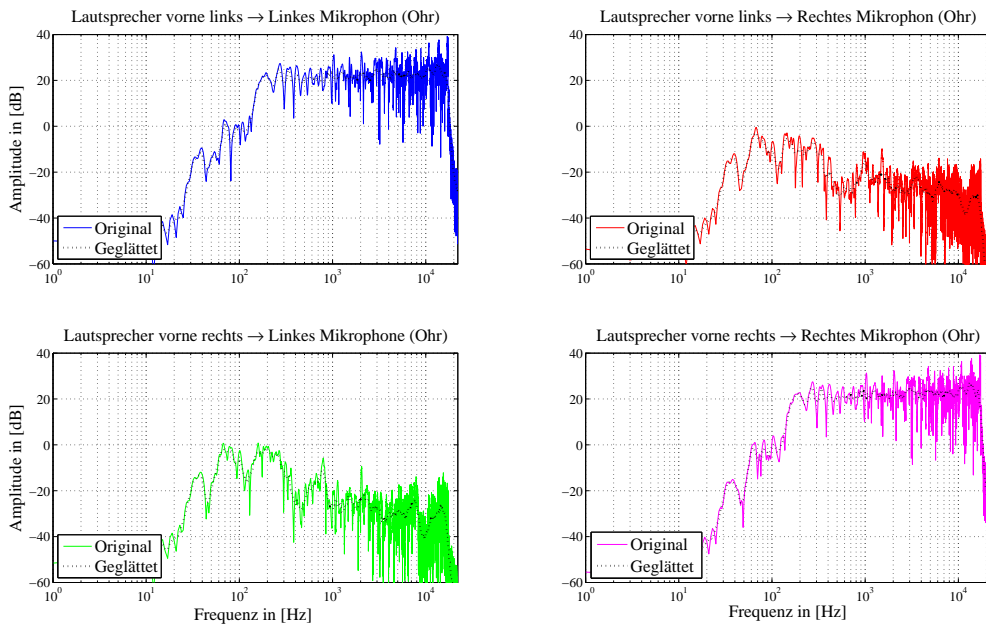
2.6 Probleme und Ausblick

Basierend auf dem in Kapitel 2.5 vorgestellten Beispiel und bei Verwendung der in Abschnitt 2.5.4 vorgestellten, modifizierten, vollständigen inversen Filter $\mathbf{C}(j\omega)$ ergeben sich in einer Umgebung von 5 [cm] bzw. 10 [cm] die in Abbildung 2.15 dargestellten Ergebnisse.

Wie man unschwer aus Abbildung 2.15 erkennen kann, bricht die Übersprechdämpfung



(a) Impulsantworten



(b) Betragsfrequenzgänge

Abbildung 2.14: Darstellung der Impulsantworten und der zugehörigen Betragsfrequenzgänge zwischen den beiden verwendeten Lautsprechern FL und FR und der beiden Mikrophone der binauralen Messungen (linkes und rechtes Ohrsignal), bei Verwendung des inversen Filter aus Abbildung 2.12(b).

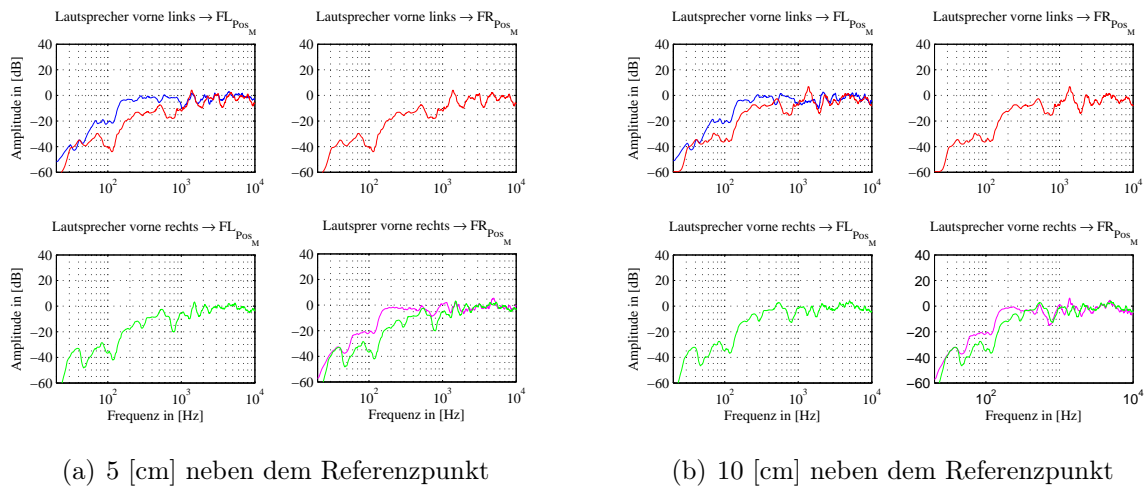


Abbildung 2.15: Darstellung der Betragsfrequenzgänge nach Anwendung der vollständigen, inversen Filter $C(j\omega)$, bei Verwendung eines 2×2 -Systems. Zur leichteren Beurteilung der resultierenden Übersprechdämpfung wurden jeweils die Verläufe der Betragsfrequenzgänge der Kreuzpfade, welche das Übersprechen beinhalten, in die direkten Pfade kopiert.

oberhalb von $f \approx 1$ [kHz] bereits bei geringer Entfernung vom Referenzpunkt vollständig ein. Unterhalb dieser Frequenz kann selbst in einer Entfernung von 10 [cm] neben dem Referenzpunkt noch eine Übersprechdämpfung erreicht werden, die jedoch im Vergleich zu den in Abbildung 2.14 dargestellten Verhältnissen am Referenzpunkt stark reduziert ist. Ferner ist aus Abbildung 2.15 ersichtlich, dass mit zunehmender Entfernung vom Referenz- bzw. Messpunkt die Übersprechdämpfung kontinuierliche abnimmt. Außerdem wird deutlich, dass die erreichbare Übersprechdämpfung mit steigender Frequenz abnimmt, wobei, wie bereits erwähnt, ab einer Frequenz von $f \approx 1$ [kHz] faktisch keine Übersprechdämpfung mehr zu verzeichnen ist. Man kann also nur in einem reduzierten Spektralbereich von $f < 1$ [kHz] von Verhältnissen sprechen, die mehr oder weniger robust gegenüber räumlichen Veränderungen sind.

Im Gegensatz zur Übersprechdämpfung haben unterschiedliche Abstände zum Referenzpunkt nur einen marginalen Einfluss auf die Klangfarbe.

Diese Ergebnisse verdeutlichen, dass bereits ein einfaches 2×2 -System auf die in diesem Kapitel aufgezeigte Art und Weise nicht problemlos eingesetzt werden kann.

Es existieren viele Arbeiten, die sich mit der Lösung dieses Problems beschäftigen, z.B. [Mou94, KNHOB98, Gar00, Sof02, LPY08, JMKM11], um nur einige zu nennen, wobei keines der dabei vorgestellten Verfahren in der Lage ist, das Problem grundlegend zu lösen.

Dennoch existieren Beispiele die aufzeigen, dass transaurale Methoden auch in der Praxis erfolgreich eingesetzt werden können [ALS05, PMV11]. Der Erfolg dieser Systeme begründet sich dabei stets in der Verwendung eines sogenannten *head-* oder *moti-*

on tracker¹⁶ und der Verwendung von richtungsabhängigen HRTFs¹⁷, die zumeist sogar individualisiert, d.h. von der jeweiligen Testperson stammen, als Grundlage für die Berechnung der notwendigen inversen Filter. Auf diese Weise lässt sich eine größtmögliche Unabhängigkeit gegenüber Positionsänderungen im Raum erreichen, allerdings mit den Einschränkungen, dass:

- ein System zur Positionsbestimmung im Raum vorhanden sein muss, dessen Sender man an einer definierten Stelle am Körper, z.B. am Kopf (*head tracker*) tragen muss sowie
- eine Datenbank mit HRTFs vorhanden sein muss, die im Idealfall sogar individualisiert ist, auf deren Basis dann
- die inversen Filter in Echtzeit berechnet werden müssen.

2.7 Zusammenfassung

In diesem Kapitel wurde das Prinzip der inversen Filterung als Möglichkeit zur Raumverzerrung bzw. Auralisation anhand eines *transaural-Stereo* Systems behandelt. Dabei wurde deutlich, dass neben der Art wie die Lautsprecher am Abhörort aufgestellt werden vorrangig die gewählte Form der Matrixinvertierung sowie die Regularisierung beim Filterentwurf eine erfolgreiche Realisierung maßgeblich beeinflussen. Die Berechnung der inversen Filter mit Hilfe der adjungierten Matrix erwies sich dabei aufgrund ihrer Eigenschaften die CTC-Filter maximal minimalphasig zu belassen, als eine sehr gute Option. Anhand eines Beispiels, welches auf realen RIR-Messungen in einem Fahrzeug basierte, konnte einerseits die Wirkungsweise der mit Hilfe der adjungierten Matrix entworfenen CTC-Filter gezeigt werden, andererseits aber auch die verbleibende Problematik der Realisierung des Kompensationsfilters $G(j\omega)$ verdeutlicht werden. Eine mögliche Lösung dieses Problems wurde in Kapitel 2.5.4 erarbeitet, welche zu folgender Erfindungsmeldung führte [Chr12c]. In diesem Zusammenhang wurden auch noch zwei weitere Erfindungsmeldungen [Chr12a, Chr12b] eingereicht. Die erste setzte sich dabei mit einem Gesamtkonzept zur Realisierung *individueller Klangzonen* mit Hilfe einer Kombination aus den im Fahrzeug bereits vorhandenen Systemlautsprechern und zusätzlichen in den Kopfstützen verbauten Lautsprechern auseinander. Die andere beschäftigte sich hingegen mit einem gewinnbringenden Einsatz gerichteter Lautsprecher (z.B. im Dachhimmel verbaute EDPL¹⁸s). Eine Diskussion über verbleibende Probleme und der daraus resultierenden Konsequenzen beim Einsatz inverser Filter beendete dieses Kapitel, wobei deutlich wurde, dass ein

¹⁶Mit Hilfe eines *head-* oder *motion tracker* kann die momentane (Kopf)Position im dreidimensionalen Raum festgestellt werden. Dabei existieren unterschiedliche Varianten, von denen die Besten imstande sind, die Position mit einer Genauigkeit von $\leq 1^\circ$ zu bestimmen.

¹⁷HRTF = engl. **H**ead **R**elated **T**ransfer **F**unction

¹⁸EDPL = engl. **E**lectro **D**ynamic **P**lanar **L**oudspeaker (elektrodynamischer Flachlautsprecher)

erfolgreicher (breitbandiger) Einsatz nur durch eine Individualisierung (*head-tracker*, individualisierte HRTFs) sowie durch eine Begrenzung der Ordnung des zu invertierenden Systems (Konditionierung der zu invertierenden Matrix) gelingen kann.

Kapitel 3

Ambisonic 1. Ordnung

Bei *Ambisonic* handelt es sich um ein System, das bereits Anfang der 70er Jahre von Michael Gerzon [Ger74] entwickelt wurde. Das Prinzip basiert auf der Tatsache, dass sich ein beliebiges Wellenfeld in mehrere Kugelflächenfunktionen, sogenannten *sphärischen Harmonischen*, ähnlich wie bei der Fourieranalyse, zerlegen lässt. Dabei kommt ein mehrkanaliges Aufnahme- und Wiedergabesystem, welches über eine frei konfigurierbare Codier- und Decodiervorschrift miteinander verbunden ist, zum Einsatz. Für die Aufnahme wird ein spezielles Mikrofon verwendet, bei dem vier jeweils senkrecht zueinander angeordnete Mikrophone mit sub-kardioider Charakteristik auf engen Raum miteinander mechanisch verbunden sind. Als Wiedergabesystem kann im Prinzip eine beliebige Anzahl von Lautsprechern verwendet werden, die sich auf einer Kreislinie (2D-Fall) oder auf einer Kugeloberfläche (3D-Fall) befinden und idealerweise symmetrisch angeordnet sind. Der Mittelpunkt des Kreises bzw. der Kugel stellt die Position dar, an welcher die gewünschte Approximation des Wellenfeldes wirken soll. Wie man bereits erkennt, ist das Ambisonicverfahren in der Lage sowohl 2D- als auch 3D-Wellenfelder zu approximieren, wobei uns auch hier vorrangig der 2D-Fall interessiert.

Nach [Ger74] darf ein Decoder (Codier-/Decodiersystem) dann als Ambisonic bezeichnet werden, wenn auf einer zentralen Abhörposition folgendes gilt:

- Der Winkel des decodierten Schallschnelle- (\mathbf{v}) und Schallenergievektors (\mathbf{e}) über alle Frequenzen übereinstimmt.
- Bei tiefen Frequenzen (unterhalb von $f \approx 400$ [Hz]) die Amplitude (Länge) des Schallschnellevektors \mathbf{v} zu 1 für alle reproduzierten, horizontalen Abstrahlwinkel (Azimut) wird.
- Bei mittleren und hohen Frequenzen (700 [Hz] $\leq f \leq 4000$ [Hz]) die Amplitude (Länge) des Schallenergievektors \mathbf{e} über alle Azimutwinkel maximiert, d.h. idealerweise ebenfalls zu 1 wird.

Diese Bedingungen werden erfüllt, wenn es sich bei der aufgenommenen Welle (Referenzwelle) um eine *ebene Welle* handelt und von den Lautsprechern ebenfalls eine *ebene Welle*

reproduziert wird, was der Fall ist, wenn

- sich die Schallquelle in ausreichender Entfernung zum Hörer befindet,
- die Lautsprecher weit genug vom Hörer entfernt sind.

Des Weiteren beschränkt sich das Ambisonicverfahren auf die Analyse und Synthese von Direktschallen, wie sie im Freifeld bzw. in reflexionsarmen Räumen vorkommen. Etwaig vorhandene Reflexionen (Nachhall) werden nicht berücksichtigt. Wie oben angedeutet, spielt beim Ambisonicverfahren sowohl der Schallschnellevektor \mathbf{v} , welcher im unteren Frequenzbereich dominiert und somit eine Approximation der ITD darstellt, als auch der Schallenergievektor \mathbf{e} , der vorrangig im mittleren und oberen Spektralbereich verwendet wird und somit ein Maß für die ILD ist, eine tragende Rolle beim Entwurf des Decoders. Werden bei der Analyse *differentielle Mikrophone* 1. Ordnung, genauer gesagt Mikrophone mit *bidirektionaler Richtcharakteristik*, die auch als *Achtcharakteristik*¹ bezeichnet werden, verwendet, wird das Schallfeld bei der Analyse in vier unterschiedliche Signale (W, X, Y, Z) codiert, die in der Studioteknik unter der Bezeichnung *B-Format-Signale* bekannt sind.

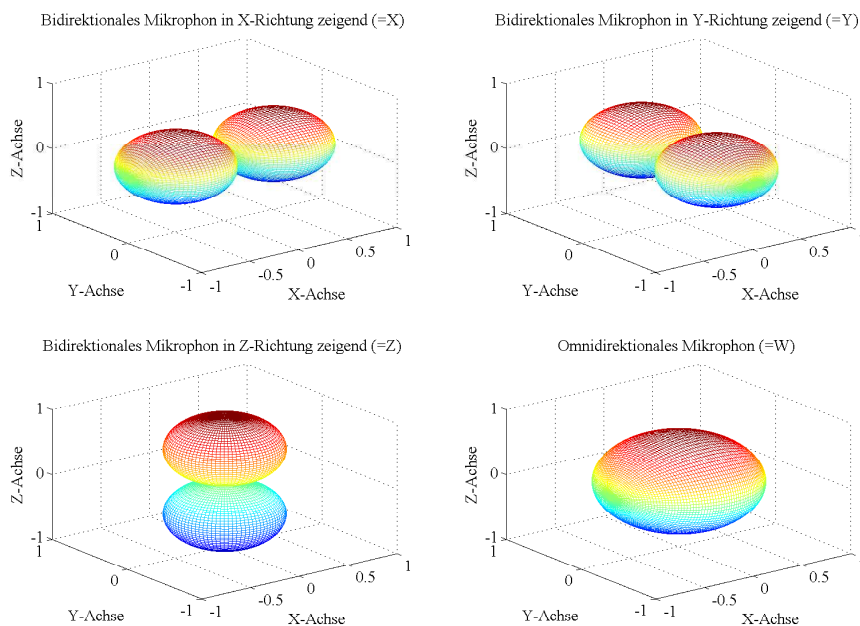


Abbildung 3.1: 3D Darstellungen der rotationssymmetrischen Richtdiagramme der Mikrophone die zur Aufzeichnung der Ambisonicsignale 1. Ordnung (X, Y, Z, W) verwendet werden.

¹Der Begriff *Achtcharakteristik* leitet sich von der Form der Richtcharakteristik ab, da diese der Form der Zahl 8 gleicht, also zwei gegenüberliegende, symmetrische Pole besitzt, weshalb Mikrophone mit einer solchen Richtcharakteristik auch unter dem Begriff *Dipolmikrophone* bekannt sind.

Hierbei bezeichnet W ein Signal, das von einem *omnidirektionalen Mikrophon*, d.h. von einem Mikrophon mit *Kugelcharakteristik*, stammt, welches sich in der Mitte des Mikrophonarrays zur Aufnahme der Ambisonicsignale befindet. Da das W -Signal von einem *Schalldruckmikrophon*, also einem Mikrophon ohne Richtwirkung stammt, repräsentiert es den *Schalldruckverlauf* bzw. die Komponente 0. Ordnung des Schallfeldes am Abhörort. Die verbleibenden B-Format-Signale X , Y und Z stammen dagegen von differentiellen Mikrophonen mit Achtcharakteristik, die entsprechend den gleichnamigen Achsen des kartesischen Koordinatensystems senkrecht zueinander angeordnet sind. Mit ihnen kann die Schallschnelle \mathbf{v} , d.h. die Schalldruckdifferenz bzw. die Komponente 1. Ordnung des Schallfeldes, in den drei Dimensionen (\mathbf{v}_x , \mathbf{v}_y , \mathbf{v}_z) bestimmt werden, was die Basis für eine dreidimensionale Ortsbestimmung darstellt. Des Weiteren ist das Ambisonicverfahren nicht auf die Verwendung Mikrophone 1. Ordnung festgelegt, sondern kann auch mit Mikrophonen höherer Ordnung betrieben werden, wobei dann mehr als nur vier Signale entstehen ([Son03]).

Erwähnenswert ist auch, dass es beim Ambisonicverfahren keine Abhängigkeit zwischen der Analyse und der Synthese gibt, d.h. ein analysiertes Signal lässt per se noch keine Aussage über das Synthesesystem zu. Bei der Synthese gilt einzig zu beachten, dass das System überbestimmt sein sollte, damit keine Richtungsinformationen verloren gehen. Das heißt die Anzahl der Lautsprecher L muss mindestens gleich oder, wie wir später noch erfahren werden, besser noch größer als die Anzahl der Ambisonickanäle sein, damit die in ihnen enthaltenen Richtungsinformationen erhalten bleiben. Außerdem werden beim Ambisonicverfahren nur drei der vier Kanäle benötigt, wenn man im Wiedergabesystem lediglich an einer 2D-Reproduktion, d.h. an einer Reproduktion in der Fläche und nicht an einer 3D-Reproduktion, also an einer Reproduktion im Raum, interessiert ist. Üblicherweise ist man an einer Reproduktion in der Horizontalebene interessiert, wobei dann der Z -Kanal beim Wiedergabesystem entfallen kann.

Das Ambisonicverfahren bietet zudem die Möglichkeit künstlich generierte Signale, die entsprechend codiert wurden, über das Synthesesystem abzuspielen. Mit Hilfe der Formeln in Gleichung 3.1 ([Wig04]) lassen sich die virtuellen B-Format-Signale 1. Ordnung (W, X, Y, Z) anhand einer beliebig vorgegebenen Raumrichtung ($\Theta_{SRC}, \varphi_{SRC}$), an welcher die virtuelle Schallquelle positioniert sein soll, wie folgt berechnen:

$$\begin{aligned}
 W &= \frac{1}{\sqrt{2}} s(t), \\
 X &= \cos(\Theta_{SRC}) \cos(\varphi_{SRC}) s(t), \\
 Y &= \sin(\Theta_{SRC}) \cos(\varphi_{SRC}) s(t), \\
 Z &= \sin(\varphi_{SRC}) s(t), \text{ mit} \\
 \Theta_{SRC} &= \text{Horizontaler Winkel (Azimut) der Schallquelle,} \\
 \varphi_{SRC} &= \text{Vertikaler Winkel (Elevation) der Schallquelle.}
 \end{aligned} \tag{3.1}$$

Für einen solchen Fall, d.h. für die Wiedergabe eines beliebigen Signals aus einer frei definierbaren Raumrichtung, was anhand eines vorgegebenen Lautsprecher-Setups erreicht werden soll, betrachten wir nachfolgend beispielhaft den Entwurf des gemäß des Ambisonic-Prinzips notwendigen Codier- und Decodiersystems. Dabei werden mit Hilfe von Gleichung 3.1 die (virtuellen) B-Format-Signale berechnet, also die Signale bestimmt, die in Wirklichkeit von den Mikrofonen mit der in Abbildung 3.1 dargestellten Charakteristik aufgezeichnet würden.

Für die Wiedergabe von B-Format-Signalen werden Richtcharakteristika virtueller Mikrophone berechnet, mit welchen die im Wiedergabesystem verwendeten Lautsprecher gespeist werden. Durch Überlagerung der zugrunde liegenden Ambisonic Richtcharakteristika 1. Ordnung ($3 \times$ bidirektional (orthogonal), $1 \times$ omnidirektional (zentral)) lassen sich beliebige virtuelle Richtcharakteristika 1. Ordnung, die zwischen einer omni- und einer bidirektionalen Richtcharakteristik liegen und in jede beliebige Raumrichtung zeigen können, erzeugen. Die nachfolgenden Formeln der Gleichung 3.2, welche aus [FATG01] stammen, verdeutlichen das Prinzip der Generierung virtueller Mikrophone 1. Ordnung auf Basis der B-Format-Signale.

$$\begin{aligned}
 g_W &= \sqrt{2}, \\
 g_X &= \cos(\Theta_S) \cos(\varphi_S), \\
 g_Y &= \sin(\Theta_S) \cos(\varphi_S), \\
 g_Z &= \sin(\varphi_S), \\
 S &= \frac{1}{2} [(2 - q) g_W W + q (g_X X + g_Y Y + g_Z Z)], \text{ mit} \\
 W, X, Y, Z &= \text{B-Format-Signale gemäß Gleichung 3.1,} \\
 S &= \text{Ansteuersignal für den Lautsprecher an der Position } (\Theta_S, \varphi_S), \\
 \Theta_S &= \text{Azimut des Lautsprechers } S, \\
 \varphi_S &= \text{Elevation des Lautsprechers } S, \\
 q &= \text{Richtfaktor (engl. } \textit{directivity factor} \text{) } (0, \dots, 2).
 \end{aligned} \tag{3.2}$$

Wie man aus Gleichung 3.2 erkennt, bestimmt die räumliche Lage des entsprechenden Lautsprechers S , gegeben durch dessen Azimut Θ_S und dessen Elevation φ_S , die Gewichtungsfaktoren (g_W, g_X, g_Y, g_Z) mit denen die B-Format-Signale aus Gleichung 3.1 (W, X, Y, Z) gewichtet werden müssen, damit ein virtuelles Mikrophon 1. Ordnung entsteht, welches exakt in die Richtung des entsprechenden Lautsprechers S zeigt.

Wie man aus den Graphen der Abbildung 3.2 sieht, können aus der Überlagerung der omnidirektionalen (W) mit der bidirektionalen Richtcharakteristik (hier Y , da $\varphi = 0^\circ$) sämtliche virtuellen Richtcharakteristika 1. Ordnung, beginnend mit der omnidirektionalen ($q = 0$), bis hin zur bidirektionalen Richtcharakteristik ($q = 2$), im fließenden Übergang durch den *Richtfaktor* q erzeugt werden.

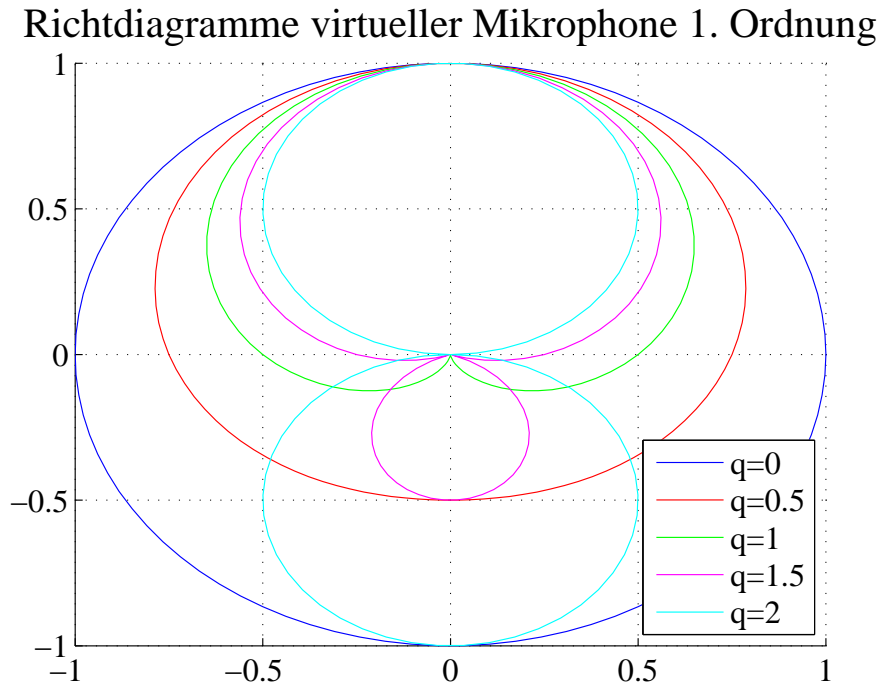


Abbildung 3.2: Richtdiagramme virtueller Mikrophone 1. Ordnung für $\varphi = 0^\circ$, variiert über den zulässigen Parameterbereich des Richtfaktors q ($0 \leq q \leq 2$).

3.1 Ambisonic Codier/Decodier-Beispiel

Nachfolgend wird anhand eines Beispiels das Prinzip des Ambisonic Codier/Decodier-Prozesses näher erläutert. Dabei wird zunächst von einem Lautsprechersetup mit insgesamt $L = 6$ Lautsprechern, die sich alle auf gleicher Höhe befinden und in gleichen Abständen in einem Kreis angeordnet sind, ausgegangen. Das heißt die L Lautsprecher besitzen alle eine Elevation von $\varphi_S = 0^\circ$ sowie einen Azimut von $\Theta_S = [0^\circ : 60^\circ : 300^\circ]$, wie aus Abbildung 3.3 ersichtlich.

Ferner wird der *Richtfaktor* zu $q = 1$ gewählt, d.h. es werden virtuelle Mikrophone mit Kardioid-Charakteristik verwendet, die gemäß Gleichung 3.2 so berechnet werden, dass je eines zu den entsprechenden Lautsprecherpositionen ($\Theta_S = [0^\circ : 60^\circ : 300^\circ]$, $\varphi_S = 0^\circ$) zeigt. Dies ist in Abbildung 3.4 dargestellt.

Darüber hinaus wird in dem Beispiel davon ausgegangen, dass sich die Quelle bei ($\Theta_{SRC} = 35^\circ$, $\varphi_{SRC} = 0^\circ$) befindet, was unter Berücksichtigung der Gewichtungsfaktoren, die sich aus Gleichung 3.1 und 3.2 ergeben, zu dem in Abbildung 3.5 dargestellten Resultat führt.

Die Größe der Richtcharakteristika der virtuellen Kardioid-Mikrophone in Abbildung 3.5 verdeutlichen, mit welcher Größe die zugehörigen Lautsprecher angesteuert werden. Wie man sieht, werden dabei die Lautsprecher, welche sich in unmittelbarer Umgebung der Signalquelle $\Theta_{SRC} = 35^\circ$, $\varphi_{SRC} = 0^\circ$ befinden ($\Theta_S = [0^\circ, 60^\circ]$, $\varphi_S = 0^\circ$)

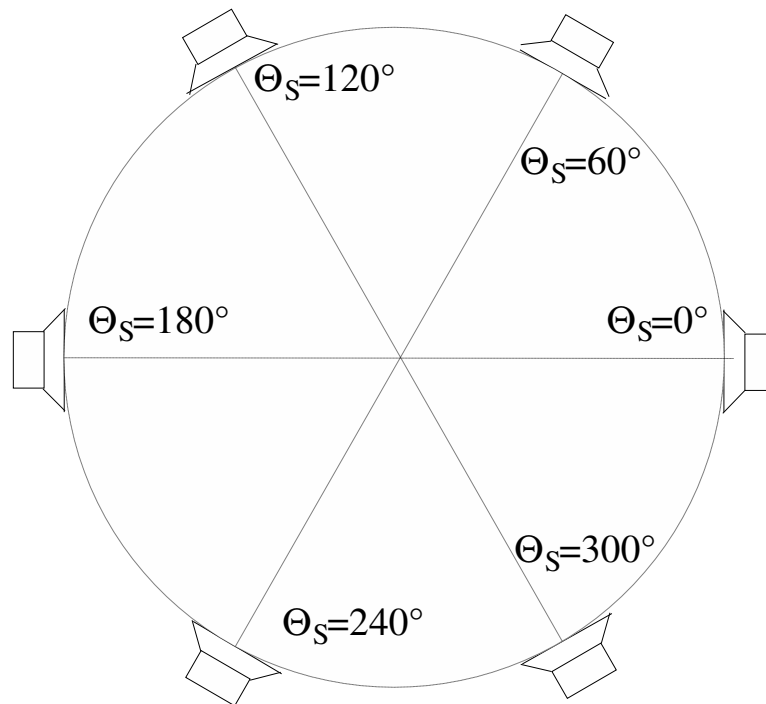


Abbildung 3.3: Gewähltes Lautsprechersetup für das Beispiel.

Richtdiagramme der virtuellen Mikrophone 1. Ordnung für 6 Lautsprecher, befindlich bei $\Theta_{\text{SRC}}=[0^\circ:60^\circ:300^\circ]$

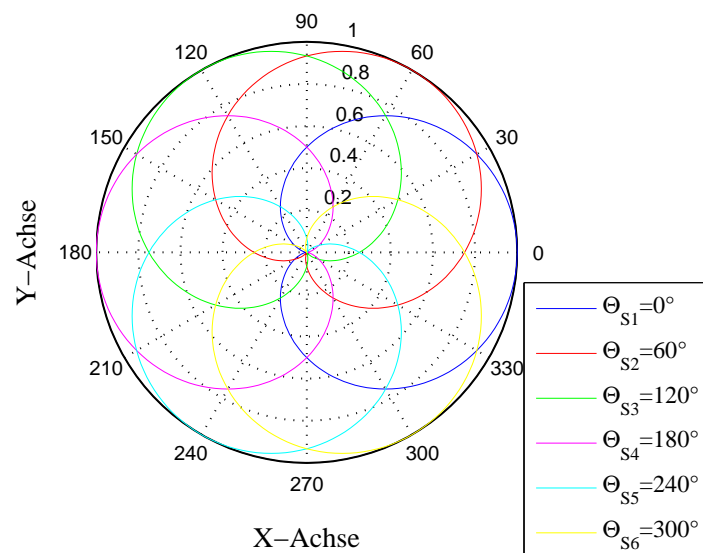


Abbildung 3.4: Virtuelle Mikrophone mit Kardiod-Charakteristik ($q = 1$), die zu den jeweiligen Lautsprecherpositionen: $\Theta_S = [0^\circ : 60^\circ : 300^\circ]$, $\varphi_S = 0^\circ$ zeigen.

Richtdiagramme der virtuellen Mikrophone 1. Ordnung für die 6 Lautsprecher, für eine Quelle bei $\Theta_{SRC}=35^\circ$ und $\phi_{SRC}=0^\circ$.

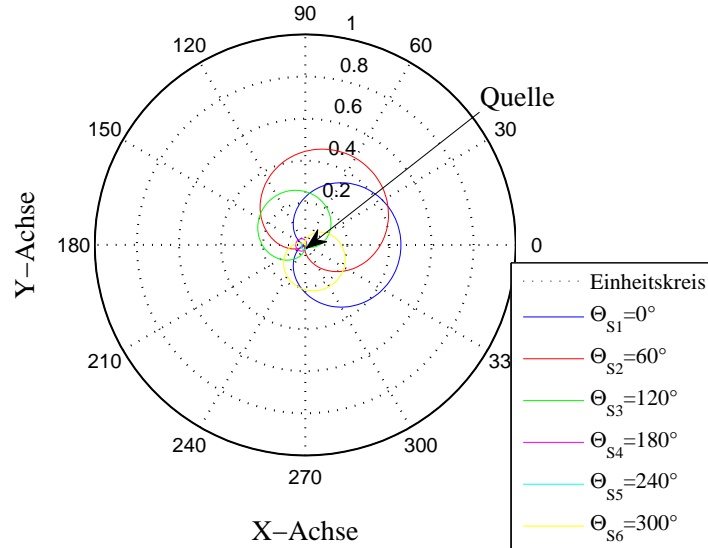


Abbildung 3.5: Richtdiagramme des Beispiels mit einer Quelle bei: $\Theta_{SRC} = 35^\circ, \varphi_{SRC} = 0^\circ$ und einem Lautsprechersetup von: $\Theta_S = [0^\circ : 60^\circ : 300^\circ], \varphi_S = 0^\circ$, zu denen je ein resultierendes virtuelles Mikrophon zeigt. (Die Komponenten der Lautsprecher bei $(\Theta_S = [180^\circ, 240^\circ], \varphi_S = 0^\circ)$ sind sehr klein und deshalb im Bild kaum zu erkennen.)

mit höheren Pegel angesteuert als jene, die sich in entgegengesetzter Richtung befinden ($\Theta_S = [180^\circ, 240^\circ], \varphi_S = 0^\circ$).

Durch Überlagerung der Richtcharakteristika sämtlicher virtuellen Mikrophone erhält man die in Abbildung 3.6 dargestellte, resultierende, d.h. wahrgenommene Richtwirkung des Quellsignals.

Auch Abbildung 3.6 lässt erkennen, dass die Horizontallokalisation erhalten bleibt, jedoch die wahrgenommene Energie aus der Empfangsrichtung etwas zu groß ($r_v > 1$) erscheint. Ferner vermittelt der *Öffnungswinkel*² der resultierenden Richtcharakteristik einen Eindruck über die Lokalisationsschärfe, welche dadurch bedingt, dass es sich hier um ein Ambisonicsystem 1. Ordnung handelt, gemäß der damit erreichbaren virtuellen Richtcharakteristika, entsprechend gering ausfällt. Außerdem ist ersichtlich, dass, obwohl in dem Beispiel eine Kardiod-Charakteristik ($q = 1$) für die virtuellen Mikrophone gewählt wurde, bei welcher aus der entgegengesetzten Empfangsrichtung keine Signale empfangen werden können, die resultierende Richtcharakteristik sehr wohl Anteile aus dieser Richtung empfängt. Dies bedeutet, dass es außerhalb der optimalen Abhörposition (Mittelpunkt des Lautsprecherrings) zu Fehlortungen kommt, die umso stärker ausfallen, je weiter man sich von diesem entfernt.

²Als *Öffnungswinkel* wird der Winkel bezeichnet der sich ergibt, wenn sich im Vergleich zur Hauptempfangsrichtung die Empfangsleistung halbiert.

Überlagerung der Richtdiagramme aller virtuellen Mikrophone 1. Ordnung für alle 6 Lautsprecher und einer Quelle bei $\Theta_{SRC}=35^\circ$ und $\phi_{SRC}=0$

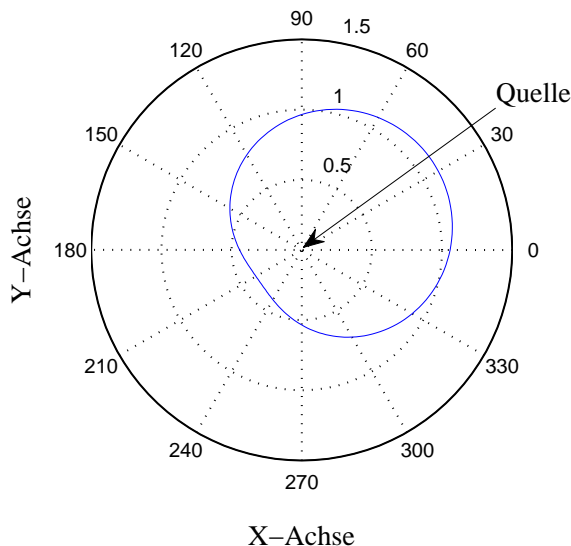


Abbildung 3.6: Resultierendes Richtdiagramm durch Überlagerung sämtlicher Richtdiagramme des Beispiels, mit einer Quelle bei: $\Theta_{SRC} = 35^\circ$, $\varphi_{SRC} = 0^\circ$ und einem Lautsprechersetup von: $\Theta_S = [0^\circ : 60^\circ : 300^\circ]$, $\varphi_S = 0^\circ$.

Die Entscheidung welche Richtcharakteristika, d.h. welcher Richtfaktor q für den Decoder verwendet werden soll, kann entweder empirisch, wie im obigen Beispiel oder theoretisch mit Hilfe eines Optimierungsalgorithmus getroffen werden. Letzteres führt uns dabei wieder zurück zu der ursprünglichen Aussage, was ein System „Ambisonic“ macht. Wie zu Beginn dieses Kapitels erwähnt, soll das Ambisonicsystem einen mehrkanaligen Decoder liefern, welcher in der Lage ist, ein gegebenes Wellenfeld so akkurat wie möglich nachzubilden. Hierfür stehen gemäß der Ambisonictheorie sowohl der Schallenergie- (\mathbf{e}) als auch der Schallschnellevektor (\mathbf{v}), in Verbindung mit dem gewählten Lautsprechersetup, für den Entwurf des Decoders zur Verfügung. Die dahinterstehenden Theorien nach Gerzon werden in Kapitel A.1.3 näher erläutert. Wendet man die beiden Theorien auf unser Beispiel an, ergeben sich für unterschiedliche, horizontale Einfallswinkel der virtuellen Schallquelle (Θ_{SRC}) die in Abbildung 3.7 dargestellten Schallschnelle- und Schallenergievektoren.

Ferner kann man aus Abbildung 3.7 erkennen, dass sich bei Verwendung von virtuellen Mikrophenen mit Kardioid-Charakteristik ($q = 1$) keine optimalen Schallschnelle- und Schallenergievektoren ergeben, was an deren Längen r_v und r_e ersichtlich ist. Erstaunlicherweise bereitet die Nachbildung der Horizontallokalisation in unserem Beispiel keine Probleme. Der Grund hierfür liegt einerseits in der Wahl des Lautsprechersetup, welches, wie aus Abbildung 3.3 ersichtlich, in unserem Beispiel eine regelmäßige, d.h. äquidistante Anordnung aufweist und andererseits in der Verwendung des gleichen Richtfaktors q für den Entwurf sämtlicher virtuellen Mikrophone.

Unabhängig von deren Anzahl L gilt allgemein für regelmäßig angeordnete Lautspre-

Schallschnelle- und Schallenergievektor bei einem Richtfaktor von $q=1$

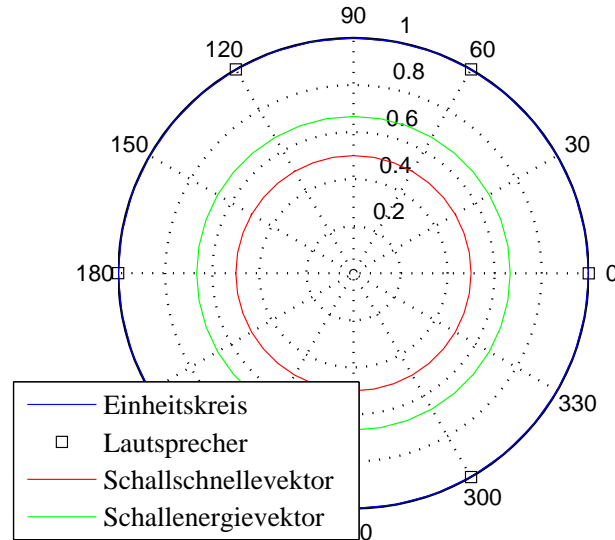


Abbildung 3.7: Schallschnelle- \mathbf{v} und Schallenergievektoren \mathbf{e} , für $q = 1$, $\Theta_{SRC} = [0^\circ, \dots, 360^\circ]$, $\varphi_{SRC} = 0^\circ$ und $\Theta_S = [0^\circ : 60^\circ : 300^\circ]$, $\varphi_S = 0^\circ$.

cheraufstellungen, auf welche virtuelle Mikrophone mit der jeweils gleichen Richtcharakteristik ($q = \text{Konstant}$) zeigen, dass:

1. es zu keiner Veränderung der Horizontallokalisation kommt, also stets $\tilde{\Theta}_{\mathbf{v}} = \tilde{\Theta}_{\mathbf{e}} = \Theta_{SRC}$ ist und
2. sowohl die Energie als auch die Leistung der Schallwelle, welche Maße für die Lautstärke der Quelle darstellen, unabhängig von deren horizontaler Lage (Θ bzw. Θ_{SRC}) sind, was anhand der kreisrunden Form sowohl von $r_{\mathbf{v}}(\Theta_{SRC})$ als auch von $r_{\mathbf{e}}(\Theta_{SRC})$ in Abbildung 3.7 zu sehen ist.

Somit reduziert sich wegen Punkt 1 der zuvor erwähnte Optimierungsalgorithmus, bei Verwendung einer regelmäßig angeordneten Lautsprecheraufstellung sowie von virtuellen Mikrophenen mit jeweils gleicher Richtcharakteristik, auf die Suche eines Richtfaktors q , welcher ausschließlich die Länge des Schallschnellevektors $r_{\mathbf{v}}(\Theta_{SRC})$ sowie des Schallenergievektors $r_{\mathbf{e}}(\Theta_{SRC})$ optimiert. Zudem genügt es, aufgrund des Punktes 2, die beiden Vektoren $r_{\mathbf{v}}(\Theta_{SRC})$ und $r_{\mathbf{e}}(\Theta_{SRC})$ lediglich bei einer beliebigen, horizontalen Schalleinfallrichtung Θ_{SRC} zu betrachten.

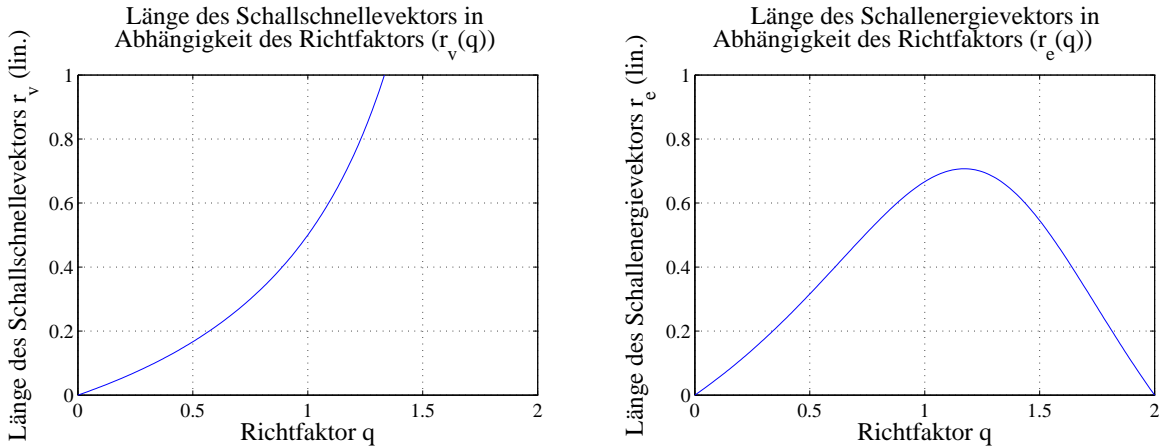
Der Schallschnellevektor $r_{\mathbf{v}}(\Theta_q)$ lässt sich für ein zweidimensionales System im Allgemeinen gemäß Gleichung A.11 unter Berücksichtigung der Gleichungen A.10 und A.9 bestimmen. Ähnliches gilt für den Schallenergievektor $r_{\mathbf{e}}(q)$. Dieser kann mit Gleichung A.14

unter Berücksichtigung der Gleichungen A.13 und A.12 berechnet werden. Somit wird deutlich, dass beide Vektoren, d.h. sowohl $r_{\mathbf{v}}(\Theta_q)$ als auch $r_{\mathbf{e}}(q)$ ausschließlich von den Verstärkungsfaktoren G_i der L Lautsprecher abhängen. Diese können gemäß Gleichung 3.2, wobei sie dort mit S bezeichnet werden, unter Berücksichtigung von Gleichung 3.1, für einen beliebigen Richtfaktors q , bestimmt werden.

Somit ergeben sich die Längen des Schallschnellevektors $r_{\mathbf{v}}(q)$ sowie des Schallenergievektors $r_{\mathbf{e}}(q)$, unter Berücksichtigung der zuvor genannten Vereinfachungen, in Abhängigkeit des Richtfaktors q zu:

$$\begin{aligned} r_{\mathbf{v}}(q) &= \frac{q}{2(2-q)}, \\ r_{\mathbf{e}}(q) &= \frac{3q + 1.5q^2}{6 - 6q + 2.25q^2}, \text{ mit} \\ r_{\mathbf{v}}(q) &= \text{Länge des Vektors } \mathbf{v} \text{ in Abhängigkeit des Richtfaktors } q, \\ r_{\mathbf{e}}(q) &= \text{Länge des Vektors } \mathbf{e} \text{ in Abhängigkeit des Richtfaktors } q. \end{aligned} \quad (3.3)$$

Variiert man den Richtfaktor q innerhalb seines gültigen Wertebereichs von $q = [0, \dots, 2]$, ergeben sich für $r_{\mathbf{v}}(q)$ sowie für $r_{\mathbf{e}}(q)$ Verläufe, wie in Abbildung 3.8 dargestellt.



(a) Länge des Schallschnellevektors $r_{\mathbf{v}}(q)$ bei unterschiedlichen Richtfaktoren $q = [0, \dots, 2]$

(b) Länge des Schallenergievektors $r_{\mathbf{e}}(q)$ bei unterschiedlichen Richtfaktoren $q = [0, \dots, 2]$

Abbildung 3.8: Darstellung der Vektorlängen $r_{\mathbf{v}}(q)$ und $r_{\mathbf{e}}(q)$ bei unterschiedlichen Richtfaktoren $q = [0, \dots, 2]$

Wie man aus Abbildung 3.8(a) sieht, existiert im Falle des Schallschnellevektors $r_{\mathbf{v}}(q)$ ein Wert $q_{\mathbf{vOpt}}$ bei dem $r_{\mathbf{v}}(q_{\mathbf{vOpt}}) = 1$ wird, was jedoch nicht für die Länge des Schallenergievektors zutrifft, wie Abbildung 3.8(b) verdeutlicht. Hier kann lediglich ein Wert für den Richtfaktor $q_{\mathbf{eOpt}}$ gefunden werden, bei dem $r_{\mathbf{e}}(q_{\mathbf{eOpt}}) = \frac{1}{\sqrt{2}}$ wird. Die jeweilig optimalen Werte, d.h. der Richtfaktor $q_{\mathbf{vOpt}}$, bei dem der Graph aus Abbildung 3.8(a)

den Wert 1 erreicht, sowie der Richtfaktor $q_{\mathbf{eOpt}}$, bei dem das Maximum des Graphen aus Abbildung 3.8(b) erreicht wird, berechnet sich dabei zu:

$$\begin{aligned} q_{\mathbf{vOpt}} &= \frac{4}{3} \approx 1.33, \\ q_{\mathbf{eOpt}} &= \frac{18 - \sqrt{162}}{4.5} \approx 1.17, \text{ mit} \\ q_{\mathbf{vOpt}} &= \text{Optimaler Richtfaktor } q \text{ des Schallschnellevektors } \mathbf{v}, \\ q_{\mathbf{eOpt}} &= \text{Optimaler Richtfaktor } q \text{ des Schallenergievektors } \mathbf{e}. \end{aligned} \quad (3.4)$$

Werden diese optimalen Werte verwendet, ergeben sich die in Abbildung 3.9 dargestellten Verhältnisse für die Schallschnelle- \mathbf{v} und Schallenergievektoren \mathbf{e} .

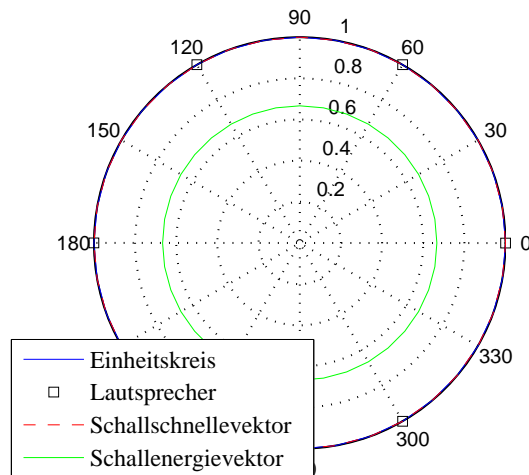
Ein Vergleich von Abbildung 3.7 ($q = 1$) mit Abbildung 3.9(a) ($q = 1.33$) bzw. Abbildung 3.9(b) ($q = 1.17$) verdeutlicht, in welchem Umfang sich die Vektorlängen durch Einsatz eines optimierten Richtfaktors vergrößern, wodurch es zu einer Verbesserung der Lokalisationsschärfe kommt. Dabei können im Sinne der Ambisonictheorie für den Schallschnellevektor \mathbf{v} durch Verwendung eines Richtfaktors von $q = 1.33$ optimale Verhältnisse erzeugt werden, was bedeutet, dass in einem Frequenzbereich von $f < 700$ [Hz] ein nahezu perfektes Abbild des Wellenfeldes möglich ist. Für den Schallenergievektor \mathbf{e} , der bekanntermaßen für Frequenzen $f > 700$ [Hz] bessere Ergebnisse in der Vorhersage der empfundenen Horizontallokalisation liefert als der Schallschnellevektor \mathbf{v} , kann der theoretische Optimalwert von $r_{\mathbf{e}} = 1$ nicht erreicht werden, da es prinzipbedingt nicht möglich ist, virtuelle Quellen mit $r_{\mathbf{e}} = 1$ zu erzeugen, da hierzu stets mehrere „echte“ Quellen (Lautsprecher) verwendet werden. Aus diesem Grund muss man sich hier mit einer Maximierung von $r_{\mathbf{e}}$ begnügen, wobei sich zeigte, dass sich das Maximum bei $r_{\mathbf{e}} = \frac{1}{\sqrt{2}}$ befindet, was nach Gleichung 3.4 durch einen Richtfaktor von $q = 1.17$ erreicht wird.

Daraus folgt, dass ein optimales Ambisonicsystem mit äquidistant angeordneten Lautsprechern, zu denen virtuelle Mikrophone mit jeweils gleicher Richtcharakteristik, wie in Abbildung 3.10 dargestellt, zeigen, für Frequenzen $f < 700$ [Hz] einen Richtfaktor von $q = 1.33$ und darüber hinaus, d.h. für Frequenzen $f > 700$ [Hz], einen Richtfaktor von $q = 1.17$ aufweisen muss. Somit muss ein frequenzabhängiger Richtfaktor $q(f)$ verwendet werden, was sich in der Praxis auf einfache Weise mit Hilfe von *shelving*-Filter³ realisieren lässt, deren Formen sich gemäß Gleichung 3.2, für das W -Signal aus $2 - q(f)$ und für die verbleibenden Signale X, Y, Z direkt aus $q(f)$ ergeben.

Das W -Signal müsste demnach mit einem anhebenden Höhen-*shelving*-Filter, welches einen Betragsfrequenzgang aufweist, wie in Abbildung 3.11(a) dargestellt und die verbleibenden X, Y, Z -Signale mit einem anhebenden Tiefen-*shelving*-Filter, wie in Abbildung 3.11(b) aufgezeigt, vor Berechnung der Decodierung gefiltert werden. Dadurch könnte auf einfache Weise ein Ambisonicsystem mit einem nahezu optimalen, *frequenzab-*

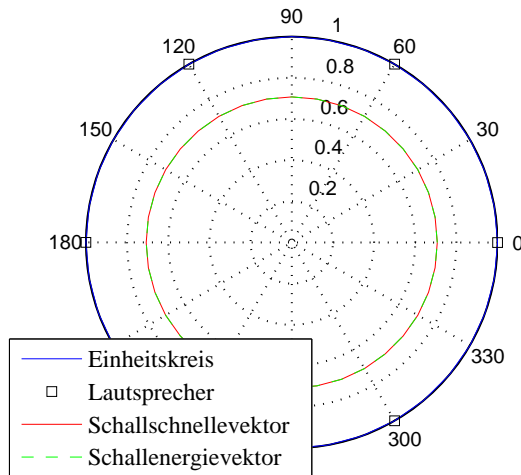
³*Shelving*-Filter werden im Deutschen auch als *Kuhschwanzfilter* bezeichnet, was jedoch nicht mehr gängige Praxis ist.

Schallschnelle- und Schallenergievektor bei einem Richtfaktor von $q=1.3333$



(a) Schallschnelle- \mathbf{v} und Schallenergievektoren \mathbf{e} , für $q = 1.33$, $\Theta_{\text{SRC}} = [0^\circ, \dots, 360^\circ]$, $\varphi_{\text{SRC}} = 0^\circ$ und $\Theta_S = [0^\circ : 60^\circ : 300^\circ]$, $\varphi_S = 0^\circ$

Schallschnelle- und Schallenergievektor bei einem Richtfaktor von $q=1.1716$



(b) Schallschnelle- \mathbf{v} und Schallenergievektoren \mathbf{e} , für $q = 1.17$, $\Theta_{\text{SRC}} = [0^\circ, \dots, 360^\circ]$, $\varphi_{\text{SRC}} = 0^\circ$ und $\Theta_S = [0^\circ : 60^\circ : 300^\circ]$, $\varphi_S = 0^\circ$

Abbildung 3.9: Darstellung der optimierten Vektoren (3.9(a): Schallschnelle \mathbf{v} mit $q_{\mathbf{v}Opt}$, 3.9(b): Schallenergie \mathbf{e} mit $q_{\mathbf{e}Opt}$)

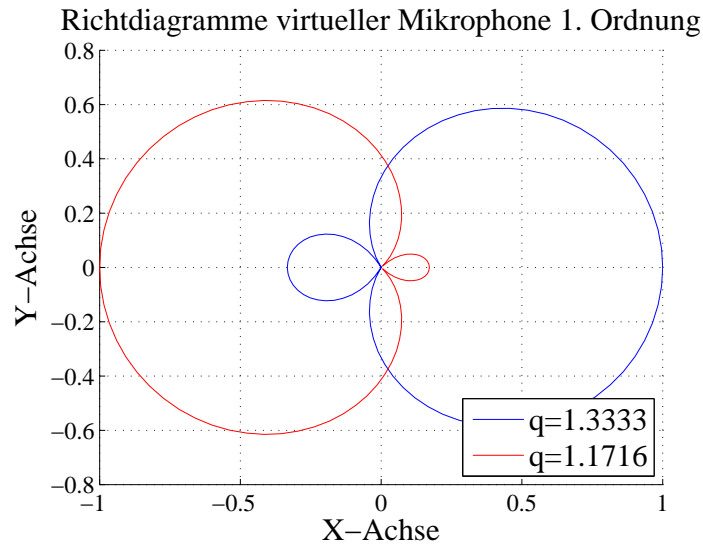
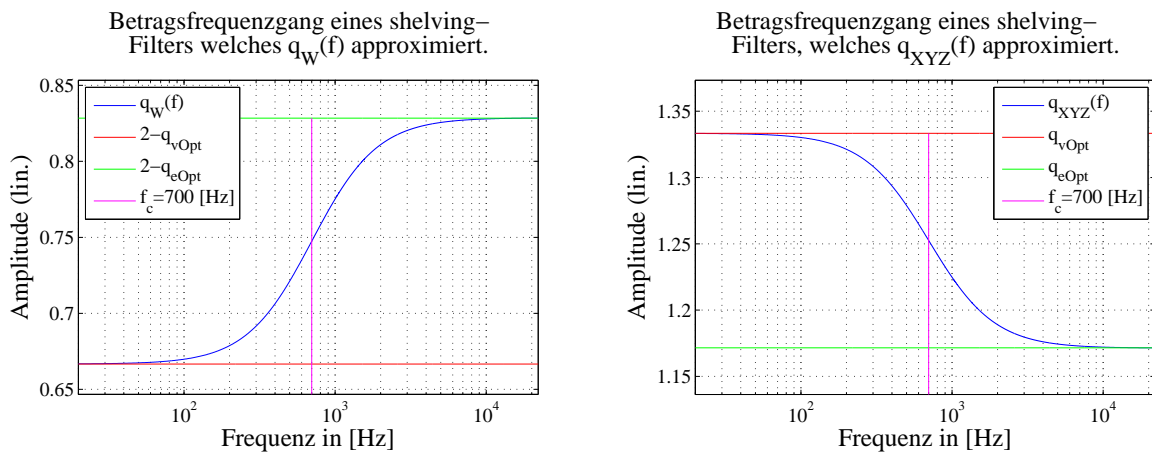


Abbildung 3.10: Richtcharakteristika, welche durch die optimierten Richtfaktoren $q = 1.33$ (Blau: Mit $\Theta_{SRC} = 0^\circ$) und $q = 1.17$ (Rot: Mit $\Theta_{SRC} = 180^\circ$) entstehen.



(a) *Shelving*-Filter für das Ambisonicsignal W , welches dessen optimalen, frequenzabhängigen Richtfaktor $q_W(f)$ approximiert

(b) *Shelving*-Filter für die Ambisonicsignale X, Y, Z , welches deren optimalen, frequenzabhängigen Richtfaktor $q_{XYZ}(f)$ approximiert

Abbildung 3.11: *Shelving*-Filter zur Approximation der frequenzabhängigen Richtfaktoren $q_W(f)$ und $q_{XYZ}(f)$.

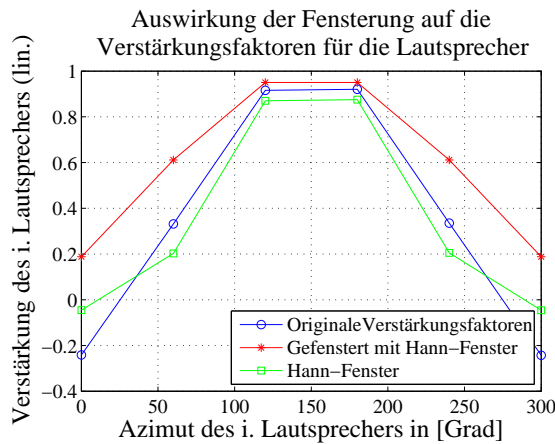
hängigen Richtfaktor realisiert werden.

3.1.1 Decodervariationen

Wie oben beschrieben, versucht ein Ambisonicdecoder das encodierte Schallfeld im Zentrum des Lautsprecher setups, d.h. im *sweet spot*, so exakt wie möglich zu reproduzieren. Da das Ambisonicsystem auf einer Rekonstruktion von Wellenfeldern unter Freifeldbedingungen basiert, trägt üblicherweise jeder der beteiligten Lautsprecher des Wiedergabesystems einen Teil dazu bei, so dass die Überlagerung aller Anteile idealerweise dem encodierten Wellenfeld, das es zu reproduzieren gilt, entspricht. Dabei kommt es sowohl zu konstruktiven als auch destruktiven Überlagerungen einzelnen Wellenfelder der beteiligten Lautsprecher. Somit lässt sich erklären wieso Decodierer für die Rekonstruktion des encodierten Wellenfeldes auch Lautsprecher verwenden, die sich in der zur Signalquelle abgewandten Hemisphäre befinden, wie beispielsweise in Abbildung 3.5 dargestellt. Dieses Prinzip funktioniert unter Einhaltung der dabei nötigen Voraussetzungen auch einwandfrei. Allerdings kommt es zu Störungen sobald im *sweet spot* ein Körper, wie z.B. der menschliche Kopf, eingebracht wird, an dem es zu Reflektionen, Abschattungen und Beugungen kommt. Dies hat zur Folge, dass es außerhalb des *sweet spot* zu Fehlortungen kommt, die umso stärker ausfallen, je weiter man sich von diesem entfernt. Dies erscheint auch logisch, da, wenn man sich neben einem Lautsprecher, der sich beispielsweise entgegengesetzt zur gewünschten Empfangsrichtung befindet, aufhält und dieser einen destruktiven Beitrag > 0 für die Rekonstruktion des Wellenfeldes liefert, dessen Anteil zumindest zu einer Störung der empfundenen Lokalisation bzw. zu einer Doppelortung führt. Da es für die Auralisation unvermeidlich ist, einen Körper im Zentrum der Lautsprecheraufstellung zu positionieren, erscheint es sinnvoll den Decoder so zu modifizieren, dass solche Störungen reduziert bzw. gänzlich vermieden werden können. Eine Lösung dieses Problems stellt der sogenannte *In-Phase-Decoder* dar ([Mal92]), bei dem lediglich Ansteuersignale für die Lautsprecher generiert werden, die sich in Phase befinden. Dadurch lassen sich destruktive Interferenzen mit den damit verbundenen Problemen vermeiden, was zu einem wesentlich größeren *sweet spot* führt. Dabei vergrößert sich jedoch der Öffnungswinkel der resultierenden Richtcharakteristik, was gleichbedeutend mit einer Verschlechterung der Lokalisationsschärfe ist. In-Phase-Decoder werden deshalb vorrangig für Wiedergabesysteme, die ein großes Publikum erreichen sollen, eingesetzt. Verallgemeinern lässt sich die Modifikation des Ambisonicdecoders in Analogie zu der Wirkung von *Fensterfunktionen*, wie sie etwa beim Entwurf von FIR⁴-Filter eingesetzt werden [Son03, Hol06]. Dazu werden die Ansteuersignale der Lautsprecher entsprechend des Verlaufes eines ausgewählten Fensters gewichtet.

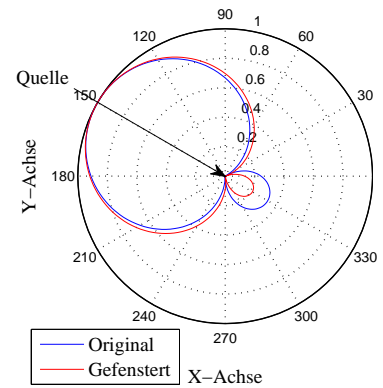
Abbildung 3.12(a) zeigt neben der verwendeten Fensterfunktion, wobei es sich hier um ein *Hann-Fenster* handelt, dessen diskrete Abtastwerte durch roten Sterne gekennzeichnet

⁴FIR = engl. *Finite Impulse Response* = Filter mit endlicher Impulsantwort



(a) Auswirkung der Fensterung auf die Gewichtungsfaktoren der Lautsprecher, für $q = 1.33$, $\Theta_{\text{SRC}} = 150^\circ$, $\varphi_{\text{SRC}} = 0^\circ$ und $\Theta_S = [0^\circ : 60^\circ : 300^\circ]$, $\varphi_S = 0^\circ$.

Überlagerung der Richtdiagramme aller virtuellen Mikrophone 1. Ordnung für alle 6 Lautsprecher und einer Quelle bei $\Theta_{\text{SRC}} = 150^\circ$ und $\varphi_{\text{SRC}} = 0^\circ$



(b) Auswirkung der Fensterung auf die resultierende Richtcharakteristik, für $q = 1.33$, $\Theta_{\text{SRC}} = 150^\circ$, $\varphi_{\text{SRC}} = 0^\circ$ und $\Theta_S = [0^\circ : 60^\circ : 300^\circ]$, $\varphi_S = 0^\circ$

Abbildung 3.12: Auswirkung der Fensterung (Hann-Fenster) der Gewichtungsfaktoren der Lautsprecher.

wurden auch noch die ursprünglichen (blaue Kreise) sowie die gefensterten (grüne Rechtecke) Verstärkungsfaktoren der jeweiligen Lautsprecher auf. Dabei erkennt man, dass die Verstärkungsfaktoren der Lautsprecher, die sich bei $\Theta_S = [0^\circ, 300^\circ]$, $\varphi_S = 0^\circ$ befinden, negative Vorzeichen besitzen und somit einen destruktiven Anteil bei der Reproduktion des encodierten Wellenfeldes beitragen. Zudem ist aus Abbildung 3.12(a) ersichtlich, dass die Fensterung die Gewichte in ihrem Betrag umso stärker reduziert je größer die Winkel-differenz zwischen dem jeweiligen Lautsprecher ($\Theta_{S_i}, \varphi_{S_i}$, mit $i = 1, \dots, L$) und der Einfallrichtung des Quellsignals ($\Theta_{\text{SRC}}, \varphi_{\text{SRC}}$) ist. Abbildung 3.12(b) verdeutlicht die Auswirkung der Fensterung der Verstärkungsfaktoren der Lautsprecher mit einem Hann-Fenster bezüglich der resultierenden Richtwirkung. Wie vermutet, bewirkt die Fensterung eine Reduktion der gegenphasigen Anteile, was einer Nebenkeulendämpfung entspricht, verursacht jedoch gleichzeitig eine Verbreiterung der Hauptkeule, so wie wir es etwa aus dem Entwurf von FIR-Filter her kennen. Auf diese Weise können Fensterfunktionen auch für die Optimierung von Ambisonicdecoder verwendet werden. Dieser fensterbasierte Decoderentwurf kann beispielsweise dazu verwendet werden, den Decoder an die vorherrschende Umgebung, in welcher die Synthese stattfinden soll, anzupassen.

3.2 Zusammenfassung

Die Intention dieses Kapitels ist dem Leser, der eventuell noch kein Vorwissen bezüglich der Ambisonictheorie hat, einen Einstieg in das Thema am einfachsten Beispiel, d.h. anhand Ambisonicsysteme 1. Ordnung zu geben. Lange Zeit konnte aufgrund der damals

eingeschränkten Aufnahmetechnik ausschließlich Ambisonicsysteme bis zu einer maximalen Ordnung von 1, zumindest für die Analyse bestehender Wellenfelder, angewendet werden.

Dabei wurde in diesem Kapitel das Grundprinzip der Ambisonictheorie nach Gerzon vorgestellt. Ferner wurde die Codierung, basierend auf der Aufnahme der B-Format-Signale (Ambisonicsignale 1. Ordnung, X, Y, Z und W), welche sich mit Hilfe eines Klangfeldmikrophons bestimmen lassen, aufgezeigt. Anhand eines Beispiels wurde neben der Codierung auch die Decodierung eines Ambisonicsystems erörtert. Dabei wurde deutlich, dass bei einem Ambisonicsystems die Analyse (Codierung eines Wellenfeldes) vollkommen unabhängig von der Synthese (Decodierung eines Wellenfeldes) ist. Die bei der Analyse entstehenden Ambisonicsignale können dabei künstlich oder anhand einer Messung mit Hilfe des Klangfeldmikrophons bestimmt werden. Zudem wurde der Einfluss des Richtfaktors q auf die resultierende Qualität der Lokalisation (Lokalisationsschärfe) bei der Decodierung behandelt, welcher die Richtcharakteristik des virtuellen Mikrophons 1. Ordnung bestimmt. Eine einfache Berechnung, wie man unter bestimmten Bedingungen optimale Werte für den Richtfaktor q erlangen kann, wurde in Gleichung 3.4 gegeben. Außerdem wurde erörtert, dass ein über der Frequenz optimal ausgelegter Richtfaktor ($q_W(f)$ für das W -Signal und $q_{XYZ}(f)$ für das X -, Y - und Z -Signal) vorteilhaft ist, wobei sich diese mit Hilfe eines einfachen *shelving*-Filter 1. Ordnung realisieren lassen. Die Vorstellung unterschiedlicher Decodervarianten (In-Phase-Decoder, Fenstermethode) beendet dieses Kapitel.

Kapitel 4

Ambisonic höherer Ordnung

Das im Kapitel 3 vorgestellte Ambisonicsystem 1. Ordnung ist in der Lage, Wellenfelder mit sphärischen Harmonischen bis zur 1. Ordnung nachzubilden. Der Grund warum man sich über lange Zeit bei der Reproduktion von Wellenfeldern auf eine derart niedrige Ordnung beschränkte lag vorrangig daran, dass in der Praxis bis vor ca. 10 Jahren keine praktikable Möglichkeit bekannt war, Kugelflächenfunktionen höherer Ordnung messtechnisch zu erfassen. Hingegen waren Aufnahmen bis zu einer Ordnung von 1 relativ problemlos möglich, da hierzu lediglich ein Mikrophon mit Kugelcharakteristik zur Aufnahme der W -Komponente (sphärische Harmonische 0. Ordnung) sowie drei orthogonal zu einander angebrachte Mikrophone mit Achtercharakteristik zur Aufnahme der X -, Y - und Z -Komponenten (Kugelflächenfunktionen 1. Ordnung) benötigt werden. Für die direkte Messung, der auch als *B-Format* bekannten Komponenten, wurden im Laufe der Zeit von mehreren Herstellern ein eigenes *Klangfeldmikrophon*¹ entwickelt, von denen in Abbildung 4.1 eines beispielhaft dargestellt ist.



Abbildung 4.1: Klangfeldmikrophon der Firma „SoundField“. [Sou12]

Auch wenn die Verwendung eines solchen Klangfeldmikrophons bequem ist, so las-

¹Klangfeldmikrophon = engl. *Soundfield microphone*

sen sich die B-Format-Signale auch durch eine einfache Kopplung mehrerer zueinander orthogonal angebrachter, omnidirektionaler Mikrophone, z.B. mit Hilfe differentieller Mikrophonarrays, nachbilden (siehe A.1). Unabhängig von der messtechnischen Erfassung der nötigen Kugelflächenfunktionen höherer Ordnung, stellt sich die Frage: „Welche Vorteile bietet ein Ambisonicsystem höherer Ordnung?“. Wie in Abschnitt 3.1 festgestellt, gibt es einen direkten Zusammenhang zwischen der Lokalisationsschärfe und der Ordnung des Ambisonicsystems. Somit kann einerseits festgehalten werden, dass ein HOA²-System eine bessere, weil *höhere Lokalisationsschärfe* besitzt. Andererseits kommt es mit steigender Ordnung zu einer stetigen *Vergrößerung des sweet spot*.

4.1 Zerlegung des Klangfeldes

In Analogie zur bekannten *Fourieranalyse*, die besagt, dass sich ein beliebiges, stetig differenzierbares Zeitsignal in orthonormale³ bzw. zumindest orthogonale, harmonische Basisfunktionen, d.h. in Sinus- und Kosinusfunktionen unterschiedlicher Amplitude, Phase und Frequenz zerlegen lässt, kann auch ein beliebiges Wellenfeld in harmonische, orthonormale Basisfunktionen, wobei es sich hier um die bereits mehrfach erwähnten, sphärischen Harmonischen ($Y_{m,n}^{\sigma}(\Theta, \varphi)$) handelt, zerlegt werden.

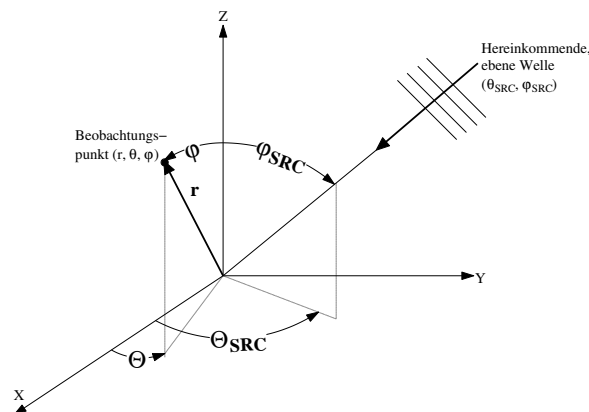


Abbildung 4.2: Koordinatensystem für das Ambisonicsystem.

Eine ebene Welle, die ein Signal mit einer Spitzenamplitude von s trägt, welche, wie in Abbildung 4.2 dargestellt, aus der Richtung, gegeben durch $\Theta_{\text{SRC}}, \varphi_{\text{SRC}}$ kommt, verursacht an einem Beobachtungspunkt \mathbf{r} zum Zeitpunkt t einen Schalldruck $p(\mathbf{r}, t)$, mit $\mathbf{r} = (r, \Theta, \varphi)$, wobei bezogen auf den Koordinatenursprung, r die Entfernung, Θ den Azimut und φ die Elevation angibt, der dem Realteil folgender Gleichung entspricht

²HOA = engl. *Higher Order Ambisoinc* = Ambisonicsystem höherer Ordnung

³Vektoren einer Menge S in einem euklidischen Vektorraum V bezeichnet man als *orthogonal*, wenn alle Vektoren $\mathbf{v}, \mathbf{v}' \in S$ mit $\mathbf{v} \neq \mathbf{v}'$ orthogonal sind. Die Menge S heißt *orthonormal*, wenn S orthogonal ist, und alle Vektoren in S (auf 1) normiert sind.

([Pol05a]):

$$\begin{aligned}
p(\mathbf{r}, t) &= s e^{(j\omega t + jkr[\cos(\varphi_{\text{SRC}})\cos(\varphi) + \sin(\varphi_{\text{SRC}})\sin(\varphi)\cos(\Theta - \Theta_{\text{SRC}}])}, \text{ mit} \\
p(\mathbf{r}, t) &= \text{Schalldruck einer ebenen Welle, mit einer Spitzenamplitude von } s, \text{ welche} \\
&\quad \text{aus der Richtung } (\Theta_{\text{SRC}}, \varphi_{\text{SRC}}) \text{ kommt, am Beobachtungspunkt } (r, \Theta, \varphi) \\
&\quad \text{zum Zeitpunkt } t, \\
k &= \text{Wellenzahl } k = \frac{2\pi f}{c} = \frac{\omega}{c}. \tag{4.1}
\end{aligned}$$

Oder, ausgedrückt im Spektralbereich ([Teu05]):

$$\begin{aligned}
P(\mathbf{k}^T \mathbf{r}, \omega) &= S(j\omega) e^{j\mathbf{k}^T \mathbf{r}}, \text{ mit} \\
\mathbf{k}^T \mathbf{r} &= kr [\cos(\varphi_{\text{SRC}})\cos(\varphi) + \sin(\varphi_{\text{SRC}})\sin(\varphi)\cos(\Theta - \Theta_{\text{SRC}})], \text{ wobei} \\
S(j\omega) &= \text{Fouriertransformierte von } s(t), \\
\mathbf{k}^T &= \text{Transponierte des Wellenvektors } \mathbf{k}. \tag{4.2}
\end{aligned}$$

Ausgehend von der *Wellengleichung*, genauer gesagt von der verlustlosen *Helmholtz-Gleichung*, gegeben durch:

$$\begin{aligned}
0 &= (\Delta + k^2) p(\mathbf{r}, t), \text{ wobei} \\
\Delta &= \text{Laplace-Operator in Kugelkoordinaten } (r, \Theta, \varphi), \text{ mit} \\
\Delta &= \nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin(\varphi)} \frac{\partial}{\partial \varphi} \left(\sin(\varphi) \frac{\partial}{\partial \varphi} \right) + \frac{1}{r^2 \sin^2(\varphi)} \frac{\partial^2}{\partial \Theta^2}, \text{ und} \\
\nabla &= \text{Nabla-Operator,} \\
p(\mathbf{r}, t) &= \text{Akustisches Schalldruckfeld,} \tag{4.3}
\end{aligned}$$

kann $p(\mathbf{r}, t)$ gemäß [Son03] in eine Reihe sphärischer Harmonischer entwickelt werden, was auch als *Fourier-Bessel-Reihe* bezeichnet wird und zu folgendem Ergebnis führt:

$$\begin{aligned}
P(\mathbf{r}, \omega) &= S(j\omega) \sum_{m=0}^{\infty} j^m j_m(kr) \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^{\sigma} Y_{m,n}^{\sigma}(\Theta, \varphi), \text{ mit} \\
j_m(kr) &= \text{Sphärische Besselfunktion 1. Art, } m. \text{ Ordnung des Arguments } kr, \\
Y_{m,n}^{\sigma}(\Theta, \varphi) &= \text{Winkelfunktion bzw. sphärische Harmonische } m. \text{ Ordnung, } n. \text{ Grades,} \\
B_{m,n}^{\sigma} &= \text{Ambisonickomponente } m. \text{ Ordnung, } n. \text{ Grades.} \tag{4.4}
\end{aligned}$$

Da die in Formel 4.4 aufgezeigte *Fourier-Bessel-Reihe* einen essentiellen Teil der HOA-Methode repräsentiert, wobei die Gesamtheit der Ambisonickomponenten $B_{m,n}^{\sigma}$ die Codierung des Wellenfeldes realisiert, wird dessen Bedeutung nachfolgend näher erläutert. Obwohl $P(\mathbf{r}, \omega)$ in Gleichung 4.4 auch mit nur einer Summe gebildet werden hätte kön-

nen, ist hier ganz bewusst eine zweifache Summierung verwendet worden. Die erste Summe enthält die *sphärische Besselfunktion* 1. Art $j_m(kr)$, die hier als *Radialfunktion* wirkt und somit das Wellenfeld um den Ursprung ($r = 0$) in Abhängigkeit sowohl von der Frequenz bzw. der Wellenzahl k der eintretenden, ebenen Welle als auch von der Entfernung zum Ursprung r beschreibt, wodurch sich auch das Argument der sphärischen Besselfunktion 1. Art erklärt.

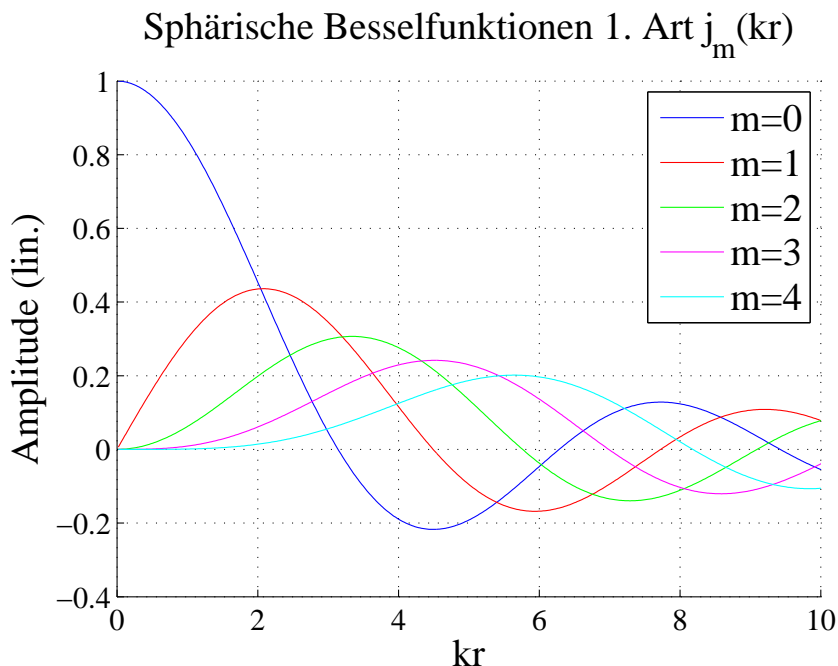


Abbildung 4.3: Sphärische Besselfunktionen 1. Art

Abbildung 4.3 veranschaulicht die sphärischen Besselfunktionen 1. Art $j_m(kr)$, für $m = [0, \dots, 4]$, d.h. für die 0.-4. Ordnung über einen Wertebereich von $kr = [0, \dots, 10]$. Für eine Frequenz von $f = 1$ [kHz], wie in Abbildung 4.3 angegeben, entspricht dies, mit $c = 343$ $\left[\frac{\text{m}}{\text{s}}\right]$ einer Entfernung vom Ursprung von $r = [0, \dots, 0.55]$ [m]. Bei näherer Betrachtung der Verläufe von $j_m(kr)$ fällt auf, dass diese mit steigender Ordnung m jeweils in größerer Entfernung zum Ursprung ihr Maximum haben. Das bedeutet, je größer die Ordnung m , desto größer wird auch der Bereich um den Ursprung in dem das Wellenfeld bestimmt werden kann. Dies ist die Erklärung für die eingangs in diesem Kapitel getätigte Aussage, dass es bei HOA-Systemen mit steigender Ordnung zu einer stetigen Vergrößerung des *sweet spot* kommt. In der Praxis kann m nicht, wie in Gleichung 4.4 angegeben, einen unendlichen Wert annehmen, sondern muss ab einer maximalen Ordnung M , wobei gilt $m \leq M$, abgebrochen werden. Wie groß dabei M gewählt werden kann hängt von mehreren Faktoren ab. Werden künstlich generierte Signale, wie im Beispiel des Kapitels 3.1, verwendet, ist M ausschließlich von der Anzahl der Lautsprecher L sowie von deren Anordnung abhängig. Sind die L Lautsprecher dabei dreidimensional auf einer Kugeloberfläche regelmäßig angeordnet, werden für die Rekonstruktion des Wel-

lenfeldes M . Ordnung $K^{3D} \geq (M + 1)^2$ Ambisonickomponenten $B_{m,n}^\sigma$, was wiederum der Anzahl der Lautsprecher L entspricht, für eine eindeutige Reproduktion des Wellenfeldes bis zur Ordnung M benötigt. Werden hingegen Lautsprecher zweidimensional in einem Kreis, wie im Beispiel des Kapitels 3.1 angeordnet, reduziert sich die, für die korrekte Nachbildung des Wellenfeldes notwendige Anzahl der Ambisonickomponenten $B_{m,n}^\sigma$ auf $K^{2D} = (2M + 1)$. Dies erklärt sich dadurch, dass nun nur noch die Ambisonickomponenten $B_{m,n}^\sigma$ nachgebildet werden müssen für die gilt $n = m$. Stehen für die Rekonstruktion etwa $L = 16$ Lautsprecher zur Verfügung, so kann bei Verwendung einer dreidimensionalen Lautsprecheraufstellung mit diesen das Wellenfeld lediglich bis zu einer maximalen Ordnung von $M_{Max}^{3D} = 3$ approximiert werden, wohingegen sich die maximale Ordnung auf einem Wert von $M_{Max}^{2D} = 7$ steigern lässt, wenn diese zweidimensional, d.h. flächig in einem Kreis, aufgestellt werden. Sollen hingegen keine künstlich generierten Signale für die Codierung verwendet werden, sondern gemessene, hängt die maximale Ordnung M zudem vom verwendeten Aufnahmesystem, d.h. vom verwendeten Mikrofonarray, ab. Das Klangfeldmikrofon ist dabei, wie bereits bekannt, in der Lage, lediglich sphärische Harmonische bis zu einer maximalen Ordnung von $M = 1$ messtechnisch zu erfassen. Somit kann ein HOA-System, welches auf Signale eines solchen Klangfeldmikrophons basiert, unabhängig von der Anzahl der Lautsprecher des Wiedergabesystems das Wellenfeld auch nur bis zu einer maximalen Ordnung von $M = 1$ nachbilden. Damit wird klar, dass eine *Auralisierung* nach der HOA-Methode nur dann funktionieren kann, wenn die Möglichkeit besteht, sphärische Harmonische höherer Ordnung aufzunehmen. Wie sich dieses Problem in der Praxis lösen lässt, wird in Kapitel 5 ausführlich behandelt.

Wenden wir uns nun aber der Interpretation der zweiten Summe der Fourier-Bessel-Reihe aus Gleichung 4.4 zu. Diese beschreibt die Zusammensetzung des Wellenfeldes im Koordinatenursprung, d.h. bei $r = 0$. Dabei können die, im Allgemeinen von der Frequenz unabhängigen Ambisonickomponenten $B_{m,n}^\sigma$ als vom Einfallswinkel unabhängige Faktoren interpretiert werden, mit denen die zugehörigen, sphärischen Harmonischen $Y_{m,n}^\sigma(\Theta, \varphi)$ gewichtet werden müssen, um das gewünschte Wellenfeld $P(\mathbf{r}, \omega)$ bei $r = 0$ erfolgreich zu rekonstruieren. In dieser Betrachtungsweise spiegelt sich nun auch die eingangs erwähnte Analogie zur Fourieranalyse, d.h. in der gewichteten Summe von orthogonalen Basisfunktionen, wider. Da die im Frequenzbereich definierten Ambisonickomponenten $B_{m,n}^\sigma$ keine Richtungsinformationen enthalten, müssen diese in den sphärischen Harmonischen $Y_{m,n}^\sigma(\Theta, \varphi)$ enthalten sein. Aus Kapitel 3 ist bereits bekannt, dass die sphärischen Harmonischen $Y_{m,n}^\sigma(\Theta, \varphi)$ die Richtcharakteristika der Aufnahmemikrophone, wie sie bereits für die 0.- und 1. Ordnung in Abbildung 3.1 dargestellt wurden, beschreiben.

Abbildung 4.4 stellt diese bis zu einer Ordnung von $M = 4$ dar, wobei jede Zeile alle von Null verschiedenen $(2M + 1)$ Winkelfunktionen beinhaltet, was bekanntermaßen zu einer Gesamtzahl von $K^{3D} = (M + 1)^2$, also für $M = 4$ von $K^{3D} = 25$, führt. Die Gesamtzahl der sphärischen Harmonischen K^{3D} erklärt sich dabei wie folgt: Für jede Ordnung m existieren $0 \leq n \leq m$ verschiedene, komplexe, sphärische Harmonische, wobei der Wert des

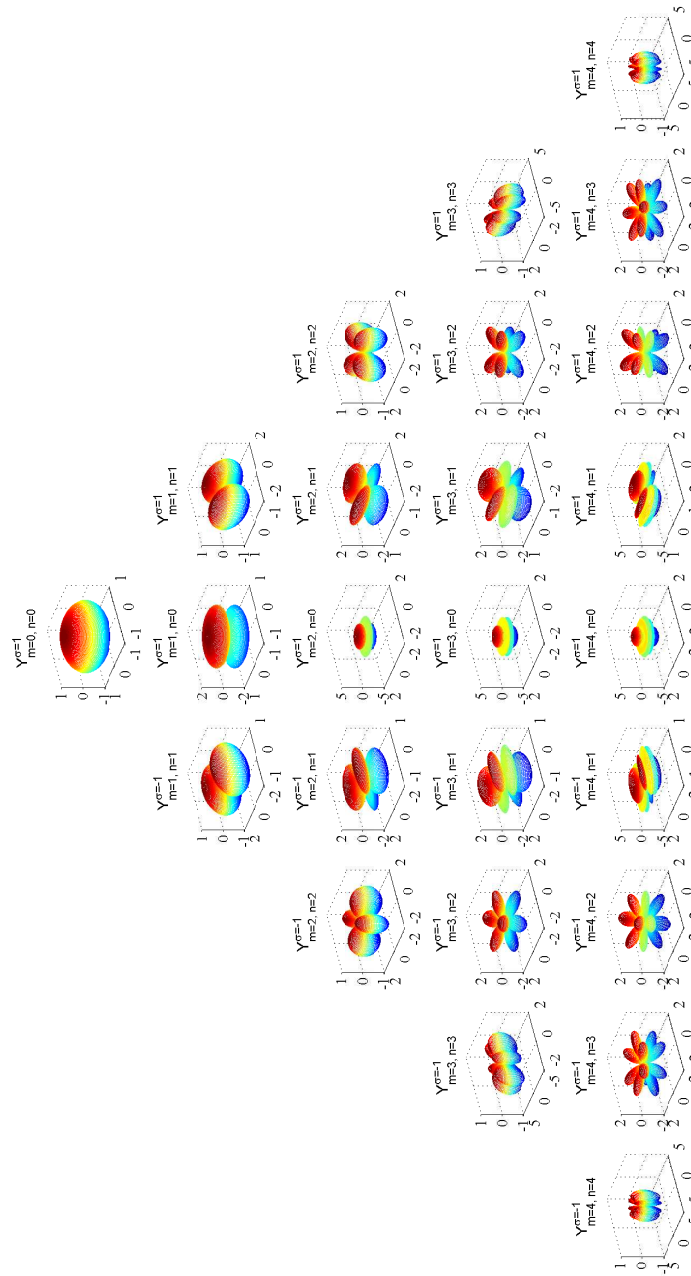


Abbildung 4.4: Graphische Darstellung der sphärischen Harmonischen $Y_{m,n}^{\sigma}(\Theta, \varphi)$ bis zur 4. Ordnung ($M = 4$)

hochgestellten Index $\sigma = \pm 1$ den Real- und Imaginärteil kennzeichnet. Dabei beschreibt $\sigma = 1$ den Real- und $\sigma = -1$ den Imaginärteil. Da $Y_{m,0}^{-1} = 0, \forall m$ gilt, verbleiben $(2M + 1)$ sphärische Harmonische für jede Ordnung m , was letztendlich in einem 3D Ambisonic-system zu einer Gesamtzahl von $K^{3D} = (M + 1)^2$ Winkelfunktionen führt. Für die in Gleichung 4.4 verwendeten Winkelfunktionen $Y_{m,n}^\sigma(\Theta, \varphi)$ gilt für den dreidimensionalen, normalisierten Fall ([Son03, Son06]):

$$\begin{aligned}
Y_{m,n}^\sigma(\Theta, \varphi) &= \sqrt{2m+1} \tilde{P}_{m,n}(\sin(\Theta)) \begin{cases} \cos(n\varphi) & \text{für } \sigma = +1 \\ \sin(n\varphi) & \text{für } \sigma = -1, \text{ mit} \end{cases} \\
\tilde{P}_{m,n}(\sin(\Theta)) &= \sqrt{\epsilon_n \frac{(m-n)!}{(m+n)!}} P_{m,n}(\sin(\Theta)), \text{ und} \\
\epsilon_n &= \begin{cases} 1, & \forall n = 0 \\ 2, & \forall n \neq 0, \text{ wobei} \end{cases} \\
P_{m,n}(z) &= (-1)^m (1-z^2)^{\frac{m}{2}} \frac{d^m}{dz^m} P_n(z), \text{ und} \\
P_n(z) &= \frac{1}{2^n n!} \frac{d^n}{dz^n} (z^2 - 1)^n, \text{ mit} \\
P_n(z) &= \text{Legendrepolynom } n. \text{ Ordnung,} \\
P_{m,n}(z) &= \text{Legendrefunktion } n. \text{ Ordnung, } m. \text{ Grades,} \\
\tilde{P}_{m,n}(z) &= \text{Schmidt-semi-normalisierte-Version der Legendrefunktion} \quad (4.5)
\end{aligned}$$

Wie man aus Gleichung 4.5 erkennen kann, ist die Abhängigkeit der sphärischen Harmonischen $Y_{m,n}^\sigma(\Theta, \varphi)$ vom horizontalen Einfallswinkel Θ durch die *Legendrefunktion* $P_{m,n}(\sin(\Theta))$ bzw. durch ihre *Schmidt-semi-normalisierte-Version* $\tilde{P}_{m,n}(\sin(\Theta))$ gegeben, wohingegen ihre Abhängigkeit vom vertikalen Einfallswinkel φ durch eine trigonometrische Funktion, genauer gesagt durch eine Kosinusfunktion ($\cos(n\varphi)$, für $\sigma = -1$) bzw. durch eine Sinusfunktion ($\sin(n\varphi)$, für $\sigma = +1$) beschrieben wird, welche, wie bereits erwähnt, den Real- bzw. den Imaginärteil des Wellenfeldes kennzeichnen. Gleichung 4.5 liefert eine allgemeine Beschreibung der sphärischen Harmonischen aus denen sich die, in der Tabelle 4.1 aufgeführten Formeln, mit denen sich die sphärischen Harmonischen $Y_{m,n}^\sigma(\Theta, \varphi)$ bis zu einer maximalen Ordnung von $M = 4$ direkt ermitteln lassen, ergeben.

Ordnung	Name	σ m, n	$Y_{m,n}^\sigma(\Theta, \varphi)$
0	W	1 0, 0	1
1	X	1 1, 1	$\sqrt{3} \cos(\Theta) \cos(\varphi)$

Tabelle 4.1: Direkte Berechnung der sphärischen Harmonischen bis zu einer Ordnung von $M = 4$.

Ordnung	Name	σ m, n	$Y_{m,n}^\sigma(\Theta, \varphi)$
1	Y	$\begin{matrix} -1 \\ 1, 1 \end{matrix}$	$\sqrt{3} \sin(\Theta) \cos(\varphi)$
1	Z	$\begin{matrix} 1 \\ 1, 0 \end{matrix}$	$\sqrt{3} \sin(\varphi)$
2	R	$\begin{matrix} 1 \\ 2, 0 \end{matrix}$	$\sqrt{5} \frac{1}{2} (3 \sin^2(\varphi) - 1)$
2	S	$\begin{matrix} 1 \\ 2, 1 \end{matrix}$	$\sqrt{5} \frac{\sqrt{3}}{2} \cos(\Theta) \sin(2\varphi)$
2	T	$\begin{matrix} -1 \\ 2, 1 \end{matrix}$	$\sqrt{5} \frac{\sqrt{3}}{2} \sin(\Theta) \sin(2\varphi)$
2	U	$\begin{matrix} 1 \\ 2, 2 \end{matrix}$	$\sqrt{5} \frac{\sqrt{3}}{2} \cos(2\Theta) \cos^2(\varphi)$
2	V	$\begin{matrix} -1 \\ 2, 2 \end{matrix}$	$\sqrt{5} \frac{\sqrt{3}}{2} \sin(2\Theta) \cos^2(\varphi)$
3	K	$\begin{matrix} 1 \\ 3, 0 \end{matrix}$	$\sqrt{7} \frac{1}{2} \sin(\varphi) (5 \sin^2(\varphi) - 3)$
3	L	$\begin{matrix} 1 \\ 3, 1 \end{matrix}$	$\sqrt{7} \sqrt{\frac{3}{8}} \cos(\Theta) \cos(\varphi) (5 \sin^2(\varphi) - 1)$
3	M	$\begin{matrix} -1 \\ 3, 1 \end{matrix}$	$\sqrt{7} \sqrt{\frac{3}{8}} \sin(\Theta) \cos(\varphi) (5 \sin^2(\varphi) - 1)$
3	N	$\begin{matrix} 1 \\ 3, 2 \end{matrix}$	$\sqrt{7} \frac{\sqrt{15}}{2} \cos(2\Theta) \sin(\varphi) \cos^2(\varphi)$
3	O	$\begin{matrix} -1 \\ 3, 2 \end{matrix}$	$\sqrt{7} \frac{\sqrt{15}}{2} \sin(2\Theta) \sin(\varphi) \cos^2(\varphi)$
3	P	$\begin{matrix} 1 \\ 3, 3 \end{matrix}$	$\sqrt{7} \sqrt{\frac{5}{8}} \cos(3\Theta) \cos^3(\varphi)$
3	Q	$\begin{matrix} -1 \\ 3, 3 \end{matrix}$	$\sqrt{7} \sqrt{\frac{5}{8}} \sin(3\Theta) \cos^3(\varphi)$
4		$\begin{matrix} 1 \\ 4, 0 \end{matrix}$	$\sqrt{9} \frac{1}{8} (35 \sin^4(\varphi) - 30 \sin^2(\varphi) + 3)$
4		$\begin{matrix} 1 \\ 4, 1 \end{matrix}$	$\sqrt{9} \sqrt{\frac{5}{8}} \cos(\Theta) \cos(\varphi) \sin(\varphi) (7 \sin^2(\varphi) - 3)$
4		$\begin{matrix} -1 \\ 4, 1 \end{matrix}$	$\sqrt{9} \sqrt{\frac{5}{8}} \sin(\Theta) \cos(\varphi) \sin(\varphi) (7 \sin^2(\varphi) - 3)$
4		$\begin{matrix} 1 \\ 4, 2 \end{matrix}$	$\sqrt{9} \frac{\sqrt{5}}{4} \cos(2\Theta) \cos^2(\varphi) (7 \sin^2(\varphi) - 1)$
4		$\begin{matrix} -1 \\ 4, 2 \end{matrix}$	$\sqrt{9} \frac{\sqrt{5}}{4} \sin(2\Theta) \cos^2(\varphi) (7 \sin^2(\varphi) - 1)$

Tabelle 4.1: Direkte Berechnung der sphärischen Harmonischen bis zu einer Ordnung von $M = 4$.

Ordnung	Name	σ m, n	$Y_{m,n}^\sigma(\Theta, \varphi)$
4		1 4, 3	$\sqrt{9\frac{\sqrt{70}}{4}} \cos(3\Theta) \cos^3(\varphi) \sin(\varphi)$
4		-1 4, 3	$\sqrt{9\frac{\sqrt{70}}{4}} \sin(3\Theta) \cos^3(\varphi) \sin(\varphi)$
4		1 4, 4	$\sqrt{9\frac{\sqrt{35}}{8}} \cos(4\Theta) \cos^4(\varphi)$
4		-1 4, 4	$\sqrt{9\frac{\sqrt{35}}{8}} \sin(4\Theta) \cos^4(\varphi)$

Tabelle 4.1: Direkte Berechnung der sphärischen Harmonischen bis zu einer Ordnung von $M = 4$.

Für den zweidimensionalen Fall berechnen sich die sphärischen Harmonischen wie folgt ([Dan00]):

$$\begin{aligned}
Y_m^\sigma(\Theta)^{2D} &= \begin{cases} \epsilon \cos(m\Theta) & \text{für } \sigma = +1 \\ \epsilon \sin(m\Theta) & \text{für } \sigma = -1, \text{ mit} \end{cases} \\
\epsilon &= \begin{cases} 1, & \text{für } m = 0 \\ \sqrt{2}, & \text{für } m > 0, \text{ wobei} \end{cases} \\
Y_m^\sigma(\Theta)^{(2D)} &= \text{Sphärische Harmonische für den 2D-Fall.} \quad (4.6)
\end{aligned}$$

Wie man aus Gleichung 4.6 erkennen kann, weisen im 2D-Fall die Winkelfunktionen $Y_m^\sigma(\Theta)^{2D}$ weder eine Abhängigkeit vom vertikalen Einfallswinkel φ noch vom Grad n auf. Neben dem Fakt, dass im 2D-Fall lediglich Komponenten berücksichtigt werden, die unabhängig vom vertikalen Einfallswinkel φ sind, d.h. bei $\varphi = 0$ nicht Null werden, was bekanntermaßen der Fall ist, wenn $n = m$, ändert sich zudem noch ihre Skalierung. Wie in [DMN03] beschrieben, lassen sich die $K^{2D} = (2M + 1)$ sphärischen Harmonischen auch aus dem allgemeinen 3D-Fall wie folgt berechnen:

$$\begin{aligned}
Y_{m,n}^\sigma(\Theta, \varphi)^{(2D)} &= \sqrt{\frac{2^{2m} m!^2}{(2m+1)!}} Y_{m,n}^\sigma(\Theta, \varphi)^{(3D)}, \text{ wobei} \\
Y_{m,n}^\sigma(\Theta, \varphi)^{(3D)} &= \text{Sphärische Harmonische für den 3D-Fall, nach Gleichung 4.5,} \\
Y_{m,n}^\sigma(\Theta, \varphi)^{(2D)} &= \text{Sphärische Harmonische für den 2D-Fall.} \quad (4.7)
\end{aligned}$$

4.2 Codierung und Decodierung

Wie bereits aus Kapitel 3 bekannt, besteht einer der Vorteile eines Ambisonicsystem darin, dass Codier- und Decodierprozess unabhängig voneinander betrachtet werden dürfen.

Hierbei versteht man unter dem Begriff *Codierprozess* die Repräsentation eines künstlichen oder gemessenen Wellenfeldes durch Überlagerung virtueller, im Raum positionierter Quellen. Dies entspricht einer *Klangfeldanalyse*, was wiederum dem Auffinden der jeweiligen Ambisonickomponenten $B_{m,n}^\sigma$ gleichkommt.

Der *Decodierprozess* beschreibt hingegen die Berechnung der Gewichte bzw. der Filterfunktionen, mit denen Signale gewichtet werden müssen, damit deren Wiedergabe über die L Lautsprecher des Wiedergabesystems möglichst exakt dem durch die Ambisonickomponenten $B_{m,n}^\sigma$ codierten Wellenfeld im Zielraum entspricht. Somit kommt der Decodierprozess eines Ambisonicsystems einer *Klangfeldsynthese* gleich, welcher der Familie der Amplituden-*panning*-Techniken zugeordnet werden kann.

4.2.1 Der Codierprozess von ebenen Wellen

Das in Kapitel 4.1 aufgezeigte Szenario aus dem letztendlich die in Gleichung 4.4 formulierte *Fourier-Bessel-Reihe* entstand, ging vom Einfall einer *ebenen Welle* aus der Raumrichtung $(\Theta_{\text{SRC}}, \varphi_{\text{SRC}})$ aus. Ein Signal $S(j\omega)$, das aus dieser Richtung erzeugt werden soll, verursacht dabei am Beobachtungspunkt \mathbf{r} , wobei es sich vorrangig um den *sweet spot*, d.h. um $r = 0$ handelt, einen Schalldruck $P(\mathbf{r}, \omega)$, der sich wie folgt, approximativ bestimmen lässt:

$$P(\mathbf{r}, \omega) = S(j\omega) \sum_{m=0}^M j^m j_m(kr) \sum_{0 \leq n \leq m, \sigma = \pm 1} Y_{m,n}^\sigma(\Theta_{\text{SRC}}, \varphi_{\text{SRC}}) Y_{m,n}^\sigma(\Theta, \varphi),$$

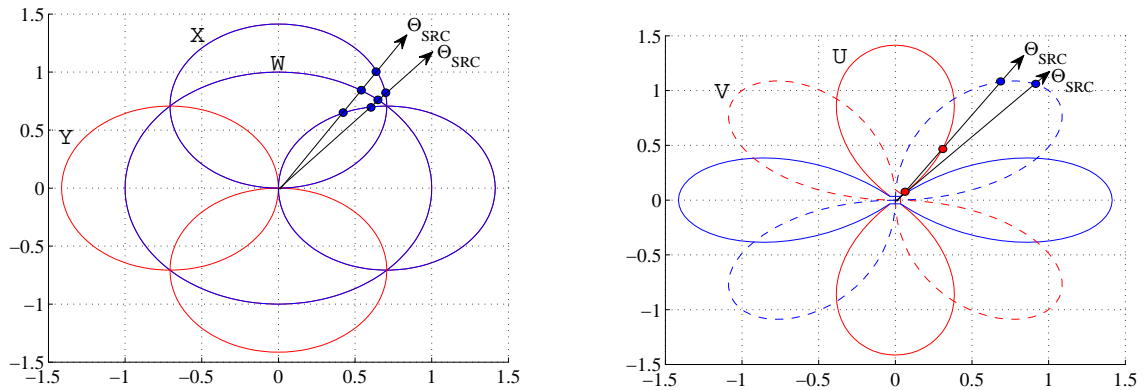
$Y_{m,n}^\sigma(\Theta_{\text{SRC}}, \varphi_{\text{SRC}})$ = Codierfunktionen des Quellsignals $S(j\omega)$, mit denen die zugehörigen sphärischen Harmonischen gewichtet werden müssen, damit der gewünschte Schalldruck $P(\mathbf{r}, \omega)$ am Ort \mathbf{r} entsteht. (4.8)

Ein Vergleich von Gleichung 4.8 mit Gleichung 4.4 offenbart, dass sich der Codierprozess in diesem Fall, d.h. bei Einfall einer ebenen Welle aus einer bestimmten Raumrichtung $(\Theta_{\text{SRC}}, \varphi_{\text{SRC}})$ wie folgt beschreiben lässt:

$$B_{m,n}^\sigma = Y_{m,n}^\sigma(\Theta_{\text{SRC}}, \varphi_{\text{SRC}}). \quad (4.9)$$

Der Codierprozess gemäß Gleichung 4.9 besagt, dass sich die Einfallrichtung einer ebenen Welle gegeben durch $(\Theta_{\text{SRC}}, \varphi_{\text{SRC}})$, welche ein Signal $S(j\omega)$ transportiert, codieren lässt, indem dieses mit den zur Einfallrichtung korrespondierenden Werten der jeweiligen sphärischen Harmonischen gewichtet wird. Die dabei entstehenden Signale $B_{m,n}^\sigma$, die den

Codierprozess eines Ambisonicsystems für ein im Raum angeordnetes Quellsignal $S(j\omega)$ vollständig beschreiben, werden auch als *Ambisonicsignale* bzw. als *Ambisonickanäle* bezeichnet.



(a) 2D-Codierung mit Hilfe sphärischer Harmonischer 0.- und 1. Ordnung.

(b) 2D-Codierung mit Hilfe sphärischer Harmonischer 2. Ordnung.

Abbildung 4.5: Horizontale Codierung unter Verwendung des B-Format-Systems, d.h. mit Hilfe der Ambisonickomponenten W , X und Y (linkes Teilbild) sowie unter alleiniger Verwendung horizontaler sphärischer Harmonischer 2. Ordnung, U und V (rechtes Teilbild), wobei die in blau bzw. rot gezeichneten Funktionen eine positive bzw. negative Verstärkung kennzeichnen (Real- und Imaginärteil).

Abbildung 4.5 verdeutlicht diesen Zusammenhang anhand des Einfalls einer ebenen Welle aus einer jeweils geringfügig unterschiedlichen Raumrichtung Θ_{SRC} . Die Schnittpunkte der sphärischen Harmonischen mit dem Vektor, welcher die horizontale Einfallrichtung kennzeichnet, entsprechen dabei den Gewichten, die in Gleichung 4.8 mit $Y_{m,n}^{\sigma}(\Theta_{\text{SRC}}, \varphi_{\text{SRC}})$ bezeichnet wurden. Dabei wird der jeweilige Verstärkungswert vom Abstand der Schnittpunkte zum Ursprung bestimmt, wohingegen die farbliche Kennung deren Phasenlage bzw. deren Vorzeichen kennzeichnet. Dabei steht die Farbe Blau für eine Phasenlage von 0° (Realteil) und Rot für negative Verstärkungswerte, d.h. für eine Phasenlage von 180° (Imaginärteil). Ein Vergleich der Verstärkungswerte, die sich beim Einfall einer ebenen Welle aus einer geringfügig unterschiedlichen Raumrichtung ergeben, offenbart, dass sich diese im Falle eines Ambisonicsystems 1. Ordnung (B-Format), wie in Abbildung 4.5(a) dargestellt, nur unwesentlich voneinander unterscheiden, was ein Indiz dafür ist, dass eine exakte Wahrnehmung der Raumrichtung schwerfällt. Verwendet man hingegen ein Ambisonicsystem 2. Ordnung, resultieren unter Verwendung der exakt gleichen Raumrichtungen Verstärkungsfaktoren, die sich wesentlich stärker voneinander unterscheiden, was Abbildung 4.5(b) verdeutlicht. Somit wird klar, dass sich die Lokalisierungsgenauigkeit sowie die Robustheit der Lokalisierung mit steigender Ordnung des Ambisonicsystems erhöht.

Für den Fall, dass nicht eine, sondern k Schallquellen $S_i(j\omega)$ ebene Wellen aus unterschiedlichen Raumrichtungen $(\Theta_{\text{SRC}_i}, \varphi_{\text{SRC}_i})$, mit $1 \leq i \leq k$ abgeben, ergibt sich für das Ambisonicsystem folgende Codierung:

$$\begin{aligned} \mathbf{B} &= \sum_{i=1}^k S_i(j\omega) Y_{m,n}^\sigma(\Theta_{\text{SRC}_i}, \varphi_{\text{SRC}_i}), \text{ mit} \\ \mathbf{B} &= [B_{0,0}^{+1}, B_{1,0}^{+1}, B_{1,1}^{+1}, B_{1,1}^{-1}, B_{2,0}^{+1}, B_{2,1}^{+1}, B_{2,1}^{-1}, B_{2,2}^{+1}, B_{2,2}^{-1}, \dots, B_{M,0}^{+1}, \dots, B_{M,M}^{-1}]^T \end{aligned} \quad (4.10)$$

Wie Gleichung 4.10 verdeutlicht, ergeben sich die resultierenden Ambisonicsignale \mathbf{B} durch einfache Überlagerung aller k Ambisonicsignale $B_{m,n}^\sigma$, die sich jeweils gemäß Gleichung 4.9 berechnen.

Zusammenfassend kann konstatiert werden, dass sich der Codierprozess, d.h. die Erzeugung der Ambisonicsignale \mathbf{B} ebener Wellen durch eine breitbandige, d.h. durch eine rein pegelmäßige Gewichtung eines Monosignals $S(j\omega)$ bzw. mehrerer Monosignale $S_i(j\omega)$ ergeben. Die Gewichte resultieren dabei aus den jeweils zu den Ambisonicsignalen $B_{m,n}^\sigma$ gehörenden sphärischen Harmonischen $Y_{m,n}^\sigma(\Theta_{\text{SRC}_i}, \varphi_{\text{SRC}_i})$ berechnet für die Raumrichtung $(\Theta_{\text{SRC}}, \varphi_{\text{SRC}})$ bzw. für die Raumrichtungen $(\Theta_{\text{SRC}_i}, \varphi_{\text{SRC}_i})$, an welcher sich die Schallquelle s bzw. die Schallquellen s_i befinden.

4.2.2 Der Codierprozess von Kugelwellen

Eine ursprüngliche Einschränkung von Ambisonicsystemen gemäß Gerzon war, dass ausschließlich von ebenen Wellen ausgegangen wurde. Da ebene Wellen jedoch definitionsgemäß eine Quelle bei $\mathbf{r} = (r, \Theta, \varphi)$, mit $r \rightarrow \infty$ besitzen, weisen diese mit steigender Entfernung keinen Pegelabfall auf. Unter dieser Annahme ist es nicht möglich Abstände von der Abhörposition zur Quelle in der Codierung bzw. Decodierung zu berücksichtigen, wobei hier üblicherweise vom Koordinatenursprung ausgegangen wird. Neben der Tatsache, dass der *Bühnenabstand*⁴ ein wichtiges Maß bei der Auralisation darstellt, spielt er vor allem bei der Decodierung, d.h. bei der Klangfeldsynthese, eine entscheidende Rolle, da sich die Quellen, d.h. die L Lautsprecher, üblicherweise in unmittelbarer Nähe zum Abhörort befinden, so dass im Normalfall nicht automatisch von *Fernfeldbedingungen* ausgegangen werden kann.

Im Allgemeinen spricht man von einem *Fernfeld* wenn folgende Bedingungen erfüllt

⁴Als *Bühnenabstand* wird in der Akustik der subjektiv empfundene Abstand zur Bühne bzw. zum Klanggeschehen bezeichnet.

sind ([Mös10]):

$$\begin{aligned}
 r &\gg l, \\
 r &\gg \lambda, \\
 \frac{r}{l} &\gg \frac{l}{\lambda}, \text{ mit} \\
 r &= \text{Abstand in [m] vom Abhörort zum Mittelpunkt der Quelle (z.B. Lautsprecher),} \\
 l &= \text{Quellenabmessung in [m], was im schlimmsten Fall der größten Seitenlänge der} \\
 &\quad \text{Quelle entspricht,} \\
 \lambda &= \text{Akustische Wellenlänge in [m], wobei gilt: } \lambda = \frac{c}{f}.
 \end{aligned} \tag{4.11}$$

Gemäß [Too08] kann approximativ von einem Fernfeld gesprochen werden, wenn der Abstand zur Quelle r zwischen 3 – 10 Mal dem maximalen Abmessungen der Quelle l entspricht, also folglich $r \geq 10l$ ist.

Signale, die von „echten“ Klangquellen wie etwa Musikinstrumenten oder Lautsprechern abgestrahlt werden, weisen unter *Freifeldbedingungen*⁵ ein Wellenfeld auf, welches sich wesentlich präziser durch *Kugelwellen*, als durch ebene Wellen darstellen lässt. In Anlehnung an Gleichung 4.2, welche den Schalldruck des Signals $S(j\omega)$, das von einer eintretenden⁶ ebenen Welle transportiert wird, am Beobachtungspunkt (r, Θ, φ) beschreibt, lässt sich dieser für Kugelwellen gemäß [Teu05] unter Vernachlässigung des Signals $S(j\omega)$ sowie des Skalierungsfaktors $\frac{1}{4\pi}$ wie folgt beschreiben:

$$\begin{aligned}
 g_k(\mathbf{r} \mid \mathbf{r}_{\text{SRC}}) &= \frac{e^{jk\|\mathbf{r}-\mathbf{r}_{\text{SRC}}\|}}{\|\mathbf{r}-\mathbf{r}_{\text{SRC}}\|}, \text{ mit} \\
 g_k(\mathbf{r} \mid \mathbf{r}_{\text{SRC}}) &= \text{Greensche Funktion,} \\
 \mathbf{r}_{\text{SRC}} &= (r_{\text{SRC}}, \Theta_{\text{SRC}}, \varphi_{\text{SRC}}) \text{ (Ortsvektor zur Quelle),} \\
 \mathbf{r} &= (r, \Theta, \varphi) \text{ (Ortsvektor zum Beobachtungspunkt),} \\
 \|x\| &= \text{2-Norm des Arguments } x.
 \end{aligned} \tag{4.12}$$

Wie man aus Gleichung 4.12⁷ erkennen kann, liefert die *Greensche Funktion*, gegeben

⁵In der Akustik spricht man von einem *Freifeld* wenn ausschließlich Direktschall und kein reflektierter Schall das Wellenfeld erzeugt.

⁶Unter einer eintretenden Welle versteht man, wie in Abbildung 4.2 dargestellt welche sich auf den Beobachtungspunkt zu bewegt.

⁷Die p -Norm eines Vektors \mathbf{x} lässt sich allgemein, für $p \geq 1$, mit $p \in \mathbb{R}$, wie folgt berechnen:

$$\begin{aligned}
 \|\mathbf{x}\|_p &:= \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}, \text{ mit} \\
 \mathbf{x} &= [x_1, x_2, \dots, x_n].
 \end{aligned}$$

durch $g_k(\mathbf{r} \mid \mathbf{r}_{\text{SRC}})$, wobei der Index k , die Abhängigkeit von der Wellenzahl k unterstreichen soll, eine Möglichkeit das Übertragungsverhalten von Kugelwellen zwischen einer sich bei \mathbf{r}_{SRC} befindlichen Quelle und dem durch \mathbf{r} festgelegten Beobachtungspunkt zu bestimmen.

Der durch Gleichung 4.12 beschriebene Schalldruck, welcher durch eine Kugelwelle von einer sich bei \mathbf{r}_{SRC} befindlichen *Punktquelle* erzeugt wird und am Beobachtungspunkt \mathbf{r} gemessen werden kann, wobei $\mathbf{r}_{\text{SRC}} > \mathbf{r}$ gelten muss, lässt sich ebenfalls, wie im Fall von ebenen Wellen (4.4), durch eine Überlagerung orthonormaler, sphärischer Basisfunktionen in eine Fourier-Bessel-Reihe wie folgt zerlegen ([Teu05]):

$$\begin{aligned} \frac{e^{j k \|\mathbf{r} - \mathbf{r}_{\text{SRC}}\|}}{\|\mathbf{r} - \mathbf{r}_{\text{SRC}}\|} &= j k \sum_{m=0}^{\infty} j_m(kr) h_m^{(1)}(k r_{\text{SRC}}) \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^{\sigma} Y_{m,n}^{\sigma}(\Theta, \varphi), \text{ mit} \\ h_m^{(1)}(x) &= j_m(x) + j y_m(x), \text{ wobei} \\ h_m^{(1)}(x) &= \text{Sphärische Hankelfunktion 1. Art, } m. \text{ Ordnung auch bekannt als} \\ &\quad \text{sphärische Besselfunktion 3. Art, } m. \text{ Ordnung (austretende Welle),} \\ y_m(x) &= \text{Sphärische Neumannfunktion } m. \text{ Ordnung, auch bekannt als} \\ &\quad \text{sphärische Besselfunktion 2. Art, } m. \text{ Ordnung,} \\ j_m(x) &= \text{Sphärische Besselfunktion 1. Art, } m. \text{ Ordnung.} \end{aligned} \quad (4.13)$$

Ausgehend von der in Gleichung 4.13 aufgezeigten Fourier-Bessel-Reihe für Kugelwellen wurde in [DMN03, Dan03] gezeigt, dass diese wie folgt codiert werden können:

$$\begin{aligned} B_{m,n}^{\sigma} &= S(j\omega) F_m(k r_{\text{SRC}}) Y_{m,n}^{\sigma}(\Theta_{\text{SRC}}, \varphi_{\text{SRC}}), \text{ mit} \\ F_m(k r_{\text{SRC}}) &= j^{-(m+1)} \frac{h_m^{(1)}(k r_{\text{SRC}})}{h_0^{(1)}(k r_{\text{SRC}})} = \sum_{n=0}^{\infty} \frac{(m+n)!}{(m-n)! n!} \left(\frac{-j c}{\omega r_{\text{SRC}}} \right)^n, \text{ wobei} \\ F_m(k r) &= F_m\left(\frac{r}{c}\right)(j\omega) = \text{Nahfeldmodellierungsfunktion.} \end{aligned} \quad (4.14)$$

Die *Nahfeldmodellierungsfunktion* $F_m(k r_{\text{SRC}})$ einer Quelle, welche sich bei \mathbf{r}_{SRC} befindet, die in der Literatur auch mit $F_m\left(\frac{r_{\text{SRC}}}{c}\right)(j\omega)$ bezeichnet wird, beschreibt den sogenannten *Nahfeldeffekt*, den Gerzon bereits in [Ger92a] erwähnte, wobei er ihn dort noch als *Bass-Boost-Effekt* bezeichnete. Dabei bezeichnet der hochgestellte Index $\left(\frac{r_{\text{SRC}}}{c}\right)$ die *Latenz*, d.h. die Verzögerungszeit vom Koordinatenursprung zur Quelle unter Berücksichtigung der Schallgeschwindigkeit c . Diese wird lediglich pro Ordnung m berechnet, d.h. sie bleibt für alle $n = (2m + 1)$ Komponenten pro Ordnung m unverändert.

Wie Abbildung 4.6 verdeutlicht, steigt der *Bass-Boost-Effekt* für Ordnungen von $m \geq 1$ mit $m \times 6 \frac{\text{[dB]}}{\text{Oktave}}$ zu tiefen Frequenzen hin an. Theoretisch könnte dieser zwar durch ein Integrationsfilter nachgebildet werden, was allerdings keine praktische Bedeutung hat, da diese instabil wären und somit nicht realisiert werden könnten. Aus diesem Grund ist

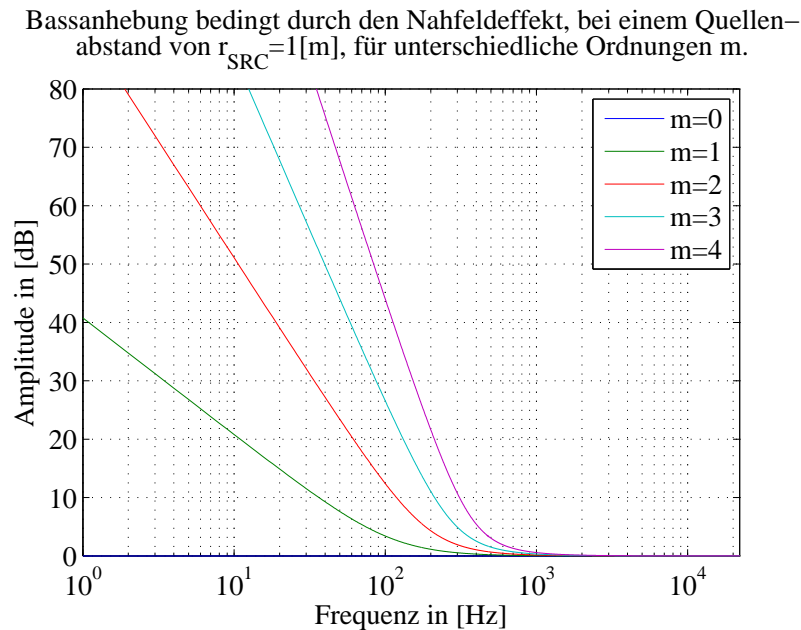


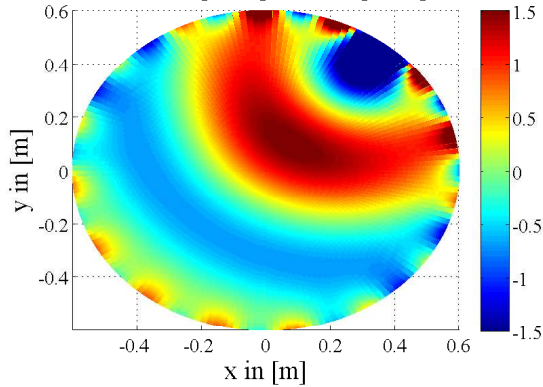
Abbildung 4.6: Nahfeldmodellierungsfunktion $F_m^{(\frac{r_{SRC}}{c})}(j\omega)$ bei $r_{SRC} = 1 [m]$. Mit änderndem Abstand zur Quelle r_{SRC} verschieben sich die Kurven dabei mit größer werdendem Abstand nach links, bzw. bei kleiner werdendem Abstand nach rechts.

die durch Gleichung 4.14 aufgezeigte Berücksichtigung des Nahfeldeffektes bei der Codierung des Wellenfeldes durch $F_m(k r_{SRC})$ lediglich von theoretischer Bedeutung. Eine praktische Bedeutung erhält diese erst bei Betrachtung der Decodierung, wie sich später noch zeigen wird.

Ungeachtet der praktischen Bedeutung ist es bereits an dieser Stelle interessant aufzuzeigen, welche Auswirkung eine *Kompensation des Nahfeldeffektes* auf das Wellenfeld hat.

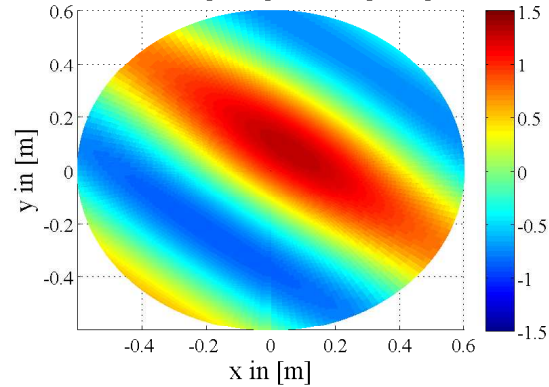
Abbildung 4.7 veranschaulicht die Wirkung des Nahfeldeffektes auf das resultierende Wellenfeld. Ohne Kompensation kommt es zu besagtem *Bass-Boost-Effekt*, also zu einer Verstärkung im unteren Spektralbereich, was aus Abbildung 4.7 anhand des höheren Maximalpegels ersichtlich ist. Zudem kommt es zu einer Akzentuierung der interauralen Pegeldifferenz (ILD), was sich durch Vergleich der Pegelverhältnisse im Zentrum des Lautsprecherrings mit und ohne Kompensation nachvollziehen lässt. Unter Verwendung eines mittleres Ohrabstandes von circa 18 [cm] ist erkennbar, dass es bedingt durch die gebogene Wellenfront im nicht-kompensierten Fall zu einer Pegelverschiebung zwischen den beiden Ohren im Vergleich zum kompensierten Fall, d.h. bei einer ebenen, bzw. geraden Wellenfront, kommt. In puncto Lokalisation kommt es hingegen im Zentrum zu keiner bzw. nur zu einer geringen Verfälschung, die sich jedoch sehr schnell verschärft je weiter man sich von diesem entfernt. Somit kommt es bei der Beurteilung von virtuellen Quellen außerhalb des Zentrums zum Teil zu starken Lokalisationsverzerrungen, wenn keine Kompensation des Nahfeldeffektes stattfindet.

Wellenfeld einer ebenen Welle ($f=500[\text{Hz}]$, $M=15$),
welche aus $\Theta=35[\text{Grad}]$ und $\phi=0[\text{Grad}]$ einfällt



(a) Wellenfeld *ohne* Kompensation des Nahfeldef-fektes

Wellenfeld einer ebenen Welle ($f=500[\text{Hz}]$, $M=15$),
welche aus $\Theta=35[\text{Grad}]$ und $\phi=0[\text{Grad}]$ einfällt



(b) Wellenfeld *mit* Kompensation des Nahfeldef-fektes

Abbildung 4.7: Darstellung des Wellenfeldes einer ebenen Welle mit $f = 500 [\text{Hz}]$ aus $\Theta_{\text{SRC}} = 35^\circ$ und $\varphi_{\text{SRC}} = 0^\circ$, welches mit 99 Lautsprechern, bis zu einer Ordnung von $M = 15$ nachgebildet wurde.

4.2.3 Der Decodierprozess

Aus Kapitel 3 ist bekannt, dass die Aufgabe der Decodierung in einem Ambisonicsystem darin besteht, einen Satz von Lautsprechersignalen bzw. Gewichtungsfaktoren zu erstellen, deren Überlagerung im *sweet spot* des im Prinzip frei wählbaren Wiedergabesystems ein zuvor codiertes, beliebiges Klangfeld möglichst originalgetreu reproduziert.

In Anlehnung an die Codierung eines künstlichen Wellenfeldes, welches sich aus der Überlagerung von k künstlichen Quellen ergibt (siehe Gleichung 4.10 für ebene Wellen), lässt sich auch eine Codierung für das Wiedergabesystem bestehend aus L Lautsprechern auf die exakt gleiche Weise aufstellen. Der einzige Unterschied besteht darin, dass nun nicht die Lagen der virtuellen Quellen für die Codierung verwendet werden, sondern die örtlichen Lagen der L Lautsprecher $(\Theta_{S_l}, \varphi_{S_l})$, mit $l = [1, \dots, L]$.

Dem Ziel der Decodierung mit dem Wiedergabesystem das gleiche entweder künstlich codierte oder mit Hilfe eines Mikrophonarrays aufgezeichnete Wellenfeld zu erzeugen, kommt man durch Gleichsetzen der beiden Codierungen näher.

$$\begin{aligned}
 \mathbf{B}_{\text{SRC}} &= \mathbf{B}_S, \text{ unter Verwendung von Gleichung 4.10 folgt} \\
 \sum_{i=1}^k S_i(j\omega) Y_{m,n}^\sigma(\Theta_{\text{SRC}_i}, \varphi_{\text{SRC}_i}) &= \sum_{l=1}^L Y_{m,n}^\sigma(\Theta_{S_l}, \varphi_{S_l}) Y_l(j\omega), \text{ wobei} \\
 \mathbf{B}_{\text{SRC}} &= \text{Ambisonicsignale der } k \text{ (virtuellen) Quellen bei} \\
 &\quad (\Theta_{\text{SRC}_i}, \varphi_{\text{SRC}_i}), \text{ mit } i = [1, \dots, k], \\
 \mathbf{B}_S &= \text{Ambisonicsignale der } L \text{ Lautsprecher bei} \\
 &\quad (\Theta_{S_l}, \varphi_{S_l}), \text{ mit } l = [1, \dots, L].
 \end{aligned} \tag{4.15}$$

Aus diesem in Gleichung 4.15 durchgeführten Gleichsetzen der beiden Codierungen, was auch unter der Bezeichnung *matching condition* bekannt ist, lässt sich nun für den allgemeinen, d.h. für den 3D-Fall, in kompakter Matrixschreibweise unter Berücksichtigung von Gleichung 4.10 folgende, als *re-encoding* bezeichnete, Vorschrift ableiten ([DMN03]):

$$\begin{aligned}
\mathbf{B}_{N \times 1} &= \mathbf{C}_{N \times L} \mathbf{Y}_{L \times 1}, \text{ mit} \\
\mathbf{Y}_{L \times 1} &= [Y_1(j\omega), Y_2(j\omega), \dots, Y_L(j\omega)]^T, \\
\mathbf{C}_{N \times L} &= \begin{pmatrix} Y_{0,0}^{+1}(\Theta_1, \varphi_1) & Y_{0,0}^{+1}(\Theta_2, \varphi_2) & \dots & Y_{0,0}^{+1}(\Theta_L, \varphi_L) \\ Y_{1,0}^{+1}(\Theta_1, \varphi_1) & Y_{1,0}^{+1}(\Theta_2, \varphi_2) & \dots & Y_{1,0}^{+1}(\Theta_L, \varphi_L) \\ Y_{1,1}^{+1}(\Theta_1, \varphi_1) & Y_{1,1}^{+1}(\Theta_2, \varphi_2) & \dots & Y_{1,1}^{+1}(\Theta_L, \varphi_L) \\ \vdots & \vdots & \ddots & \vdots \\ Y_{M,0}^{+1}(\Theta_1, \varphi_1) & Y_{M,0}^{+1}(\Theta_2, \varphi_2) & \dots & Y_{M,0}^{+1}(\Theta_L, \varphi_L) \\ \vdots & \vdots & \ddots & \vdots \\ Y_{M,M}^{-1}(\Theta_1, \varphi_1) & Y_{M,M}^{-1}(\Theta_2, \varphi_2) & \dots & Y_{M,M}^{-1}(\Theta_L, \varphi_L) \end{pmatrix}, \text{ wobei} \\
\mathbf{Y}_{L \times 1} &= \text{Spaltenvektor, welcher alle } L \text{ Lautsprechersignale beinhaltet.} \\
\mathbf{C}_{N \times L} &= \text{Inverse Decodermatrix (re-encoding-Matrix) mit } N = (M+1)^2 \text{ Zeilen und} \\
&\quad L \text{ Spalten.} \\
\mathbf{B}_{N \times 1} &= \text{Spaltenvektor, welcher alle } N = (M+1)^2 \text{ Ambisonicsignale, gemäß} \\
&\quad \text{Gleichung 4.10 beinhaltet.} \tag{4.16}
\end{aligned}$$

Durch Umstellung von Gleichung 4.16 ergibt sich nun direkt die *Decodiervorschrift* eines Ambisonicsystems, mit welcher sich die L Lautsprechersignale \mathbf{Y} wie folgt berechnen lassen:

$$\begin{aligned}
\mathbf{Y}_{L \times 1} &= \mathbf{C}_{N \times L}^{-1} \mathbf{B}_{N \times 1} = \mathbf{D}_{L \times N} \mathbf{B}_{N \times 1}, \text{ wobei} \\
\mathbf{D}_{L \times N} &= \mathbf{C}_{N \times L}^{-1} = \text{Decodermatrix.} \tag{4.17}
\end{aligned}$$

Die Decodiervorschrift nach Gleichung 4.17 besagt, dass sich die in \mathbf{Y} enthaltenen Ansteuersignale der L Lautsprecher ergeben, indem die aus der Codierung stammenden N Ambisonicsignale des Vektors \mathbf{B} mit den $L \times N$ Faktoren der Decodermatrix \mathbf{D} gewichtet werden. Diese Gewichte ergeben sich dabei aus der Inversen der *re-encoding*-Matrix \mathbf{C} , deren Inhalt durch die N sphärischen Harmonischen der L Lautsprecher, welche sich an den Orten $(\Theta_{S_l}, \varphi_{S_l})$, mit $l = [1, \dots, L]$ befinden, definiert ist.

Da die Codierung alleinig in \mathbf{B} , d.h. in den Ambisonicsignalen steckt und die Decodierung ausschließlich durch die Decodermatrix \mathbf{D} realisiert ist, wird deutlich was eingangs bereits mehrfach erwähnt wurde:

\implies Bei einem Ambisonicsystem ist der Codierprozess vom Decodierprozess von-

einander vollkommen unabhängig.

Je nachdem ob die Codierung des Wellenfeldes auf der Zerlegung durch sphärische Harmonische nach Gleichung 4.4 für ebene Wellen bzw. gemäß Gleichung 4.13 für Kugelwellen beruht, berechnet sich \mathbf{B} entweder nach Formel 4.10 oder nach Formel 4.14, wobei Letzteres, wie bereits bekannt, nicht bei der Codierung sondern vorzugsweise bei der Decodierung angewandt werden sollte, worauf nachfolgend eingegangen wird.

Damit sämtliche Richtungsinformationen des codierten Wellenfeldes \mathbf{B} erhalten bleiben, muss gemäß [Son03] die Decodiervorschrift nach Gleichung 4.17 bestimmt oder vorzugsweise überbestimmt sein, was der Fall ist, wenn

$$\begin{aligned} L &\geq N, \text{ mit} \\ N &= \begin{cases} (2M + 1), & \text{2D-Fall} \\ (M + 1)^2, & \text{3D-Fall} \end{cases} \end{aligned} \quad (4.18)$$

ist.

4.2.4 Lautsprecheraufstellung und Decodermatrix

Zur Bestimmung der Decodermatrix \mathbf{D} ist es nötig, die *re-encoding*-Matrix \mathbf{C} zu invertieren. Im Allgemeinen ist dies nur möglich, wenn es sich dabei um eine quadratische Matrix handelt, die nicht-singulär ist, d.h. vollen Rang besitzt. Dies ist der Fall, wenn $L = N$ und $L > 1$ gilt. Für alle anderen Fälle, d.h. $L > N$, wobei es sich dann um nicht-quadratische Matrizen handelt, ergibt sich \mathbf{D} aus der sogenannten *Pseudoinversen*⁸ von \mathbf{C} , die nach

⁸Die Pseudoinverse der Matrix \mathbf{A} gegeben durch $\mathbf{A}^+ \in \mathbb{R}^{n \times m}$ lässt sich mit Hilfe der SVD (engl. *Singular Value Decomposition*) in zwei orthogonale und eine diagonale Matrix zerlegen, auf deren Basis sich dann zeigen lässt, dass

$$\begin{aligned} \mathbf{A}^+ &= \begin{cases} \left(\mathbf{A}^T \mathbf{A} \right)^{-1} \mathbf{A}^T, & \text{wenn } Rg(\mathbf{A}) = n \\ \mathbf{A}^{-1}, & \text{wenn } Rg(\mathbf{A}) = m = n, \text{ wobei} \end{cases} \\ Rg(\mathbf{A}) &= \text{Rang der Matrix } \mathbf{A}. \end{aligned}$$

Dabei muss \mathbf{A}^+ stets die vier *Moore-Penrose-Bedingungen* gegeben durch:

1. $\mathbf{A} \mathbf{A}^+ \mathbf{A} = \mathbf{A}$,
2. $\mathbf{A}^+ \mathbf{A} \mathbf{A}^+ = \mathbf{A}^+$,
3. $(\mathbf{A} \mathbf{A}^+)^T = \mathbf{A} \mathbf{A}^+$,
4. $(\mathbf{A}^+ \mathbf{A})^T = \mathbf{A}^+ \mathbf{A}$,

erfüllen, weshalb die Pseudoinverse \mathbf{A}^+ auch als *Moore-Penrose-Inverse* bezeichnet wird.

Beinhaltet \mathbf{A} hingegen *komplexe Werte*, d.h. ist $\mathbf{A} \in \mathbb{C}^{n \times m}$, gelten obige Bedingungen nach wie vor, sofern die dort verwendeten *Transponierungen* (z.B. \mathbf{A}^T) allesamt durch *konjugiert komplexe Transponierungen* (z.B. $\mathbf{A}^H = (\mathbf{A}^*)^T$) ersetzt werden.

[BK99, GL96] wie folgt definiert ist:

$$\begin{aligned}\mathbf{D} &= \mathbf{C}^+ = (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T, \text{ wobei} \\ \mathbf{C}^+ &= \text{Pseudoinverse der Matrix } \mathbf{C}.\end{aligned}\tag{4.19}$$

Bekanntermaßen wird die *re-encoding*-Matrix \mathbf{C} und somit auch ihre Pseudoinverse gegeben durch die Decodermatrix \mathbf{D} alleine durch die Art der Lautsprecheraufstellung, also deren örtliche Lage $(r_{S_i}, \Theta_{S_i}, \varphi_{S_i})$, mit $i = [1, \dots, L]$ bestimmt.

Wählt man dabei eine reguläre, äquidistante Aufstellung der Lautsprecher gegeben durch:

$$\begin{aligned}r_{S_1} &= r_{S_i}, \forall i \in [1, \dots, L], \\ \Delta\Theta_S &= \Theta_{S_{i+1}} - \Theta_{S_i} = \text{Konstant}, \forall i \in [1, \dots, L-1], \\ \Delta\varphi_S &= \varphi_{S_{i+1}} - \varphi_{S_i} = \text{Konstant}, \forall i \in [1, \dots, L-1], \text{ wobei} \\ \Delta\Theta_S &= \text{Horizontale Winkeldifferenz in [rad] zwischen den beiden benachbarten} \\ &\quad \text{Lautsprechern bei } i \text{ und } i+1, \\ \Delta\varphi_S &= \text{Vertikale Winkeldifferenz in [rad] zwischen den beiden benachbarten} \\ &\quad \text{Lautsprechern bei } i \text{ und } i+1,\end{aligned}\tag{4.20}$$

was einer regelmäßigen Anordnung entlang einer Kreislinie für den 2D-Fall (regelmäßiges Polygon) bzw. auf einer Kugeloberfläche für den 3D-Fall (regulärer Polyeder ⁹) gleich kommt, vereinfacht sich die Berechnung der Pseudoinversen \mathbf{C}^+ wie folgt:

$$\begin{aligned}\mathbf{D} &= \mathbf{C}^+ = \frac{1}{N} \mathbf{C}^T, \text{ wobei} \\ N &= \text{Anzahl der Ambisonicsignale}.\end{aligned}\tag{4.21}$$

⁹Von den regulären Polyedern, wobei es sich um Körper handelt, die sich ausschließlich aus regelmäßigen, kongruenten, also deckungsgleichen Vielecken zusammensetzen, existieren insgesamt 9 Grundformen. Bei 5 von diesen handelt es sich um konvexe Grundformen, d.h. um welche mit einer nach außen gewölbten Linienführung, die auch als *platonische Körper* bezeichnet werden, wobei es sich konkret um folgende Körper handelt:

Platonischer Körper (=Konvexer Polyeder)	Tetraeder	Hexaeder	Oktaeder	Dodekaeder	Ikosaeder
Art der Seitenflächen	Gleichseitige Dreiecke	Quadrate	Gleichseitige Dreiecke	Regelmäßige Fünfecke	Gleichseitige Dreiecke
Anzahl der Flächen	4	6	8	12	20
Anzahl der Ecken	4	8	6	20	12
Anzahl der Kanten	6	12	12	30	30

4.2.5 Verallgemeinerung der Decodervariationen

Wie bereits für Ambisonicsysteme 1. Ordnung in Kapitel 3.1.1 gezeigt, lassen sich auch für HOA-Systeme Decodervariationen durch Modifikation der Decodervorschrift gemäß Gleichung 4.17 wie folgt verallgemeinern:

$$\begin{aligned}
 \mathbf{Y}_{L \times 1} &= \mathbf{D}_{L \times N} \mathbf{G}_{N \times N} \mathbf{B}_{N \times 1}, \text{ mit} \\
 \mathbf{G}_{N \times N} &= \begin{pmatrix} g_0 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & g_1 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ \vdots & 0 & g_1 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ \vdots & \vdots & 0 & g_2 & 0 & \dots & \dots & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & 0 & \ddots & 0 & \dots & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 & g_2 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \ddots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & g_M & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & g_M \end{pmatrix}, \text{ oder} \\
 \mathbf{G}_{N \times N} &= \text{diag}(\mathbf{g}) = \text{diag} \left(\left[g_0, \dots, \overbrace{g_m, \dots, g_m}^{(2m+1)}, \dots, \overbrace{g_M, \dots, g_M}^{(2M+1)} \right]^T \right), \text{ wobei} \\
 \mathbf{G}_{N \times N} &= \text{Diagonale (quadratische) Matrix der Ordnung } N, \text{ welche für den} \\
 &\text{allgemeinen 3D-Fall, pro Ordnung } m, \text{ mit } 0 \leq m \leq M, (2m+1) \\
 &\text{identische Gewichtungsfaktoren } g_m \text{ besitzt.} \tag{4.22}
 \end{aligned}$$

Für regelmäßige Lautsprecheranordnungen, d.h. für Fälle in denen die vereinfachte Berechnung der Decodermatrix \mathbf{D} nach Gleichung 4.21 gültig ist, lassen sich die Gewichte \mathbf{g} nach [DRP98, Dan00, Son03, Neu07, Neu] für die Decodervarianten:

- **Basic (Max r_v):**

Für die Optimierung der Länge des Schallschnellektors \mathbf{v} , was einer Optimierung der ITD gleichkommt (bekanntermaßen im unteren Frequenzbereich vorteilhaft),

- **Max r_e :**

Für die Optimierung der Länge des Schallenergievektors \mathbf{e} , was einer Optimierung der ILD gleichkommt (günstig für höhere Frequenzen),

- **In-Phase:**

Für die Optimierung der Größe der Fläche innerhalb des Wiedergabesystems, innerhalb welcher die Lokalisation erhalten bleibt (für große Räume bzw. bei vielen Hörern),

wie in Tabelle 4.2 angegeben berechnen.

Decodervariante	<i>Basic</i> (= Max r_v)	Max r_e	In-Phase
2D-Fall:			
Gewichte:	$\mathbf{g}_m = g_0$	$\mathbf{g}_m = g_0 \cos\left(\frac{m\pi}{2M+2}\right)$	$\mathbf{g}_m = g_0 \frac{(M!)^2}{(M+m)!(M-m)!}$
Skalierung:	$g_0 = \sqrt{\frac{N}{2M+1}} = 1$ (Energieerhaltend)	$g_0 = \sqrt{\frac{N}{M+1}}$ (Energieerhaltend) $g_0 = 1$ (Schalldruckerhaltend)	$g_0 = \frac{(2M)!(2M+1)}{4^M (M!)^2}$ (Energieerhaltend)
Energie:	$E(\mathbf{g}_m, M) = \frac{1}{N} \left(g_0^2 + 2 \sum_{m=1}^M g_m^2 \right)$		
Länge des Energievektors: r_e	$r_e = \frac{2 \sum_{m=1}^M g_{m-1} g_m}{g_0^2 + 2 \sum_{m=1}^M g_m^2}$		
r_e^{Max} :	$r_e = \frac{2M}{2M+1}$	$r_e = \cos\left(\frac{\pi}{2M+2}\right)$	$r_e = \frac{2M}{2M+1}$
Länge des Schnellektors: $r_v = \frac{\mathbf{g}_1}{g_0}$			
r_v :	$r_v = 1$	$r_v = r_e^{\text{Max}}$	$r_v = \frac{M}{M+1}$
Richtdiagramm: $\Psi(\gamma_l, M) = \frac{1}{N} \left(g_0 + 2 \sum_{m=1}^M g_m \cos(m\gamma_l) \right)$ mit (Panning) $\gamma_l = \Theta_{S_l} - \Theta_{SRC}$, $l = [1, \dots, L]$			
Richtdiagramm:	$\Psi(\gamma_l, M) = \frac{\sin\left(\frac{2M+1}{2}\gamma_l\right)}{N \sin\left(\frac{1}{2}\gamma_l\right)}$		$\Psi(\gamma_l, M) = \left(\cos\left(\frac{\gamma_l}{2}\right)\right)^{2M} = \frac{1}{2^M} \sum_{m=0}^M \binom{M}{m} \cos^m(\gamma_l)$
3D-Fall:			
Gewichte:	$\mathbf{g}_m = g_0$	$\mathbf{g}_m = g_0 \mathbf{g}'_m$	$\mathbf{g}_m = g_0 \frac{M!(M+1)}{(M+m+1)!(M-m)!}$
Skalierung:	$g_0 = \frac{\sqrt{N}}{M+1} = 1$ (Energieerhaltend)	$g_0 = \sqrt{\frac{1}{E(\mathbf{g}'_m, M)}}$, mit $\mathbf{g}'_m = P_m(r_e^{\text{Max}})^{10}$ (Energieerhaltend)	$g_0 = \frac{\sqrt{N(2M+1)}}{M+1}$ (Energieerhaltend)
Energie:	$E(\mathbf{g}_m, M) = \frac{1}{N} \left(\sum_{m=0}^M (2m+1) g_m^2 \right)$		
Länge des Energievektors: r_e	$r_e = \frac{2 \sum_{m=1}^M m g_{m-1} g_m}{\sum_{m=0}^M (2m+1) g_m^2}$		
r_e^{Max} :	$r_e = \frac{M}{M+1}$	$r_e = P_{M+1}^{-1}(0)^{11}$	$r_e = \frac{M}{M+1}$
Länge des Schnellektors: $r_v = \frac{\mathbf{g}_1}{g_0}$			
r_v :	$r_v = 1$	$r_v = r_e^{\text{Max}}$	$r_v = \frac{M}{M+2}$
Richtdiagramm: $\Psi(\gamma_l, M) = \frac{1}{N} \left(\sum_{m=0}^M (2m+1) g_m P_m(\cos(\gamma_l)) \right)$, mit (Panning) $\gamma_l = \angle(\mathbf{r}_{S_l} - \mathbf{r}_{SRC}) = \arccos(\mathbf{r}_{S_l} \cdot \mathbf{r}_{SRC})$, $l = [1, \dots, L]$			
Richtdiagramm:	$\Psi(\gamma_l, M) = \frac{\sin\left(\frac{2M+1}{2}\gamma_l\right)}{N \sin\left(\frac{1}{2}\gamma_l\right)}$		$\Psi(\gamma_l, M) = \left(\cos\left(\frac{\gamma_l}{2}\right)\right)^{2M} = \frac{1}{2^M} \sum_{m=0}^M \binom{M}{m} \cos^m(\gamma_l)$

Tabelle 4.2: Berechnungsvorschriften für die Decodervarianten *Basic* (Max r_v), *Max r_e* und *In-Phase*, welche ausschließlich für regelmäßige Lautsprecheranordnungen gültig sind.

Neben den Gewichten \mathbf{g}_m , deren Berechnungsvorschriften unter der Rubrik „Gewichte“ für die drei betrachteten Decodervarianten *Basic* ($Maxr_v$), $Maxr_e$ und *In-Phase* sowohl für den 2D- als auch für den 3D-Fall in Tabelle 4.2 zu finden sind, enthält diese auch noch weitere Größen auf die nachfolgend genauer eingegangen wird.

Bei der Berechnung der Gewichte \mathbf{g}_m wird der Wert g_0 aufgeführt, wobei es sich um einen positiven Wert handelt, der zur Skalierung bzw. Normierung der jeweiligen Berechnungsvorschrift für \mathbf{g}_m dient. Je nach verwendeter Decodervariante kann die Skalierung dabei energieerhaltend oder schalldruck- bzw. pegelerhaltend ausgelegt werden. Im Falle des In-Phase-Decoders wird die Decodierung so bestimmt, dass das Richtungsempfinden über einen möglichst maximalen räumlichen Bereich erhalten bleibt, da bei dieser Decodervariante rein konstruktive Signale, d.h. solche mit gleicher Phasenlage verwendet werden. Somit realisiert der In-Phase-Decoder ein Amplituden-*panning*, d.h. es werden nur von jenen Lautsprechern Signale mit nennenswerter Amplitude abgestrahlt, die sich auch physikalisch in Richtung der gewünschten Quelle befinden.

Die Gleichung zur Berechnung der Gesamtenergie des Schallfeldes E definiert sich anhand der Gewichte \mathbf{g}_m sowie durch die Ordnung M des Ambisonicsystems. Sofern eine energieerhaltende Skalierung für g_0 verwendet wird, wird diese zu $E(\mathbf{g}_m, M) = 1$.

In der nächsten Zeile folgt die Formel zur Berechnung der Länge des Energievektors r_e , die unabhängig von der verwendeten Decodervariante allgemeine Gültigkeit besitzt und sich ebenfalls einzig durch die eingestellten Gewichte \mathbf{g}_m sowie durch die Ordnung M berechnet. Sie spielt eine wichtige Rolle in der qualitativen Beurteilung des synthetisierten Wellenfeldes, wobei sie idealerweise eine Länge von $r_e = 1$ aufweisen soll, was auch für r_v zutrifft. Nachfolgend sind Formeln zur Berechnung von r_e^{Max} aufgeführt, mit denen sich die maximal mögliche Länge des Energievektors \mathbf{e} bei Verwendung der drei unterschiedlichen Decodervarianten sowohl für den 2D- als auch für den 3D-Fall ermitteln lässt. Diese ergeben sich, in dem man die zuvor vorgestellte Formel zur allgemeinen Berechnung der Länge des Energievektors r_e partiell nach \mathbf{g}_m ableitet.

Als letzte Größe beinhaltet Tabelle 4.2 unter der Rubrik „Richtdiagramm“ eine Gleichung zur allgemeinen Berechnung der sich ergebenden *panning*-Funktion $\Psi(\gamma_l, M)$, die sich durch Kombination der Codiervorschrift nach Formel 4.10 und Gleichung 4.23 ergibt,

$$\mathbf{D}_{L \times N} = \frac{1}{N} \mathbf{C}_{N \times L}^T \mathbf{G}_{N \times N}, \quad (4.23)$$

¹⁰ $P_m(r_e^{\text{Max}})$ = Legendrepolynom des Arguments r_e^{Max} , welches für die Ordnungen $m = [1, \dots, M]$ zu berechnen ist.

¹¹ $P_{M+1}^{-1}(0)$ = Größte Nullstelle des Legendrepolynoms ($M+1$). Ordnung in einem Wertebereich von ± 1 . Somit gilt: $P_{M+1}^{-1}(0) = \max\{\text{roots}(P_{M+1}(x))\}$, mit $x \in [-1, \dots, 1]$. Alternativ lässt sich die größte Nullstelle eines Legendrepolynoms M . Ordnung aber auch anhand des größten Eigenwerts einer Matrix bestimmen, in welcher lediglich die obere und untere Nebendiagonale folgende, von Null verschiedene Werte aufweisen: $\frac{m}{\sqrt{4m^2-1}}$, mit $m = [1, \dots, M-1]$, die auch als *Jacobimatrix* bezeichnet wird.

was eine Verallgemeinerung der vereinfachten Decodierung nach Gleichung 4.21 darstellt. Auch hier existieren Spezialfälle für die Decodervarianten *Basic* und *In-Phase*, welche unterhalb der allgemein gültigen Formel angegeben sind. Mit ihnen lassen sich deren *panning*-Funktionen wesentlich effizienter berechnen als bei Verwendung der allgemeinen Formel. Dies ist vor allem dann interessant, wenn es darum geht, Ambisonicsysteme für virtuelle Quellen in der Praxis zu erstellen, deren räumliche Lage man in Echtzeit verändern können soll.

Mit Hilfe der *panning*-Funktion $\Psi(\gamma_l, M)$ lassen sich demnach die Ansteuersignale der L Lautsprecher direkt, d.h. ohne zuerst die Codierung und Decodierung berechnen zu müssen, für reguläre Lautsprecheranordnungen und virtuelle Quellen, die ebene Wellen aussenden wie folgt berechnen:

$$\begin{aligned}
\mathbf{Y} &= \mathbf{S} \Psi(\gamma_{l,i}, M), \text{ wobei} \\
\mathbf{S} &= [S_1(j\omega), \dots, S_k(j\omega)] = \text{Vektor, welcher die Signale der } k \text{ virtuellen Quellen} \\
&\quad \text{beinhaltet,} \\
\gamma_{l,i} &= \text{Differenz des Winkels zwischen dem Vektor zum Lautsprecher } l \\
&\quad (\mathbf{r}_{S_l} = f(r_{S_l}, \Theta_{S_l}, \varphi_{S_l})), \text{ mit } l \in [1, \dots, L] \text{ und dem zur Schallquelle } i \\
&\quad (\mathbf{r}_{\text{SRC}_i}, (\mathbf{r}_{\text{SRC}_i} = f(\Theta_{\text{SRC}_i}, \varphi_{\text{SRC}_i}))), \text{ mit } i \in [1, \dots, k].
\end{aligned} \tag{4.24}$$

Wie aus Gleichung 4.24 hervorgeht, bezeichnet der Vektor $\gamma_{l,i}$ die Winkeldifferenz zwischen dem Vektor zum l . Lautsprecher \mathbf{r}_{S_l} und dem Vektor zur i . Quelle $\mathbf{r}_{\text{SRC}_i}$, was allgemein wie folgt beschrieben werden kann:

$$\begin{aligned}
\gamma_{l,i} &= \begin{cases} \Theta_{S_l} - \Theta_{\text{SRC}_i}, & \text{2D-Fall} \\ \angle(\mathbf{r}_{S_l} - \mathbf{r}_{\text{SRC}_i}) = \arccos(\mathbf{r}_{S_l} \cdot \mathbf{r}_{\text{SRC}_i}), & \text{3D-Fall, mit} \end{cases} \\
\mathbf{r}_{S_l} \cdot \mathbf{r}_{\text{SRC}_i} &= \text{Skalarprodukt zwischen den Vektoren } \mathbf{r}_{S_l} \text{ und } \mathbf{r}_{\text{SRC}_i}, \\
l &= [1, \dots, L], \\
i &= [1, \dots, k].
\end{aligned} \tag{4.25}$$

Abbildung 4.8 veranschaulicht die *panning*-Funktionen, d.h. die Gewichte für die L Lautsprecher sowohl im kartesischen als auch im polaren Koordinatensystem anhand eines Beispiels. Hierbei bezeichnen die blauen Balken positive und die roten Balken negative Gewichte, also solche mit umgekehrten Vorzeichen, wobei die Länge der Balken den Betrag der Gewichte widerspiegeln. Die Funktion an denen die Balken angrenzen entsteht, wenn unendlich viele Lautsprecher für die Decodierung verwendet werden. Anhand der Darstellungen in Polarkoordinaten ist es ein Leichtes zu erkennen, dass diese Funktionen Richtdiagrammen, wie sie beim Entwurf von Mikrophonarrays verwendet werden, entsprechen. Folglich lassen sich die *panning*-Funktionen und somit auch der gesamte Codier- und Decodierprozess auch als Richtdiagramm eines virtuellen Mikrophons interpretieren, wo-

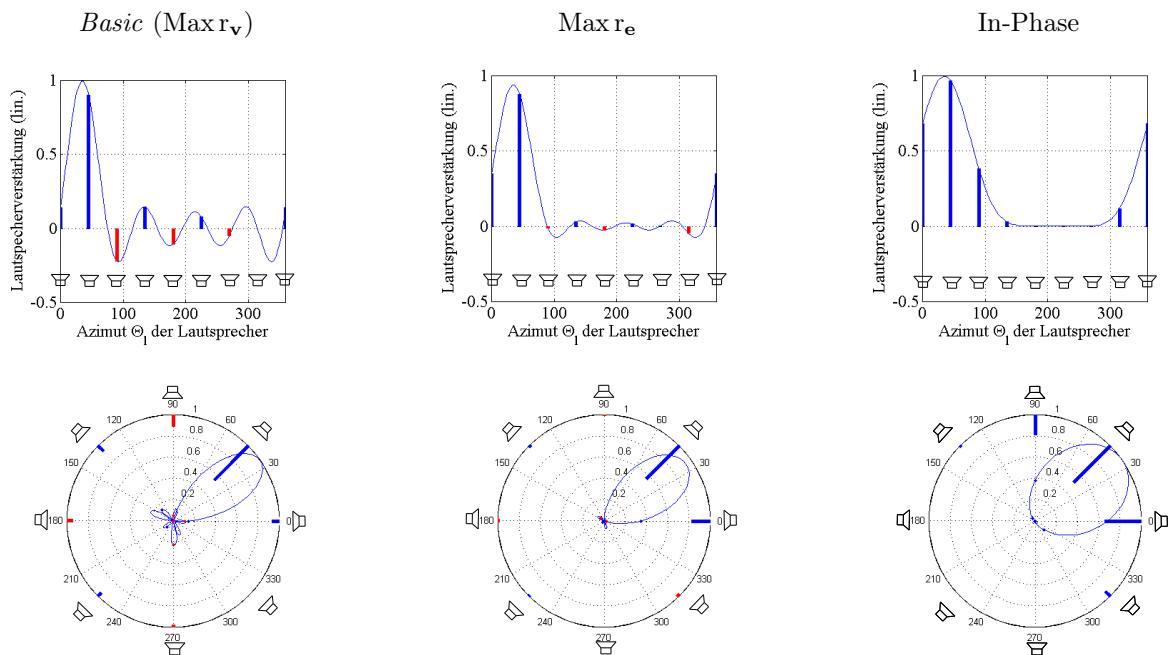


Abbildung 4.8: Panning-Funktionen $\Psi(\gamma_l, M)$ der drei Decodervarianten *Basic*, *Maxr_e* und *In-Phase* für ein System 4. Ordnung ($M = 4$), anhand eines Beispiels mit einer regulären, zweidimensionalen Lautsprecheraufstellung mit $L = 8$ Lautsprechern, die sich an den Orten $\mathbf{r}_{S_l} = (\Theta_{S_l} = [0^\circ : 45^\circ : 315^\circ], \varphi_{S_l} = 0^\circ)$ befinden und von dem Einfall einer ebenen Welle, deren Quelle sich am Ort $\mathbf{r}_{SRC} = (\Theta_{SRC} = 35^\circ, \varphi_{SRC} = 0^\circ)$ befindet, ausgegangen wird. Anhand der Darstellung der panning-Funktionen im Polarkoordinatensystem erkennt man, dass diese auch als Richtdiagramm eines virtuellen Mikrophons, welches in Richtung der Quelle zeigt, aufgefasst werden können.

mit sich auch der in Tabelle 4.2 verwendete Begriff „Richtdiagramm“ für die Beschreibung der *panning*-Funktionen erklärt.

Ein Vergleich der Richtcharakteristika der virtuellen Mikrophone, die sich bei Gebrauch der drei Decodervarianten ergeben, offenbart, dass die Decodervariante *Basic*, also jene, welche die Länge des Schallschnellektors r_v optimiert, den geringsten Öffnungswinkel, jedoch dafür die größten Nebenkeulen besitzt. Die Decodervariante r_e besitzt hingegen einen Öffnungswinkel, der zwar geringfügig größer ausfällt, jedoch dafür deutlich geringere Nebenzipfel zeigt. Der In-Phase-Decoder hingegen zeigt keine negativen Anteile mehr auf, weshalb hier auch keine roten Balken, d.h. *keine Nebenkeulen* mehr zu erkennen sind. Diesem Vorteil steht allerdings ein deutlich größerer Öffnungswinkel entgegen.

Aus dem Richtdiagramm des virtuellen Mikrophons lässt sich auf einfache und plausible Weise eine Aussage über die akustische Qualität eines Ambisonicsystems abschätzen. Dabei gibt der Öffnungswinkel Auskunft über die Lokalisationsschärfe und die Nebenkeulen etwas über die Größe des Raumes, innerhalb dessen das akustische Abbild Gültigkeit besitzt. Dies kann auch als *räumliche Stabilität* des Klangbildes bezeichnet werden.

Weiter Klarheit darüber, wie sich ein Ambisonicsystem anhand von Richtcharakteristika virtueller Mikrophone interpretieren lässt, verschafft Abbildung 4.9. Hier sind auf der linken Seite die 2D-Richtcharakteristika der sphärischen Harmonischen $Y_{m,n}^{\sigma=1}$, mit $m = n$ der Ordnungen $m = [0, \dots, 4]$ aufgezeigt, aus deren Überlagerungen je nach verwendeter Decodervorschrift (*Basic*, $Max r_e$, In-Phase), die jeweils rechts dargestellten Richtcharakteristika resultieren. Bei genauerer Betrachtung der Resultate, die sich für ein Ambisonicsystem 1. Ordnung ergeben (siehe Kapitel 3), erkennt man, insbesondere durch einen Vergleich mit Abbildung 3.2, wie und warum sich die unterschiedlichen Richtcharakteristika ergeben. So zeigt sich, dass eine gewichtete Überlagerung der 0. mit der 1. sphärischen Harmonischen gemäß der Gleichungen 3.1 und 3.2 bei Verwendung des Richtfaktors $q = 1.33$ die Länge des Schallschnellektors v ebenso maximiert, wie bei Anwendung der Formel 4.22 bzw. Formel 4.24 unter Berücksichtigung der in Tabelle 4.2 aufgeführten (zweidimensionalen) *Basic*-Decodervariante. Analog gilt dies auch bei Verwendung eines Richtfaktors $q = 1.17$ sowie der Decodervariante $Max r_e$ für die Maximierung der Länge der Schallenergievektors e sowie mit $q = 1$ und des *In-Phase*-Decoders für die Optimierung der örtlichen Stabilität. Dabei sind beim Vergleich der beiden Varianten im Falle der Maximierung der Länge des Schallenergievektors e leichte Abweichungen zu verzeichnen, deren Ursprung jedoch nicht weiter verfolgt wurde, da dies für die Zielsetzung dieser Arbeit nicht von Bedeutung ist.

Anhand des Vergleichs der Ergebnisse eines Ambisonicsystems 1. Ordnung bei Anwendung unterschiedlicher Decodervarianten wird deutlich, dass nun mit Betrachtung von Ambisonicsystemen höherer Ordnung eine Verallgemeinerung des Codier- und Decodierprozesses erreicht wurde, der unabhängig von der gewählten Ordnung ist.

Auch bei den Bildern in Abbildung 4.9 kennzeichnen rote Linien innerhalb der Darstellungen Anteile der sphärischen Harmonischen mit umgekehrtem Vorzeichen. Ein weiterer

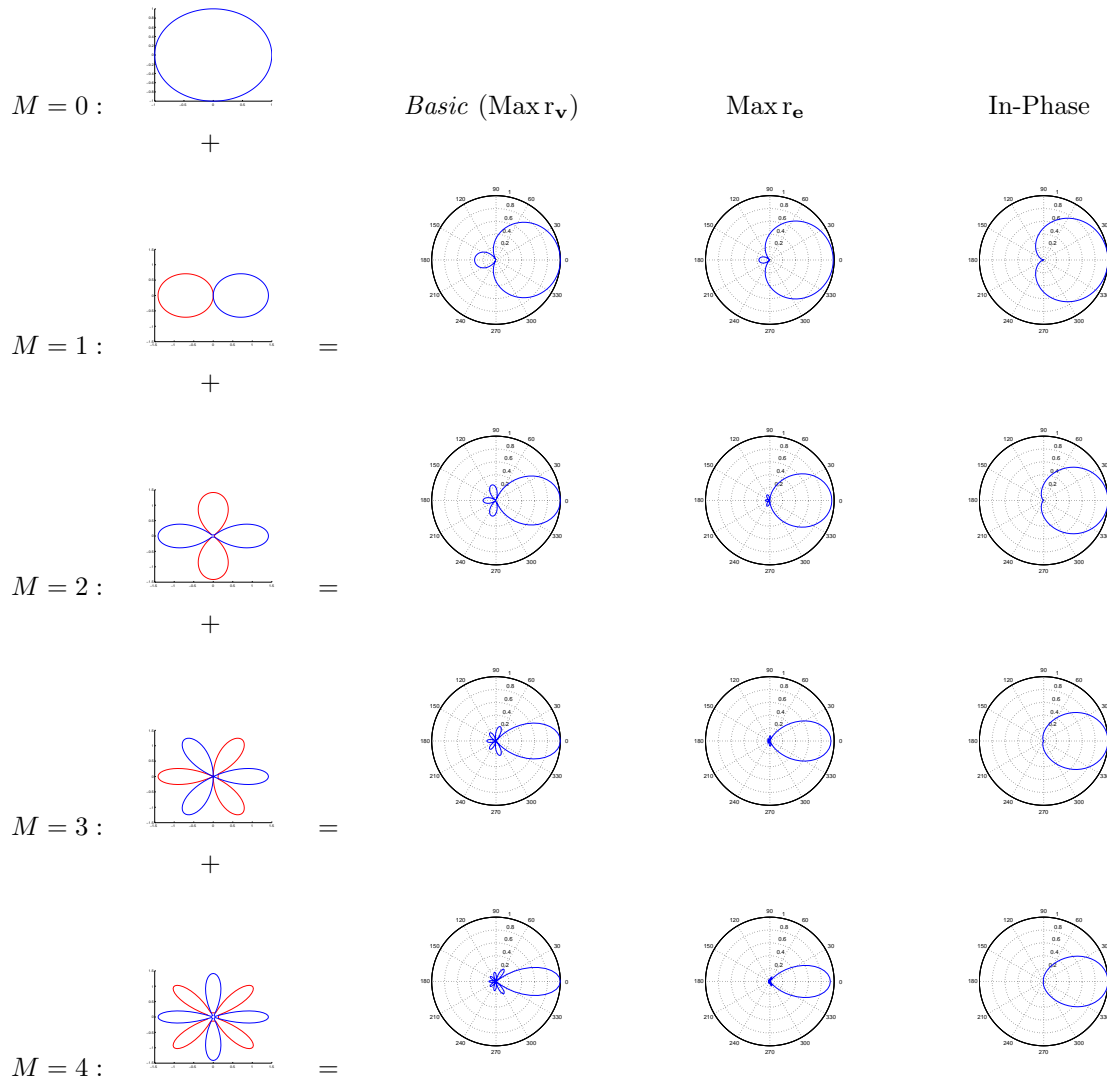


Abbildung 4.9: Interpretation der panning-Funktion $\Psi(\gamma_l, M)$, d.h. der kombinierten Codierung und Decodierung als Richtdiagramm eines virtuellen Mikrophons, für die drei Decodervarianten *Basic*, $\text{Max } r_e$ und *In-Phase* für die Ordnungen $M = [0, \dots, 4]$ anhand eines Beispiels, bei dem einerseits vom Einfall einer ebenen Welle ausgegangen wird, wobei sich die Quelle am Ort $\mathbf{r}_{SRC} = (\Theta_{SRC} = 0^\circ, \varphi_{SRC} = 0^\circ)$ befindet und eine reguläre, zweidimensionale Lautsprecheraufstellung mit $L = 9$ Lautsprechern verwendet wird, wobei sich diese an den Orten $\mathbf{r}_{S_l} = (\Theta_{S_l} = [0^\circ : 40^\circ : 320^\circ], \varphi_{SRC} = 0^\circ)$ befinden.

Aspekt, den Abbildung 4.9 offenbart, ist, dass mit steigender Ordnung der Öffnungswinkel abnimmt sowie die Anzahl der Nebenkeulen zunimmt, wobei sich deren Amplitude und somit Einfluss stetig reduziert. Somit wird deutlich, dass sich der spektrale Einsatzbereich der unterschiedlichen Decodervarianten mit zunehmender Ordnung hin zu höheren Frequenzen erweitert, da mit zunehmender Frequenz eine immer stärker fokussierende, räumliche Filterung benötigt wird, die erst mit steigender Ordnung realisiert werden kann. In Kapitel 3.1 wurde für ein System 1. Ordnung, gemäß Gerzon, eine Übergangsfrequenz zwischen dem *Basic*- und dem *Maxr_e*-Decoder von $f \approx 700$ [Hz] angegeben.

Bestimmung der Übergangsfrequenzen unterschiedlicher Decodervarianten

Wie kommt man aber auf diese Übergangsfrequenz und wie verändert sich diese mit steigender Ordnung? Dieser Frage wurde in [Pol96] nachgegangen und in [BDM06] verallgemeinert. Dabei wird von der sogenannten *Interferenzbreite* gesprochen, welche die Störung des synthetisierten Wellenfeldes widerspiegelt. Sie ist abhängig vom minimalen Abstand der Lautsprecher untereinander. Geht man etwa von einer regulären, zweidimensionalen Lautsprecheranordnung mit $L = 2M + 2 = N + 1$ Lautsprechern aus, die Daniel im Übrigen als optimal bezeichnet, so ergibt sich eine Lücke zwischen zwei benachbarten Lautsprechern von $\frac{2\pi}{L} = \frac{2\pi}{2M+2}$, die gemäß [BDM06] eine frequenzabhängige Interferenzbreite von:

$$\begin{aligned} \Lambda(f) &= \frac{\text{Wellenlänge in [m]}}{\text{Lautsprecherabstand (normiert)}} = \frac{\left(\frac{c}{f}\right)}{2 \sin\left(\frac{\pi}{(2M+2)}\right)}, \text{ wobei} \\ \Lambda(f) &= \text{Interferenzbreite in [m] bei einer Frequenz } f \text{ in [Hz]}, \end{aligned} \quad (4.26)$$

ergibt.

Geht man von einer Abhörfläche bzw. einem Abhörraum mit Radius R aus, die bzw. der sich im Zentrum eines regulären Wiedergabesystems befindet, wobei man sinnigerweise R so wählt, dass damit der Ohrabstand bzw. der Radius eines menschlichen Kopfes simuliert wird¹², so kann das Wellenfeld unter Berücksichtigung einer gewissen Sicherheitsreserve R_m an den Ohren des Zuhörers bis zu einer oberen Grenzfrequenz von f_{Lim} ,

¹²Ein erwachsener Mann hat in etwa einen Kopfumfang von $U_{\text{Kopf}} \approx 56$ [cm], was in etwa einen Radius von $R_{\text{Kopf}} \approx 8.9$ [cm] entspricht.

korrekt rekonstruiert werden, wenn gilt:

$$\begin{aligned} \Lambda(f) &= 2(R + R_m), \text{ wobei} \\ R &= \text{Radius der Abhörfläche bzw. des Abhörtraumes in [m], welcher sich um das} \\ &\quad \text{Zentrum des Abhörsystems herum befindet,} \\ R_m &= \text{Sicherheitsreserve in [m], die zu } R \text{ addiert werden muss, um Störungen} \\ &\quad \text{bei Überlagerung der einzelnen Wellenfelder der } L \text{ Lautsprecher zu} \\ &\quad \text{vermeiden.} \end{aligned} \tag{4.27}$$

Ferner stellte sich heraus, dass R_m indirekt proportional zur Ordnung M ist, wobei folgende, approximative Abhängigkeit gefunden werden konnte:

$$R_m = \frac{R}{M}. \tag{4.28}$$

Setzt man Formel 4.28 in Gleichung 4.27 ein, so ergibt sich daraus unter Berücksichtigung von Gleichung 4.26 die Grenzfrequenz f_{Lim} zu:

$$f_{\text{Lim}} \approx \frac{c M}{4 R (M + 1) \sin\left(\frac{\pi}{(2M+2)}\right)} \approx \frac{c M}{2 \pi R}, \tag{4.29}$$

deren gerundete Werte für die Ordnungen $M = [1, \dots, 4]$ aus Tabelle 4.3 entnommen werden können. Veranschaulichen lassen sich die Übergangsfrequenzen f_{Lim} durch Betrachtung

Ordnung M	1	2	3	4
f_{Lim} in [Hz]	700	1300	1900	2500
α_E in [°]	45	30	22.5	18

Tabelle 4.3: Übergangsfrequenzen zwischen dem $\text{Max}r_v$ - und $\text{Max}r_e$ -Decoder bei unterschiedlichen Ordnungen ($M = [1, \dots, 4]$) mit Angabe der zugehörigen Lokalisationsunschärfe α_E .

ung des sogenannten *truncation error* $e_M(kr)$, also dem Fehler der entsteht, wenn die Fourier-Bessel-Reihe zur Berechnung des Schallfeldes $P(\mathbf{r}, \omega)$ nach Gleichung 4.4 nicht, wie theoretisch gefordert bis zu einer unendlichen Ordnung, sondern lediglich bis zu einer bestimmten oberen Ordnung M berechnet wird. Bezeichnet man das so approximierte

Schallfeld als $P_M(\mathbf{r}, \omega)$ ergibt sich der *truncation error* allgemein zu:

$$e_M(kr) = \frac{\iint_S |p(\Theta, \varphi, kr) - \hat{p}_M(\Theta, \varphi, kr)|^2 dS}{\iint_S |p(\Theta, \varphi, kr)|^2 dS}, \text{ mit}$$

$$p(\Theta, \varphi, kr) = \sum_{m=0}^{\infty} j^m j_m(kr) \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^\sigma Y_{m,n}^\sigma(\Theta, \varphi), \text{ und}$$

$$\hat{p}_M(\Theta, \varphi, kr) = \sum_{m=0}^M j^m j_m(kr) \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^\sigma Y_{m,n}^\sigma(\Theta, \varphi), \text{ wobei}$$

$$\int_S \dots dS = \text{Integral über die gesamte Kugeloberfläche } S, \text{ mit } \Theta = [0, \dots, 2\pi]$$

und $\varphi = [0, \dots, \pi]$,

e_M = Fehler, der sich durch eine Approximation des Wellenfeldes bis zu einer maximalen Ordnung von M ergibt.

(4.30)

Beschränkt man sich hingegen bei der Betrachtung auf ebene Wellen, vereinfacht sich Gleichung 4.30 gemäß [BDM06] zu:

$$e_M(kr) = 1 - \sum_{m=0}^M (2m+1) (j_m(kr))^2. \quad (4.31)$$

Abbildung 4.10 veranschaulicht nun den Effekt, den ein Abbruch der Berechnung des Wellenfeldes einer ebenen Welle mit Hilfe der Fourier-Bessel-Reihe nach einer bestimmten Ordnung M bewirkt.

Weiter verdeutlicht Abbildung 4.10, dass bei $10 \log_{10} (|e_M(kr)|) \approx -14$ [dB], was in etwa einem Fehler von $e_M(kr) \approx 4\%$ entspricht, die Projektionen der Schnittpunkte der Verläufe $10 \log_{10} (|e_M(kr)|)$ mit dieser Fehlerschwelle (4%-Linie) auf die X-Achse, also der resultierenden kr -Werte, in etwa der Ordnung M entsprechen, also $kr \approx M$ gilt. Somit kann für eine bestimmte Wellenzahl k bzw. für eine bestimmte Frequenz f und einen gegebenen Radius r , an welchen das Wellenfeld betrachtet werden soll, folgende einfache Formel angegeben werden:

$$M = \lceil kr \rceil, \text{ wobei}$$

$$\lceil \cdot \rceil = \text{Aufrunden zum nächst größeren Integerwert.} \quad (4.32)$$

Mit deren Hilfe lässt sich die minimale Ordnung M bestimmen, damit das Schallfeld innerhalb der gegebenen Grenzen in ausreichender Genauigkeit approximiert werden kann.

¹³NMSE = engl. *Normalized Mean Squared Error*

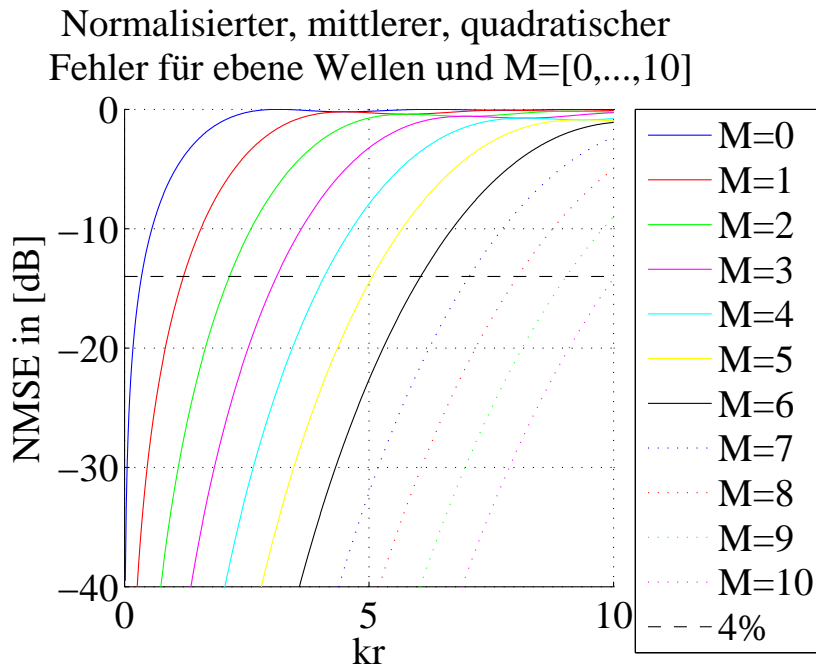


Abbildung 4.10: Normalisierter, quadratischer Fehler ($NMSE^{13}$) des truncation error ($= 10 \log_{10}(|e_M(kr)|)$) bei der Beschreibung des Wellenfeldes einer ebenen Welle bis zu einer maximalen Ordnung von $M = [0, \dots, 10]$

Geht man nun von einer gegebenen Beobachtungsfläche mit Radius r aus, welche z.B. der Abhörfläche entspricht, also einen Radius von $r = R$ aufweist, welcher sich am Ohrabstand orientiert und setzt zudem eine bestimmte Ordnung M voraus, so lassen sich aus Gleichung 4.32 nun die bereits bekannten Grenzfrequenzen f_{Lim} wie folgt berechnen:

$$f_{\text{Lim}} = \frac{M c}{2 \pi R}. \quad (4.33)$$

Ein Vergleich von Gleichung 4.33 mit Gleichung 4.29 verdeutlicht, dass die approximative Bestimmung von f_{Lim} aus Gleichung 4.29 ihren Ursprung in der Betrachtung des *truncation error* hat. Für den oben geschilderten Fall, d.h. für $r = R = 0.089$ [m] und einer Ordnung von $M = 4$, ergibt sich mit Hilfe von Formel 4.33, d.h. bei einem *truncation error* von $e_M(kr) \approx 4\%$, eine Grenzfrequenz von $f_{\text{Lim}} = 2432$ [Hz]. Dies entspricht etwa dem in Tabelle 4.3 angegebenen Wert.

Neben den Übergangsfrequenzen f_{Lim} beinhaltet Tabelle 4.3 noch die Werte α_E . Dabei handelt es sich bei α_E um ein Maß für die *Lokalisationsunschärfe*, was auch als Qualitätsmaß für die Lokalisation angesehen werden kann, welches sich aus der maximalen Länge des Schallenergievektors $\mathbf{r}_e^{\text{Max}}$, deren Formel aus Tabelle 4.2 zu entnehmen ist, wie folgt berechnet:

$$\alpha_E = \arccos(\mathbf{r}_e^{\text{Max}}). \quad (4.34)$$

Setzt man beispielsweise für r_e^{Max} die Formel des $\text{Max}r_e$ Decoders für den 2D-Fall aus Tabelle 4.2 ein, so ergibt sich eine Lokalisationsunschärfe von:

$$\alpha_E = \frac{\pi}{(2M + 2)}. \quad (4.35)$$

Anders ausgedrückt α_E beschreibt den *localization blurr*, also die Abweichung vom tatsächlichen Ort der Quelle, der sich ergibt, wenn sich diese genau zwischen zwei Lautsprechern, die in einem Abstand von $\pm\alpha_E$ um die Quelle aufgestellt sind, befindet.

Weitere Decodervarianten

Neben den drei hier behandelten Decodervarianten existieren noch weitere. Dabei wären neben der in Kapitel 3.1.1 bereits vorgestellten Fenstermethode (*windows-applied*-Decoder) auch noch der in [Pol96] vorgestellte *Super-Minimal*-Decoder zu nennen. Für den 2D-Fall mit Ordnung $M = \frac{(L-1)}{2}$ berechnet sich dessen Richtdiagramm $\Psi_{\text{SuperMinimal}}^{2D}(\gamma_l, M)$, d.h. die Gewichte der L Lautsprecher gemäß [Son03] anhand der sogenannten ASINC¹⁴-Funktionen. Diese stellen nach Poletti einen allgemeinen Ansatz für die Bestimmung optimaler Codierfunktionen unabhängig von der gewählten Anzahl der Lautsprecher dar, solange sich diese symmetrisch entlang einer Kreislinie befinden, also eine regelmäßige Lautsprecheranordnung vorherrscht.

Die ASINC-Funktionen und somit die Gewichte der L Lautsprecher berechnen sich dabei wie folgt:

$$\text{ASINC}_L(\gamma) = \begin{cases} \frac{\sin\left(\frac{L\gamma}{2}\right)}{L \sin\left(\frac{\gamma}{2}\right)} = \text{CSINC}_L(\gamma) = \frac{2}{L} \sum_{l=1}^{\frac{(L-1)}{2}} \cos(l\gamma), & L = \textit{gerade} \\ \frac{\sin\left(\frac{(L-1)\gamma}{2}\right)}{L \sin\left(\frac{\gamma}{2}\right)} + \frac{\cos(L\gamma)}{L} = \text{CSINC}_{L-1}(\gamma) + \frac{\cos(L\gamma)}{L} \\ = \frac{1}{L} \left(1 + 2 \sum_{l=1}^{\frac{(L-1)}{2}} \cos(l\gamma) \right), & L = \textit{ungerade} \end{cases}$$

$$\text{CSINC} = \text{Zirkulare SINC-Funktion (engl. } \textit{Cyclic SINC}). \quad (4.36)$$

Abbildung 4.11 verdeutlicht eine Besonderheit dieses Decoders, nämlich, dass dessen Gewichte, die gemäß der ASINC-Funktionen verteilt sind, einem Amplituden-*panning* entsprechen, bei dem für den Fall, dass die Quellrichtung exakt mit dem Ort eines Lautsprechers l zusammenfällt, nur dieser und sonst keiner der verbleibenden Lautsprecher angesteuert wird und somit das Gewicht $g_l = 1$ erhält. In den übrigen Fällen tragen, wie schon bei den vorherigen Decodervarianten, stets alle Lautsprecher durch konstruktive- oder destruktive Interferenz, zur Bildung des Wellenfeldes bei, wobei sich deren Gewichte, unabhängig vom Winkel γ , stets zu $\sum_{l=1}^L g_l = 1$ summieren. Die Gewichte des zweidimensionalen *Super-Minimal*-Decoders ergeben sich dabei durch Aufsuchen des kleinsten, mitt-

¹⁴ASINC = engl. *Angular SINC*

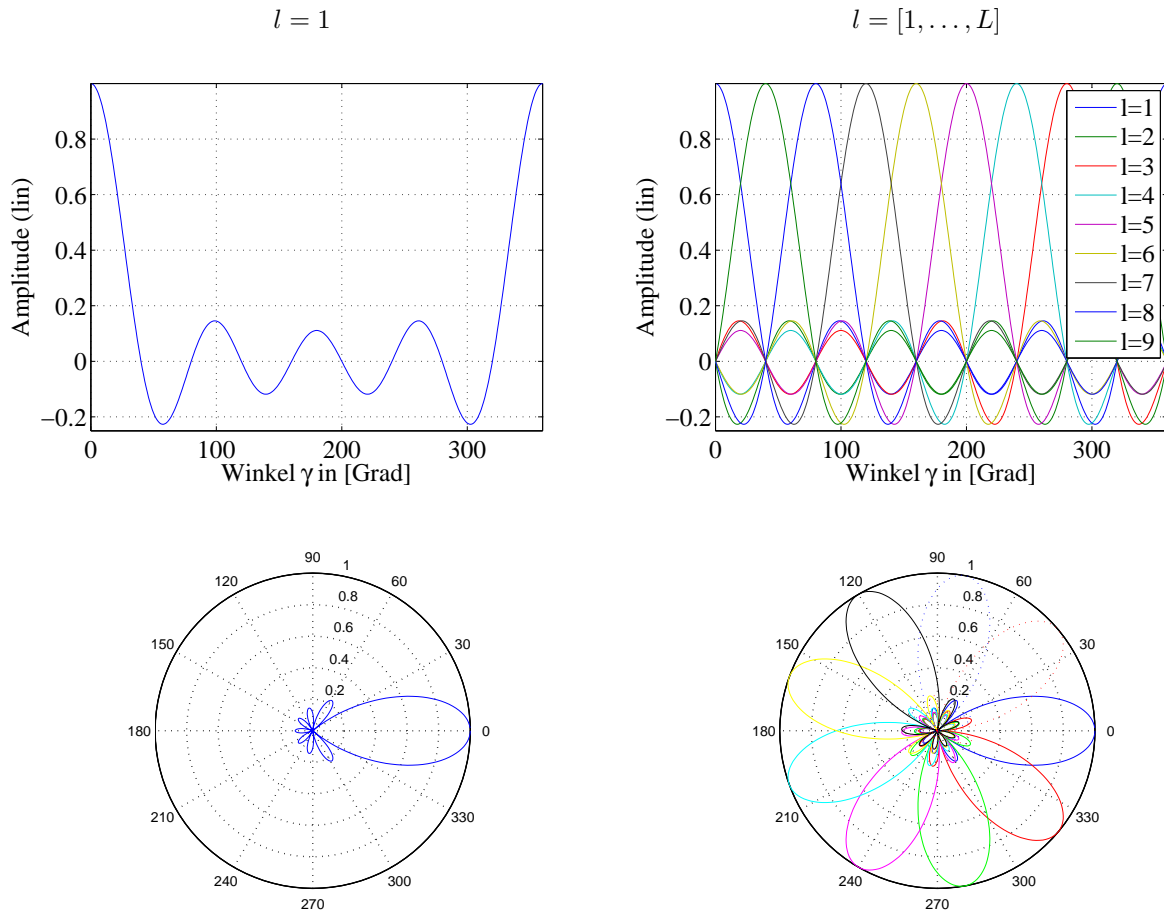


Abbildung 4.11: Richtdiagramme der virtuellen Mikrophone (panning-Funktionen) des Super-Minimal-Decoders für den zweidimensionalen Fall $\Psi_{SuperMinimal}^{2D}(\gamma_l, M)$, wobei von einem Schalleinfall bei $\mathbf{r}_{SRC} = (\Theta_{SRC} = 0^\circ, \varphi_{SRC} = 0^\circ)$ sowie von einer regulären Lautsprecheraufstellung mit $L = 9$ Lautsprechern ausgegangen wird, die sich an den Orten $\mathbf{r}_{S_i} = (\Theta_{S_i} = [0^\circ : 40^\circ : 320^\circ], \varphi_{SRC} = 0^\circ)$ befinden, was definitionsgemäß einer Ordnung von $M = \frac{(L-1)}{2} = 4$ entspricht.

leren, quadratischen Fehlers unter Verwendung eines LMS¹⁵-basierten Optimierungsalgorithmus zwischen dem Wellenfeld einer ebenen Welle entlang einer Kreislinie mit einem bestimmten Radius r um den Ursprung und dem resultierenden, überlagerten Wellenfeld der L Lautsprecher, die an N Punkten ausgewertet werden.

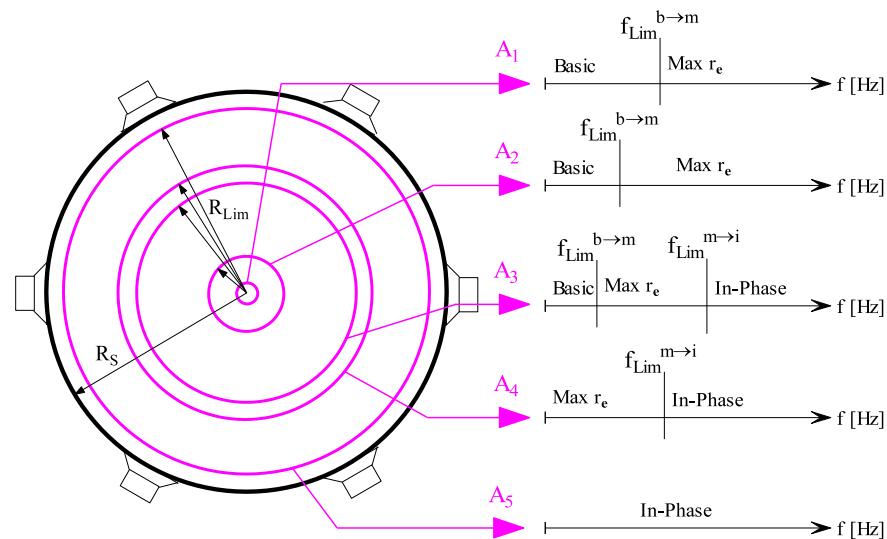


Abbildung 4.12: Bevorzugte Abhörregionen innerhalb des Wiedergabesystems, bei Verwendung unterschiedlicher Kombinationen der Decodervarianten *Basic*, *Max r_e* und *In-Phase*, über der Frequenz.

Abbildung 4.12 veranschaulicht den Zusammenhang zwischen den Decodervarianten *Basic*, *Max r_e* und *In-Phase* und der sich dabei ergebenden nutzbaren Abhörfläche, die sich gemäß der Definition von Ambisonicsystemen stets um das Zentrum aufspannt. Dabei wird deutlich, dass der *Basic*-Decoder in die kleinste bevorzugte Fläche A_1 und der *In-Phase*-Decoder in die größte Fläche A_5 resultiert, wobei sich die Fläche des *Max r_e* -Decoders dazwischen befindet. Konnten in Tabelle 4.3 noch Werte für die Übergangsfrequenzen f_{Lim} zwischen dem *Basic*- und dem *Max r_e* -Decoder für unterschiedliche Ordnungen M angegeben werden, die in Abbildung 4.12 mit $f_{Lim}^{b \rightarrow m}$ bezeichnet sind, so existieren derzeit noch keine Daten bzw. Gleichungen für die Übergangsfrequenzen zwischen dem *Max r_e* - und dem *In-Phase*-Decoder, welche in Abbildung 4.12 unter der Bezeichnung $f_{Lim}^{m \rightarrow i}$ auftauchen. Wie schon in Kapitel 3.1 angesprochen, können die frequenzabhängigen Gewichtungsfaktoren durch phasenkompensierte *shelving*-Filter, wie etwa in Abbildung 3.11 dargestellt, realisiert werden, wobei der Übergang möglichst glatt erfolgen sollte.

Abschließend veranschaulicht Abbildung 4.13 anhand der Darstellung synthetisierter Wellenfelder bei Verwendung von Ambisonicsystemen unterschiedlicher Ordnung ($M = [0, 1, 2, 3, 4, \infty]$), wie sich die nutzbare bzw. bevorzugte Abhörfläche über der Ordnung M verändert. Dabei wurde stets von einer Quelle bei

¹⁵LMS = engl. *Least Mean Square*

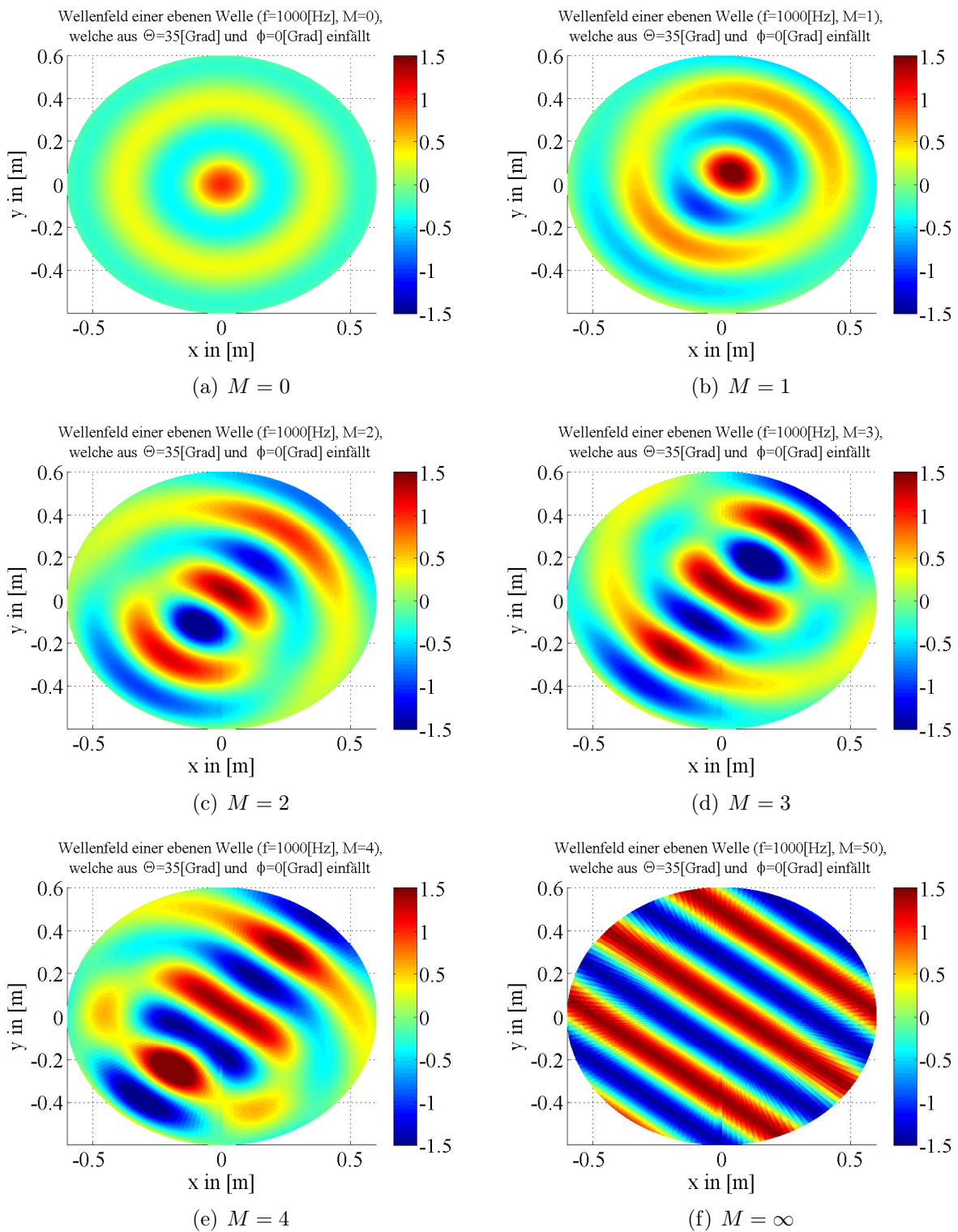


Abbildung 4.13: Rekonstruktion des Wellenfeldes einer ebenen Welle mit $f = 1 [\text{kHz}]$ aus $\Theta_{\text{SRC}} = 35^\circ$ und $\varphi_{\text{SRC}} = 0^\circ$ bis zu einer maximalen Ordnung von $M = [0, 1, 2, 3, 4, \infty]$.

$\mathbf{r}_{\text{SRC}} = (\Theta_{\text{SRC}} = 35^\circ, \varphi_{\text{SRC}} = 0^\circ)$ ausgegangen, die eine ebene Welle mit einer Frequenz von $f = 1$ [kHz] aussendet sowie von einem Lautsprecheraufstellradius von $R_S = 0.6$ [m] und der alleinigen Verwendung des *Basic*-Decoders. Idealerweise sollte das synthetisierte Wellenfeld, wie in Abbildung 4.13(f) dargestellt, d.h. bei Anwendung einer Ordnung von $M = \infty$, aussehen, was natürlich nicht möglich ist und hier nur zur Orientierung dienen soll. Das System 0. Ordnung kann dieses Optimum lediglich im Zentrum an einem infinitesimal kleinen Punkt erreichen und ist deshalb ausschließlich von theoretischer Bedeutung. Mit steigender Ordnung erkennt man, wie sich das synthetisierte Wellenfeld in einem immer größer werdendem Bereich um das Zentrum dem Optimum nähert, wodurch abermals die Bedeutung von Ambisonicsystemen höherer Ordnung unterstrichen wird. Während Systeme 1. Ordnung, die in der Vergangenheit ja von besonderem Interesse waren, bei der hier betrachteten Frequenz von $f = 1$ [kHz] imstande sind das Wellenfeld mit einem Radius von $r \approx 10$ [cm] in guter Näherung zu approximieren, erweitert sich dieser Bereich bei einer Ordnung von $M = 4$, was im späteren Verlauf für uns noch von besonderer Bedeutung sein wird, zu $r \approx 30$ [cm]. Dies ist bereits deutlich größer als der menschliche Kopf, wodurch selbst bei dieser Frequenz Kopfbewegungen erlaubt sind, ohne dabei einen falschen akustischen Eindruck zu erhalten.

Bezüglich der Höhe der Ordnung M wurde in [Far08] festgestellt, dass Quellen, wie sie in der Praxis vorkommen (wie z.B. Lautsprecher), sich bereits recht gut mit HOA-Systemen 2. Ordnung approximieren lassen. Systeme 3. Ordnung sind dagegen bereit imstande die Richtwirkung der menschlichen HRTF in ausreichender Genauigkeit nachzubilden. Somit kann festgehalten werden, dass zum Zwecke der akustischen Dokumentation ein HOA-System 4. Ordnung mehr als ausreichend sein dürfte.

4.2.6 Nahfeldeffekt bei der Decodierung

Der bereits aus der Codierung bekannte Nahfeldeffekt, welcher in Kapitel 4.2.2 in Gleichung 4.14 vorgestellt wurde, muss auch bei der Decodierung berücksichtigt werden. Dabei kann hier auf dessen Anwendung nicht verzichtet werden, da das verwendete Wiedergabesystem üblicherweise Lautsprecherabstände zur bevorzugten Abhörposition, wobei es sich bei HOA-Systemen bekanntermaßen um das Zentrum handelt, verwendet werden, die so gering sind, dass sich die jeweiligen Lautsprecher normalerweise in dessen Nahfeld befinden. Somit müssen die Wellen, die von den Lautsprechern ausgesendet werden, als Kugelwellen interpretiert werden und nicht, wie in ursprünglichen Ambisonicsystemen der Fall, als ebene Wellen.

Im Gegensatz zur Berücksichtigung des Nahfeldeffektes bei der Codierung, wobei instabile Filter für dessen Kompensation verwendet werden müssten, entstehen bei der Kompensation des Nahfeldeffektes bei der Decodierung M stabile und somit realisierbare Filter $\frac{1}{F_m^{(j\omega)}\left(\frac{r_S}{c}\right)}$, welche vom Abstand r_S vom Koordinatenursprung zu den regulär aufgestellten Lautsprechern abhängen.

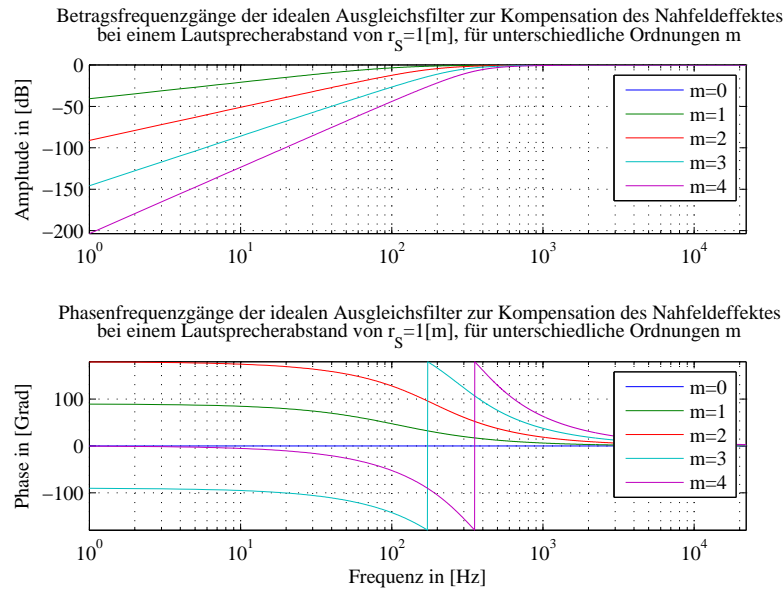


Abbildung 4.14: Bodediagramm der Ausgleichsfilter $\frac{1}{F_m\left(\frac{r_S}{c}\right)(j\omega)}$ zur Kompensation des Nahfeldeffektes bei der Decodierung für einen Lautsprecherabstand von $r_S = 1$ [m], für die 0.-4. Ordnung.

Abbildung 4.14 stellt diese beispielhaft für einen Lautsprecherabstand von $r_S = 1$ [m] vom Koordinatenursprung für die 0.-4. Ordnung dar. Wie man erkennen kann, weisen die resultierenden Filter $\frac{1}{F_m\left(\frac{r_S}{c}\right)(j\omega)}$ einen inversen Verlauf zu den Kompensationsfiltern der Codierung auf, was ein Vergleich mit Abbildung 4.6 verdeutlicht. Dadurch erklärt sich deren Stabilität und somit auch deren Umsetzbarkeit.

Die Decodiervorschrift nach Gleichung 4.22 wandelt sich unter Berücksichtigung von

Formel 4.23 sowie der Kompensationsfilter des Nahfeldeffektes zu:

$$\mathbf{Y}_{L \times 1} = \mathbf{D}_{L \times N} \mathbf{F}_m(k r_S)_{N \times N} \mathbf{B}_{N \times 1}, \text{ mit}$$

$$\mathbf{F}_m(k r_S)_{N \times N} = \text{diag} \left(\left[\begin{array}{c} \frac{1}{F_0^{(\frac{r_S}{c})}(j\omega)}, \dots, \overbrace{\frac{1}{F_m^{(\frac{r_S}{c})}(j\omega)}}^{(2m+1)}, \dots, \frac{1}{F_m^{(\frac{r_S}{c})}(j\omega)}, \dots, \\ \overbrace{\frac{1}{F_M^{(\frac{r_S}{c})}(j\omega)}}^{(2M+1)}, \dots, \frac{1}{F_M^{(\frac{r_S}{c})}(j\omega)} \end{array} \right]^T \right), \text{ wobei}$$

$$\mathbf{F}_m(k r_S)_{N \times N} = \text{Diagonale (quadratische) Matrix der Ordnung } N, \text{ welche für den}$$

allgemeinen 3D-Fall pro Ordnung m , mit $0 \leq m \leq M$, $(2m + 1)$

identische Kompensationsfilter $\frac{1}{F_m^{(\frac{r_S}{c})}(j\omega)}$ besitzt. (4.37)

Die Wirkung von Formel 4.37 ist dabei die Gleiche wie bei der Codierung, was durch Abbildung 4.7 veranschaulicht wurde, weshalb an dieser Stelle auf die Darstellung des sich ergebenden Wellenfeldes mit und ohne Verwendung der Kompensationsfilter $\frac{1}{F_m^{(\frac{r_S}{c})}(j\omega)}$ verzichtet wird.

In [Dan00] wurde zudem dazu übergegangen, den Nahfeldeffekt der Codierung und der Decodierung zu kombinieren, wodurch eine Vorschrift zur Codierung des Abstandes virtueller Quellen, dem sogenannten *distance coding* entstand, welches sich wie folgt durch das *distance coding*-Filter $H_m^{\text{NFC}}(\frac{r_{\text{SRC}}}{c}, \frac{r_S}{c})(j\omega)$ beschreiben lässt:

$$H_m^{\text{NFC}}(\frac{r_{\text{SRC}}}{c}, \frac{r_S}{c})(j\omega) = \frac{F_m^{(\frac{r_{\text{SRC}}}{c})}(j\omega)}{F_m^{(\frac{r_S}{c})}(j\omega)}, \text{ wobei}$$

$$H_m^{\text{NFC}}(\frac{r_{\text{SRC}}}{c}, \frac{r_S}{c})(j\omega) = \text{Distance coding-Filter, welches den Nahfeldeffekt des}$$

Wiedergabesystems, bei dem die L Lautsprecher regulär in

einem Abstand von r_S vom Zentrum entfernt angeordnet

sind, so kompensiert (NFC¹⁶), dass die

virtuelle Quelle (SRC¹⁷) in einem gewünschten Abstand von r_{SRC}

wahrgenommen wird. (4.38)

Mit Hilfe der *distance coding*-Filter nach Gleichung 4.38 lässt sich nun eine umsetzbare und somit praktikable Codiervorschrift gemäß Gleichung 4.14 aufstellen, in der lediglich

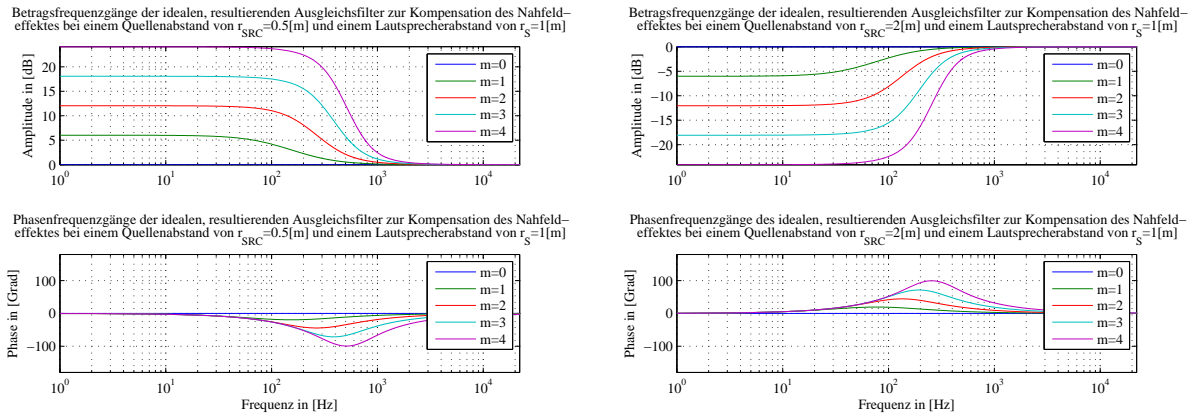
¹⁶NFC = engl *Near Field Coding*

¹⁷SRC = engl. *Source*

das Filter zur Kompensation des Nahfeldeffektes der Codierung $F_m^{(\frac{r_{\text{SRC}}}{c})}(j\omega)$ durch das *distance coding*-Filter $H_m^{\text{NFC}}(\frac{r_{\text{SRC}}}{c}, \frac{r_S}{c})(j\omega)$ zu ersetzen ist. Dadurch wird bereits bei der Erstellung der Codierung der Einfluss des Nahfeldeffektes der Decodierung berücksichtigt. Somit entsteht die nachfolgend aufgeführte, neue, realisierbare Codiervorschrift, welche in der Lage ist, den Nahfeldeffekt in der Praxis zu kompensieren:

$$\begin{aligned} \tilde{B}_{m,n}^{\sigma \text{NFC}}\left(\frac{r_S}{c}\right) &= S(j\omega) H_m^{\text{NFC}}\left(\frac{r_{\text{SRC}}}{c}, \frac{r_S}{c}\right)(j\omega) Y_{m,n}^{\sigma}(\Theta_{\text{SRC}}, \varphi_{\text{SRC}}) \\ &= \frac{1}{F_m^{(\frac{r_S}{c})}(j\omega)} B_{m,n}^{\sigma}, \text{ mit } B_{m,n}^{\sigma} \text{ nach Gleichung 4.14 und} \\ \tilde{B}_{m,n}^{\sigma \text{NFC}}\left(\frac{r_S}{c}\right) &= \text{Nahfeldkompensierte Ambisonicsignale, die zusätzlich vom} \\ &\quad \text{Lautsprecherradius } r_S \text{ abhängig sind.} \end{aligned} \quad (4.39)$$

Das *distance coding*-Filter $H_m^{\text{NFC}}(\frac{r_{\text{SRC}}}{c}, \frac{r_S}{c})(j\omega)$ ist dabei sowohl vom Lautsprecher-radius r_S als auch vom gewünschten Abstand der virtuellen Quelle r_{SRC} abhängig. Je nach Verhältnis zwischen r_S und r_{SRC} nimmt $H_m^{\text{NFC}}(\frac{r_{\text{SRC}}}{c}, \frac{r_S}{c})(j\omega)$ dabei für innenliegende Quellen, d.h. für $r_{\text{SRC}} < r_S$, die Form eines anhebenden- bzw. für außenliegende Quellen, also für $r_{\text{SRC}} > r_S$, die eines absenkenden Tiefen-*shelving*-Filters der Ordnung $m \times 20 \log_{10}\left(\frac{r_S}{r_{\text{SRC}}}\right)$ [dB] an, was durch Abbildung 4.15 veranschaulicht wird.



(a) *Distance coding*-Filter $H_m^{\text{NFC}}(\frac{r_{\text{SRC}}}{c}, \frac{r_S}{c})(j\omega)$ der Ordnungen $m = [0, \dots, 4]$ für eine innenliegende Quelle mit $r_{\text{SRC}} = 0.5$ [m] und $r_S = 1.0$ [m].

(b) *Distance coding*-Filter $H_m^{\text{NFC}}(\frac{r_{\text{SRC}}}{c}, \frac{r_S}{c})(j\omega)$ der Ordnungen $m = [0, \dots, 4]$ für eine außenliegende Quelle mit $r_{\text{SRC}} = 2.0$ [m] und $r_S = 1.0$ [m].

Abbildung 4.15: *Distance coding*-Filter $H_m^{\text{NFC}}(\frac{r_{\text{SRC}}}{c}, \frac{r_S}{c})(j\omega)$ der Ordnungen $m = [0, \dots, 4]$ für innen- und außenliegende Quellen.

Bei der neuen Codiervorschrift nach Gleichung 4.39, die auch als *NFC-HOA*¹⁸, bezeich-

¹⁸NFC-HOA = engl. *Near Field Compensated - Higher Order Ambisonics*

net wird, ist es notwendig, bereits bei der Codierung zumindest den Aufstellradius des Wiedergabesystems r_S zu kennen, wobei die dadurch bedingte Signallaufzeit $\frac{r_S}{c}$ als Referenzparameter bei der Codierung wirkt. Ist diese jedoch zum Zeitpunkt der Codierung noch nicht bekannt, kann die neue Codiervorschrift dennoch verwendet werden. Dazu ist lediglich ein beliebiger „Referenzlautsprecherradius“ $r_{S_{\text{Ref}}}$ bei der Codierung zu verwenden, welcher bei der Decodierung durch Einbringen eines weiteren *distance coding*-Filters $H_m^{\text{NFC}}\left(\frac{r_{S_{\text{Ref}}}}{c}, \frac{r_S}{c}\right)(j\omega)$ kompensiert werden kann. Dieser vollzieht eine Transformation von $r_{S_{\text{Ref}}}$ auf r_S .

$$\begin{aligned} \tilde{B}_{m,n}^{\sigma \text{NFC}}\left(\frac{r_S}{c}\right) &= H_m^{\text{NFC}}\left(\frac{r_{S_{\text{Ref}}}}{c}, \frac{r_S}{c}\right)(j\omega) \tilde{B}_{m,n}^{\sigma \text{NFC}}\left(\frac{r_{S_{\text{Ref}}}}{c}\right), \text{ wobei} \\ H_m^{\text{NFC}}\left(\frac{r_{S_{\text{Ref}}}}{c}, \frac{r_S}{c}\right)(j\omega) &= \text{Distance coding-Filter zur Transformation der Nahfeld-} \\ &\text{kompensation vom Referenzlautsprecherradius } r_{S_{\text{Ref}}} \text{ auf} \\ &\text{den tatsächlichen Lautsprecherradius } r_S. \end{aligned} \quad (4.40)$$

Dadurch kann wieder die für Ambisonicsysteme so charakteristische und vorteilhafte Eigenschaft der Unabhängigkeit zwischen der Codierung und der Decodierung hergestellt werden. Mit Hilfe der Nahfeldkompensation bzw. der *distance coding*-Filter gelingt es, einen weiteren beschränkenden Faktor des ursprünglichen Ambisonicsystems nach Gerzon nämlich den Bühnenabstand bzw. die Distanz zum Klanggeschehen beim Aufnahme/Wiedergabe System zu berücksichtigen bzw. zu verallgemeinern. Die Notwendigkeit dabei reguläre Lautsprecheraufstellungen sowohl für den 2D- als auch für den 3D-Fall verwenden zu müssen, verbleibt jedoch und ist auch nicht trivial zu lösen, wobei diese Bedingung zumindest zum Zwecke der hier behandelten Auralisierung keine all zu große Einschränkung darstellt und somit in dieser Arbeit nicht weiter behandelt wird. In die Codierung lassen sich neben den bereits beschriebenen Methoden auch noch weitere, das Klangfeld gezielt verändernde Operationen, wie etwa die *Rotation*, die *Fokussierung*, die auch als *Zoom* bezeichnet wird, das *Spiegeln*, was auch unter dem Begriff *W-panning* bekannt ist sowie einige andere Kombinationen daraus einbringen. Da es jedoch nicht Ziel dieser Arbeit ist, künstliche Klangfelder zu erzeugen, sondern bestehende in einem beliebigen Raum so gut wie möglich nachzubilden, wird an dieser Stelle auf eine weitere Erörterung dieser Operationen verzichtet. Der interessierte Leser sei jedoch an folgende Arbeiten verwiesen: [Dan00, Wig04, Hol06]. Ein weiterer Vorteil dieser Codierung ist, dass sie nicht nur auf virtuelle Quellen beschränkt ist, sondern auch auf jedes natürliche Klangfeld, welches z.B. mit Hilfe von Mikrofonarrays aufgezeichnet wurde, angewendet werden kann. Wie sich im nachfolgendem Kapitel noch zeigen wird, ergeben sich durch Anwendung des oben beschriebenen Prinzips der Nahfeldkompensation auch bei der Aufzeichnung mit Hilfe sphärischer Mikrofonarrays realisierbare Kompensationsfilter.

4.3 Zusammenfassung

In diesem Kapitel wurde erklärt, wie ein Wellenfeld mit Hilfe von sphärischen Harmonischen, welche die Basisfunktionen eines Ambisonicsystems bilden, zerlegt bzw. mit Hilfe der Fourier-Bessel-Reihe synthetisiert werden kann. Neben der Erläuterung der Wirkungsweise der Ambisonickomponenten $B_{m,n}^\sigma$ wurde in diesem Abschnitt zudem die Codierung und Decodierung eines Ambisonicsystems für höhere Ordnungen verallgemeinert. Zudem wurde die Ambisonictheorie durch Berücksichtigung des Nahfeldeffektes mit Hilfe der Nahfeldmodellierungsfunktion $F_m^{\frac{r}{c}}(j\omega)$ nach Daniel erweitert. Mit deren Hilfe kann nun der von Gerzon als *Bass-Boost* bezeichnete Effekt in der Codierung und Decodierung entsprechend berücksichtigt werden. Somit lassen sich nun auch Nahfeldbedingungen wie etwa innenliegende Quellen korrekt reproduzieren. Aus diesem Grund ist es mit Hilfe der NFC-HOA nun auch möglich, unterschiedliche Bühnenabstände zu codieren. Zudem wurde in diesem Kapitel die Unabhängigkeit zwischen der Codierung und der Decodierung von Ambisonicsystemen bewiesen. Ferner wurden unterschiedliche Decodervarianten vorgestellt und deren Vor- und Nachteile diskutiert. Außerdem wurde gezeigt, dass sich *panning*-Funktionen auch als Richtdiagramme virtueller Mikrophone interpretieren lassen, was zu einer anschaulichen und somit einfachen Beurteilung der kombinierten Codierung und Decodierung führt. Zudem wurde erörtert, in welchen Spektralbereichen welche Decodervarianten sinnvoll sind und wie man dabei die Übergangsfrequenzen bestimmen kann. Auch eine Berechnungsvorschrift für die sich ergebende Lokalisationsunschärfe α_E , welche von der Ordnung M abhängt, wurde aufgezeigt. Es wurde festgestellt, dass höhere Ordnungen M nicht den räumlichen Gültigkeitsbereich, welcher sich um das Zentrum ($r = 0$) aufspannt, erhöht, sondern auch die Lokalisationsunschärfe α_E reduziert sowie den spektralen Einsatzbereich nach oben hin erweitert. Die Beschreibung der Wirkungsweise des *distance coding* beendet dieses Kapitel.

Kapitel 5

HOA-Codierung natürlicher Klangfelder mit Hilfe von Mikrophonarrays

5.1 Allgemeine Fourier-Bessel-Reihe

Bevor mit der Beschreibung eines Systems, mit dem Ambisonickomponenten höherer Ordnung $B_{m,n}^\sigma$ gemessen werden können, begonnen werden kann, gilt es zunächst sich abermals mit akustischen Grundlagen zur Herleitung eines Schalldruckfeldes $p(\mathbf{r}, t)$ zu beschäftigen. In Kapitel 4.1 wurde bereits auf die Herleitung der *Fourier-Bessel-Reihe* eingegangen, die im Wesentlichen die Zerlegung eines Wellenfeldes an einem beliebigen Beobachtungs- bzw. Messpunkt \mathbf{r} im Raum beschreibt und zu der in Gleichung 4.4 aufgezeigten Lösung führte.

Dabei wurde davon ausgegangen, dass es keine innenliegenden Quellen gab, d.h. $R_1 = 0$ war, deren Anteile sich von innen nach außen bewegten, sondern sich das Wellenfeld am Messpunkt \mathbf{r} ausschließlich anhand der Überlagerung der Anteile der außenliegenden Quellen, welche den gesamten Raum durchdringen, ergab. Im Allgemeinen können jedoch, wie in Abbildung 5.1 dargestellt, neben den außenliegenden Quellen ($r > R_2$) auch noch innenliegende Quellen ($r < R_1$) bei der Gestaltung des resultierenden Wellenfeldes am Messpunkt \mathbf{r} beteiligt sein, deren Anteile bedingt durch den üblicherweise geringen Abstand zum Messpunkt \mathbf{r} als Kugelwellen aufgefasst werden müssen. Dies führt schließlich zu der *allgemeinen Fourier-Bessel-Reihe*, welche das Wellenfeld im Freifeld, das durch die

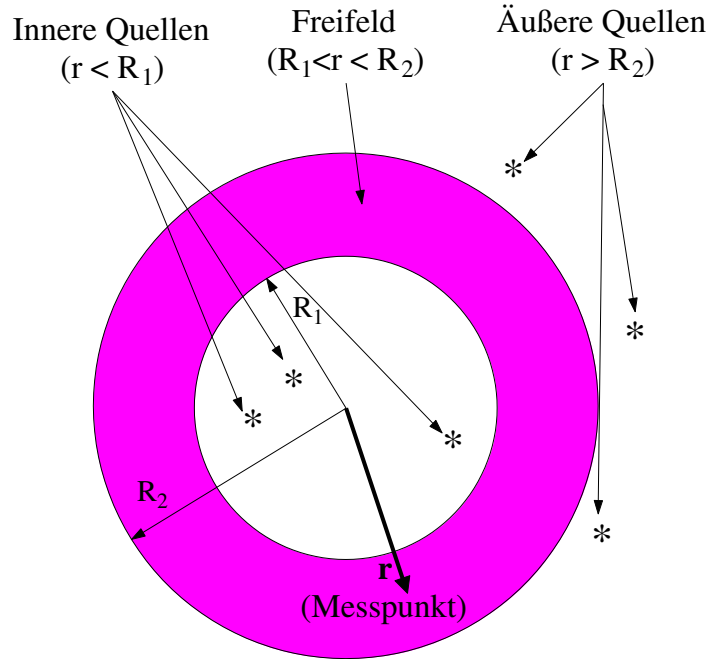


Abbildung 5.1: Messung des Schalldruckfeldes $p(\mathbf{r}, t)$ am Messpunkt \mathbf{r} , wobei im Raum eine allgemeine Freifeldbedingung vorherrschen soll, d.h. es existieren sowohl außen- als auch innenliegende Quellen, deren Anteile sich entsprechen am Beobachtungspunkt \mathbf{r} überlagern.

Kugelschale mit dem Radius $R_1 < r < R_2$ definiert ist, beschreibt.

$$P(\mathbf{r}, \omega) = S(j\omega) \left(\sum_{m=0}^{\infty} j^m j_m(kr) \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^{\sigma} Y_{m,n}^{\sigma}(\Theta, \varphi) + \sum_{m=0}^{\infty} j^m h_m^{(2)}(kr) \sum_{0 \leq n \leq m, \sigma = \pm 1} A_{m,n}^{\sigma} Y_{m,n}^{\sigma}(\Theta, \varphi) \right), \text{ mit}$$

$B_{m,n}^{\sigma}$ = Gewichtungskoeffizienten (Ambisonikkoeffizienten), welche zu der sphärischen Besselfunktion 1. Art $j_m(kr)$ gehören und das durchdringende Wellenfeld (ebene Wellen) beschreiben,

$A_{m,n}^{\sigma}$ = Gewichtungskoeffizienten, welche zu der sphärischen Hankelfunktion 2. Art $h_m^{(2)}(kr)$ gehören und das eintretende Wellenfeld (Kugelwellen) beschreiben. (5.1)

Der 1. Teil der allgemeinen Fourier-Bessel-Reihe nach Gleichung 5.1 gegeben durch die Ambisonikkomponenten $B_{m,n}^{\sigma}$, welche die Gewichte für die sphärischen Harmonischen $Y_{m,n}^{\sigma}(\Theta, \varphi)$ darstellen sowie der dazugehörigen sphärischen Besselfunktionen 1. Art $j_m(kr)$, welche die radiale Abhängigkeit repräsentieren, beschreibt das durchdringende Wellenfeld

der außenliegenden Quellen, die hier als Lieferanten ebener Wellen betrachtet werden. Der 2. Teil von Gleichung 5.1 beschreibt hingegen anhand der Gewichtungsfaktoren $A_{m,n}^\sigma$ für die Winkelfunktionen $Y_{m,n}^\sigma(\Theta, \varphi)$ sowie der dazugehörigen sphärischen Hankelfunktionen 2. Art $h_m^{(2)}(kr)$, welche ebenfalls die radiale Abhängigkeit repräsentieren, das eintretende Wellenfeld der innenliegenden Quellen, von denen Kugelwellen ausgesendet werden.

An dieser Stelle sei angemerkt, dass es sich beim Ambisonicverfahren um ein System handelt, welches vom Zentrum aus betrachtet wird, wodurch definitionsgemäß $R_1 = 0$ gilt. Deshalb ist für Ambisonicsysteme die Beschreibung des Schalldruckfeldes $p(\mathbf{r}, t)$ am Beobachtungspunkt \mathbf{r} nach Gleichung 4.4 für ebene Wellen bzw. nach Gleichung 4.13 für Kugelwellen vollständig. Anders ausgedrückt, durch die zentrale Betrachtungsweise des Ambisonicverfahrens werden die Gewichtungsfaktoren $A_{m,n}^\sigma$ stets zu 0.

Der hier vorgestellte, erweiterte Ansatz dient hingegen lediglich als Vorbereitung für die nachfolgenden Herleitungen der Formeln zur Messung der HOA-Komponenten mit Hilfe von Mikrophonarrays.

5.2 Mikrophonarrays

Die Untersuchung von unterschiedlichen Mikrophonarrays bezüglich ihrer Tauglichkeit zur Messung von Wellenfeldern bzw. deren Parametern war und ist nach wie vor ein reges Forschungsgebiet. In [Hul04] wurden etwa lineare, kreuzförmige und zirkulare Mikrophonarrays, bestückt mit Schalldruck- bzw. mit Schalldruck- und Schallschnellempfänger, auf ihre Verwendbarkeit zur Auralisation unter Verwendung der Wellenfeldsynthese (WFS¹) für ein zweidimensionales System untersucht. Die dabei erzielten Ergebnisse lassen sich wie folgt zusammenfassen:

- Lineare Arrays, welche alleinig mit Schalldruck- bzw. Schallschnellempfänger bestückt sind, können nicht für Auralisationszwecke eingesetzt werden, da mit ihnen keine Unterscheidung zwischen Wellenfeldern, die von vorne oder von hinten auf das Array eintreffen, vollzogen werden kann.
- Lineare Arrays, welche sowohl Schalldruck- als auch Schallschnellempfänger besitzen, sind in der Lage, zwischen von vorne oder von hinten ankommenden Wellenfeldern zu unterscheiden. Allerdings sind auch sie nicht in der Lage Wellenfelder aufzuzeichnen, die in ihrer *endfire*²-Ausrichtung auf das Array auftreffen. Somit können lineare Arrays lediglich in ihrer *broadside*³-Ausrichtung die gewünschten Ergebnisse liefern. Aus diesem Grund muss auch diese Variante für die Auralisation als unbrauchbar bezeichnet werden.
- Kreuzförmige Arrays, bestehend aus zwei senkrecht zueinander angebrachten linearen Arrays, die jeweils mit Schalldruck- und Schallschnellempfänger bestückt sind,

¹WFS = engl. *Wave Field Synthesis* (Wellenfeldsynthese)

können das zuvor geschilderte Problem der vom Einfallswinkel abhängigen Winkelauflösung, wenn das Array zunehmend in *endfire*-Orientierung zum einfallenden Wellenfeld steht, beheben. Allerdings müssen dann vom Einfallswinkel des Wellenfeldes abhängige Gewichtungsfaktoren bei der Überlagerung der beiden linearen Arrays eingesetzt werden, da es ansonsten bei einem Einfallswinkel z.B. von $\Theta_{\text{SRC}} = 45^\circ$ zu einer Verdoppelung des empfangenen Signals kommt, da in diesem Fall beide linearen Arrays zu jeweils gleichen Teilen einen Signalanteil beisteuern. Zudem treten an den Enden des kreuzförmigen Arrays Artefakte auf, die theoretisch vermieden werden könnten, wenn das Array unendlich groß wäre. Allerdings lassen sich diese Artefakte durch eine räumliche Fensterung, d.h. durch eine unterschiedliche Gewichtung der Mikrophone z.B. gemäß eines Hanning-Fensters reduzieren, wobei sich dann aber die Größe des synthetisierten Wellenfeldes verkleinert. Zusammenfassend kann somit gesagt werden, dass kreuzförmige Arrays im Prinzip für eine 2D-Auralisation verwendet werden können, wenn entsprechende Maßnahmen zur Vermeidung der ansonsten auftretenden Artefakte getätigt werden.

- Zirkulare Arrays, die mit Schalldruck- und Schallschnellempfänger bestückt sind, haben im Vergleich zu kreuzförmigen Arrays keine Enden, weshalb es bei deren Verwendung zu keiner Ausbildung von Artefakten kommt. Somit kann das Wellenfeld, welches von dem kreisrunden Array eingeschlossen wird, auch vollständig rekonstruiert werden. Außerhalb des Mikrophonarrays ist das Wellenfeld jedoch bei Verwendung der WFS welche auf dem *Kirchhoff-Helmholtz-Integral* beruht, gleich Null. Somit erklärt sich, warum ein WFS-basiertes Auralisationssystem üblicherweise über ein zirkulares Mikrophonarray mit großem Radius verfügt, welches idealerweise genauso groß wie der Radius des Abhörsystems, d.h. des Lautsprecherrings, ist. Hierin ist auch begründet, warum WFS-basierte Auralisationssysteme heutzutage im Gegensatz zu früheren Anordnungen oft eine zirkulare Form aufweisen, obwohl bedingt durch das Kirchhoff-Helmholtz-Integral diese im Gegensatz zu Ambisonicsystemen nicht zwingend vorgeschrieben wäre.

Im Prinzip stellt also ein zirkulares Array für den 2D-Fall ein geeignetes Instrument zur Messung des Wellenfeldes dar. Leider handelt es sich bei der Aufnahme eines realen Wellenfeldes nicht um ein zwei- sondern um ein dreidimensionales Wellenfeld. Dessen vertikal verlaufenden Anteile stören gemäß [Pol05a] die Aufzeichnung der gesuchten zirkularen Harmonischen, da das zirkulare Mikrophonarray, insbesondere bei tiefen Frequenzen, ungewollte, vertikale Nebenkeulen aufweist, mit denen diese Anteile des Wellenfeldes zwangsläufig mit gemessen werden, was letztendlich das Ergebnis verfälscht.

²Ein lineares Mikrophonarray, welches auf dem Empfang von Wellen, welche parallel zu der Linie auf der die Mikrophone aufgereiht sind optimiert wurde, bezeichnet man als *endfire*-Array.

³Lineare Mikrophonarrays, welche für den Empfang von Wellen, die senkrecht zu der Achse auf welcher die Mikrophone angebracht sind, optimiert wurden, bezeichnet man als *broadside*-Arrays.

Um dieses Problem zu lösen, wurde in [ME08] ein zirkulares Array vorgestellt, welches in der Mitte ein zusätzliches Mikrofon positioniert hat, mit dessen Hilfe sich die vertikalen Komponenten des Wellenfeldes, also diejenigen Anteile des Wellenfeldes, die senkrecht auf das zirkulare Array auftreffen, bis zu einer Ordnung von $M = 2$ kompensieren lassen. In [BA05] wurde dieser Ansatz verallgemeinert in dem anstelle eines einzelnen, zentral positionierten Mikrophons ein zweites, zirkulares Mikrofonarray zentral platziert wurde. Je nach Größe (mechanische Abmaße) des zusätzlichen zirkularen Arrays sowie dessen Anzahl an Mikrofonen können die Fehler, verursacht durch das Wellenfeld, welches die zirkularen Mikrofonarrays vertikal durchdringt, bis zu einer beliebigen Ordnung kompensiert werden.

Eine andere Möglichkeit das Problem der Vieldeutigkeit zirkularer Mikrofonarrays bei Vermessung dreidimensionaler Wellenfelder zu lösen, bietet die Verwendung von sphärischen, d.h. kugelförmigen Mikrofonarrays. Diese sind in der Lage nicht nur die zirkularen, d.h. die zwei-, sondern die sphärischen, also die dreidimensionalen Harmonischen messtechnisch zu erfassen, wodurch Vieldeutigkeiten erst gar nicht auftreten können. Ein weiterer Vorteil bei Verwendung eines kugelförmigen Mikrofonarrays besteht darin, dass die maximale Flexibilität beim Wiedergabesystem erhalten bleibt, d.h. die Lautsprecheranordnung kann sowohl zwei- als auch dreidimensional sein. Ein Nachteil besteht jedoch darin, dass die Anzahl Q der benötigten Mikrophone, für die gilt:

$$\begin{aligned}
 Q &\geq N \begin{cases} N = (2M + 1) & \text{für 2D-Fall} \\ N = (M + 1)^2 & \text{für 3D-Fall} \end{cases}, \text{ wobei} \\
 Q &= \text{Anzahl der Mikrophone,}
 \end{aligned} \tag{5.2}$$

im 3D-Fall wesentlich höher ausfällt, um die entsprechenden harmonischen Komponenten bis zur gewünschten Ordnung M zu bestimmen, als im 2D-Fall. Für $M = 4$ würden beispielsweise im 2D-Fall lediglich $Q \geq (2M + 1) = 9$ Mikrophone genügen, um die zirkularen Harmonischen bis zur Ordnung $M = 4$ vollständig zu bestimmen, während im 3D-Fall hierzu bereits $Q \geq (M + 1)^2 = 25$ Mikrophone benötigt werden, um die sphärischen Harmonischen bis $M = 4$ zu messen.

5.2.1 Kugelförmige Mikrofonarrays

Allgemein lassen sich die Schalldrücke mit Hilfe der Fourier-Bessel-Reihe bis zu einer Ordnung von M anhand von Q Mikrofonen, welche an den Positionen (Θ_q, φ_q, a) , mit $[1 \leq q \leq Q]$ auf einer Kugel mit einem Radius a angebracht sind, wie folgt beschreiben

([MD04]):

$$\begin{aligned}
 p_a(\Theta_q, \varphi_q) &= \sum_{m=0}^M W_m(ka) \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^\sigma Y_{m,n}^\sigma(\Theta_q, \varphi_q), \text{ wobei} \\
 p_a(\Theta_q, \varphi_q) &= \text{Schalldruck auf einer Kugel mit Radius } a \text{ an der Position } (\Theta_q, \varphi_q), \\
 &\text{an welcher sich das } q. \text{ Mikrophon befindet,} \\
 W_m(ka) &= \text{Radiale Funktion}^4 \text{ } m. \text{ Ordnung des Arguments } ka, \text{ wobei } k \text{ die} \\
 &\text{Wellenzahl und } a \text{ den Radius vom Mittelpunkt zur Kugeloberfläche,} \\
 &\text{also den Kugelradius } (r = a) \text{ bezeichnen.} \tag{5.3}
 \end{aligned}$$

Oder in Matrixschreibweise:

$$\begin{aligned}
 \begin{pmatrix} p_a(\Theta_1, \varphi_1) \\ \vdots \\ p_a(\Theta_q, \varphi_q) \\ \vdots \\ p_a(\Theta_Q, \varphi_Q) \end{pmatrix} &= \begin{pmatrix} Y_{0,0}^{+1}(\Theta_1, \varphi_1) & Y_{1,0}^{+1}(\Theta_1, \varphi_1) & \dots & Y_{m,n}^\sigma(\Theta_1, \varphi_1) \\ \vdots & \vdots & \vdots & \vdots \\ Y_{0,0}^{+1}(\Theta_q, \varphi_q) & Y_{1,0}^{+1}(\Theta_q, \varphi_q) & \dots & Y_{m,n}^\sigma(\Theta_q, \varphi_q) \\ \vdots & \vdots & \vdots & \vdots \\ Y_{0,0}^{+1}(\Theta_Q, \varphi_Q) & Y_{1,0}^{+1}(\Theta_Q, \varphi_Q) & \dots & Y_{m,n}^\sigma(\Theta_Q, \varphi_Q) \end{pmatrix} \times \\
 &\begin{pmatrix} W_0(ka) & 0 & \dots & 0 \\ 0 & W_1(ka) & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & W_M(ka) \end{pmatrix} \begin{pmatrix} B_{0,0}^{+1} \\ B_{1,0}^{+1} \\ \vdots \\ B_{m,n}^\sigma \end{pmatrix}, \tag{5.4}
 \end{aligned}$$

bzw. in Kurzform:

$$\begin{aligned}
 \mathbf{p}_{a_{Q \times 1}} &= \mathbf{Y}_{Q \times N} \mathbf{W}_{N \times N} \mathbf{B}_{N \times 1}, \text{ mit} \\
 \mathbf{W}_{N \times N} &= \text{diag}(\mathbf{W}_m(ka)) \\
 &= \text{diag} \left(\left[\begin{array}{c} W_0(ka), \dots, \overbrace{W_m(ka), \dots, W_m(ka)}^{(2m+1)}, \dots, \\ \overbrace{W_M(ka), \dots, W_M(ka)}^{(2M+1)} \end{array} \right]^T \right). \tag{5.5}
 \end{aligned}$$

Durch Umformung von Formel 5.5 können somit die Ambisonickomponenten bis zur

⁴Eine Funktion, deren Argument nur von einer Größenordnung abhängt, nennt man radial.

Ordnung M wie folgt aus den Q Mikrophonsignalen berechnet werden:

$$\begin{aligned}
 \mathbf{B} &= \mathbf{W}^{-1} (\mathbf{Y}^T \mathbf{Y})^{-1} \mathbf{Y}^T \mathbf{p}_a, \text{ mit} \\
 \mathbf{Y}^+ &= (\mathbf{Y}^T \mathbf{Y})^{-1} \mathbf{Y}^T = \text{Pseudoinverse von } \mathbf{Y}, \text{ folgt:} \\
 \mathbf{B} &= \text{diag}(\mathbf{W}_m^{-1}(ka)) \mathbf{Y}^+ \mathbf{p}_a, \text{ wobei} \\
 \text{diag}(\mathbf{W}_m^{-1}(ka)) &= \text{diag}(\mathbf{EQ}_m(ka)) = \text{diag} \left(\left[\begin{array}{cccc} 1 & & & \\ & \overbrace{1 \cdots 1}^{(2m+1)} & & \\ & & \overbrace{1 \cdots 1}^{(2m+1)} & \\ & & & \overbrace{1 \cdots 1}^{(2m+1)} \end{array} \right] \right) \text{ und} \\
 &\quad \left(\left[\begin{array}{cccc} & & & \\ & \overbrace{1 \cdots 1}^{(2M+1)} & & \\ & & \overbrace{1 \cdots 1}^{(2M+1)} & \\ & & & \overbrace{1 \cdots 1}^{(2M+1)} \end{array} \right] \right) \text{ und} \\
 \text{diag}(\mathbf{EQ}_m(ka)) &= \text{Diagonale Matrix, welche die radialen } \textit{equalizing}\text{-Funktionen} \\
 &\quad \text{EQ}_m(ka), \text{ mit } 0 \leq m \leq M \text{ enth\u{a}lt.} \tag{5.6}
 \end{aligned}$$

Graphisch l\u00e4sst sich die Messung der N Ambisonickomponenten der Ordnung M , wie in Abbildung 5.2 aufgezeigt, darstellen.

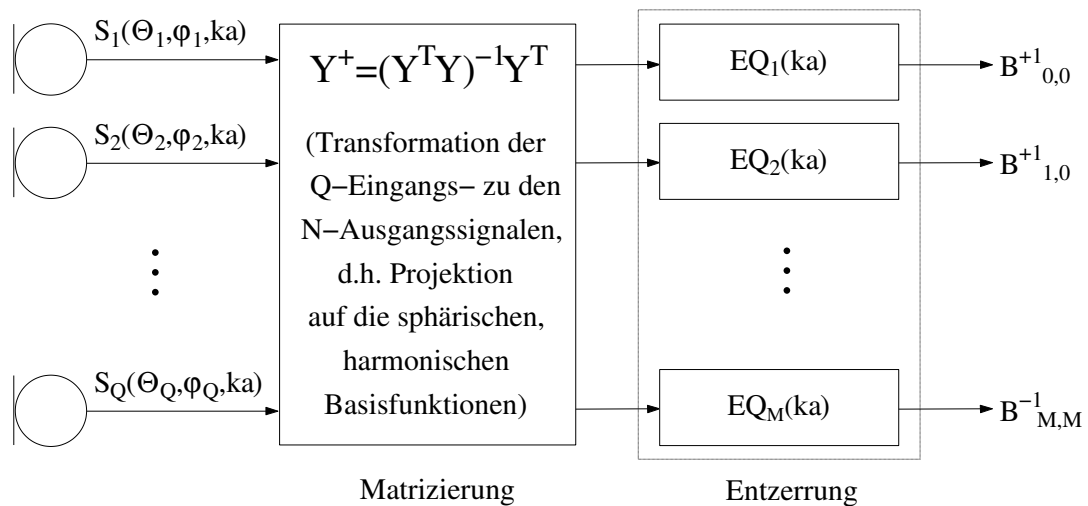


Abbildung 5.2: Vorgehensweise bei der Messung der N Ambisonickomponenten $B_{m,n}^\sigma$ bis zur M . Ordnung.

Der Raum und somit auch die sph\u00e4rischen Harmonischen werden durch die Q Mikrophone an den Stellen (Θ_q, φ_q) , mit $[1 \leq q \leq Q]$ abgetastet, wobei auch hier die *Orthonor-*

malitätsbedingung für sphärische Harmonische gegeben durch:

$$\begin{aligned} \frac{1}{Q} Y_{m,n}^{\sigma}(\Theta_q, \varphi_q)^T Y_{m',n'}^{\sigma'}(\Theta_q, \varphi_q) &= \delta_{m,m'} \delta_{n,n'} \delta_{\sigma,\sigma'}, \text{ mit} \\ 0 \leq m, m' \leq M, 0 \leq n, n' \leq M, \sigma, \sigma' = \pm 1, \text{ und} \\ \delta_{i,j} &= \begin{cases} 1 & \text{wenn } i = j \\ 0 & \text{wenn } i \neq j \end{cases}, \text{ wobei} \\ \delta_{i,j} &= \text{Kronecker-Delta-Funktion,} \end{aligned} \tag{5.7}$$

eingehalten werden muss, um die N Ambisonickomponenten der gewünschten Ordnung M bestimmen zu können. Die benötigte, minimale Anzahl von Q Mikrofonen kann dabei mit Hilfe von Gleichung 5.2 bestimmt werden.

In der Praxis sollten jedoch mehr Mikrophone, als durch Gleichung 5.2 ermittelt, verwendet werden, so dass stets gelten sollte: $Q > N$. Damit verschiebt sich die Eckfrequenz f_{Alias} , ab welcher das räumliche Aliasing gegeben durch:

$$\begin{aligned} f_{\text{Alias}} &= \frac{c}{2 d_{\text{Mic}}} = \frac{c}{2 a \gamma_{\text{Mic}}}, \text{ wobei} \\ f_{\text{Alias}} &= \text{Räumliche Aliasingfrequenz in [Hz],} \\ \gamma_{\text{Mic}} &= \text{Winkeldifferenz zwischen zwei benachbarten Mikrofonen in [rad],} \\ d_{\text{Mic}} &= \text{Abstand zwischen zwei benachbarten Mikrofonen in [m],} \\ a &= \text{Kugelradius in [m],} \\ a \gamma_{\text{Mic}} &= \text{Länge des Kreisbogens, welcher } d_{\text{Mic}} \text{ entspricht,} \end{aligned} \tag{5.8}$$

wirkt, weiter nach oben, wodurch sich der nutzbare Frequenzbereich hin zu höheren Frequenzen erweitert.

5.3 Untersuchung unterschiedlicher sphärischer Mikrophonarrays

In diesem Kapitel soll vorrangig die Frage beantwortet werden, wie man die bis dato noch nicht näher spezifizierten, radialen Funktionen $W_m(ka)$ bzw. deren Inversen, d.h. die radialen *equalizing*-Funktionen $EQ_m(ka)$ bestimmen kann.

Diese sind einerseits von der Richtcharakteristik der verwendeten Mikrophone, wobei ausschließlich omnidirektionale und gerichtete Mikrophone 1. Ordnung behandelt werden und andererseits von der Art ihrer Anbringung abhängig, wobei auch hier zwischen den folgenden zwei Formen unterschieden wird:

1. Eine kugelförmige Anbringung der im Idealfall im Raum freischwebenden Mikrophone, was im Weiteren als *offene Kugel* bezeichnet wird.

2. Eine Anbringung der Mikrophone auf der Oberfläche einer soliden Kugel, was im Weiteren mit dem Begriff *geschlossene Kugel* versehen wird.

5.3.1 Offene Kugel mit omnidirektionalen Mikrofonen

Ein Vergleich von Gleichung 5.3 mit Gleichung 4.4 liefert sofort die Antwort auf die Frage, wie die gesuchten radialen Funktionen $W_m(ka)$ aussehen, wenn eine offene Kugel, welche mit omnidirektionalen, d.h. mit Schalldruckmikrofonen, bestückt ist, verwendet wird.

In diesem Fall ergeben sich die radialen Funktionen zu:

$$W_m(ka) = j^m j_m(ka). \quad (5.9)$$

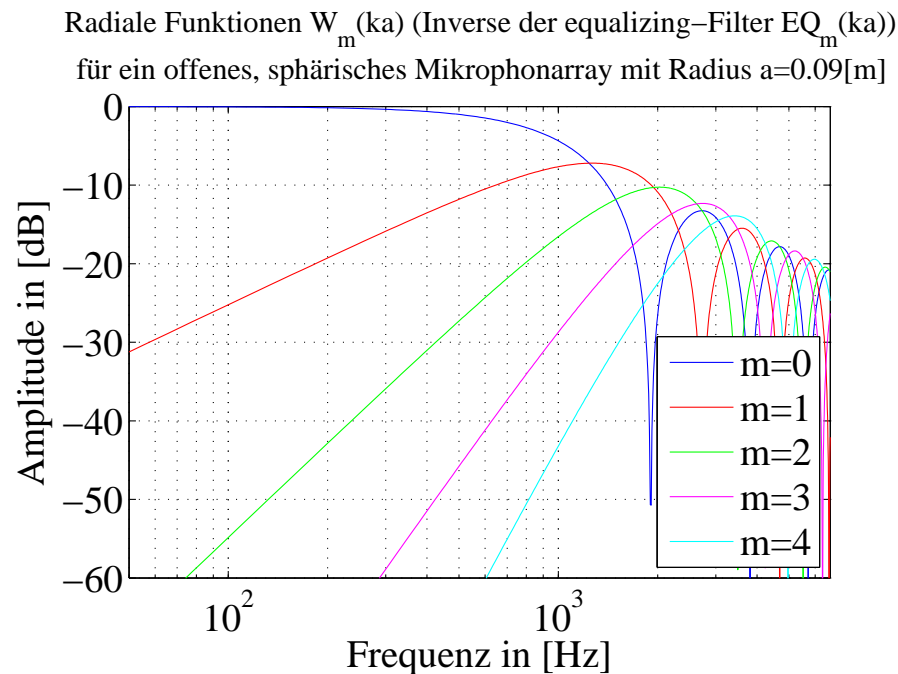


Abbildung 5.3: Radiale Funktionen $W_m(ka)$, dargestellt bis zu einer Ordnung von $M = 4$ und einem Kugelradius von $a = 0.09$ [m], für ein offenes Kugelmikrophon, welches aus omnidirektionalen Mikrofonen besteht.

Wie aus Abbildung 5.3 zu sehen ist, weisen die radialen Funktionen $W_m(ka)$ Nullstellen auf, sodass zumindest bei diesen Frequenzen und in der Praxis in deren Umgebung, kein *equalizing* mit $EQ_m(ka) = \frac{1}{W_m(ka)}$ möglich ist.

5.3.2 Offene Kugel mit gerichteten Mikrofonen 1. Ordnung

Eine Möglichkeit dieses Problem zu lösen besteht darin, anstelle von omnidirektionalen, gerichtete Mikrophone 1. Ordnung zu verwenden, wie sie beispielsweise im Anhang A.1 beschrieben werden.

Prinzipiell ergibt sich das Signal $s(\Theta, \varphi, kr)$, welches von einem gerichteten Mikrofon 1. Ordnung aufgezeichnet wird, aus der anteiligen Überlagerung des Schalldrucks $p(\Theta, \varphi, kr)$ sowie der Schallschnelle $v(\Theta, \varphi, kr)$, die auch als Partikelgeschwindigkeit ($v(\Theta, \varphi, kr) = \frac{\partial p(\Theta, \varphi, kr)}{\partial kr}$) bezeichnet wird, wie folgt:

$$\begin{aligned}
 s(\Theta, \varphi, kr) &= \alpha p(\Theta, \varphi, kr) + (1 - \alpha) \rho c \frac{\partial p(\Theta, \varphi, kr)}{\partial kr}, \text{ wobei} \\
 p(\Theta, \varphi, kr) &= \text{Schalldruck resultierend aus der Schalleinfallrichtung} \\
 &\quad \text{gegeben durch } (\Theta, \varphi, kr), \\
 \frac{\partial p(\Theta, \varphi, kr)}{\partial kr} &= v(\Theta, \varphi, kr) = \text{Schallschnelle resultierend aus der} \\
 &\quad \text{Schalleinfallrichtung gegeben durch } (\Theta, \varphi, kr), \\
 \rho &= \text{Dichte des Mediums (Luft) in } \left[\frac{\text{kg}}{\text{m}^3} \right], \\
 \rho c &= Z = \text{Akustische (Feld)Impedanz in } \left[\frac{\text{N s}}{\text{m}^3} \right], \\
 s(\Theta, \varphi, kr) &= \text{Mikrophonsignal resultierend aus dem Schalleinfall, gegeben durch} \\
 &\quad (\Theta, \varphi, kr), \\
 \alpha &= \text{Richtcharakteristikparameter für den gilt: } 0 \leq \alpha \leq 1. \tag{5.10}
 \end{aligned}$$

Setzt man nun Gleichung 5.10 in die Fourier-Bessel-Reihe gemäß Gleichung 4.4 ein, ergibt sich diese zu:

$$\begin{aligned}
 P(\mathbf{r}, \omega) &= S(j\omega) \sum_{m=0}^{\infty} j^m (\alpha j_m(kr) + (1 - \alpha) j'_m(kr)) \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^{\sigma} Y_{m,n}^{\sigma}(\Theta, \varphi), \text{ mit} \\
 j'_m(kr) &= 1. \text{ Ableitung der sphärischen Besselfunktion 1. Art,} \\
 &\quad m. \text{ Ordnung des Arguments } kr. \tag{5.11}
 \end{aligned}$$

Gleichung 5.11 beschreibt eine verallgemeinerte Form der Fourier-Bessel-Reihe, welche für unterschiedlich gerichtete Mikrofone 1. Ordnung, deren Richtcharakteristik durch den Richtcharakteristikparameter α bestimmt wird, gültig ist. Ein Vergleich von Gleichung 4.4 mit Gleichung 5.11 offenbart, dass Letztere in die zuerst genannte Gleichung übergeht, wenn $\alpha = 1$ gewählt wird. Somit wird deutlich, dass die ursprüngliche Fourier-Bessel-Reihe von einer Verwendung von omnidirektionalen, d.h. von ungerichteten Schalldruckmikrofonen, ausgeht.

Die radialen Funktionen ergeben sich somit zu:

$$W_m(ka) = j^m (\alpha j_m(ka) + (1 - \alpha) j'_m(ka)). \tag{5.12}$$

Bei Verwendung von Kardioid-Mikrofonen, d.h. von differentiellen Mikrofonen 1.

Ordnung mit einem Richtcharakteristikparameter von $\alpha = \frac{1}{2}$, ergeben sich die in Abbildung 5.4 bis zu einer Ordnung von $M = 4$ dargestellten Verläufe für die radialen Funktionen $W_m(ka)$.

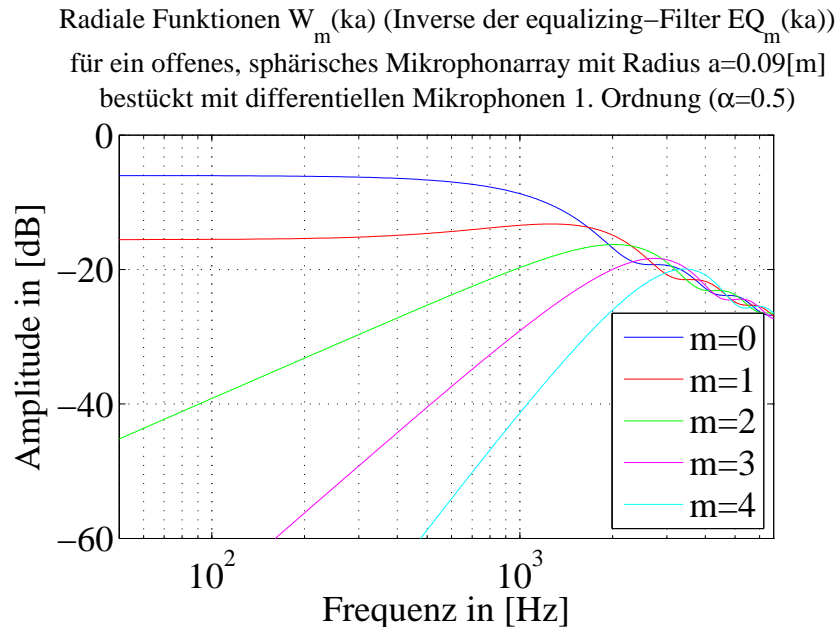


Abbildung 5.4: Radiale Funktionen $W_m(ka)$, dargestellt bis zu einer Ordnung von $M = 4$ und einem Kugelradius von $a = 0.09$ [m], für ein offenes Kugelmikrofon, welches aus gerichteten Mikrofonen 1. Ordnung mit Kardioid-Charakteristik ($\alpha = 0.5$) besteht.

Wie die Verläufe aus Abbildung 5.4 verdeutlichen, ist man durch Verwendung von gerichteten Mikrofonen 1. Ordnung anstelle von reinen Schalldruckmikrofonen in der Tat in der Lage, die in den radialen Funktionen auftretenden, störenden Nullstellen zu vermeiden, was diese Art von Kugelmikrophonarray wesentlich praxistauglicher macht. Leider existiert auch hier ein Problem, welches sich in der Praxis nicht bzw. nur mit sehr großem Aufwand beheben lässt. Dabei handelt es sich um die Tatsache, dass die Richtcharakteristik differentieller Mikrophone (z.B. Kardioid-Charakteristik bei $\alpha = \frac{1}{2}$) nicht wie gefordert über der Frequenz konstant bleibt, sondern zu tiefen Frequenzen hin stetig abnimmt. Ursache hierfür ist, dass zu tiefen Frequenzen hin die Phasendifferenz, bei konstantem Abstand der beiden Mikrophone, welche das differentielle Mikrofon formen, kontinuierlich abnimmt und somit die Anforderungen sowohl an die mechanische als auch an die elektrische Präzision der Bauteile bzw. der Anordnung steigt. Da diesem prinzipbedingtem Problem kaum etwas entgegen zu setzen ist, resultiert eine frequenzabhängige Richtcharakteristik die, wie zuvor bereits erwähnt, zu tiefen Frequenzen hin stetig abnimmt.

Nichtsdestotrotz gibt es Hersteller wie z.B. die Firma *Acoustic Camera*, die solche offenen Kugelmikrophone vertreiben.



Abbildung 5.5: *Offenes, sphärisches Kugelmikrophon der Firma „Acoustic Camera“ mit einem Radius von $a = 0.175$ [m], welches aus $Q = 32$ differentiellen Mikrofonen 1. Ordnung besteht. [IIE]*

Das in Abbildung 5.5 dargestellte Array wird, wie der Name des Herstellers bereits vermuten lässt, dazu verwendet, akustische Ereignisse über ein Kamerabild zu legen und diese somit „sichtbar“ zu machen. Laut Hersteller ist dieses Mikrophonarray dabei in der Lage, Schallereignisse in einem Spektralbereich von $f \approx [1.0, \dots, 10.0]$ [kHz], welche sich in einem Abstand von $r \approx [0.3, \dots, 1.5]$ [m] von diesem entfernt befinden, korrekt zu erfassen. Solche Systeme werden vor allem dann eingesetzt, wenn es darum geht, Lärmquellen, z.B. Klapper- oder Quietschgeräusche im Inneren eines Automobils, schnell und eindeutig aufzuspüren. Wie aufgezeigt wurde, könnte ein solches Kugelmikrophon aber auch für Auralisationssysteme verwendet werden.

5.3.3 Geschlossene Kugel mit omnidirektionalen Mikrofonen

Eine alternative Möglichkeit zur Verwendung der offenen Kugel mit gerichteten Mikrofonen 1. Ordnung wurde z.B. in [Mey01, ME02, MA03, ME04, MD04, Pol05b, Raf05, BDM06, ZR07] mit Hilfe einer soliden, geschlossenen Kugel, auf deren Oberfläche omnidirektionale Mikrophone, also reine Schalldruckmikrophone, verwendet werden, eingehend untersucht.

Die zugrunde liegende Idee lässt sich dabei wie folgt beschreiben: Auf der Oberfläche einer geschlossenen, soliden Kugel herrscht stets eine Schallschnelle von $v_a(\Theta_q, \varphi_q, ka) = 0$. Folglich definiert sich das Schallfeld ausschließlich anhand der Schalldruckverteilung

$p_a(\Theta_q, \varphi_q, ka)$ auf der Kugeloberfläche, die sich mit einfachen Schalldruckmikrofonen messtechnisch erfassen lässt. Mathematisch lässt sich die zugrunde liegende Bedingung von $v_a(\Theta_q, \varphi_q, ka) = 0$ erreichen, indem man innen- und außenliegende Schallquellen entsprechend überlagert (siehe hierzu Abbildung 5.1). Dabei kommt den innenliegenden Schallquellen die Aufgabe zu, das an der Oberfläche der soliden Kugel auftretende *Streufeld*⁵ nachzubilden.

Mit Hilfe der allgemeinen Fourier-Bessel-Reihe gemäß Gleichung 5.1 sowie der Annahme, dass die außenliegenden Quellen Lieferanten ebener Wellen sind, ergibt sich das Wellenfeld der außenliegenden Quellen, welches sich nach innen, d.h. auf die Kugeloberfläche hin, bewegt⁶, gemäß Gleichung 4.4 zu:

$$p_{\text{Inc}}(\Theta_q, \varphi_q, ka) = \sum_{m=0}^{\infty} j^m j_m(ka) \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^{\sigma} Y_{m,n}^{\sigma}(\Theta_q, \varphi_q). \quad (5.13)$$

Ferner muss gelten dass die Schallschnelle auf der Kugeloberfläche, d.h. bei $r = a$ zu Null wird, was sich mathematisch wie folgt ausdrücken lässt:

$$\begin{aligned} v_{\text{Inc}}(\Theta_q, \varphi_q, ka) + v_{\text{Scat}}(\Theta_q, \varphi_q, ka) &= 0 \iff \\ v_{\text{Scat}}(\Theta_q, \varphi_q, ka) &= -v_{\text{Inc}}(\Theta_q, \varphi_q, ka). \end{aligned} \quad (5.14)$$

Leitet man Gleichung 5.13 nach r bzw. a ab, ergibt sich unter Vernachlässigung der Konstante 4π [Teu05] gemäß Gleichung 5.14, $v_{\text{Scat}}(\Theta_q, \varphi_q, ka)$ zu:

$$v_{\text{Scat}}(\Theta_q, \varphi_q, ka) = -k \sum_{m=0}^{\infty} j^m j'_m(ka) \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^{\sigma} Y_{m,n}^{\sigma}(\Theta_q, \varphi_q). \quad (5.15)$$

Mit Hilfe der *eulerschen Gleichung*

$$v(\Theta, \varphi, kr) = \frac{1}{j \rho c k} \nabla p(\Theta, \varphi, kr), \quad (5.16)$$

die einen Zusammenhang zwischen der Schallschnelle $v(\Theta, \varphi, kr)$ und dem Schalldruck $p(\Theta, \varphi, kr)$ herstellt sowie der Tatsache, dass sich die Schallschnelle wie auch der Schalldruck mit Hilfe einer Doppelsumme wie folgt darstellen lässt:

$$\begin{aligned} v(\Theta, \varphi, kr) &= \sum_{m=0}^{\infty} \sum_{0 \leq n \leq m, \sigma = \pm 1} v_{m,n}^{\sigma}(kr) Y_{m,n}^{\sigma}(\Theta, \varphi), \text{ wobei} \\ v_{m,n}^{\sigma}(kr) &= \text{Schallschnellekoeffizienten für die korrespondierenden, sphärischen} \\ &\quad \text{Harmonischen,} \end{aligned} \quad (5.17)$$

⁵Streufeld = engl. *Scattered field* (\implies Index **Scat**)

⁶Eingehendes Wellenfeld = engl. *Incoming wavefield* (\implies Index **Inc**)

kann folgender Zusammenhang zwischen dem Schalldruck und der Schallschnelle auf der Oberfläche einer soliden Kugel angegeben werden (näheres hierzu siehe [Teu05]):

$$p(\Theta, \varphi, kr) = j \rho c \sum_{m=0}^{\infty} \frac{h_m^{(2)}(kr)}{h_m^{(2)'}(kr)} \sum_{0 \leq n \leq m, \sigma = \pm 1} v_{m,n}^{\sigma}(kr) Y_{m,n}^{\sigma}(\Theta, \varphi), \text{ wobei}$$

$$h_m^{(2)}(kr) = 1. \text{ Ableitung der sphärischen Hankelfunktion 2. Art, } m. \text{ Grades}$$

$$\text{des Arguments } kr. \quad (5.18)$$

Durch Einsetzen von Gleichung 5.15 in Gleichung 5.17, wobei hier $v(\Theta, \varphi, kr)$ durch $v_{\text{Scat}}(\Theta_q, \varphi_q, ka)$ ersetzt wird, ergeben sich die Schallschnellekoeffizienten $v_{m,n}^{\sigma}(ka)$ einer einfallenden, ebenen Welle aus den korrespondierenden Ambisonickoeffizienten $B_{m,n}^{\sigma}$ zu:

$$v_{m,n}^{\sigma}(ka) = -k j^m j'_m(ka) B_{m,n}^{\sigma}. \quad (5.19)$$

Setzt man nun Gleichung 5.19 in Gleichung 5.18 ein, ergibt sich der Schalldruck $p_{\text{Scat}}(\Theta_q, \varphi_q, ka)$, welcher aus dem Schallfeld der innenliegenden Quellen, die Kugelwellen abstrahlen resultiert und auf der Kugeloberfläche, d.h. bei $r = a$ an den Positionen (θ_q, φ_q) , an denen sich die q Schalldruckmikrophone befinden, gemessen werden kann, unter Vernachlässigung der, aus Gleichung 5.16 bekannten Konstante $j \rho c k$ sowie unter abermaliger Vernachlässigung der Konstanten 4π , zu:

$$p_{\text{Scat}}(\Theta_q, \varphi_q, ka) = - \sum_{m=0}^{\infty} \frac{j^m j'_m(ka) h_m^{(2)}(ka)}{h_m^{(2)'}(ka)} \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^{\sigma} Y_{m,n}^{\sigma}(\Theta_q, \varphi_q). \quad (5.20)$$

Überlagert man nun die Wellenfelder, genauer gesagt die daraus resultierenden Schalldrücke, der innen- und außenliegenden Quellen, d.h. Gleichung 5.13 und Gleichung 5.20, ergibt sich der auf der soliden Kugeloberfläche resultierende Schalldruck zu:

$$p(\Theta_q, \varphi_q, ka) = p_{\text{Inc}}(\Theta_q, \varphi_q, ka) + p_{\text{Scat}}(\Theta_q, \varphi_q, ka) \Rightarrow$$

$$p(\Theta_q, \varphi_q, ka) = \sum_{m=0}^{\infty} j^m \left[j_m(ka) - \frac{j'_m(ka) h_m^{(2)}(ka)}{h_m^{(2)'}(ka)} \right] \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^{\sigma} Y_{m,n}^{\sigma}(\Theta_q, \varphi_q). \quad (5.21)$$

Somit ergeben sich die radialen Funktionen $W_m(ka)$ bei Verwendung einer soliden Kugel, welche mit omnidirektionalen Mikrophenen besetzt ist, zu:

$$W_m(ka) = j^m \left[j_m(ka) - \frac{j'_m(ka) h_m^{(2)}(ka)}{h_m^{(2)'}(ka)} \right]. \quad (5.22)$$

Mit Hilfe der *Wronskian-Beziehung*, für die gilt:

$$j_m(kr) h_m^{(2)'}(kr) - j_m'(kr) h_m^{(2)}(kr) = \frac{j}{(kr)^2}, \quad (5.23)$$

lässt sich Gleichung 5.22 wie folgt angeben:

$$W_m(ka) = \frac{j^{(m-1)}}{(ka)^2 h_m^{(2)'}(ka)}, \quad (5.24)$$

wodurch sich, eingesetzt in Gleichung 5.21, diese wie folgt vereinfacht:

$$p(\Theta_q, \varphi_q, ka) = \sum_{m=0}^{\infty} \frac{j^{(m-1)}}{(ka)^2 h_m^{(2)'}(ka)} \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^{\sigma} Y_{m,n}^{\sigma}(\Theta_q, \varphi_q). \quad (5.25)$$

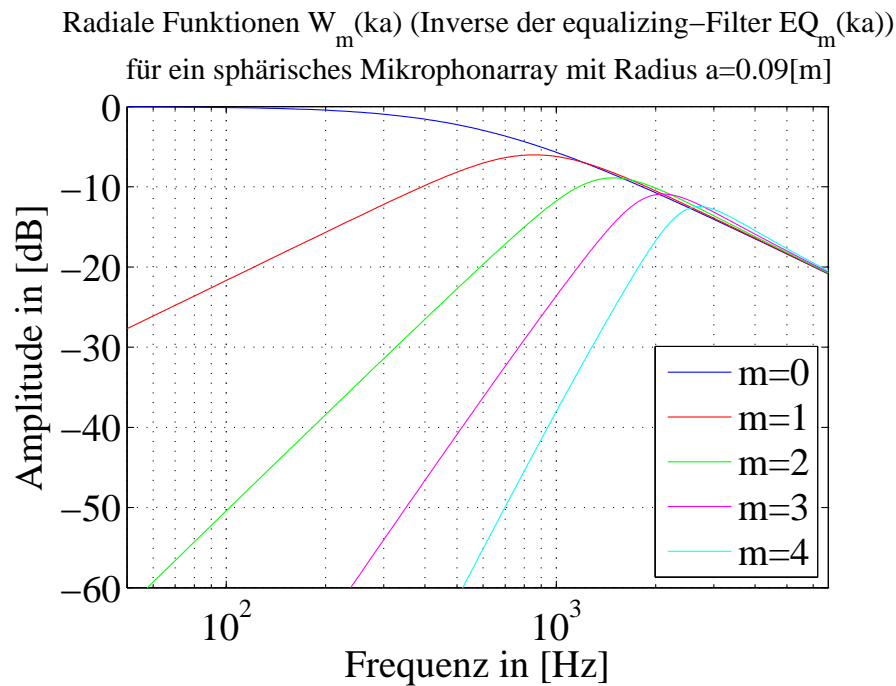


Abbildung 5.6: Radiale Funktionen $W_m(ka)$, dargestellt bis zu einer Ordnung von $M = 4$ und einem Kugelradius von $a = 0.09$ [m], für ein geschlossenes Kugelmikrophon, welches aus omnidirektionalen Mikrophonen besteht.

Die graphische Darstellung von Gleichung 5.22 bzw. von Gleichung 5.24, für einen Kugelradius von $a = 0.09$ [m], in einem Frequenzbereich von $f \in [50, \dots, 6700]$ [Hz], bis zu einer Ordnung von $M = 4$ zeigt, dass die resultierenden Funktionsverläufe sehr homogen ausfallen und ebenfalls keine Nullstellen aufweisen, obwohl nun ausschließlich omnidirektionale Mikrophone verwendet wurden.

Ein Vergleich der Verläufe von Abbildung 5.4 mit Abbildung 5.6, was zur Vereinfach-

chung nochmals in Abbildung 5.7 dargestellt wurde, zeigt, dass Letztere im Vergleich zu den Erstgenannten einen um eine Größenordnung steileren Verlauf aufweisen, d.h. die Steigung z.B. der radialen Funktion 3. Ordnung weist bei Verwendung der geschlossenen Kugel mit omnidirektionalen Mikrofonen die gleiche Steigung auf als die radiale Funktion 4. Ordnung bei Verwendung einer offenen Kugel mit gerichteten Mikrofonen 1. Ordnung (hier $\alpha = \frac{1}{2}$).

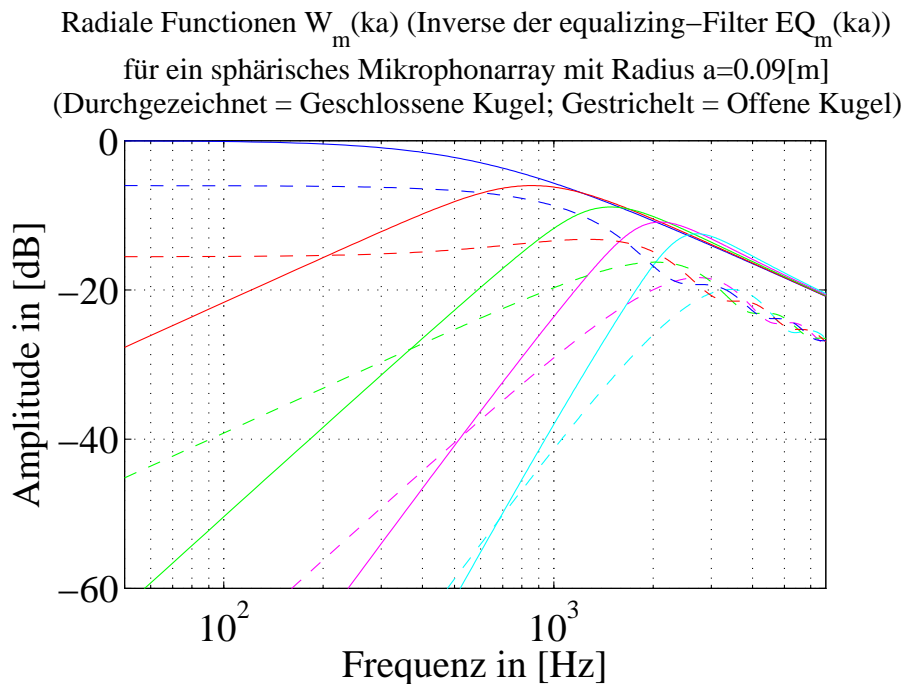


Abbildung 5.7: Radiale Funktionen $W_m(ka)$, dargestellt bis zu einer Ordnung von $M = 4$ und einem Kugelradius von $a = 0.09 [m]$, für ein geschlossenes Kugelmikrophon, welches aus omnidirektionalen Mikrofonen besteht (durchgezeichnete Linien) sowie für ein offenes Kugelmikrophon mit gerichteten Mikrofonen 1. Ordnung ($\alpha = \frac{1}{2}$) (gestrichelte Linien).

Somit ist bei Verwendung der geschlossenen Kugel mit omnidirektionalen Mikrofonen hin zu tieferen Frequenzen in diesem Beispiel für Frequenzen von $f < 200 [Hz]$, für $M = 1$, ein wesentlich stärkeres *equalizing* nötig, wodurch die damit erzielte Richtwirkung für diesen Spektralbereich geschmälert wird. Oberhalb von $f \approx 700 [Hz]$, für $M = 4$, dreht sich jedoch der Effekt um, sodass in dem Beispiel spätestens ab dieser Frequenz, bei einer maximalen Ordnung von $M = 4$, das geschlossene Kugelmikrophon, welches mit omnidirektionalen Mikrofonen bestückt ist, dem offenen Kugelmikrophon, welches mit Kardiod-Mikrofonen 1. Ordnung ausgestattet ist, überlegen ist.

Hinsichtlich der bereits erwähnten Probleme differentieller Mikrophone 1. Ordnung, ihre Richtwirkung bei tiefen Frequenzen aufrecht zu erhalten, ist es schwierig an dieser Stelle eine Aussage über die tatsächlichen Vor- und Nachteile des einen oder anderen Kugelmikrophons zu tätigen.

Seit nicht all zu langer Zeit kann nun auch ein geschlossenes Kugelmikrofon kommerziell erworben werden, wobei dem Autor dieser Arbeit derzeit ausschließlich die Firma *MH Acoustics* bekannt ist, die ein solches Mikrophonarray unter dem Markennamen *Eigenmike*[®] vertreibt.



Abbildung 5.8: Geschlossenes, sphärisches Kugelmikrofon der Firma *MH Acoustics* mit einem Radius von $a = 0.042$ [m], welches aus $Q = 32$ omnidirektionalen Mikrofonen besteht, die quasi-regulär angeordnet sind. [CCFS10]

Das in Abbildung 5.8 dargestellte Mikrophonarray dient laut Hersteller dazu, in einem zweigliedrigen Prozess einen 3D-*beamformer* zu realisieren, der eine beliebige Form aufweisen und in eine beliebige Raumrichtung deuten kann. Dazu werden im ersten Schritt sogenannte *Eigenbeams* erzeugt, wodurch sich auch der Markenname *Eigenmike*[®] begründet, was im Wesentlichen nichts anderem als der Generierung der Ambisonickomponenten $B_{m,n}^{\sigma}$ entspricht, die z.B. in Abbildung 4.4 vollständig bis zu einer Ordnung von $M = 4$ aufgezeigt wurden. Diese sogenannten *Eigenbeams*, bei denen es sich sozusagen um *beamformer* mit definierten Richtcharakteristika handelt, die in exakt vorgegebenen Richtungen zeigen, werden nun in einen zweiten Schritt miteinander kombiniert, um resultierend einen 3D-*beamformer* mit gewünschter Richtcharakteristik zu erhalten, welcher in eine beliebige Raumrichtung zeigen kann. Dies ist dabei nur möglich, da die *Eigenbeams*, d.h. die sphärischen Harmonischen, eine orthonormale Beziehung zu einander aufweisen, wodurch sie sich in exakt vorgegebener Weise miteinander kombinieren lassen. Um diese Aufgabenstellung zu gewähren, besteht dieses Mikrophonarray nicht nur aus Mikrofonen, die in geeigneter Art auf einer Kugeloberfläche angeordnet sind, sondern aus einem kompletten Messsystem, d.h. aus einer entsprechenden Anzahl von Mikrofonverstärkern und A/D-Wandlern. Ferner verfügt es über eine synchrone Signalverarbeitung, sodass letztendlich am PC⁷ über ein VST⁸-Plugin bis zu 16 verschiedene Richtkeulen entworfen und separat gesteuert werden können, deren resultierende akustische Signale dann aufgezeichnet oder weiterverarbeitet werden. Dabei ist dieses Messsystem in der Lage Schallereignisse in einem Spektralbereich von $f \approx [0.03, \dots, 20.0]$ [kHz] zu erfassen, wobei fairerweise erwähnt werden muss, dass hin zu tiefen Frequenzen die maximale Richtwirkung nicht

⁷ *Personal Computer* (PC)

⁸ *Virtual Studio Technology*

mehr vollständig eingehalten werden kann und hin zu hohen Frequenzen sich das räumliche Aliasing mehr und mehr störend bemerkbar macht. Letztendlich bleibt ein nutzbarer Frequenzbereich gemäß den Ergebnissen von [CCFS10] von $f \approx [0.1, \dots, 8.0]$ [kHz] übrig, was aber die Qualität bzw. das Potential dieses Mikrophonarrays nicht schmälert, sondern lediglich relativiert.

5.3.4 Geschlossene Kugel mit Kardioid-Mikrofonen 1. Ordnung

In [Ple09] wird vorgestellt, wie man gerichtete Mikrophone 1. Ordnung mit Kardioid-Charakteristik mit einer soliden Kugel kombinieren kann.

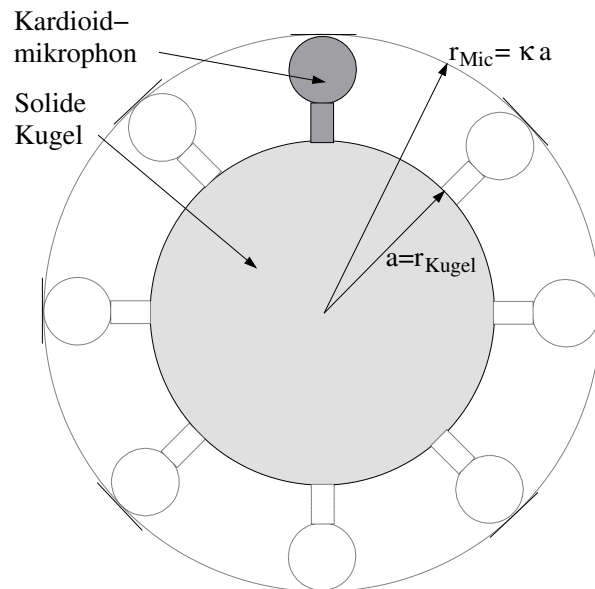


Abbildung 5.9: Schematische Darstellung die aufzeigt wie sich differentielle Mikrophone (z.B. mit Kardioid-Charakteristik) auf einer soliden Kugel mit Radius a anordnen lassen, um schließlich ein kugelförmiges Mikrophone Array mit Radius r_{Mic} zu erhalten.

Im Prinzip stellt die Forderung, gerichtete Mikrophone auf der Oberfläche einer soliden Kugel anzubringen, eine unlösbare Aufgabe dar, da differentielle Mikrophone, wie Abbildung A.1 verdeutlicht, für eine einwandfreie Funktionsweise zwei Öffnungen benötigen, die üblicherweise gegenüberliegend angeordnet sind. Somit würde bei einem differentiellen Mikrophone, welches direkt auf der Kugeloberfläche angebracht würde, eine der beiden Öffnungen verschlossen werden, wodurch sich zwangsläufig eine omnidirektionale Richtcharakteristik einstellt. Aus diesem Grund ist es, wie Abbildung 5.9 zeigt, unumgänglich differentielle Mikrophone in einem gewissen Abstand zur Oberfläche der soliden Kugel anzubringen, um deren Richtwirkung aufrechtzuerhalten. Der äußere Radius wird dabei mit r_{Mic} bezeichnet und kann auch in Abhängigkeit zum Radius der soliden Kugel a wie

folgt ausgerückt werden:

$$\begin{aligned}
 r_{\text{Mic}} &= \kappa a, \text{ wobei} \\
 r_{\text{Mic}} &= \text{Äußerer Radius des Mikrophonarrays in [m]} \\
 \kappa &= \text{Skalierungsfaktor für den gilt: } \kappa \geq 1.
 \end{aligned}
 \tag{5.26}$$

Die radialen Funktionen $W_m(ka)$ lassen sich dabei gemäß [Ple09] wie folgt angeben:

$$W_m(ka) = j^m \left[j_m(k\kappa a) - j j'_m(k\kappa a) + (j h_m^{(2)'}(k\kappa a) - h_m^{(2)}(k\kappa a)) \left(\frac{j'_m(ka)}{h_m^{(2)'}(ka)} \right) \right].
 \tag{5.27}$$

Bei Verwendung von Kardiod-Mikrofonen sowie einen Skalierungsfaktor von $\kappa = 1.5$ ergeben sich die in Abbildung 5.10 bis zu einer Ordnung von $M = 4$, dargestellten Verläufe für die radialen Funktionen $W_m(ka)$.

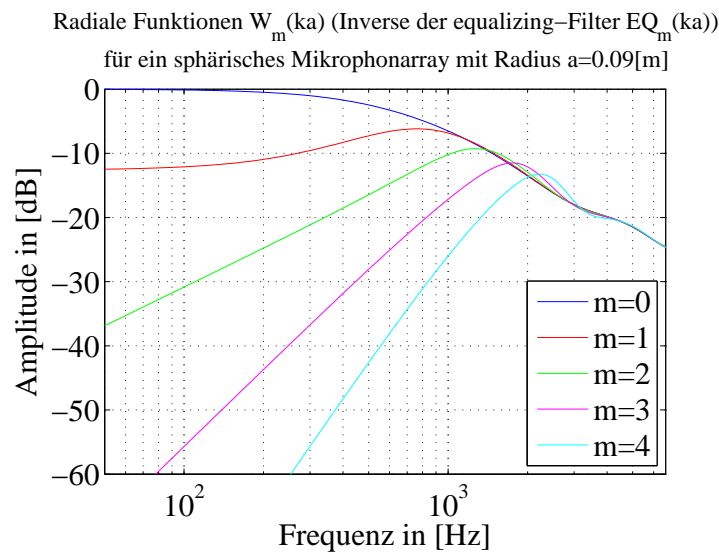


Abbildung 5.10: Radiale Funktionen $W_m(ka)$, dargestellt bis zu einer Ordnung von $M = 4$ und einem Kugelradius von $a = 0.09$ [m], für ein geschlossenes Kugelmikrofon, welches mit Kardiod-Mikrofonen aufgebaut ist, deren äußerer Radius eine Größe von $r_{\text{Mic}} = \kappa a = 1.5 \cdot 0.09$ [m] = 0.135 [m] aufweist.

Der Vergleich der Verläufe der radialen Funktionen $W_m(ka)$ der bis dato vorgestellten, praxistauglichen, kugelförmigen Mikrophonarrays, welche in Abbildung 5.11 bis zu einer Ordnung von $M = 4$ dargestellt sind, offenbart, dass mit der Verwendung von gerichteten, anstelle von omnidirektionalen Mikrofonen und einer soliden Kugel, vor allem im unteren Spektralbereich eine zum Teil deutliche Verbesserung erreicht werden kann. Diese fällt im Vergleich zu der bereits in Kapitel 5.3.3 vorgestellten Gegenüberstellung des geschlossenen Kugelmikrophons, welches mit omnidirektionalen Mikrofonen bestückt ist und einem

offenen Kugelmikrophon, welches mit gerichteten Mikrofonen bestückt ist, nun noch deutlicher aus.

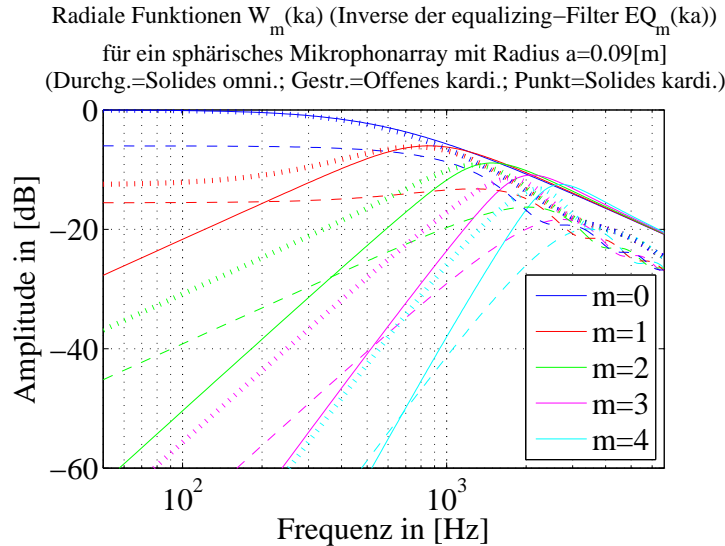


Abbildung 5.11: Radiale Funktionen $W_m(ka)$, dargestellt bis zu einer Ordnung von $M = 4$ und einem Kugelradius von $a = 0.09 [m]$, für ein geschlossenes Kugelmikrophon, welches aus omnidirektionalen Mikrofonen besteht (durchgezeichnete Linien), für ein offenes Kugelmikrophon mit gerichteten Mikrofonen 1. Ordnung ($\alpha = \frac{1}{2}$) (gestrichelte Linien) sowie für ein geschlossenes Kugelmikrophon, welches mit Kardioid-Mikrofonen aufgebaut ist, deren äußerer Radius eine Größe von $r_{Mic} = \kappa a = 1.5 \cdot 0.09 [m] = 0.135 [m]$ aufweist (gepunktete Linien).

Ein genauer Blick auf Abbildung 5.11 lässt dabei erkennen, dass nun die geschlossene Kugel, welche mit Mikrofonen mit Kardioid-Charakteristik bestückt ist, für Ordnungen von $M \geq 1$, ab einer bestimmten Frequenz f , welche mit steigender Ordnung ebenfalls ansteigt, gegenüber dem Mikrophonarray, bei welchem auf einer geschlossenen Kugel omnidirektionale Mikrofone angebracht sind, eine schlechtere Performance aufweist, wobei für dieses Beispiel folgende Übergangsfrequenzen f angegeben werden können:

Ordnung	f in [Hz]
$M = 1$	744
$M = 2$	1289
$M = 3$	1822
$M = 4$	2356

Dieser vermeintliche Vorteil des geschlossenen Kugelmikrophons mit omnidirektionalen Mikrofonen bei höheren Frequenzen, gegenüber dem, welches mit gerichteten Mikrofonen bestückt ist, revidiert sich, wenn eine kleinere, solide Kugel verwendet wird, so dass dessen Außenradius r_{Mic} genauso groß ist, wie der Radius der soliden Kugel auf dem die omnidirektionalen Mikrofone angebracht sind.

Radiale Funktionen $W_m(ka)$ (Inverse der equalizing-Filter $EQ_m(ka)$) für ein sphärisches Mikrophonarray mit einem äußeren Radius von $a=0.09$ [m] (Gurchg.=Solides omni.; Gestr.=Offenes kardi.; Punkt=Solides kardi.)

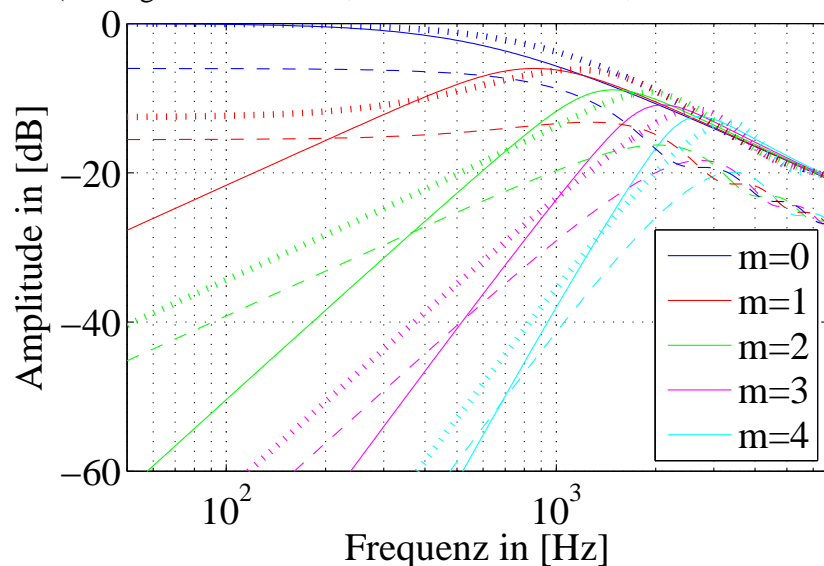


Abbildung 5.12: Radiale Funktionen $W_m(ka)$, dargestellt bis zu einer Ordnung von $M = 4$ und einem Kugelradius von $a = 0.09$ [m], für ein geschlossenes Kugelmikrophon, welches aus omnidirektionalen Mikrofonen besteht (durchgezeichnete Linien), für ein offenes Kugelmikrophon mit gerichteten Mikrofonen 1. Ordnung ($\alpha = \frac{1}{2}$) (gestrichelte Linien) sowie für ein geschlossenes Kugelmikrophon, welches mit Kardioid-Mikrofonen aufgebaut ist, deren äußerer Radius eine Größe von $r_{Mic} = \kappa a = 1.5 \cdot 0.06$ [m] = 0.09 [m] aufweist (gepunktete Linien).

Dies ist der Fall, wenn in unserem Beispiel bei einem gleich bleibenden Skalierungsfaktor von $\kappa = 1.5$ eine solide Kugel mit einem Radius von $a = 0.06$ [m] verwendet wird. Aus den sich dabei einstellenden Verhältnissen, die aus Abbildung 5.12 entnommen werden können, ist deutlich zu erkennen, dass nun kaum mehr Spektralbereiche existieren, in denen das solide Kugelmikrophon mit den omnidirektionalen Mikrofonen dem mit den gerichteten Mikrofonen überlegen ist. Stattdessen weist dieses nach wie vor im unteren Spektralbereich eine überlegene Performance auf ohne dafür einen Nachteil im oberen Spektralbereich in Kauf nehmen zu müssen, da hier beide Mikrophonarrays auf die gleichen Werte zulaufen.

Ein solches Kugelmikrophonarray kann zumindest nach dem derzeitigen Kenntnisstand des Autors dieser Arbeit nicht käuflich erworben werden, sondern müsste, wie in [Ple09] aufgezeigt, selbst angefertigt werden.

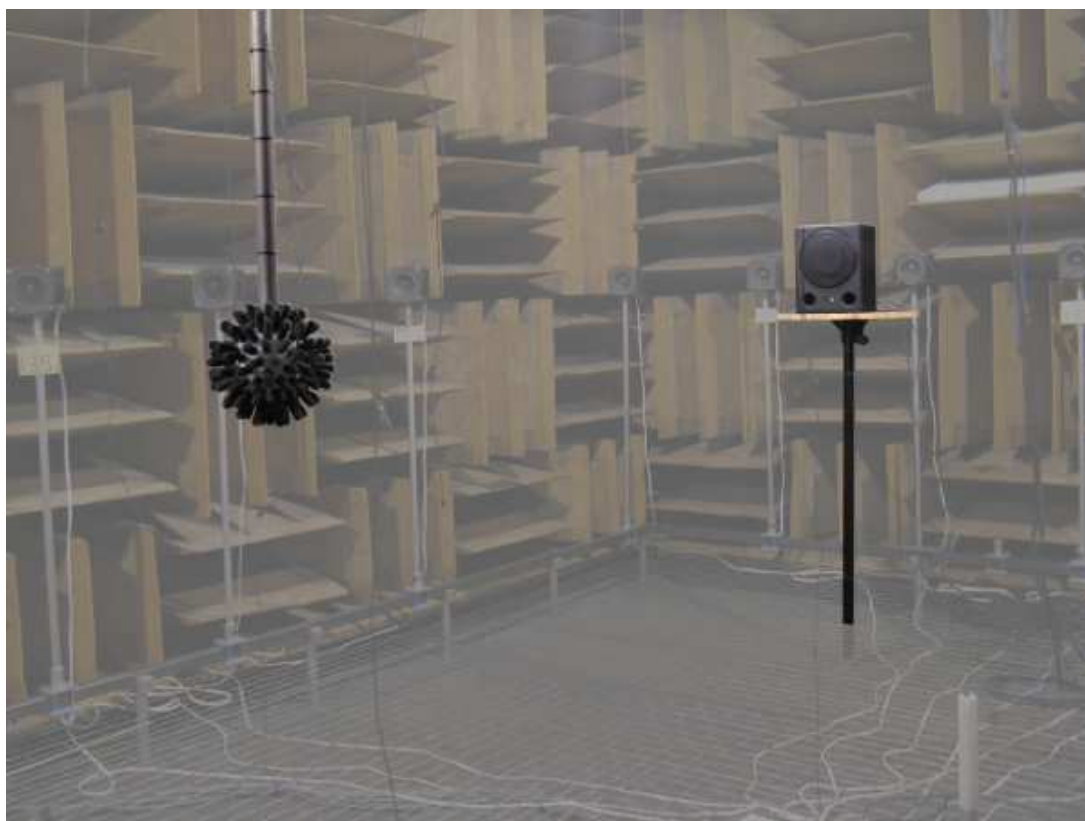


Abbildung 5.13: Geschlossenes, sphärisches Kugelmikrophon aus [Ple09] mit einem inneren Radius von $a = 0.045$ [m] sowie einem äußeren bzw. Mikrophonradius von $r_{Mic} = 0.069$ [m], welches aus $Q = 120$ Mikrofonen mit Kardioid-Charakteristik aufgebaut ist, die quasi-regulär auf einem Ikosaeder⁹ (platonischer Körper) angeordnet sind, wobei jeweils 6 Mikrophone möglichst regelmäßig auf jeder seiner 20 Flächen angebracht wurden.

⁹ $\epsilon\acute{\iota}\kappa\omicron\sigma\iota$, gesprochen „ikosi“ (Griechisch) = Zwanzig

5.3.5 Zusammenfassung

Die anfängliche Untersuchung zur Klärung der Frage, welche Mikrophonarrays sich gut für Auralisationszwecke eignen, zeigte, dass sich im 2D-Fall lediglich das kreuzförmige sowie das zirkulare Array, welche jeweils sowohl mit Schalldruck- als auch mit Schallschnellempfänger bestückt sein müssen, eignen, wobei sich deutliche Vorteile zu Gunsten des zirkularen Arrays abzeichneten. Ferner zeigte sich, dass sich zum Zwecke der *akustischen Dokumentation* kein zweidimensionales Mikrophonarray eignet, da es zwangsweise zu Fehlern im gemessenen Wellenfeld kommt sobald dieses Komponenten besitzt, die sich in der dritten Dimension bewegen. Diese Fehler lassen sich auf unterschiedliche Arten beheben, wobei sich als flexibelste die Verwendung von dreidimensionalen Mikrophonarrays herausstellte.

Aufgrund der bereits im 2D-Fall erkennbaren Vorteile eines zirkularen Arrays beschränkte sich die nachfolgende Untersuchung auf dessen dreidimensionales Pendant, d.h. auf kugelförmige Mikrophonarrays. Dabei wurden sowohl Aussagen über die minimal notwendige Anzahl Q von Mikrophonen gemacht, die ein Kugelmikrophon besitzen muss, um das Wellenfeld bis zu einer bestimmten maximalen Ordnung M messtechnisch erfassen zu können, als auch über die Anbringungsorte der Mikrophone, die so gewählt werden müssen, dass die *Orthonormalitätsbedingung* stets erfüllt wird.

Die detaillierte Untersuchung von 4 unterschiedlichen kugelförmigen Mikrophonarrays ergab, dass die Verwendung einer offenen Kugel, welche mit omnidirektionalen Mikrophonen bestückt ist, ungeeignet ist, da die dabei notwendigen M *equalizing*-Filter unendlich hohe Werte aufweisen müssten und somit nicht realisierbar sind. Wie sich herausstellte, kann dieses Problem jedoch durch Verwendung von gerichteten Mikrophonen 1. Ordnung anstelle der zuvor untersuchten omnidirektionalen Mikrophone behoben werden. Ersetzt man die offene Kugel durch eine geschlossene, d.h. durch eine solide Kugel auf deren Oberfläche ausschließlich omnidirektionale Mikrophone angebracht sind, ist es möglich, ein praxistaugliches Mikrophonarray zu erstellen. Dieses weist jedoch im unteren Spektralbereich diverse Schwächen auf, da die hierfür benötigten M *equalizing*-Filter zum Teil sehr hohe Verstärkungswerte besitzen. Dies führt zu einer Verstärkung des *Eigenrauschens* und beschränkt somit die Verwendung der Ambisonickomponenten höherer Ordnung mit abnehmender Frequenz zunehmend, was schließlich zu einer sich zu tiefen Frequenzen hin verstärkenden Reduktion der Richtwirkung des Mikrophonarrays führt. Ab einer bestimmten Frequenz ist jedoch, gleiche Größen vorausgesetzt, das geschlossene dem offenen Mikrophonarray überlegen. Somit ist das geschlossene Mikrophonarray auch hinsichtlich der angesprochenen Schwierigkeiten die differentielle Mikrophone 1. Ordnung im unteren Frequenzbereich besitzen, mit denen das offene Array bestückt ist, gegenüber diesem zu bevorzugen. Als letztes wurde ein unorthodoxer Arraytyp vorgestellt, der versucht das Beste aus beiden Welten in sich zu vereinen, d.h. die gute Performance im unteren Spektralbereich durch Verwendung von gerichteten Mikrophonen 1. Ordnung sowie die positiven Eigenschaften im oberen Spektralbereich, welche die geschlossene Kugel, die

mit omnidirektionalen Mikrofonen versehen ist, zeigt. Wählt man dabei den Außenradius dieses Mikrofonarrays so, dass dieser mit dem Radius der geschlossenen Kugel, die mit reinen Schalldruckempfängern bestückt ist ($a = r_{\text{Mic}}$) übereinstimmt, so zeigt sich in der Tat eine Überlegenheit dieses Konzeptes.

Betrachtet man sich die Ergebnisse, die durch diese beiden Arrays tatsächlich erzielt wurden (geschlossene Kugel mit omnidirektionalen Mikrofonen: [CCFS10]; geschlossene Kugel mit Kardiod-Mikrofonen: [Ple09]) darf bezweifelt werden, ob sich der Mehraufwand, der auf jeden Fall bei der Erstellung dieses Mikrofonarrays notwendig ist, gerechtfertigt ist. Über den Grund des schlechten Abschneidens dieses unorthodoxen Arraytyps kann jedoch nur spekuliert werden. Ob es an den Problemen, welche die Kardiod-Mikrophone bei tiefen Frequenzen haben bzw. an der mechanischen Anordnung, die sich letztendlich sehr stark vom Ideal, d.h. von einer soliden Kugel, an der in einem gewissen Abstand Mikrophone mit Kardiod-Charakteristik angebracht sind, ohne dass dabei eine Verbindung zwischen der Oberfläche und den Mikrofonen herrscht, abwich oder ob es etwa an einer Kombination aus beiden lag, konnte abschließend nicht geklärt werden. Allerdings wird auch in [DMN03] darüber gemutmaßt, dass in der Praxis differentielle Mikrophone 1. Ordnung, z.B. mit Kardiod-Charakteristik, dazu neigen, bei tiefen Frequenzen omnidirektional zu werden und/oder dabei ein schlechteres Rauschverhalten aufweisen. Zudem wurde darauf hingewiesen, dass selbst, wenn die differentiellen Mikrophone 1. Ordnung ein perfektes Verhalten aufweisen sollten, akustische Störungen hervorgerufen z.B. durch Beugungseffekte an der ausladenden, mechanischen Struktur, welche, wie oben aufgezeigt, unvermeidlich ist, letztendlich zu Abweichungen vom theoretischen Modell führen.

Solche Probleme können bei Verwendung einer soliden Kugel, auf deren Oberfläche omnidirektionale Mikrophone angebracht werden, weitestgehend vermieden werden. Aus diesem Grund wird im weiteren Verlauf dieses Kapitels als einziger Vertreter der kugelförmigen Mikrofonarrays dieses weiterverfolgt. Obwohl, wie bereits aufgezeigt wurde, ein solches Mikrofonarray bereits käuflich erhältlich ist, wird nachfolgend auf den praktischen Entwurf eines solchen Mikrofonarrays detailliert eingegangen, einerseits weil das sogenannte *Eigenmike*[®] der Firma *MH Acoustics* sehr teuer ist, andererseits auch deshalb, da sich nur durch einen eigenen Aufbau eines solchen Mikrofonarrays wichtige Fragestellungen ergeben. Schließlich wird sich zeigen, dass sich dabei neue Möglichkeiten beim Entwurf eines solchen Mikrofonarrays ergeben, mit denen sich dessen Performance weiter steigern lässt.

5.4 Praktischer Entwurf eines geschlossenen, kugelförmigen Mikrophonarrays zur Messung von HOA-Komponenten

In Kapitel 5.2.1 wurden mehrere Kugelmikrophone hinsichtlich ihrer Verwendbarkeit zur Messung, Analyse und Codierung von Wellenfeldern untersucht.

Dabei wurde unter anderem festgestellt, dass die messtechnische Erfassung der einzelnen Raummoden, die im weiteren Verlauf kurz als Moden bezeichnet werden, aus deren Überlagerung sich letztendlich das gesamte Wellenfeld zusammensetzt, mit Hilfe von Mikrophonarrays erfolgen kann. Die Art des Mikrophonarrays (geschlossen, offen, etc.), deren Größe sowie der räumlichen Anordnung und die Richtcharakteristik der dabei verwendeten Mikrophone bestimmen die Form der radialen Funktionen $W_m(ka)$. Mit Anwendung der zu dem jeweilig verwendeten Mikrophonarray passenden, radialen Funktionen $W_m(ka)$ ist es möglich, Moden bis zu einer bestimmten Ordnung M messtechnisch zu erfassen und somit das Wellenfeld, zumindest bis zu dieser physikalisch vorgegebenen Grenze zu bestimmen.

In Abbildung 5.11 sind deren Verläufe bis zu einer Ordnung von $M = 4$ unter Verwendung der jeweils gleichen Geometrie zusammengefasst.

Die Verläufe der einzelnen Moden kennzeichnen dabei, wie sich das Wellenfeld über der Frequenz zusammensetzt. Dabei fällt auf, dass prinzipiell mit steigender Frequenz zunehmend Moden höherer Ordnung an der Bildung des Wellenfeldes beteiligt sind.

Betrachtet man sich dieses Prinzip am Beispiel des geschlossenen, kugelförmigen Mikrophonarrays erkennt man, dass dieses bis zu einer Frequenz von $f \approx 120$ [Hz] lediglich durch die Mode der Ordnung $m = 0$ gebildet wird, unter der Annahme, dass die Wirkung jeder Mode die mehr als 20 [dB] unterhalb der Grundmode, d.h. der Mode mit der Ordnung $m = 0$ liegt, unbedeutend bei der Bildung des Wellenfeldes ist. Die radialen Funktionen eines geschlossenen kugelförmigen Mikrophonarrays wurden bereits in Abbildung 5.6 bis zu einer Ordnung von $M = 4$ dargestellt und in Abbildung 5.14, unter Berücksichtigung der gleichen Rahmenbedingungen, nochmals bis zu einer Ordnung von $M = 10$ veranschaulicht.

Oberhalb dieser Frequenz weist die Mode der Ordnung $m = 1$ dann einen Pegel auf, der zu hoch ist, um bei der Bildung des Wellenfeldes vernachlässigt werden zu können. Selbiges gilt ab einer Frequenz von $f \approx 500$ [Hz] für die Mode der Ordnung $m = 2$, ab einer Frequenz von $f \approx 930$ [Hz] für die Mode der Ordnung $m = 3$ und so weiter.

Da das hier untersuchte Mikrophonarray aus praktischen Gründen¹⁰ eine Mikropho-

¹⁰Als praktische Gründe, welche die Gesamtanzahl der verwendeten Mikrophone Q beschränken, können folgende genannt werden:

- Mechanik:

Insbesondere im Fall von offenen Kugelmikrophonarrays kann es zu Störungen des Wellenfeldes bedingt durch die Mechanik (Chassis) des Mikrophonarrays kommen. Theoretische Voraussetzun-

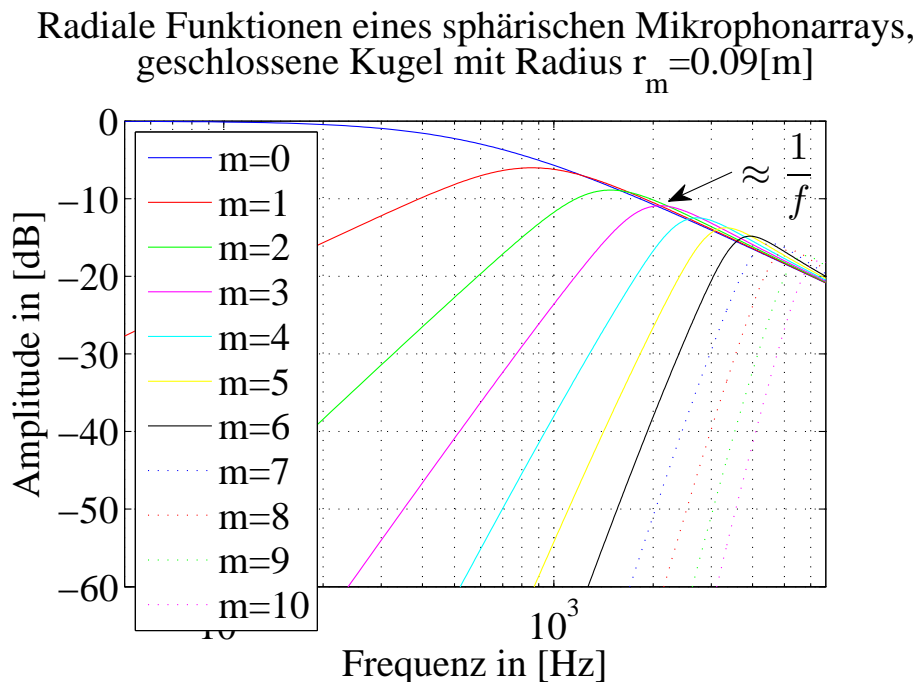


Abbildung 5.14: Radiale Funktionen $W_m(ka)$, dargestellt bis zu einer Ordnung von $M = 10$ und einem Kugelradius von $a = 0.09$ [m], für ein geschlossenes Kugelmikrophon, welches aus omnidirektionalen Mikrofonen besteht.

nanzahl von $Q \leq 32$ ausweisen soll, kann, wie bereits gezeigt, das Wellenfeld lediglich bis zu einer maximalen Ordnung von $M = 4$ analysiert werden. Somit ist klar, dass ober-

gen für eine korrekte Erfassung des Wellenfeldes würden dabei verletzt und somit besonders für hohe Frequenzen unbrauchbar (vergleiche Kapitel 5.3.4).

- Anzahl der Kanäle:

Das zugrunde liegende Messsystem muss in der Lage sein, sämtliche Mikrophone des Mikrophonarrays *zeitgleich* zu erfassen, da ansonsten eine messtechnische Bestimmung der jeweiligen Moden nicht möglich ist. Dies stellt diverse Anforderungen an die verwendete Hardware (z.B. Bussystem zu Synchronisierung einzelner Hardwarekomponenten) und Software (Menge der prozessierbaren Ein- und Ausgangskanäle).

- Bandbreite:

Je nachdem, wie die verwendete Hardware an den verwendeten Computer angeschlossen ist, besteht eine mehr oder weniger große Bandbreite zur Signalübertragung zur Verfügung. In Abhängigkeit der gewählten Abtastrate und der Anzahl der Ein- und Ausgangskanäle kann diese durchaus zu einem Flaschenhals werden.

- Datenmenge/Signalverarbeitung:

Selbst wenn die verwendete Hardware, Software sowie das verwendete Bussystem eine höhere Anzahl an Kanälen zulassen würde, müssen diese noch entsprechend in Echtzeit verarbeitet werden. Selbst die derzeitig erhältlichen Prozessoren können dabei schnell an ihre Leistungsfähigkeit gelangen insbesondere, wenn man deren Mehrkerntechnik nicht schon bei der Programmierung berücksichtigt. Moderne, leistungsstarke Graphikprozessoren (GPU) stellen dabei eine echte Alternative dar, falls es hierbei zu Engpässen kommen sollte.

halb einer bestimmten Grenzfrequenz, Moden höherer Ordnung, die vom kugelförmigen Mikrophonarray nicht mehr erfasst werden können, so stark werden, dass sich dadurch Vieldeutigkeiten bei der Wellenfeldanalyse ergeben. Höhere Moden ähneln dann in ihrem Verlauf niedrigeren Moden und können nicht mehr eindeutig vom gesamten Wellenfeld extrahiert werden.

Diese obere Grenzfrequenz ergibt sich durch das räumliche Aliasing und kann gemäß Gleichung 5.8 berechnet werden.

Bei dem Fehler, der sich durch die Beschränkung auf eine maximale Ordnung M ergibt, handelt es sich um den bereits in Formel 4.30 allgemein und in Formel 4.31 unter Freifeldbedingungen¹¹ beschriebenen und in Abbildung 4.10 dargestellten *tuncation error*.

5.4.1 Codierung eines natürlichen Klangfeldes

Wie in Kapitel 5.2.1 aufgezeigt, lassen sich gemäß der Gleichungen 5.3, 5.4 und 5.5 die Schalldrücke im Raum an beliebigen Positionen (Θ_q, φ_q) auf einer Kugeloberfläche bestimmen.

Diese Gleichungen beschreiben dabei den Vorgang, der sich bei der Messung eines beliebigen Schallfeldes abspielt. Auf Grundlage von Gleichung 5.5 lässt sich dieser wie folgt interpretieren:

1. Die N sphärischen Harmonischen, welche die *Moden des Wellenfeldes* repräsentieren, werden an den Positionen auf der Kugeloberfläche (Θ_q, φ_q) , mit $q = [1, \dots, Q]$, d.h. an den Orten, an denen sich die Q Mikrophone befinden, abgetastet. Mathematisch repräsentieren die Moden die *Eigenwerte des Wellenfeldes*, wodurch sich auch der Begriff *Eigenbeams*, welcher z.B. beim Entwurf modaler *beamformer* verwendet wird (siehe Kapitel A.4), erklärt. Bei Betrachtung von Abbildung 4.4, welche die $N = 25$ *Eigenbeams*, d.h. sämtliche Eigenwerte des Wellenfeldes bis zu einer Ordnung von $M = 4$, darstellt, lässt sich die in Gleichung 5.5 durch $\mathbf{Y}_{Q \times N}$ beschriebene Abtastung der jeweiligen Eigenbeams gut vorstellen.
2. Der Einfluss der Arraygeometrie (Radius a der Kugel), die Art der verwendeten Kugel (offen oder geschlossen) sowie die Richtcharakteristika der verwendeten Mikrophone (omnidirektional, direktional), welche allesamt in den M frequenzabhängigen Radialfunktionen $W_m(ka)$ bzw. in der Diagonalmatrix $\mathbf{W}_{N \times N}$ beinhaltet sind, werden durch Multiplikation mit den abgetasteten N sphärischen Harmonischen $\mathbf{Y}_{Q \times N}$ in selbigen berücksichtigt. Bei Verwendung der geschlossenen Kugel mit omnidirektionalen Mikrophenen bestimmen sich die Radialfunktionen $W_m(ka)$ gemäß Gleichung 5.22 bzw. nach Gleichung 5.24 und sind für einen Kugelradius von $a = 0.09$ [m] und einer Ordnung von $M = 10$ in Abbildung 5.14 graphisch dargestellt.

¹¹Gemäß [EKM03] kann bei Mikrophonarrays von Fernfeldbedingungen gesprochen werden, wenn gilt: $r_S \geq 10a$

3. Schließlich werden die durch die radialen Funktionen $\mathbf{W}_{N \times N}$ modifizierten, N sphärischen Harmonischen $\mathbf{Y}_{Q \times N}$ mit Hilfe der Ambisonickomponenten $\mathbf{B}_{N \times 1}$ zusammengefasst. Diese definieren dabei das Klangfeld und sind alleinig von den Quellsignalen sowie deren örtlichen Lagen $(\Theta_{\text{SRC}}, \varphi_{\text{SRC}})$ abhängig, wie z.B. für Freifeldbedingungen in Gleichung 4.10 aufgezeigt. Somit ergeben sich die Q Schalldrucksignale $\mathbf{p}_{a_{Q \times 1}}$, welche von den Mikrofonen, die sich im Raum an den Positionen (Θ_q, φ_q) auf einer Kugeloberfläche mit Radius a befinden.

Da sich diese Arbeit mit der Auralisation beschäftigt, wobei ein vermessenenes Wellenfeld an einem beliebigen Ort möglichst exakt reproduziert werden soll, sind wir vorrangig an der Analyse des selbigen, d.h. an der Bestimmung dessen Ambisonickomponenten $\mathbf{B}_{N \times 1}$, interessiert. Theoretisch lassen sich diese aus den gemessenen Q Mikrophonsignalen (Θ_q, φ_q) gemäß Gleichung 5.6 berechnen. Dieser Vorgang, der in Abbildung 5.2 bereits graphisch verdeutlicht wurde, wird nachfolgend schrittweise zum besseren Verständnis interpretiert.

1. Die Matrix $\mathbf{Y}_{N \times Q}^+$, bei welcher es sich um die Pseudoinverse der Matrix $\mathbf{Y}_{Q \times N}$ handelt, gewichtet und fasst die gemessenen Q Mikrophonsignale $\mathbf{p}_{a_{Q \times 1}}$ so zusammen, dass an ihrem Ausgang die Signale der $N = (M + 1)^2$ sphärischen Harmonischen $Y_{m,n}^\sigma(\Theta_q, \varphi_q)$, mit $0 \leq m \leq M$, $0 \leq n \leq m$, $\sigma = \pm 1$ resultieren. Somit stellt $\mathbf{Y}_{N \times Q}^+$ die *Übergangs- bzw. Transformationsmatrix vom Raum-Zeit- in den Wellenfeldbereich* dar. Sofern sich die Q Mikrophone an regelmäßigen Abtastpunkten (Θ_q, φ_q) auf der Kugeloberfläche befinden, wenn also die Orthonormalitätsbedingung gemäß Gleichung 5.7 durch diese Positionen exakt erfüllt wird, darf die pseudoinverse Matrix $\mathbf{Y}_{N \times Q}^+$ durch $\mathbf{Y}_{N \times Q}^T$, d.h. durch die Transponierte der Matrix $\mathbf{Y}_{Q \times N}$, ersetzt werden. Ist dies der Fall, ergeben sich die Gewichtungsfaktoren der Übergangsmatrix $\mathbf{Y}_{N \times Q}^+$ für alle Q Mikrofonpositionen, pro sphärische Harmonische aus den Längen der Vektoren vom Ursprung zu den jeweiligen Oberflächenwerten des n . Eigenbeams, die an den Anbringungsstellen der Mikrophone, d.h. an den Orten (Θ_q, φ_q) abgetastet werden.
2. Am Ausgang der Transformationsmatrix $\mathbf{Y}_{N \times Q}^+$ stehen somit die Ausgangssignale der N Eigen-beamformer $Y_{m,n}^\sigma(\Theta_q, \varphi_q)$ zur Verfügung. Um die gesuchten Ambisonickomponenten $\mathbf{B}_{N \times 1}$ und somit eine kompakte Beschreibung des Wellenfeldes zu erhalten, müssen die N Eigenbeams frequenzunabhängig sein und dabei gleichstarke Ausgangswerte liefern. Um dies zu erreichen, müssen die Ausgänge der Eigenbeams mit den inversen der Radialfunktionen $\mathbf{W}_{N \times N}^{-1}$ gewichtet werden. Dadurch wird die Art, wie sich das Wellenfeld auf der Oberfläche des verwendeten Mikrophonarrays zusammensetzt, wie z.B. aus Abbildung 5.14 ersichtlich, berücksichtigt.

Somit lassen sich mit Hilfe der Gleichung 5.6 HOA-Komponenten $\mathbf{B}_{N \times 1}$ bis zu einer maximalen Ordnung von $M = \lfloor \sqrt{Q} - 1 \rfloor$, wobei $\lfloor \dots \rfloor$ das Abrunden auf den nächsten

Ganzzahlwert (*integer*-Wert)¹² bezeichnet, bestimmen.

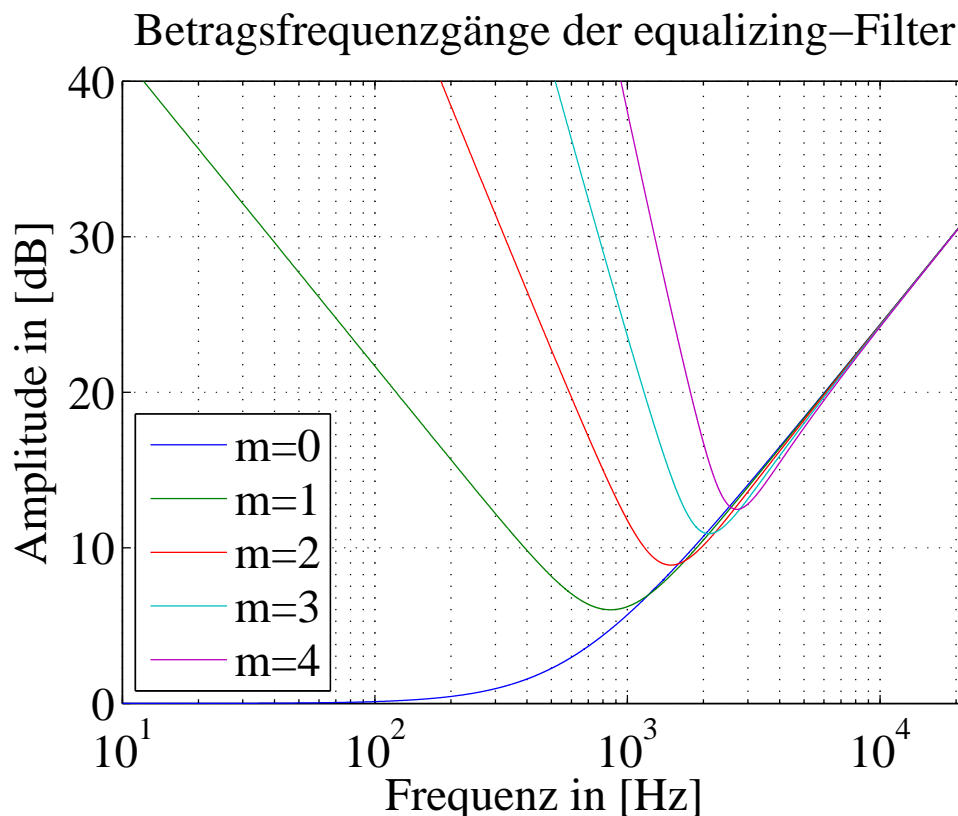


Abbildung 5.15: Equalizing-Funktionen $EQ_m(ka)$ dargestellt bis zu einer Ordnung von $M = 4$ und einem Kugelradius von $a = 0.09$ [m] für ein geschlossenes Kugelmikrophon, welches aus omnidirektionalen Mikrophonen besteht.

Die in Gleichung 5.6 angegebene LS¹³-Lösung, welche den kleinsten MSE¹⁴ erzeugt, kann jedoch so nicht direkt verwendet werden. Wie Abbildung 5.15 verdeutlicht, kommt es immer dann, wenn $W_m(ka)$ sehr kleine Werte annimmt, was bei hohen Ordnungen m und niedrigen Frequenzen der Fall ist, zu Instabilitäten bedingt durch das hier sehr hoch verstärkte Eigenrauschen der Sensoren bzw. des gesamten Messsystems bzw. anderer Fehler wie etwa der *Positionierungsfehler der Sensoren* oder *Abweichungen in deren Übertragungsfunktionen*. Durch diese Instabilitäten wird das Ergebnis letztendlich unbrauchbar.

Durch Einführung einer *Regularisierung*, d.h. durch eine Beschränkung der radialen equalizing-Funktionen $EQ_m(ka)$ mit Hilfe der Regularisierungsfunktionen $T_m(ka)$ etwa auf eine maximal erlaubte Verstärkung, lässt sich dieses Problem z.B. mit Hilfe der so-

¹²Ganzzahl = engl. *Integer*

¹³LS = engl. *Least Square* = Kleinstes (Fehler)Quadrat

¹⁴MSE = engl. *Mean Squared Error* = Mittlerer quadratischer Fehler

nannten *Tikhonov-Filter* nach [BDM06] wie folgt beheben:

$$\begin{aligned}
 EQ_m(ka) &= \frac{T_m(ka)}{W_m(ka)}, \text{ mit} \\
 T_m(ka) &= \frac{|W_m(ka)|^2}{|W_m(ka)|^2 + \epsilon^2}, \text{ und} \\
 \epsilon &\geq 0 \begin{cases} \epsilon = 0 & \text{LS-Lösung, d.h. keine Regularisierung (maximale Richtwirkung)} \\ \epsilon > 0 & \text{Regularisierung, maximal bis zum } \textit{delay-and-sum beamformer}^{15} \end{cases}, \\
 T_m(ka) &= \text{Regularisierungsfunktionen, mit } m = [0, \dots, M], \text{ berechnet z.B. nach} \\
 &\quad \text{Tikhonov,} \\
 \epsilon &= \text{Regularisierungsparameter } \epsilon \geq 0.
 \end{aligned} \tag{5.28}$$

Visualisierung des Einflusses der Regularisierungsfilter $T_m(kR)$, in Abhängigkeit von $W_m(kR)$

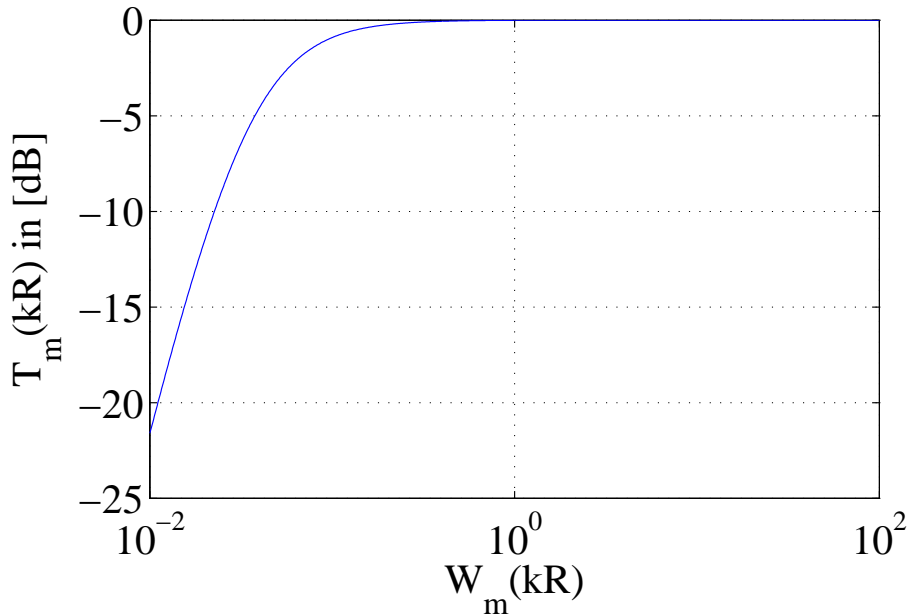


Abbildung 5.16: Darstellung der Tikhonov-Regularisierungsfilter $T_m(ka)$ in Abhängigkeit von $|W_m(ka)|$.

Abbildung 5.16 veranschaulicht dabei die Abhängigkeit der Regularisierungsfilter $T_m(ka)$, welche gemäß Gleichung 5.28 nach Tikhonov berechnet wurden, von den Be-

¹⁵Ein *beamformer*, bei dem die einzelnen Mikrophone lediglich verzögert (*delay*) und anschließend aufsummiert (*sum*) werden, stellt einerseits den robustesten, andererseits aber auch den Repräsentanten mit dem geringsten Bündelungsmaß aller möglichen *beamformer* dar. In der Literatur wird ein solcher meist als *delay-and-sum beamformer* bezeichnet.

tragsfrequenzgängen der radialen Filter $|W_m(ka)|$.

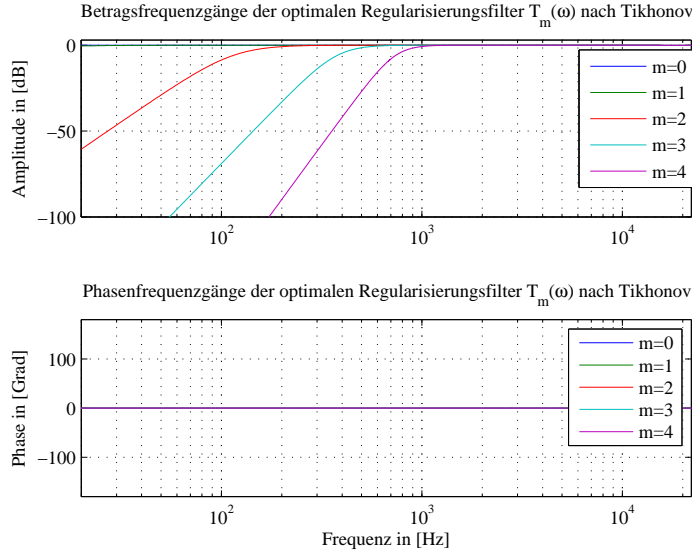


Abbildung 5.17: Tikhonov-Regularisierungsfiler $T_m(ka)$ dargestellt bis zu einer Ordnung von $M = 4$, welche die radialen equalizing-Funktionen $EQ_m(ka)$, die für ein geschlossenes Kugelmikrophon, welches aus omnidirektionalen Mikrofonen besteht und einen Radius von $a = 0.09$ [m] hat, berechnet wurden, sodass diese nach Anwendung der Regularisierungsfiler, welche mit $\epsilon = 0.0332$ bestimmt wurden, eine maximale Anhebung von ≈ 24 [dB] aufweisen.

Wendet man die in Abbildung 5.17 dargestellten Tikonov-Regularisierungsfiler auf die aus Abbildung 5.15 bekannten, unbeschränkten *equalizing*-Funktionen an, resultieren daraus die in Abbildung 5.18 aufgezeigten, regularisierten *equalizing*-Funktionen $EQ_m(ka)$.

Wie man aus Abbildung 5.18 erkennen kann, ist die Aufgabe der Regularisierungsfiler $T_m(ka)$ die maximale Anhebung der theoretisch optimalen *equalizing*-Filer $EQ_m(ka) = \frac{1}{W_m(ka)}$ auf eine obere Schranke, welche durch den Regularisierungsparameter ϵ bestimmt wird, zu begrenzen.

Die in Gleichung 5.28 aufgezeigte Berechnung, der nach Tikhonov regularisierten M *equalizing*-Filer $EQ_m(ka)$, lassen sich zudem wie folgt vereinfachen:

$$EQ_m(ka) = \frac{W_m(ka)^*}{|W_m(ka)|^2 + \epsilon^2}, \text{ wobei}$$

$$W_m(ka)^* = \text{Konjugiert komplexe der Matrix } W_m(ka). \quad (5.29)$$

5.4.2 Optimale Wahl des Regularisierungsparameters ϵ

Die Wahl eines guten bzw. möglichst optimalen Regularisierungsparameters ϵ ist keine triviale Aufgabe. In [BDM06] wird eine Möglichkeit zur Bestimmung des Regularisie-

rungsparameters ϵ in Abhängigkeit des Sensorrauschens σ_q sowie der Anzahl der Sensoren Q beschrieben, welche sich mathematisch wie folgt zusammenfassen lässt:

$$\epsilon = \frac{1 - \sqrt{1 - \frac{1}{s^2}}}{1 + \sqrt{1 - \frac{1}{s^2}}}, \text{ mit}$$

$$s = \sqrt{Q} 10^{\left(\frac{\sigma_q}{20}\right)}, \text{ wobei}$$

Q = Anzahl der verwendeten Mikrophone (Sensoren),
 s = Maximaler, linearer Verstärkungsfaktor der Tikhonov-Filter,
 σ_q = Maximales Sensorrauschen in [dB] (Wert der aus dem Datenblatt des Mikrophons entnommen werden kann).

(5.30)

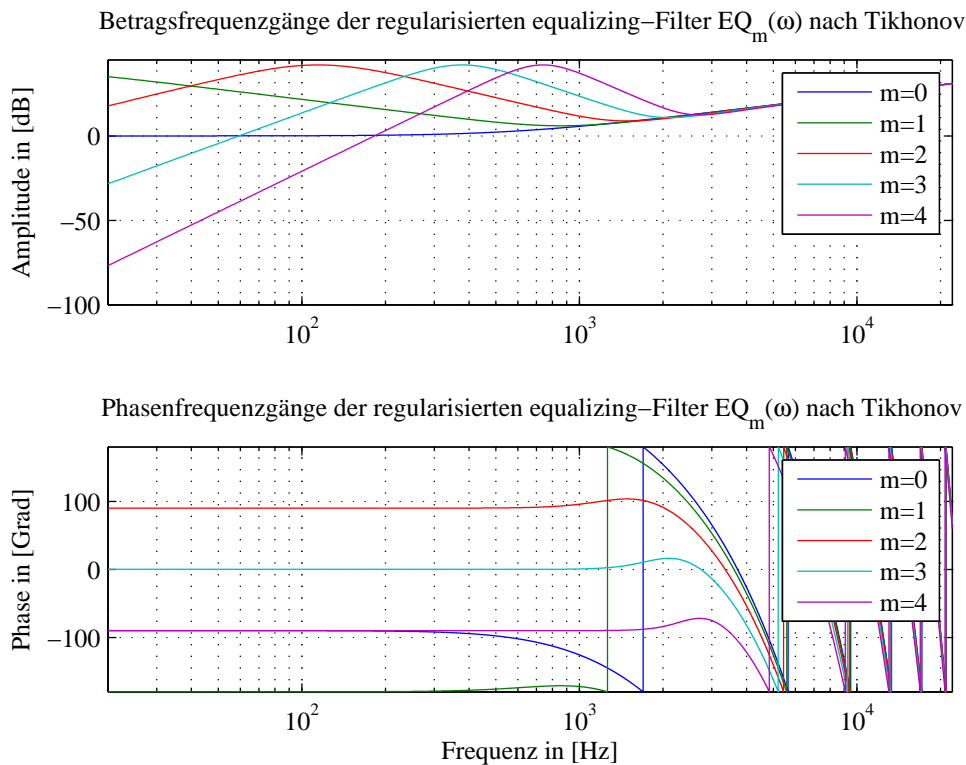


Abbildung 5.18: Nach Tikhonov regularisierte equalizing-Funktionen $EQ_m(ka)$ dargestellt bis zu einer Ordnung von $M = 4$ und einem Kugelradius von $a = 0.09$ [m], für ein geschlossenes Kugelmikrofon, welches aus omnidirektionalen Mikrofonen besteht.

Ein maximaler Sensorrauschwert von $\sigma_q = -6$ [dB] ergibt dabei einen Wert für den Regularisierungsparameter von $\epsilon = 0.0332$, auf dem die in Abbildung 5.18 dargestellten

regularisierten *equalizing*-Filter $EQ_m(ka)$, basieren. Durch diese Art der Bestimmung des Regularisierungsparameters ϵ wird gewährleistet, dass bei der Bestimmung der HOA-Komponenten $B_{m,n}^\sigma$ das maximale Sensorrauschen von σ_q nicht überschritten wird. Ferner erkennt man bei aufmerksamer Betrachtung von Gleichung 5.30, dass der lineare Verstärkungsfaktor s mit wachsender Anzahl der Sensoren Q steigt, was, unter Annahme gleichen Sensorrauschens σ_q , wiederum zu einer Verringerung des Regularisierungsparameters ϵ führt und sich somit der theoretisch optimalen Lösung, die bei $\epsilon = 0$ erreicht wird, nähert. Somit bestätigt Gleichung 5.30 die allgemein gültigen Tatsache, dass sich durch eine Erhöhung der Anzahl der Sensoren das SNR¹⁶ des Gesamtsystems verbessert. Obwohl die in Gleichung 5.30 vorgestellte Methode auf physikalischen Gegebenheiten beruht, wird diese selbst in [BDM06], aus welcher sie entnommen wurde, nicht als optimal bezeichnet, sodass auch dort vorgeschlagen wird, den Regularisierungsparameter ϵ im Allgemeinen empirisch zu bestimmen.

Ein optimaler Regularisierungsparameter ϵ muss, wie auch in [BDM06] richtig erkannt, Rauschen vermeiden, ohne dabei die relevanten Informationen zu zerstören. Wie Eingangs in diesem Kapitel erwähnt, existieren beim Entwurf der (regularisierten) *equalizing*-Filter $EQ_m(ka)$ Parallelen zum Entwurf von *beamforming*-Filter. Auch dort herrschen im Prinzip die gleichen Probleme mit den Sensoren, d.h. mit dem Sensorrauschen, den Abweichungen in deren Übertragungsfunktionen sowie mit den Variationen ihrer Anbringungsorte. Eine beim Entwurf von robusten *beamforming*-Filter verwendete Methode ist, deren *Suszeptibilität* $K(\omega)$ (siehe Kapitel A.1.2) auf einen frequenzunabhängigen, konstanten, maximalen Wert zu begrenzen, was für Mikrophon-ULAs¹⁷ z.B. in [Dör98] detailliert aufgezeigt wurde.

Der Entwurf optimaler, robuster *beamforming*-Filter für modale *beamformer*, die auf einem *sphärischen Mikrophonarray* beruhen, ist hingegen, z.B. aus [EKM03, ME04, Teu05, LD07, SYS09, YSS⁺11, SYS11, YSM⁺11] bekannt. Dabei wurde nicht, wie bei dem in [Dör98] beschriebenen Verfahren, beim Entwurf robuster ULAs die Suszeptibilität als Grundlage der Optimierung verwendet, sondern dessen Kehrwert WNG¹⁸, den es dann zu minimieren gilt. Bei beiden Verfahren, d.h. unabhängig ob man nun über der Frequenz die Suszeptibilität auf einen konstanten Wert maximiert oder das WNG minimiert, wurden von mehreren Autoren unabhängig voneinander dann die besten, weil praxistauglichsten, Ergebnisse erzielt, wenn bei der Optimierung das Bündelungsmaß (DI¹⁹) maximiert und gleichzeitig eine vorgegebene maximale (Suszeptibilität) bzw. minimale (WNG), frequenzunabhängige Schranke nicht über- bzw. unterschritten wurde.

In [Raf05] wurde eine detaillierte Fehleranalyse für sphärische Mikrophonarrays durchgeführt, wobei Kugelmikrophone mit einer unterschiedlichen Anzahl an Mikrofonen so-

¹⁶SNR = engl. *Signal to Noise Ratio* = Signal-Rausch-Abstand

¹⁷ULA = engl. *Uniform Line Array* = Lineares (Mikrophon) Array

¹⁸WNG = engl. *White-Noise-Gain* = Wert der angibt, um wie viel unkorreliertes (weißes) Rauschen, wie etwa das Sensorrauschen, verstärkt wird.

¹⁹DI = engl. *Directivity Index* = Bündelungsmaß

wie verschiedenen, räumlichen Abtastungen, d.h. Positionierungskonzepten der Mikrophone, miteinander hinsichtlich ihrer Aliasingfehler, Positionierungsfehler sowie Fehler durch das Messrauschen (verursacht durch das Eigenrauschen der Sensoren) verglichen wurden.

Die dort erzielten Erkenntnisse lassen sich dabei wie folgt zusammenfassen:

- Unabhängig von der Art der untersuchten räumlichen Abtastung (*equiangle*²⁰, *gaussian*²¹, *uniform*²²) ergab sich bei gleicher Anzahl Q von Mikrophonen ein identischer, spektraler Verlauf des WNG. Somit kann festgehalten werden, dass für das WNG nicht das Positionierungskonzept für die Mikrophone, sondern ausschließlich deren Anzahl Q von Bedeutung ist.
- Die Aliasingfehler aller drei untersuchten Positionierungskonzepte zeigten einerseits mit steigender Frequenz einen Anstieg, was so auch zu erwarten war. Andererseits traten hier zum Teil deutliche Unterschiede zwischen den jeweils verwendeten räumlichen Abtastungen auf, sofern wieder eine gleiche Anzahl Q an Mikrophonen verwendet wurde. Dabei schnitt die *equiangle*-Abtastung am schlechtesten und die reguläre (*uniform*) Mikrophonanordnung am besten ab. Dies zeigt, dass eine reguläre Anbringung der Mikrophone, vor allem im oberen Spektralbereich, hinsichtlich des Aliasingfehlers deutliche Vorteile gegenüber anderen Positionierungskonzepten bei einer gleichen Anzahl von Q Mikrophonen bietet.
- Bei der Untersuchung der Positionierungsfehler wurde unabhängig vom untersuchten Positionierungskonzept stets von einem gleichverteilten Fehler von $|\Delta| \leq 0.005$ [rad] $\approx 0.3^\circ$, was bei einem Kugelradius von $a = 0.2$ [m] in etwa einem Positionierungsfehler von $\approx \pm 1$ [mm] entspricht, ausgegangen. Unter diesen Voraussetzungen ergab sich abermals unabhängig von der verwendeten räumlichen Abtastung bei gleicher Mikrophonanzahl Q ein maximaler Fehler von ≤ -40 [dB], der im Vergleich, vor allem zum Fehler des Messrauschens, als sehr gering einzustufen ist. Dies bedeutet, dass Positionierungsfehler bei Betrachtung des gesamten Fehlers so gut wie vernachlässigt werden können, sofern eine ausreichende, mechanische Präzision bei der Positionierung der Mikrophone gewährleistet werden kann, was in der Praxis bedingt durch eine maschinelle Fertigung üblicherweise kein Problem darstellt. Zudem zeigte sich, dass der spektrale Verlauf der Positionierungsfehler dem des Messrauschens ähnelt, was die Verwendung des WNG als alleiniges Robustheitsmaß für das zu untersuchende System bestärkt.

Zusammenfassend kann also konstatiert werden, dass das WNG ein Maß dafür ist, dass, abgesehen vom Aliasingfehler, die Robustheit des Systems gegenüber sämtlichen Fehlern, wie etwa den Positionierungsfehlern der Mikrophone im Array, den *beamforming*-Filtern

²⁰ *Equiangle* = engl. gleiche Winkel

²¹ *Gaussian* = engl. Gauß- bzw. Normalverteilt

²² *Uniform* = engl. Regulär bzw. regelmäßig

sowie dem Eigenrauschen der Sensoren, die im Allgemeinen als voneinander unabhängig betrachtet werden, darstellt. Werte von $WNG > 0$ [dB] kennzeichnen dabei eine Unterdrückung von unkorrelierten Fehlern und weisen somit auf ein robustes Systemverhalten hin, wohingegen Werte von $WNG < 0$ [dB] eine Verstärkung selbiger beschreibt, was auf ein instabiles Systemverhalten schließen lässt.

5.5 Beschreibung des Optimierungsalgorithmus

Um möglichst optimal regularisierte *beamforming*-Filter berechnen zu können, ist es unerlässlich zu wissen, wie sich das WNG berechnen lässt. Wie sich später noch zeigen wird, ist neben der Berechnung des $WNG_{\text{dB}}(\omega)$ auch noch die Berechnung des Array-Gewinns $G_{\text{dB}}(\omega)$ für die Bestimmung optimal regularisierter *beamforming*-Filter nötig.

Die Berechnung dieser beiden Größen wird nachfolgend, sozusagen als Vorbereitung für die Entwicklung eines geeigneten Optimierungsalgorithmus, aufgezeigt.

5.5.1 Berechnung des *white noise gain* ($WNG_{\text{dB}}(\omega)$)

Wie im vorherigen Kapitel erörtert, eignet sich das WNG als Maß für die Beurteilung der Robustheit und somit der Qualität eines Mikrophonarrays. Das WNG kann dabei wie folgt berechnet werden.

1. Allgemeine Definition des WNG:

$$\begin{aligned}
 WNG(\omega) &= \frac{|d(\Theta_0, \varphi_0, \omega)|^2}{\frac{1}{Q^2} \sum_{q=1}^Q |H_q(\Theta_q, \varphi_q, \omega)|^2} \\
 &= \frac{\text{Energie des Nutzsignals am Arrayausgang}}{\text{Energie des Rauschsignals am Arrayausgang}}, \text{ wobei} \\
 d(\Theta_0, \varphi_0, \omega) &= \text{Ausgangssignal des } \textit{beamformer}, \text{ welcher in die} \\
 &\quad \text{Richtung } (\Theta_0, \varphi_0) \text{ zeigt,} \\
 H_q(\Theta_q, \varphi_q, \omega) &= \text{Rauschsignal des } q. \text{ Mikrophons des } \textit{beamformer}, \\
 &\quad \text{verursacht durch dessen Eigenrauschen.} \tag{5.31}
 \end{aligned}$$

2. Berechnung von $d(\Theta_0, \varphi_0, \omega)$, d.h. des Nutzsignals am Arrayausgang, bzw. des Signals, das am *beamformer*-Ausgang ansteht, wenn dieser in die gewünschte Blick-

richtung (Θ_0, φ_0) zeigt:

$$\begin{aligned}
 d(\Theta_0, \varphi_0, \omega) &= \frac{1}{N} \sum_{m=0}^M \sum_{0 \leq n \leq m, \sigma = \pm 1} T_m(\omega) Y_{m,n}^\sigma(\Theta_0, \varphi_0), \text{ wobei} \\
 N &= (M+1)^2 = \text{Anzahl der sphärischen Harmonischen,} \\
 Y_{m,n}^\sigma(\Theta_0, \varphi_0) &= \text{Sphärische Harmonische } m. \text{ Ordnung, } n. \text{ Grades, die sich} \\
 &\quad \text{aus der Blickrichtung des } \textit{beamformer} (\Theta_0, \varphi_0) \text{ ergeben,} \\
 T_m(\omega) &= \text{Regularisierungsfiler nach Tikhonov, welche sich mit} \\
 &\quad \text{Gleichung 5.28 berechnen lassen.} \tag{5.32}
 \end{aligned}$$

3. Berechnung von $H_q(\Theta_q, \varphi_q, \omega)$, d.h. des Rauschsignals des q . Arraymikrophons über der Frequenz, verursacht durch das Eigenrauschen des Sensors:

$$\begin{aligned}
 H_q(\Theta_q, \varphi_q, \omega) &= \frac{1}{N^2} \sum_{m=0}^M \sum_{0 \leq n \leq m, \sigma = \pm 1} EQ_m(\omega) Y_{m,n}^\sigma(\Theta_q, \varphi_q), \text{ wobei} \\
 Y_{m,n}^\sigma(\Theta_q, \varphi_q) &= \text{Sphärische Harmonische } m. \text{ Ordnung, } n. \text{ Grades, die sich aus} \\
 &\quad \text{der Position des } q. \text{ Mikrophons } (\Theta_q, \varphi_q), \text{ ergeben,} \\
 EQ_m(\omega) &= \textit{Equalizing-Filer} \text{ pro Ordnung } m, \text{ berechnet z.B. nach} \\
 &\quad \text{Gleichung 5.29.} \tag{5.33}
 \end{aligned}$$

4. Berechnung des frequenzabhängigen WNG in [dB] ($WNG_{\text{dB}}(\omega)$):

$$WNG_{\text{dB}}(\omega) = 10 \log_{10}(WNG(\omega)). \tag{5.34}$$

Der maximal erreichbare WNG-Wert kann für einen modalen *beamformer* nach [EKM03] wie folgt angegeben werden:

$$WNG_{\text{Max}} = 10 \log_{10}(Q). \tag{5.35}$$

In unserem Fall, d.h. für $Q = 32$ Mikrophone, entspricht dies einem maximalen WNG-Wert von etwas mehr als $WNG_{\text{dBMax}}(\omega) \approx 15$ [dB]. Wie zuvor erwähnt, wurden bei einigen Untersuchungen festgestellt, dass dann die besten Ergebnisse erzielt werden, wenn bei der Optimierung das *Bündelungsmaß*, welches auch unter dem Begriff *Array-Gewinn* $G(\omega)$ bekannt ist, maximiert wird, bei gleichzeitiger Einhaltung eines minimal erlaubten WNG-Wertes, wobei dafür in [EKM03] für die praktische Anwendung ein Wert von $WNG_{\text{dBMin}} = -10$ [dB] vorgeschlagen wird. Obwohl, wie später noch gezeigt werden wird, es in dem hier verwendeten Optimierungsverfahren nicht erforderlich ist, den Array-Gewinn $G(\omega)$ explizit zu berechnen, sollen die dafür notwendigen Schritte dem interessierten Leser nicht vorenthalten werden, da insbesondere die hierfür notwendigen Formeln

für die Erzeugung einiger graphischer Darstellungen benötigt werden.

5.5.2 Berechnung des Array-Gewinns ($G_{dB}(\omega)$)

Der Array-Gewinn stellt ein Maß für die akustische SNR-Verbesserung dar, die auf der Richtwirkung des *beamformer*, für Signale, die aus dessen Blickrichtung auf das Array einfallen, basiert.

1. Allgemeine Definition des Array-Gewinns $G(\omega)$:

$$\begin{aligned}
 G(\omega) &= \frac{|\Psi(\Theta_0, \varphi_0, \omega)|^2}{\frac{1}{4\pi} \int_{\Theta=0}^{2\pi} \int_{\varphi=0}^{\pi} |\Psi(\Theta, \varphi, \omega)|^2 \sin(\Theta) d\Theta d\varphi} \\
 &= \frac{\text{Energie bei Schalleinfall aus der Blickrichtung } (\Theta_0, \varphi_0)}{\text{Energie bei omnidirektionalen Schalleinfall}}, \text{ wobei} \\
 \Psi(\Theta, \varphi, \omega) &= \text{Richtdiagramm des Mikrophonarrays.} \tag{5.36}
 \end{aligned}$$

2. Berechnung von $\Psi(\Theta_0, \varphi_0, \omega)$, d.h. des Richtdiagramms bei Schalleinfall aus der Blickrichtung (Θ_0, φ_0) :

$$\begin{aligned}
 \Psi(\Theta_0, \varphi_0, \omega) &= \frac{1}{N} \sum_{m=0}^M \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^{\sigma}(\Theta, \varphi, \omega) Y_{m,n}^{\sigma}(\Theta_0, \varphi_0), \text{ mit} \\
 B_{m,n}^{\sigma}(\Theta, \varphi, \omega) &= \text{diag}(\mathbf{E}\mathbf{Q}_m(\omega)) \mathbf{Y}^+ \mathbf{p}_a, \text{ wobei} \\
 B_{m,n}^{\sigma}(\Theta, \varphi, \omega) &= \text{Spektrale Ambisonicsignale, welche das Wellenfeld codieren,} \\
 &\quad \text{(siehe hierzu die Gleichungen 5.3-5.6),} \\
 \mathbf{E}\mathbf{Q}_m(\omega) &= \text{Regularisierungsfiler nach Gleichung 5.16 bzw.} \\
 &\quad \text{Gleichung 5.29.} \tag{5.37}
 \end{aligned}$$

3. Berechnung von $\Psi(\Theta, \varphi, \omega)$, d.h. des Richtdiagramms bei omnidirektionalen Schalleinfall:

$$\Psi(\Theta, \varphi, \omega) = \frac{4\pi}{Q} \frac{1}{N} \sum_{m=0}^M \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^{\sigma}(\Theta, \varphi, \omega) Y_{m,n}^{\sigma}(\Theta, \varphi). \tag{5.38}$$

4. Berechnung des frequenzabhängigen Array-Gewinns in [dB] ($G_{dB}(\omega)$):

$$G_{dB}(\omega) = 10 \log_{10}(G(\omega)). \tag{5.39}$$

Der maximal erreichbare Array-Gewinn kann für einen modalen *beamformer* nach [Elk00] wie folgt angegeben werden:

$$G_{\text{dBMax}}(\omega) = 20 \log_{10}(M + 1). \quad (5.40)$$

In unserem Fall, d.h. für $M = 4$ entspricht dies einem maximalen Array-Gewinn von $G_{\text{dBMax}}(\omega) \approx 14$ [dB].

5.5.3 Optimierungsalgorithmus

Nachdem nun alle Größen, die für die Bestimmung des optimalen Regularisierungsparameter $\epsilon(\omega)$ und den daraus resultierenden, gesuchten, optimalen *equalizing*-Filter $EQ_m(\omega)$, bekannt sind, kann der dafür nötige Ablauf beschrieben werden, was mit Hilfe des in Abbildung 5.19 aufgezeigten Flussdiagramms versucht wird.

Zum besseren Verständnis werden nachfolgend die in Abbildung 5.19 dargestellten Schritte des Optimierungsprozesses zusätzlich in Worten beschrieben.

1. Initialisierung:

Bevor der hier vorgestellte Optimierungsprozess gestartet werden kann, müssen diverse Größen festgelegt werden, wobei es sich um folgende Parameter handelt:

- *White noise gain* $WNG_{\text{dB}}(\omega)$:
 - $WNG_{\text{dBMin}} = -10$ [dB]: Minimale WNG-Schwelle in [dB], die vom regularisierten, modalen *beamformer* nicht unterschritten werden darf.
 - $\Delta WNG_{\text{dB}} = 0.5$ [dB]: Offsetwert in [dB], um welchen die minimale WNG-Schwelle WNG_{dBMin} beim Entwurf maximal über- bzw. unterschritten werden darf.
- Regularisierungsparameter $\epsilon(\omega)$:
 - $\epsilon_{\text{Max}} = 1$: Obere Grenze, bis zu welcher der Regularisierungsparameter $\epsilon(\omega)$ maximal erhöht werden darf.
 - $\Delta\epsilon = 1e^{-5}$: Schrittweite, mit welcher der Regularisierungsparameter $\epsilon(\omega)$ verändert, d.h. erhöht bzw. erniedrigt, werden darf.
- Frequenz ω :
 - $\omega = 2\pi \cdot 1$ [Hz]: Startwert, d.h. Kreisfrequenz ω in $[\frac{1}{s}]$, ab welcher mit der Optimierung begonnen werden soll.
 - $\Delta\omega = 2\pi \cdot 1$ [Hz]: Inkrement in $[\frac{1}{s}]$, um welchen die Kreisfrequenz ω erhöht werden soll, wenn die Optimierung bei der aktuellen Kreisfrequenz ω erfolgreich beendet wurde.
 - $\omega_{\text{Max}} = \pi f_A$ [Hz]²³: Maximale Kreisfrequenz in $[\frac{1}{s}]$, bis zu welcher optimiert werden soll.

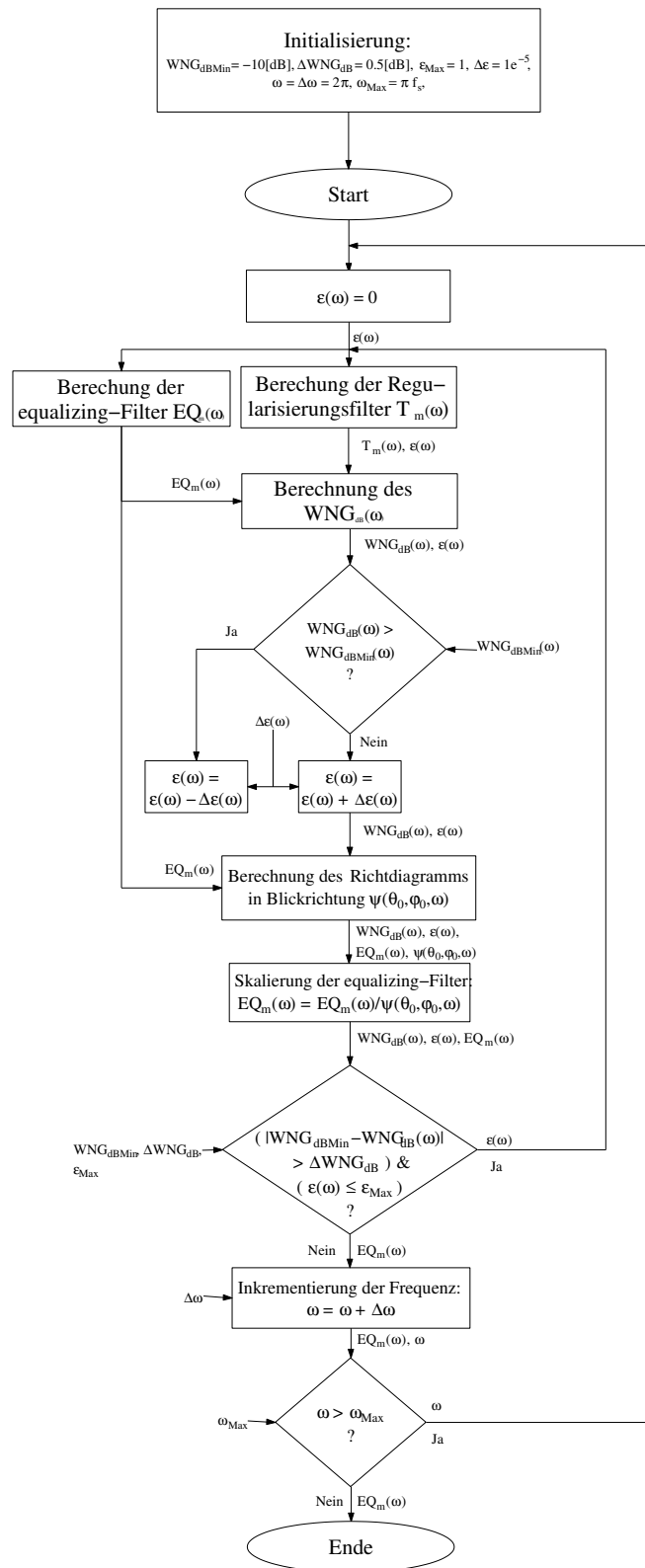


Abbildung 5.19: Flussdiagramm, welches den Ablauf zur Erzeugung der optimalen equalizing-Filter $EQ_m(\omega)$, dargestellt.

2. Festlegen des Startwertes des Regularisierungsparameters $\epsilon(\omega) = 0$ für die aktuell anstehende Kreisfrequenz ω .

Die Regularisierung bietet die Möglichkeit, mit Hilfe des Regularisierungsparameters $\epsilon(\omega)$ ein robustes Systemverhalten zu erreichen.

Dabei „erkauft“ man sich gewissermaßen mehr Robustheit, d.h. ein höheres $WNG_{\text{dB}}(\omega)$, zum Preis einer geringer werdenden Richtwirkung in Blickrichtung $\psi(\theta_0, \varphi_0, \omega)$, d.h. eines kleiner werdenden Array-Gewinns $G_{\text{dB}}(\omega)$.

Wird also, wie hier vorgeschlagen, als Startwert ein Regularisierungsparameter von $\epsilon(\omega) = 0$ verwendet, startet die Optimierung stets beim maximalen Bündelungsmaß $G_{\text{dBMax}}(\omega)$ und wird dann durch den steigenden Regularisierungsparameter solange verringert, bis der gewünschte WNG Schwellwert gegeben durch WNG_{dBMin} nicht mehr unterschritten wird.

Somit erklärt sich die am Ende des Kapitels 5.5.1 getätigte Aussage, dass eine explizite Berechnung des Array-Gewinns $G_{\text{dB}}(\omega)$ in dem hier vorgestellten Optimierungsverfahren nicht nötig ist.

3. Berechnung des WNG-Wertes $WNG_{\text{dB}}(\omega)$:
 - (a) Berechnung der Regularisierungsfiler $T_m(\omega)$ nach Gleichung 5.28 unter Verwendung des aktuellen Regularisierungsparameters $\epsilon(\omega)$.
 - (b) Berechnung der *equalizing*-Filter $EQ_m(\omega)$ nach Gleichung 5.28 oder Gleichung 5.29 unter Verwendung der aktuellen Regularisierungsfiler $T_m(\omega)$ oder des aktuellen Regularisierungsparameters $\epsilon(\omega)$.
 - (c) Berechnung des momentanen WNG-Wertes $WNG_{\text{dB}}(\omega)$, unter Verwendung der aktuellen Regularisierungsfiler $T_m(\omega)$ und *equalizing*-Filter $EQ_m(\omega)$ gemäß der Gleichungen 5.31-5.34.
4. Vergleich des aktuellen WNG-Wertes $WNG_{\text{dB}}(\omega)$ mit der eingangs festgelegten, minimalen Schwelle WNG_{dBMin} , die dieser nicht unterschreiten darf:

$$\epsilon(\omega) = \begin{cases} \epsilon(\omega) - \Delta\epsilon, & \text{wenn } WNG_{\text{dB}}(\omega) > WNG_{\text{dBMin}}, \\ \epsilon(\omega) + \Delta\epsilon, & \text{sonst.} \end{cases}$$

5. Skalierung der *equalizing*-Filter $EQ_m(\omega)$, sodass das Richtdiagramm des modalen *beamformer* bei Schalleinfall aus der Blickrichtung frequenzunabhängig stets einen Wert von $\Psi(\Theta_0, \varphi_0, \omega) = 0$ [dB] aufweist:

²³Im Flussdiagramm aus Abbildung 5.19 wird hier die Bezeichnung f_s verwendet.

- (a) Berechnung von $\Psi(\Theta_0, \varphi_0, \omega)$, d.h. des Richtdiagramms des modalen *beamformer* bei Schalleinfall aus der Blickrichtung nach Gleichung 5.37 unter Verwendung der aktuellen *equalizing*-Filter $EQ_m(\omega)$ aus Schritt 3b.
- (b) Skalierung der *equalizing*-Filter: $EQ_m(\omega) = \frac{EQ_m(\omega)}{\Psi(\Theta_0, \varphi_0, \omega)}$.
6. Abfrage, ob der aktuelle WNG-Wert $WNG_{dB}(\omega)$ die vorgegebene WNG-Schwelle WNG_{dBMin} erreicht hat ($|WNG_{dBMin} - WNG_{dB}(\omega)| > \Delta WNG$) und ob der aktuelle Regularisierungsparameter $\epsilon(\omega)$ seinen maximalen Wert ϵ_{Max} erreicht hat ($\epsilon(\omega) \leq \epsilon_{Max}$).

Wenn beide Bedingungen erfüllt werden, d.h. wenn: $(|WNG_{dBMin} - WNG_{dB}(\omega)| > \Delta WNG) \wedge (\epsilon(\omega) \leq \epsilon_{Max})$, dann ist der Optimierungsprozess noch nicht beendet, weshalb zum Schritt 3, mit dem aktuellen Regularisierungsparameter $\epsilon(\omega)$, zurückgekehrt werden muss.

Andernfalls, d.h. wenn der Optimierungsprozess für die aktuell anstehende Kreisfrequenz ω erfolgreich beendet wurde, indem entweder der aktuelle WNG-Wert auf dem vorgesehenen, minimalen Schwellwert begrenzt wurde oder der aktuelle Regularisierungsparameter seinen maximalen Wert erreicht hat, darf die Kreisfrequenz um $\Delta\omega$ inkrementiert und dann zum nächsten Schritt 7 weiter gegangen werden.

7. Abfrage, ob die aktuelle Kreisfrequenz ω ihren maximalen Wert, gegeben durch ω_{Max} , erreicht hat.

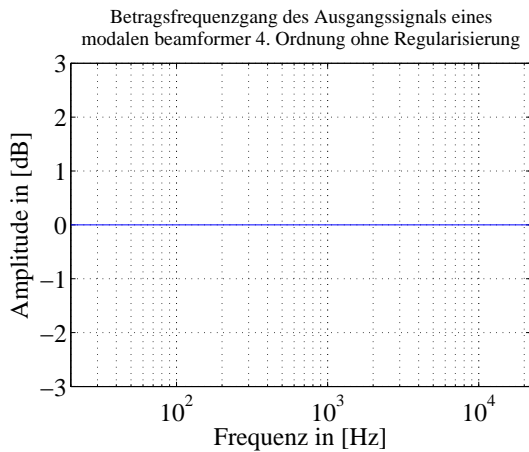
Wenn gilt: $\omega \leq \omega_{Max}$, d.h. wenn die maximale Kreisfrequenz ω_{Max} noch nicht erreicht wurde, wird unter Verwendung der aktuellen Kreisfrequenz ω zu Schritt 2 zurückgekehrt.

Andernfalls, d.h. wenn für alle Frequenzen das *equalizing*-Filter $EQ_m(\omega)$ erfolgreich optimiert wurde, können diese ausgegeben werden, wodurch der Optimierungsprozess beendet wird.

5.5.4 Simulationsergebnisse bei Anwendung unterschiedlicher *equalizing*-Filter

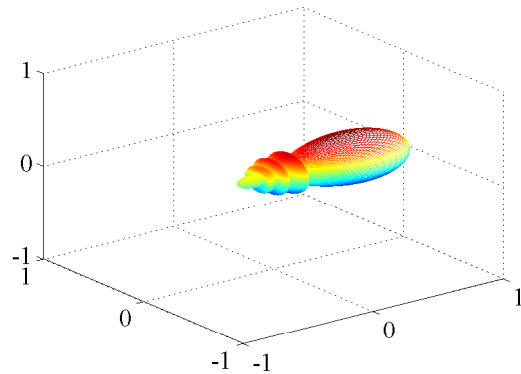
Wie die Darstellungen aus Abbildung 5.20 zeigen, sollte ein kugelförmiges Mikrophonarray idealerweise keine spektrale Abhängigkeit aufweisen. Das bedeutet, dass die Empfangsstärke in Blickrichtung über der Frequenz konstant bei $\Psi(\Theta_0, \varphi_0, \omega) = 0$ [dB] liegt, was aus Teilbild 5.20(a) ersichtlich ist. Zudem ändert sich die Richtcharakteristik über der Frequenz nicht, was Teilbild 5.20(d) verdeutlicht.

Berechnet man gemäß der in Abschnitt 5.5.1 angegebenen, Vorschrift das WNG für unterschiedliche Ordnungen m , resultieren für das hier beispielhaft verwendete Kugelmikrophon, die in Abbildung 5.21 dargestellten, frequenzabhängigen Verläufe $WNG_{dB}(\omega)$. Dabei ist zu erkennen, dass bei einer Ordnung von $m = 0$ die höchste Robustheit im



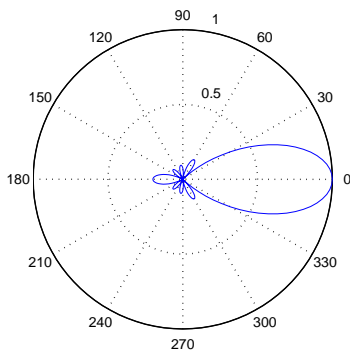
(a) Betragsfrequenzgang des idealen *beamformer* in Blickrichtung.

3D Darstellung eines nicht-regularisierten, modalen beamformer 4. Ordnung, welcher nach $\Theta=0^\circ$, $\phi=0^\circ$, bei einer Frequency von $f=100$ [Hz], zeigt



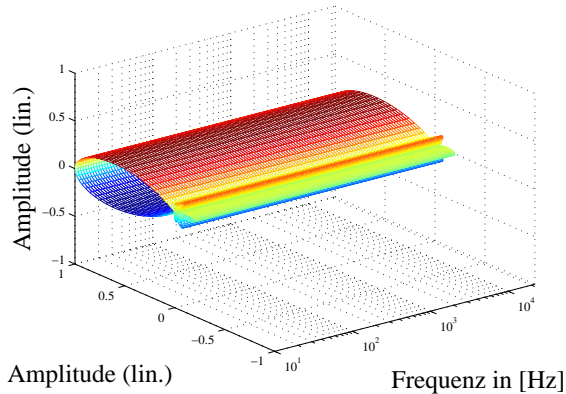
(b) 3D-Darstellung des idealen *beamformer* in Blickrichtung, bei einer Frequency von $f = 100$ [Hz].

Polardiagramm des beamformer ohne Regularisierung bei einer Frequency von $f=100$ [Hz]



(c) Polardiagramm des idealen *beamformer* in Blickrichtung, bei einer Frequency von $f = 100$ [Hz].

Zylindrische 3D-Darstellung der Richtdiagramme des nicht-regularisierten, modalen beamformer 4. Ordnung über der Frequency



(d) Zylindrische 3D-Darstellung der Polardiagramme des idealen *beamformer*, über den kompletten Audiobereich von $20 \text{ [Hz]} \geq f \geq \frac{f_A}{2} \text{ [Hz]}$.

Abbildung 5.20: Darstellung des idealen, nicht regularisierten beamformer, bei dem die verwendeten equalizing-Filter $EQ_m(\omega)$ nach Gleichung 5.6 berechnet wurden.

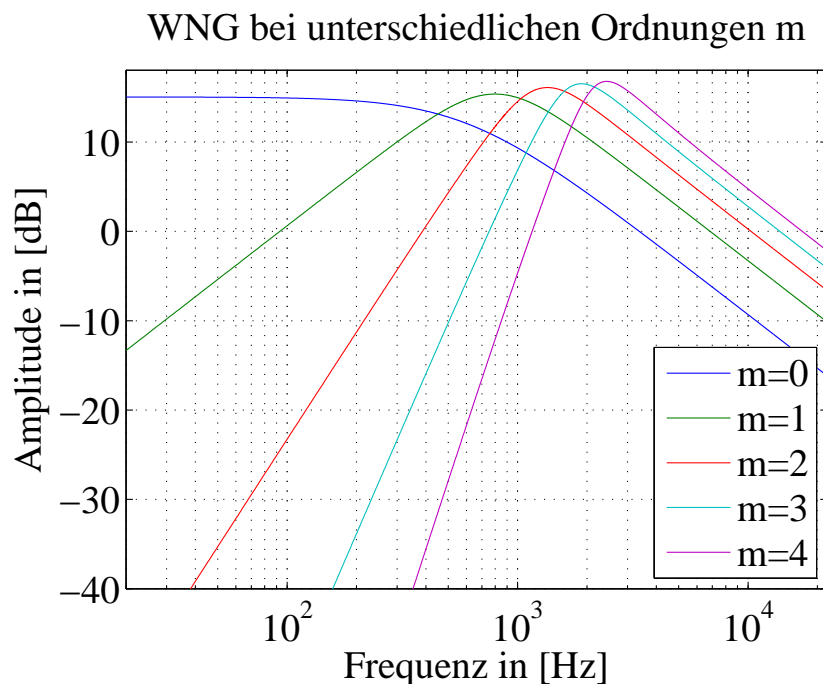
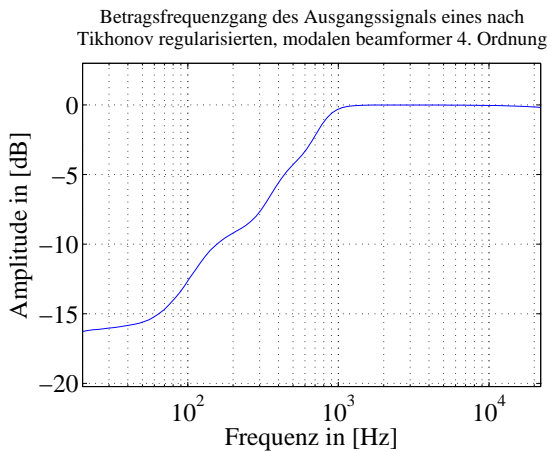


Abbildung 5.21: WNG über der Frequenz $WNG_{dB}(\omega)$, bei Verwendung des idealen beamformer unterschiedlicher Ordnung ($m = [0, \dots, 4]$).

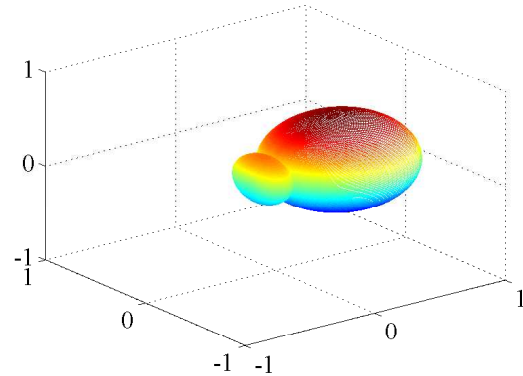
unteren Spektralbereich resultiert, die dann mit steigender Frequenz kontinuierlich abnimmt. Die Begründung hierfür ist, dass bei höheren Frequenzen, wie Abbildung 5.14 zeigt, mehr und mehr Moden bei der Zusammensetzung des Wellenfeldes beteiligt sind, die bei Verwendung eines Systems 0. Ordnung nicht mehr berücksichtigt werden, wodurch der Rekonstruktionsfehler steigt. Somit erklärt sich, warum ein System 0. Ordnung im Vergleich zu Systemen höherer Ordnung bei höheren Frequenzen ein geringeres WNG aufweist oder umgekehrt, warum Systeme höherer Ordnung bei höheren Frequenzen zunehmend stabiler werden. Ferner erkennt man, dass zumindest bei logarithmischer Betrachtung der Spektralbereich, in denen hohe WNG-Werte erreicht werden, sich mit steigender Ordnung nach oben verschiebt und zunehmend schmalbandiger wird, dafür jedoch das Maximum des WNG-Verlaufes leicht ansteigt. Somit veranschaulicht Abbildung 5.21, dass sich die Robustheit eines Kugelmikrophons hin zu tiefen Frequenzen steigern lässt, indem schrittweise die Ordnung verringert wird.

Wendet man die in den Gleichungen 5.28-5.29 vorgestellten Regularisierungsfiler nach Tikhonov, die hier mit einem Regularisierungsparameter von $\epsilon = 0.0039$, was einem maximalen Sensorrauschen von $\sigma_q = 3$ [dB] entspricht, an, erhält man die in Abbildung 5.22 dargestellten Ergebnisse. Dabei zeigt Teilbild 5.22(a) auf, wie sich der Betragsfrequenzgang in Blickrichtung durch die Regularisierungsfiler verändert. Wie man sieht, nimmt hin zu tieferen Frequenzen die Empfangsstärke in Blickrichtung kontinuierlich mit der Beschränkung auf niedrigere Ordnungen ab oder anders ausgedrückt, ab der Frequenz, ab welcher keine Regularisierung, d.h. keine Reduzierung der Ordnung M , für welche der



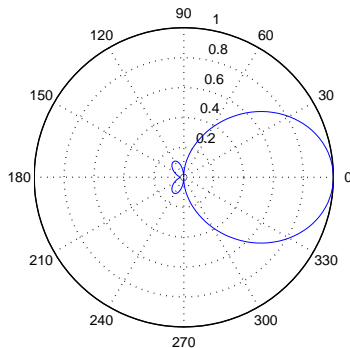
(a) Betragsfrequenzgang des *beamformer* in Blickrichtung.

3D-Darstellung eines nach Tikhonov regularisierten, modalen beamformer 4. Ordnung, welcher nach $\Theta=0^\circ$, $\phi=0^\circ$, bei einer Frequenz von $f=100$ [Hz], zeigt



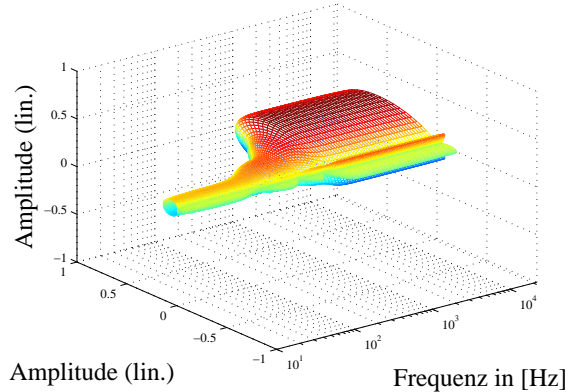
(b) 3D-Darstellung des *beamformer* in Blickrichtung, bei einer Frequenz von $f = 100$ [Hz].

Polardiagramm eines nach Tikhonov regularisierten, modalen beamformer 4. Ordnung bei einer Frequenz von $f=100$ [Hz]



(c) Polardiagramm des *beamformer* in Blickrichtung, bei einer Frequenz von $f = 100$ [Hz].

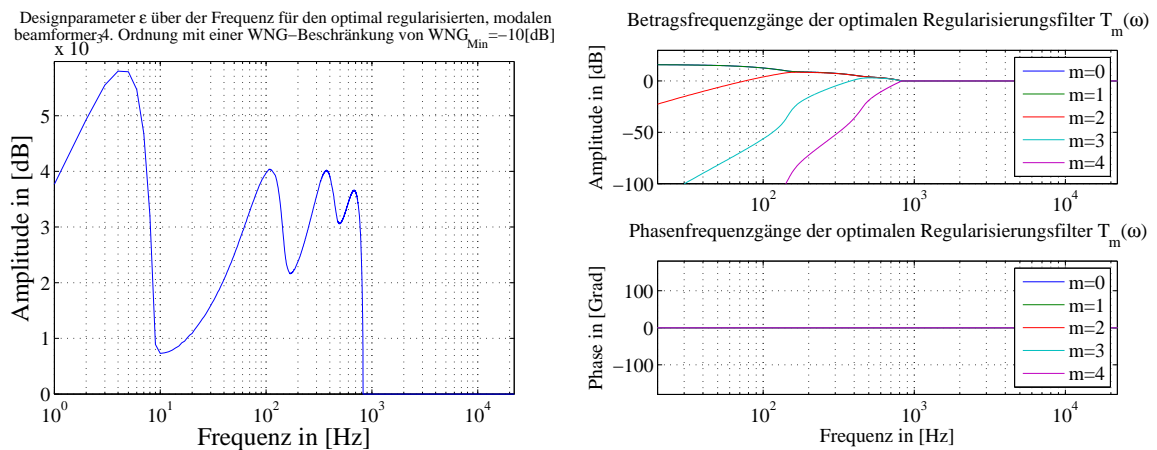
Zylindrische 3D-Darstellung der Richtdiagramme eines nach Tikhonov regularisierten, modalen beamformer 4. Ordnung über der Frequenz



(d) Zylindrische 3D-Darstellung der Polardiagramme des *beamformer*, über den kompletten Audiobereich von 20 [Hz] $\geq f \geq \frac{f_A}{2}$ [Hz].

Abbildung 5.22: Darstellung des mit Tikhonov-Filter mit $\epsilon = 0.0039$ ($\sigma_q = 3$ [dB]) regularisierten, *beamformer*, bei dem die equalizing-Filter $EQ_m(\omega)$ nach Gleichung 5.28 bzw. 5.29 berechnet wurden.

beamformer entworfen wurde, mehr nötig ist, was in unserem Beispiel bei einer Frequenz von $f \approx 1$ [kHz] der Fall ist, weist auch der Betrag in Blickrichtung einen konstanten Wert von $\Psi(\Theta_0, \varphi_0, \omega) = 0$ [dB] auf. Die Teilbilder 5.22(b) und 5.22(c) stellen das Richtverhalten des regularisierten *beamformer* bei einer Frequenz von $f = 100$ [Hz] im drei- als auch im zweidimensionalen Bereich dar. Dabei offenbart sich der Tribut, den man durch die Regularisierung in Form eines zu tiefen Frequenzen stetig abnehmenden Bündelungsmaßes zahlen muss. Die zylindrische Darstellung des Richtverhaltens des regularisierten *beamformer*, welche in Teilbild 5.22(d) ersichtlich ist, liefert zum einen einen Einblick, wie sich der Pegel in Blickrichtung über der Frequenz ändert, zum anderen kann man aus ihr aber auch die Frequenzabhängigkeit des Bündelungsmaßes zumindest teilweise entnehmen.



(a) Optimierter, frequenzabhängiger Regularisierungsparameter $\epsilon(\omega)$ zur Einhaltung einer minimalen WNG-Schwelle von $WNG_{\text{dBMin}} = -10$ [dB].

(b) Bodediagramme der optimierten Regularisierungsfiler $T_m(\omega)$, die neben der Einhaltung der WNG-Schwelle dafür sorgen, dass sich in Blickrichtung ein konstanter Betragsfrequenzgang von $\Psi(\Theta_0, \varphi_0, \omega) = 0$ [dB] einstellt.

Abbildung 5.23: Darstellung des optimierten Regularisierungsparameters $\epsilon(\omega)$ sowie der Bodediagramme der optimierten Regularisierungsfiler $T_m(\omega)$.

Obwohl die Ergebnisse der einfachen, d.h. mit einem frequenzunabhängigen Regularisierungsparameter ϵ entworfenen, in Abbildung 5.22 angewendeten Tikhonov-Filer schon ziemlich gut sind, wollen wir dennoch das maximale aus einem gegebenen System herausholen. Aus diesem Grund ist eine Optimierung notwendig. Wird dabei das, in Kapitel 5.5.3 vorgestellte Optimierungsverfahren angewendet, resultiert bei Verwendung einer WNG-Schwelle von $WNG_{\text{dBMin}} = -10$ [dB] der in Abbildung 5.23(a) dargestellte, frequenzabhängige Verlauf für den Regularisierungsparameter $\epsilon(\omega)$. Wie man sieht entspricht dabei der zuvor angewendete, frequenzunabhängige Regularisierungsparameter von $\epsilon = 0.0039$ bei Vernachlässigung des untersten Spektralbereiches in etwa dem Maximum des in Abbildung 5.23(a) aufgezeigten Verlaufes. Somit ist klar, dass der optimierte, frequenzabhängige Regularisierungsparameter $\epsilon(\omega)$ in einen *beamformer* mit einem höheren Bündelungsmaß im tieffrequenten Bereich resultiert, da bekanntermaßen ein kleiner werdender Regula-

risierungsparameter hin zum idealen *beamformer* tendiert. Abbildung 5.23(b) stellt die Bodediagramme der optimierten Regularisierungsfiler $T_m(\omega)$ dar, die überlagert mit den in Abbildung 5.3 aufgezeigten, idealen *equalizing*-Filter zu den in Abbildung 5.24 ersichtlichen, optimalen, regularisierten *equalizing*-Filter $EQ_m(\omega)$ werden. Wie man sieht, handelt es sich bei den optimalen Regularisierungsfiler $T_m(\omega)$ um *nullphasige Filter*, die gemäß der Ausgestaltung des Optimierungsalgorithmus so ausgelegt wurden, dass in Blickrichtung stets ein Betrag von 0 [dB] erreicht wird, deren Wirksamkeit ein Blick auf Abbildung 5.25(a) bestätigt.

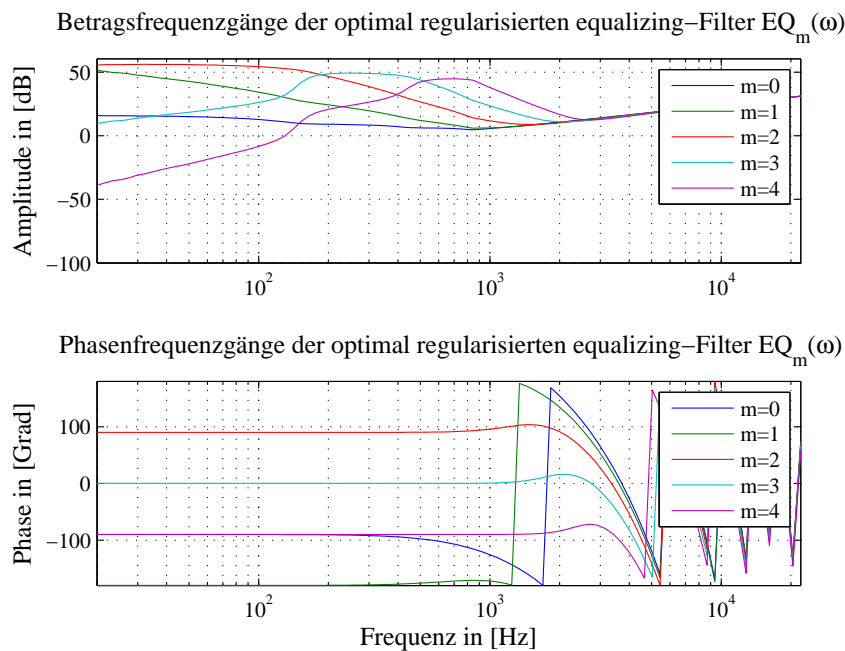
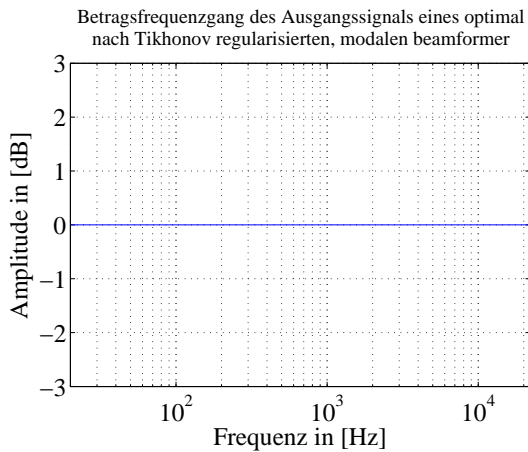


Abbildung 5.24: Bodediagramm der optimalen, regularisierten *equalizing*-Filter $EQ_m(\omega)$ der Ordnungen $m = [0, \dots, 4]$.

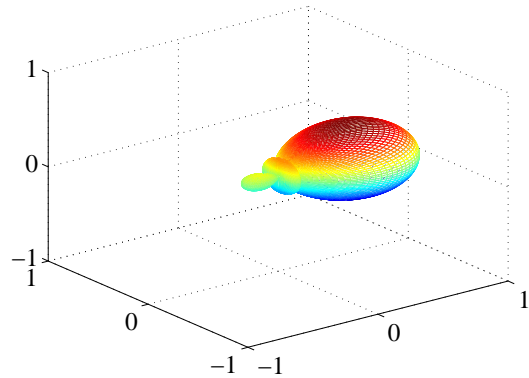
Ein Vergleich der Richtcharakteristika des *beamformer* bei $f = 100$ [Hz], welcher mit einem frequenzunabhängigen Regularisierungsparameter ϵ entworfen wurde, also von Abbildung 5.22(b) und Abbildung 5.22(c) mit dem *beamformer*, welcher mit dem optimierten, frequenzabhängigen Regularisierungsparameter $\epsilon(\omega)$ aus Abbildung 5.23(a) konzipiert wurde, dessen Ergebnisse in den Abbildungen 5.25(b) und 5.25(c) dargestellt sind, ergibt, dass durch die Optimierung vor allem im unteren Spektralbereich eine bessere Richtwirkung erzielt werden konnte. Dieses Ergebnis spiegelt sich jedoch nicht bei Betrachtung der im unteren Teilbild von Abbildung 5.27 dargestellten Array-Gewinne $G(\omega)$ ²⁴ der beiden *beamformer* wider. Die Lösung dieses Problems findet sich bei genauerer Betrachtung z.B. in den, in Abbildung 5.22(c) und Abbildung 5.25(c) dargestellten, Richtdiagrammen. Aus

²⁴Der Array-Gewinn $G(\omega)$ wird in Abbildung 5.27 mit *DI*, was für engl. *Directivity Index* steht, bezeichnet.



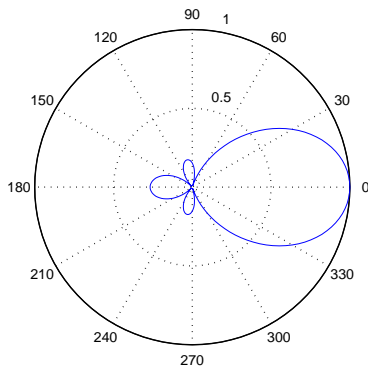
(a) Betragsfrequenzgang des *beamformer* in Blickrichtung.

3D-Darstellung eines optimal nach Tikhonov regulierten beamformer 4.Ordnung, welcher nach $\Theta=0^\circ, \phi=0^\circ$, bei einer Frequenz von $f=100$ [Hz], zeigt



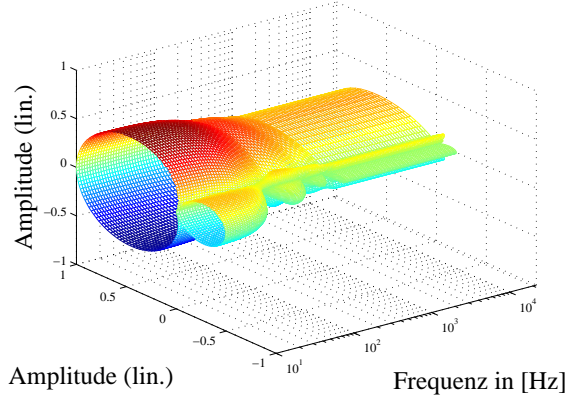
(b) 3D-Darstellung des *beamformer* in Blickrichtung, bei einer Frequenz von $f = 100$ [Hz].

Polardiagramm des beamformer bei einer Frequenz von $f=100$ [Hz]



(c) Polardiagramm des *beamformer* in Blickrichtung, bei einer Frequenz von $f = 100$ [Hz].

Zylindrische 3D-Darstellung eines optimal nach Tikhonov regulierten, modalen beamformer 4. Ordnung über der Frequenz



(d) Zylindrische 3D-Darstellung der Polardiagramme des *beamformer*, über den kompletten Audiobereich von 20 [Hz] $\geq f \geq \frac{f_A}{2}$ [Hz].

Abbildung 5.25: Darstellung des mit den optimierten Tikhonov-Filter regulierten beamformer, welche nach dem in Abschnitt 5.5 vorgestellten Verfahren entworfen wurden.

diesen ist zu erkennen, dass der optimierte *beamformer* zwar eine schmalere, in Blickrichtung zeigende Hauptkeule hat, aber dafür eine zusätzliche Nebenkeule aufweist, die in die entgegengesetzte Richtung deutet, welche der *beamformer* aus Abbildung 5.22 nicht aufweist. Somit erklärt sich, dass resultierend beide *beamformer* in etwa einen ähnlichen Array-Gewinn $G(\omega)$ bei dieser Frequenz aufweisen. Nun da der optimierte *beamformer* einen konstanten Betrag von 0 [dB] in Blickrichtung über der Frequenz aufweist, kann aus der zylindrischen Darstellung der Polardiagramme über der Frequenz, welche in Abbildung 5.25(d) dargestellt ist, problemlos entnommen werden, wie sich die Richtcharakteristik über der Frequenz bedingt durch die Regularisierung ändert.

Draufsicht auf die zylindrische 3D-Darstellung eines optimal nach Tikhonov regularisierten, modalen *beamformer* 4. Ordnung

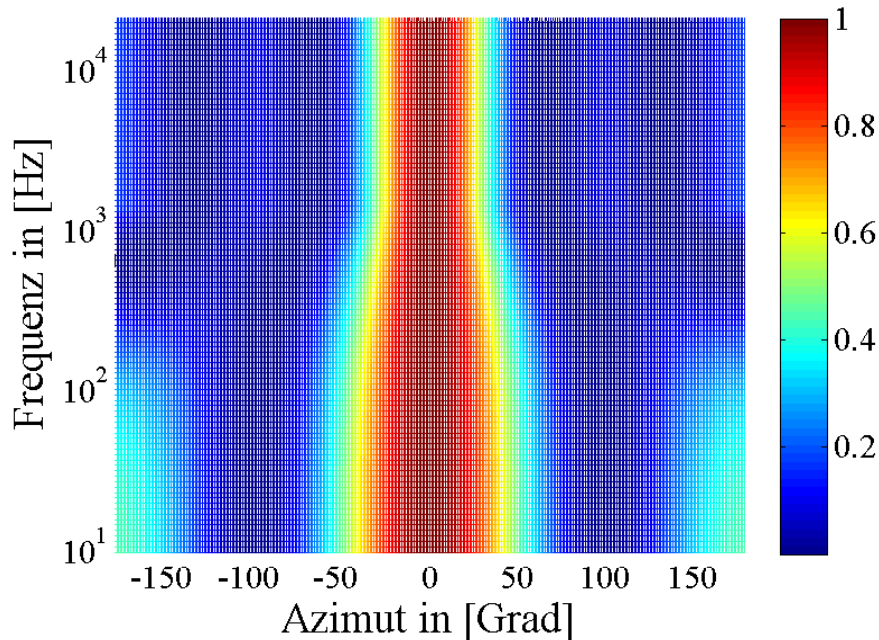


Abbildung 5.26: Draufsicht auf die zylindrische 3D-Darstellung eines optimal nach Tikhonov regularisierten, modalen *beamformer* 4. Ordnung.

Noch einfacher lässt sich das Verhalten der Richtcharakteristik über der Frequenz des optimal, nach Tikhonov regularisierten *beamformer* aus Abbildung 5.26 erkennen, welche die Draufsicht auf dessen zylindrische 3D-Darstellung beinhaltet. Im vorherigen, in Abbildung 5.25(d) aufgezeigten Fall, d.h. beim *beamformer*, welcher mit einem konstanten Regularisierungsparameter ϵ entworfen wurde, konnte dies nur bei genauer Betrachtung erahnt werden.

Zur einfacheren Beurteilung wurden sämtliche dreidimensionalen, zylindrischen Darstellungen der Richtcharakteristika über eine logarithmische Frequenzachse dargestellt.

Zudem verdeutlicht Abbildung 5.27 wie sich einerseits das WNG $WNG_{dB}(\omega)$ und andererseits der Array-Gewinn $G(\omega)$ der drei hier behandelten *beamformer* voneinander unterscheiden. Dabei ist aus dem oberen Teilbild von Abbildung 5.27 ersichtlich, dass mit Hil-

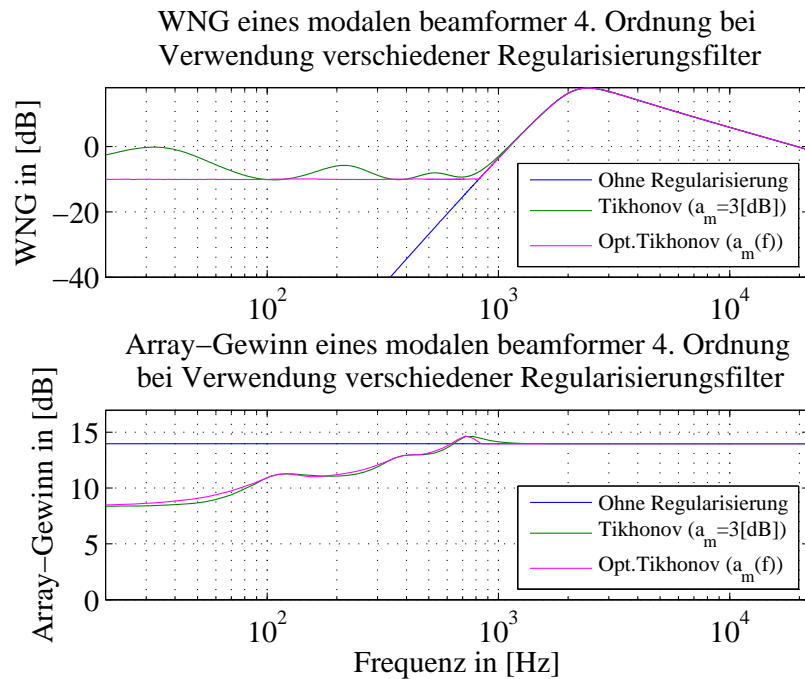


Abbildung 5.27: WNG $WNG_{dB}(\omega)$ und Array-Gewinn $G(\omega)$, des idealen, des mit einem Tikhonov-Filters mit $\epsilon = 0.0039$ entworfenen sowie des, mit dem optimierten Tikhonov-Filter regularisierten beamformer 4. Ordnung.

fe der Optimierung das gewünschte Ziel, ein minimales WNG von $WNB_{dB_{\min}} = -10$ [dB] nicht zu unterschreiten, voll erreicht wurde. Ferner ist zu erkennen, dass der frequenzunabhängige Regularisierungsparameter ϵ so gewählt wurde, dass das Minimum des hierbei resultierenden WNG-Verlaufes die gegebene WNG-Schwelle von $WNB_{dB_{\min}} = -10$ [dB] ebenfalls nicht unterschreitet, wodurch ein einfacher und fairer Vergleich beider Entwürfe ermöglicht wird. Außerdem verdeutlicht ein Vergleich der WNG- mit den Array-Gewinn-Verläufen der unterschiedlichen Entwurfsarten auf sehr einfache, weil offensichtliche Weise, wie sich eine Steigerung der Robustheit, d.h. eine Beschränkung des WNG, auf das Richtverhalten, d.h. auf das Bündelungsmaß, auswirkt.

Zusammenfassend kann konstatiert werden, dass sowohl die Regularisierung nach [BDM06], d.h. die Verwendung eines über der Frequenz konstanten Regularisierungsparameters ϵ , brauchbar, jedoch ein über der Frequenz optimierter Regularisierungsparameter $\epsilon(\omega)$ zu bevorzugen ist. Letzteres führt zu einem *beamformer*, der in der Lage ist, eine vorgegebene Robustheit, gegeben durch eine konstante WNG-Schelle, über der Frequenz einzuhalten, der einen frequenzunabhängigen, konstanten Wert in Blickrichtung aufweist und der über eine optimierte Richtcharakteristik in Blickrichtung verfügt, wobei sich dabei der resultierende Array-Gewinn $G(\omega)$ nicht maßgeblich vom vorherigen Entwurf mit fixem ϵ unterscheidet.

5.6 Praxisrelevante Faktoren die beim Bau eines Kugelmikrophonarrays beachtet werden sollten

5.6.1 Positionierung der Mikrophone auf der Kugeloberfläche

Es existieren viele Möglichkeiten, Mikrophone auf einer Kugeloberfläche zu positionieren. Erwähnenswert sind dabei eine winkelgleiche (*equiangle*), eine normalverteilte (*gaussian*) sowie eine regelmäßige (*uniform*) Anbringung der Sensoren, wobei unabhängig von der gewählten Anbringungsstrategie, stets die Orthonormalitätsbedingung gegeben durch Gleichung 5.7 möglichst exakt eingehalten werden muss. Untersuchungsergebnisse aus [Raf05] verdeutlichen jedoch, dass eine reguläre Anbringung zu bevorzugen ist, da diese, bei gleicher Anzahl an Mikrophenen im Gegensatz zu den beiden anderen Methoden (winkelgleiche und normalverteilte Anbringung) zu einem wesentlich geringeren Aliasingfehler führt. In eine ähnliche Richtung deuten auch die Ergebnisse aus [BDM06]. Dort wurde aufgezeigt, dass eine winkelgleiche $4(M+1)^2$ eine normalverteilte $2(M+1)^2$, jedoch eine regelmäßige Anbringung nur $(M+1)^2$ Sensoren²⁵ benötigt, um HOA-Komponenten bis zu einer Ordnung von M messen zu können. Somit lässt sich durch eine reguläre Anbringung die Anzahl der Mikrophone auf ein Minimum reduzieren, was hinsichtlich der Kosten, aber auch der Handhabbarkeit nicht zu unterschätzen ist. Leider existieren nur fünf Arten für eine exakte, reguläre Abtastung einer Kugeloberfläche. Dabei definieren sich die Abtastpunkte, d.h. die regulären Positionen der Sensoren auf der Kugeloberfläche, anhand der Flächenmittelpunkte der fünf *platonischen Körper*²⁶: *Tetraeder*(4), *Hexaeder*(6), *Oктаeder*(8), *Dodekaeder*(12), *Ikosaeder*(20), wobei die Ziffer in Klammern die Anzahl der Flächen des konvexen Polyeders und somit die Anzahl der verwendbaren Mikrophone, angibt. Weiterführende Details bezüglich platonischer und anderer quasi-regulärer Körper können in Kapitel A.5 nachgelesen werden.

Fehler bei der Positionierung der Mikrophone auf der Kugeloberfläche

Fehler bei der Positionierung der Sensoren können bedingt durch eine im Normalfall problemlos erreichbare, hohe mechanische Präzision im Vergleich zu anderen Fehlerarten, wie etwa deren Eigenrauschen, Aliasing oder aufgrund unterschiedlicher Übertragungsfunktionen der Sensoren, wie bereits in Kapitel 5.4.2 gezeigt, nahezu vernachlässigt werden.

²⁵Wie im Anhang A.5 verdeutlicht, ist diese Aussage nur approximativ und nicht zwingend gültig, d.h. eine Verwendung von Q regulär angebrachten Sensoren garantiert nicht, dass damit sphärische Harmonische bis zu einer Ordnung von $M = \lfloor \sqrt{Q} - 1 \rfloor$ bestimmt werden können. Meist trifft dies zwar zu, letztendliche Gewissheit erlangt man jedoch erst durch Überprüfung der verwendeten Anbringungsorte mit Hilfe der Orthonormalitätsbedingung.

²⁶Da die Flächenmittelpunkte der platonischen Körper die Eckpunkte ihrer dualen Körper kennzeichnen, können auch diese alternativ für die Bestimmung der Positionen der Sensoren auf der Kugeloberfläche verwendet werden.

5.6.2 Festlegung des Kugelradius

Die Größe des Mikrofonarrays, welche durch den Radius a der soliden Kugel bestimmt wird, kann nur als Kompromiss aufgefasst werden.

Einerseits gilt es, das räumliche Aliasing, gegeben durch Gleichung 5.8, zu berücksichtigen, welches anhand des Mikrofonabstandes d_{Mic} bzw. deren minimalen Winkeldifferenz γ , eine obere Grenzfrequenz f_{Alias} festlegt, bis zu welcher das Mikrofonarray eingesetzt werden kann, ohne dabei mit Artefakten rechnen zu müssen. Der Mikrofonabstand d_{Mic} bzw. die Winkeldifferenz zwischen zwei benachbarten Mikrofonen γ sollte hierbei möglichst gering sein, damit das Mikrofonarray auch bei hohen Frequenzen noch einwandfrei arbeiten und somit eingesetzt werden kann. Andererseits ist aus Kapitel 5.4.1 bekannt, dass der Verlauf der *equalizing*-Filter mit steigender Ordnung hin zu tiefen Frequenzen stetig zunimmt, wodurch das Eigenrauschen der Mikrophone sowie der nachfolgenden verwendeten Verarbeitungskette (Mikrophonvorverstärker, AD-Wandler, usw.) verstärkt wird. Bekanntermaßen ist dabei eine Extrahierung der HOA-Komponenten $B_{m,n}^\sigma$ nicht mehr möglich, sofern das auf diese Weise verstärkte Rauschsignal eine gewisse Schwelle überschreitet.

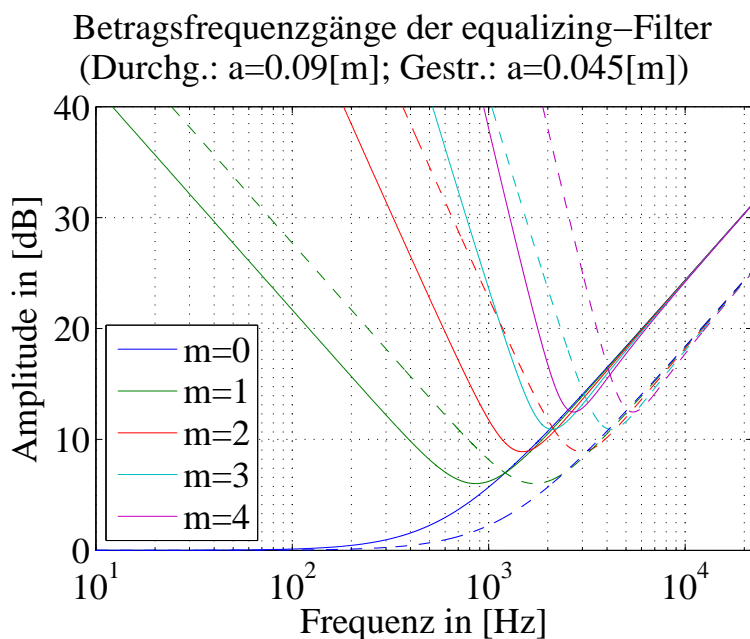


Abbildung 5.28: Vergleich der theoretischen *equalizing*-Funktionen $EQ_m(\omega)$ bis zu einer Ordnung von $M = 4$, bei Verwendung einer soliden Kugel mit einem Kugelradius von $a = 0.09 [\text{m}]$ (durchgezeichnete Linien) mit einer Kugel mit einem Kugelradius von $a = 0.045 [\text{m}]$ (gestrichelte Linien).

Je größer der Radius a der soliden Kugel gewählt wird, umso geringer fällt die durch die *equalizing*-Filter $EQ_m(\omega)$, mit $m = [0, \dots, M]$ hervorgerufenen Verstärkung im unteren Spektralbereich aus, was Abbildung 5.28 verdeutlicht, wodurch sich der Einsatzbe-

reich des Kugelmikrophons zu tiefen Frequenzen hin erweitert.

Ferner muss sich der Radius a des Mikrophonarrays zwischen $a = 0$ (kein räumliches Aliasing) und dem Abstand vom Mittelpunkt zum Rand des für die Decodierung verwendeten Lautsprecherrings $a = r_S$ (minimal mögliche Rauschverstärkung) befinden. Beide Werte sind dabei nur von theoretischer Bedeutung, da es einerseits nicht möglich ist, ein Mikrofon mit einem infinitesimal kleinen Radius herzustellen und andererseits die Abmessungen des Mikrophonarrays nur so groß werden darf, dass es problemlos z.B. in Autoinnenräumen eingesetzt werden kann.

Gemäß der Aufgabenstellung gilt es ein möglichst exaktes Abbild eines aufgenommen Schallfeldes zu erzeugen, welches an einem beliebigen Zielort möglichst präzise zu reproduzieren ist, wobei sowohl der Quell- als auch der Zielraum stets von einem Menschen, z.B. von einem Akustiker, beurteilt wird. Aufgrund der aus [CS10] bekannten Ergebnisse bezüglich der Einflussnahme des menschlichen Körpers auf die Lokalisierung wurde der Entschluss gefasst, dass der menschliche Körper im Allgemeinen bzw. der menschliche Kopf im Speziellen bei jeder dem Menschen betreffenden Messung berücksichtigt werden soll. Aus diesem Grund orientiert sich die Größe des hier verwendeten Mikrophonarrays am typischen Kopfradius eines erwachsenen Mannes, der nach [Pol05b] mit $a \approx 0.087$ [m] angegeben werden kann, sodass letztendlich die Wahl auf eine „solide“ Styroporkugel mit einem Radius von $a = 0.09$ [m] als Basis für das Kugelmikrofon fiel. Eine solche Kugel ist kostengünstig, leicht erhältlich und lässt sich zudem auf einfache Weise bearbeiten.

Für die Messungen des Schallfeldes wird das Kugelmikrophonarray auf eine kopflose Puppe anstelle des menschlichen Kopfes montiert, wodurch intrinsisch sämtliche Beugungen, Brechungen und Reflexionen des menschlichen Körpers berücksichtigt werden und die Messung dadurch einen höheren Praxisbezug erfährt. Zudem ist bekannt, dass sich der für die akustische Wahrnehmung (Klangfarbe, Lokalisierung, Umhülltheit, Bühnenbreite, usw.) mit Abstand relevanteste Spektralbereich unterhalb von $f \in [2, \dots, 4]$ [kHz] befindet. Da Fehler, die oberhalb dieses Bereiches auftreten als weniger störend empfunden werden, spricht auch dieser Aspekt für die Wahl des Kugelradius von $a = 0.09$ [m] bei welchen die räumliche Aliasingfrequenz f_{Alias} , gemäß Gleichung 5.8 mit $f_{\text{Alias}} \approx 2.9$ [kHz]²⁷ angegeben werden kann. Somit ist klar, dass der für die akustische Wahrnehmung mit Abstand wichtigste Spektralbereich durch eine solide Kugel mit Kopfdurchmesser d_{Head} ²⁸ angegeben werden kann.

²⁷Grundlage für die Berechnung ist die Verwendung des in Kapitel A.5.7 vorgestellten Ikosaederstumpfes. Nach [EKM03] weisen dabei die Mikrophone einen Abstand von $d_{\text{Mic}} = 0.65 a$ bzw. von $d_{\text{Mic}} = 0.73 a$ auf, wobei für die Berechnung vom besten anzunehmenden Fall (engl. *best case*), d.h. von einem Mikrofonabstand von $d_{\text{Mic}} = 0.65 a$ ausgegangen wurde. Bei Verwendung des schlechtest anzunehmenden Fall (engl. *worst case*) ergibt sich hingegen eine räumliche Aliasingfrequenz von $f_{\text{Alias}} \approx 2.58$ [kHz].

²⁸Der Kopfdurchmesser eines durchschnittlichen, erwachsenen Mannes beträgt $d_{\text{Head}} \approx 0.175$ [m].

5.6.3 Möglichkeiten zur Erhöhung der räumlichen Abtastfrequenz

Nun da wir wissen, dass ein solides Kugelmikrophonarray mit $Q = 32$ Mikrofonen, welches auf einem Ikosaederstumpf, wie in Kapitel A.5.7 angegeben, basiert, dessen Größe sich an der Geometrie des menschlichen Kopfes orientiert, also einen Radius von $a = 0.09$ [m] hat, eine maximale, räumliche Abtastfrequenz von $f_{\text{Alias}} \approx 2.5$ [kHz] aufweist, werden nachfolgend Möglichkeiten diskutiert, wie sich diese, unter Beibehaltung der vorgegebenen Geometrie des Arrays, zu höheren Frequenzen hin erweitern lässt.

Erhöhung der Mikrofonanzahl

Als wohl einfachste Möglichkeit bietet es sich an die Anzahl der Sensoren zu erhöhen. Wie etwa in [Pol05a] erwähnt und bereits in Kapitel 5.4.2 diskutiert, führt eine Erhöhung der Anzahl der Sensoren, bei unveränderter Geometrie, zu einer Verringerung des Abstandes des Sensoren zueinander und somit gemäß Gleichung 5.8 zu einer Erhöhung der räumlichen Aliasing Frequenz f_{Alias} . Allerdings bedeutet dies auch einen entsprechenden Mehraufwand sowohl bei der Erstellung des Kugelmikrophonarrays, als auch der notwendigen Signalverarbeitung, da hierdurch die Anzahl der Eingangskanäle entsprechend steigt.

Räumliche Tiefpassfilterung

Ferner ist es in diesem Zusammenhang hilfreich, zunächst noch einmal einen Blick auf Abbildung 5.14 zu werfen, in welcher die radialen Funktionen $W_m(ka)$ bis zu einer Ordnung von $m = 10$ für das hier behandelte Mikrophonarray aufgezeigt werden. Hierbei fällt auf, dass hin zu höheren Frequenzen der Einfluss sphärischer Harmonischer höherer Ordnung stetig zunimmt. Gelingt es diesen Einfluss zu unterbinden, was einer *räumlichen Tiefpassfilterung* gleichkommt, kann auch der spektrale Einsatzbereich des Mikrophonarrays zu höheren Frequenzen hin erweitert werden. Nachfolgend werden mehrere Methoden zur räumlichen Tiefpassfilterung aufgeführt. Eine fundierte, theoretische Untersuchung bezüglich des Entwurfs räumlicher Tiefpassfilter ist zudem in [RWB07] zu finden.

- Verwendung von Mikrofonen mit einer großflächigen Membran

Wie bereits in [ME02, EKM03] erwähnt und in [DE08] näher untersucht, kommt es bei Verwendung von Mikrofonen mit großer Membranfläche, aufgrund der dort stattfindenden Integralwirkung, zur Ausbildung einer räumlichen Tiefpassfilterung. Um dabei jedoch einen merklichen Erfolg zu gewähren, müsste die Membranfläche enorm gesteigert werden, was in der Praxis so nicht ohne weiteres machbar ist. Aus diesem Grund wird in [EKM03] auf sogenannte *patch*-Mikrophone eingegangen. Diese bestehen lediglich aus einer druckempfindlichen Membran, die in beliebiger Größe auf einer nahezu beliebigen Oberfläche aufgebracht werden kann. Leider

weisen die derzeit erhältlichen Materialien, mit denen ein solches *patch*-Mikrofon realisiert werden könnte, wie etwa das in [LP99] vorgestellte EMFi²⁹-Material, mit $50 \text{ [dBA]} @ A = 3.1 \text{ [cm]}^2$ noch ein sehr hohes Eigenrauschen auf. Bei großflächigen, „klassischen“ Mikrofonen, wie etwa bei Elektretmikrofonen gilt dies jedoch nicht. Hier weisen im Allgemeinen großflächige Mikrophone tendenziell ein geringeres Eigenrauschen als Mikrophone mit kleinerer Membranfläche auf. Allerdings weisen großflächige Mikrophone einen nach oben hin eingeschränkten spektralen Einsatzbereich auf, was es bei der Wahl der Mikrophone zu berücksichtigen gilt.

- Verwendung von mehreren Mikrofonen pro Polyederfläche

Bis dato wurde davon ausgegangen, dass jeweils nur ein Sensor, egal ob nun mit großer oder kleiner Membranfläche, im Zentrum der jeweiligen Polyederflächen positioniert wird. Wie jedoch im vorherigen Punkt erwähnt, ist es im Sinne des Rauschens besser, Mikrophone mit kleiner Membranfläche zu verwenden, obwohl eine große Fläche benötigt wird um die räumliche Aliasingfrequenz zu erhöhen. Die Membranfläche lässt sich allerdings auch erhöhen, indem im bzw. um das Zentrum der jeweiligen Polyederflächen mehrere Mikrophone mit kleinerer Membranfläche positioniert werden, was beispielsweise bereits in [Hul04] für ein zirkulares Mikrophonarray praktiziert wurde. Hierbei kamen insgesamt 288 Mikrophone zum Einsatz, von denen jeweils $Q_p = 24$ mit Hilfe einer aktiven, analogen Summierschaltung zusammengefasst wurden, wodurch lediglich $Q = 12$ zu verarbeitende Ausgangskanäle verblieben.³⁰ Durch diese Maßnahme lässt sich nicht nur wie gewünscht, die Membranfläche und somit die räumliche Aliasingfrequenz erhöhen, sondern auch, wie in Kapitel 5.4.2 aufgezeigt, das Rauschen reduzieren, wobei sich die Rauschreduktion, im Vergleich zu einem einzelnen Sensor, wie folgt berechnet:

$$10 \log_{10} (Q_p) \text{ in [dB]}. \quad (5.41)$$

Das reduzierte Rauschen wiederum führt zu einer Entspannung beim Entwurf der Regularisierungsfiler, wodurch sich der spektrale Einsatzbereich des Mikrophonarrays auch zu tiefen Frequenzen hin erweitert. Zudem bewirkt die Mittelwertbildung über die Anzahl der Mikrophone pro Polyederfläche, dass ein Ausgleich der Exemplarstreuungen der resultierenden Eingangskanäle, d.h. der Mittelwerte, entfallen kann, da sich diese nun wesentlich geringer voneinander unterscheiden als die, der ansonsten verwendeten, einzelnen Mikrophone. Zusammenfassend stellt diese Methode zwar eine gewisse Herausforderung bei der Erstellung des Mikrophonarrays dar, bietet ansonsten jedoch eine durchaus praktikable Lösung zur Performancesteigerung nicht nur im oberen, sondern auch im unteren Spektralbereich. Darüber hinaus führt diese Maßnahme zu keiner Erhöhung der zu bearbeitenden Eingangskanäle sowie der damit einhergehenden Signalverarbeitung.

- Modifikation der Kugeloberfläche

In [DE08] wurde der Frage nachgegangen, inwieweit Dellen in der Kugeloberfläche zu einer räumlichen Tiefpassfilterung führen. Die dabei zugrundeliegende Idee ist so simpel wie genial: Ein am Fusspunkt einer Vertiefung, die z.B. die Form einer Satellitenschüssel, d.h. einer Parabolantenne, aufweist, welche auf der Kugeloberfläche eingebracht wird, befindliches Mikrofon, kann nur Signale empfangen, die entweder im Verhältnis zum Durchmesser der Delle eine wesentlich größere Wellenlänge besitzen oder, wenn die Wellenlänge kleiner als der Durchmesser der Delle ist, aus Raumrichtungen, die senkrecht zum Fusspunkt der Delle, an dem sich das Mikrofon befindet, kommen. Die dabei in [DE08] erzielten Ergebnisse waren sehr vielversprechend und überstiegen bei weitem die Ergebnisse die mit einer vergrößerten Membranfläche erzielt werden konnten, was ebenfalls zumindest theoretisch untersucht wurde. Obwohl die Untersuchungen mit einer deformierten Kugeloberfläche auf Messungen basierten, da sich solche Modifikationen der Kugeloberfläche nicht oder nur sehr schwer simulatorisch erfassen lassen, fehlte der Untersuchung dennoch ein echter Praxisbezug, da sich die Untersuchungen auf eine einzige auf einer Styroporkugel eingebrachte, halbkugelförmige Delle, an deren Fusspunkt ein Schalldruckmikrofon angebracht war beschränkte. Dieses wurde in einer Schirmkammer vermessen und anschließend basierend auf diesen Messergebnissen für die Simulation auf die komplette Kugeloberfläche künstlich erweitert. Im Prinzip wäre dieses Vorgehen in Ordnung, auch wenn die Oberfläche nicht 32 sondern nur 1 Delle besessen hätte. Allerdings wurde der Durchmesser der Delle so gewählt, dass auf der Oberfläche in Wirklichkeit keine 32 Dellen dieses Ausmaßes Platz gefunden hätten, ohne dass sich deren Radien überschneiden hätten. Um ein aussagekräftiges, belastbares Ergebnis, welches auch in der Praxis bestand hat, zu erhalten, wurde ein Kugelmikrophonarray nach diesem Prinzip entworfen und vermessen. Die dabei erzielten Ergebnisse bestätigten im Wesentlichen die in [DE08] aufgezeigten Erkenntnisse. Durch diese einfache Maßnahme konnte die räumliche Abtastfrequenz von anfänglich $f_{\text{Alias}} \approx 3$ [kHz] in etwa verdoppelt, d.h. auf $f_{\text{Alias}} \approx 6$ [kHz], gesteigert werden, ohne dabei die Anzahl der Mikrophone oder die nachfolgende Signalverarbeitung zu erhöhen. Aus diesem Grund ist diese rein mechanisch erreichbare Maßnahme wohl die effizienteste Methode, die räumliche Aliasingfrequenz bei unveränderter Geometrie des Mikrophonarrays zu steigern. Für das Einbringen der Dellen auf die Kugeloberfläche erwies sich die Verwendung einer Styroporkugel ebenfalls als nützlich. Mit Hilfe einer Halbkugel aus Stahlblech mit einem Durchmesser von $\varnothing = 8$ [cm], welche heiß gemacht wurde, konnten unter Verwendung einer großen Standbohrmaschine mit Anschlag, welcher auf 1.1 [cm] eingestellt wurde mit einer ausreichend hohen mechanischen Präzision, Dellen mit einem Durchmesser von $\varnothing \approx 5.5$ [cm] in den Grundkörper, d.h. in die Styroporkugel eingedampft werden. Dabei verblieb ein minimaler Abstand von 0.35 [cm] zwischen zwei benachbarten Dellen gewahrt, sodass sich diese nicht überschneiden.

Beschränkung der Ordnung

Wie bereits aus [EM08] bekannt, bestätigt ein Blick auf Abbildung A.26(b), welche die Orthonormalitätsfehlermatrix des verwendeten Ikosaederstumpfes darstellt, dass die Gefahr, dass Moden höherer Ordnung durch andere Moden gestört werden können größer ist, als bei Moden niedrigerer Ordnung.

Während sphärische Harmonische 4. Ordnung bereits durch Moden 6. Ordnung gestört werden können, trifft dies bei sphärischen Harmonischen 1. Ordnung erst bei Moden 9. Ordnung zu.

Aus Abbildung 5.14, welche die radialen Funktionen über der Frequenz, bis zu einer Ordnung von $m = 10$ darstellt, ist zu entnehmen, dass die Mode 9. Ordnung erst ab einer Frequenz von $f > 4.1$ [kHz] ausreichend stark wird, um die Messung zu stören.

Somit lässt sich das bereits im unteren Spektralbereich angewendete Prinzip auch im oberen Spektralbereich nutzen, d.h. der störende Einfluss von Moden höherer Ordnung lässt sich vermeiden, indem mit steigender Frequenz die Ordnung der zu messenden sphärischen Harmonischen sukzessive reduziert wird.

Auf diese Weise, d.h. durch eine schrittweise Reduktion der Richtwirkung des ursprünglichen *beamformer* 4. Ordnung, welcher bekanntermaßen eine Aliasingfrequenz von $f_{\text{Alias}} \approx 2.5$ [kHz] hat, auf die Richtwirkung eines *beamformer* 1. Ordnung, lässt sich die räumliche Aliasingfrequenz auf $f_{\text{Alias}} \approx 4.1$ [kHz] steigern.

Beschränkt man sich gar auf die Messung der sphärischen Harmonischen 0. Ordnung, bei der es zu keiner Störung durch Moden höherer Ordnung mehr kommt, lässt sich der Frequenzbereich sogar bis zur Nyquist-Frequenz ($f_{\text{Alias}} = \frac{f_A}{2}$ [Hz]) erweitern.

Nutzung der natürlichen Abschattung einer soliden Kugel

Die in Abbildung 5.29 dargestellten Richtdiagramme entsprechen den am Äquator einer soliden Kugel auftretenden Schalldrücken bei unterschiedlichen Frequenzen, wenn diese von einer Schallquelle, die sich in einem Abstand von $r_{\text{SRC}} = 1$ [m] an dem Ort ($\Theta_{\text{SRC}} = 0^\circ, \varphi_{\text{SRC}} = 0^\circ$) befindet, beschallt wird. Dabei wurden sphärische Harmonische bis zu einer Ordnung von $m = 85$ berücksichtigt, um möglichst das vollständige Wellenfeld zu berücksichtigen. Eine weitere Steigerung der Ordnung war nicht möglich, da es bei Anwendung von noch höheren Ordnungen zu numerischen Problemen in der Simulation kam. Allerdings scheint die verwendete Ordnung ausreichend hoch zu sein, um die Richtwirkung, welche aufgrund der Abschattung der soliden Kugel entsteht, zu verdeutlichen.

Die Abschattung der soliden Kugel kann dabei durchaus mit der des menschlichen Kopfes verglichen werden, da beide in etwa die gleiche Geometrie aufweisen, was ein

²⁹Bei dem EMFi Material (engl. *Electro Mechanical Film*) handelt es sich um ein geladenes, zelluläres Polymer, welches piezoelektrische Eigenschaften aufweist.

³⁰ Q_p bezeichnet die Anzahl der Mikrophone pro Polyederfläche (*patch* \rightarrow Index p)

Richtdiagramme einer soliden Kugel mit Radius $r_m = 0.09[m]$,
 die sich aufgrund der natürlichen Abschattung ergeben

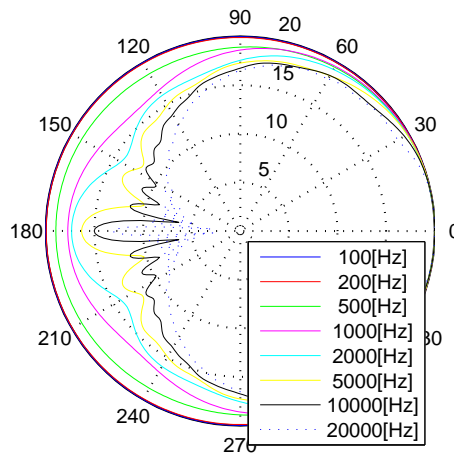


Abbildung 5.29: Polardiagramm, welches die Richtwirkung aufgrund der natürlichen Abschattung einer soliden Kugel mit einem Radius von $a = 0.09 [m]$ bei unterschiedlichen Frequenzen in [dB] aufzeigt.

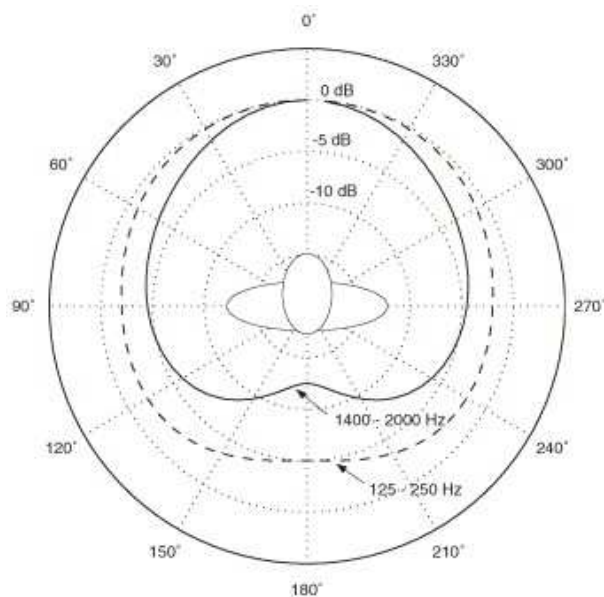


Abbildung 5.30: Abschattungseffekt des menschlichen Kopfes. [SH06]

Vergleich von Abbildung 5.29 mit Abbildung 5.30 verdeutlicht. Ferner fällt auf, dass der Abschattungseffekt der soliden Kugel bereits bei einer Frequenz von $f \in [2, \dots, 5]$ [kHz] als gut bezeichnet werden kann, wie ein Vergleich mit dem Verlauf z.B. bei $f = 10$ [kHz] zeigt. Somit wird deutlich, dass durch Ausnutzung des Abschattungseffektes ab einer Frequenz von $f = 4.1$ [kHz] durch Reduktion der Ordnung auf $m = 0$ das Mikrophon Array auch bis zur Nyquist-Frequenz betreiben kann, ohne dabei mit Fehlern, bedingt durch das räumliche Aliasing rechnen zu müssen. Dies ist jedoch nur möglich, wenn eine *solide* Kugel und keine offene oder halboffene Kugel verwendet wird, wodurch sich die Wahl des Grundkörpers abermals als vorteilhaft erwies.

5.6.4 Empfindlichkeit bezüglich der Exemplarstreuungen der Sensoren

In [Pol05a] wurde festgestellt, dass vor allem im unteren Spektralbereich, d.h. wenn die Wellenlänge Λ im Verhältnis zum Mikrophonabstand d_{Mic} groß ist und dementsprechend bereits geringe Phasenunterschiede über die Qualität des Mikrophonarrays bzw. des *beamformer* entscheiden, eine Anpassung der im Array verwendeten Mikrophone insbesondere bei tiefen Frequenzen notwendig ist. Damit die Qualität des resultierenden *beamformer* gesteigert werden kann, müssen sämtlichen modalen Keulen, deren Richtwirkungen die Form der verwendeten sphärischen Harmonischen haben, nach Anwendung der *equalizing*-Filter einen flachen Frequenzverlauf aufweisen. Dies ist jedoch nur möglich, wenn einerseits die einzelnen Sensoren einen flachen Frequenzgang aufweisen und sich andererseits die Übertragungsfunktionen der einzelnen Sensoren nicht bzw. nur unwesentlich, voneinander unterscheiden. Aus diesem Grund können zum einen vermessene, sogenannte „gematchte“³¹ Mikrophone verwendet werden oder die verwendete Mikrophone kalibriert, d.h. mit Hilfe von entsprechenden Ausgleichsfilter, auf elektronischem Wege, aneinander angepasst werden, wobei selbstverständlich auch eine Kombination der beiden Verfahren möglich ist.

Durchführung der Mikrophonkalibrierung

Die für die Kalibrierung, d.h. für die Bestimmung der *equalizing*-Filter pro Mikrophon, benötigten Messungen sollten in einer Schirmkammer durchgeführt werden, wobei ein möglichst breitbandiger Lautsprecher, der vor allem im unteren Spektralbereich über einen ausreichenden Pegel verfügt, verwendet werden sollte.

Dieser wird zunächst mit einem *Referenzmikrophon*³², welches sich an einem bestimm-

³¹Der umgangssprachliche Begriff „gematcht“ leitet sich aus dem englischen Wort *matched* ab, was hier die Bedeutung „anpassen“ oder „angleichen“ hat und beschreibt den Vorgang, dass zwei oder mehrere Objekte aus einer größeren Menge entnommen werden, deren Eigenschaften idealerweise identisch zueinander sind.

³²Als Referenzmikrophon wird hier ein Mikrophon bezeichnet, von dem ausgegangen wird, das es zumindest über den gewünschten Spektralbereich ein flaches Verhalten aufweist, von dem es nicht bzw.

ten Punkt im Raum, z.B. in einem Abstand von $d = 1$ [m], auf Achse zum verwendeten Lautsprecher befindet, vermessen.

Basierend auf der dabei resultierenden Übertragungsfunktion, welche im Wesentlichen die spektralen und zeitlichen Eigenschaften des verwendeten Lautsprechers beinhaltet, wird ein Ausgleichsfilter entworfen, sodass der verwendete Lautsprecher nach Anwendung dieses Ausgleichsfilters einen flachen Frequenzgang besitzt und somit zu einem *Referenzlautsprecher* wird.

Alle weiteren Messungen werden nun mit Hilfe des Referenzlautsprechers durchgeführt, d.h. das Messsignal wird, bevor es über den Lautsprecher abgespielt wird mit Hilfe des Ausgleichsfilters modifiziert, sodass der resultierende Frequenzgang am vorgegebenen Messpunkt bekannt, d.h. flach, ist.

Unter Verwendung des Referenzlautsprechers werden nun sämtliche Mikrophone des kugelförmigen Mikrophonarrays vermessen, wobei zum einen mit Hilfe eines Lasers sowie eines Meterstabs, an dem eine Wasserwaage angebracht ist, dafür gesorgt wird, dass deren örtliche Lage stets der des anfänglich verwendeten Referenzmikrophons entspricht und zum anderen für jedes Arraymikrophon, die speziell für dieses vorgesehene Signalverarbeitungskette bestehend aus Mikrophonvorverstärker und Soundkarte verwendet wird, damit auch deren Unterschiede beim Erstellen der Kalibrierfilter sofort mit erfasst und somit berücksichtigt werden.

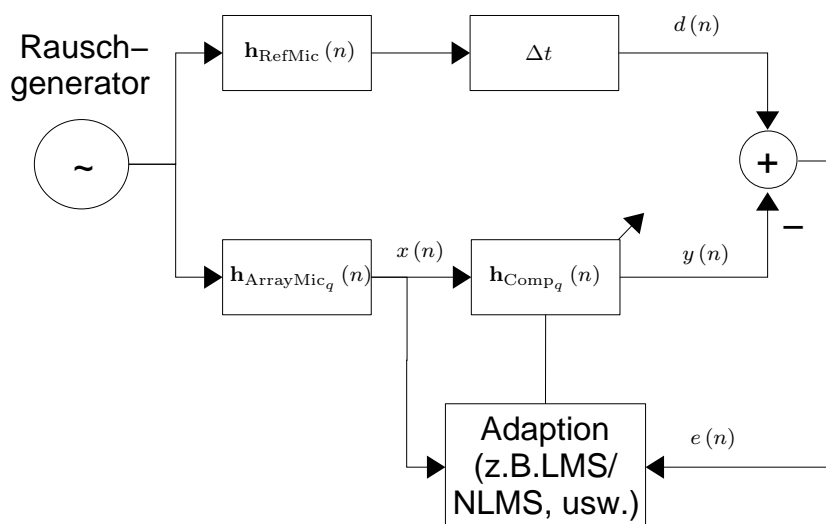


Abbildung 5.31: Bestimmung der Kalibrierfilter mit Hilfe eines adaptiven Filters.

Die Q Kalibrier- bzw. Kompensationsfilter $\mathbf{h}_{\text{Comp}_q}$ entstehen dann, indem die RIR vom Referenzlautsprecher zu jedem der Q Arraymikrophone $\mathbf{h}_{\text{ArrayMic}_q}$, mit $1 \leq q \leq Q$, auf die RIR vom Referenzlautsprecher zum Referenzmikrophon $\mathbf{h}_{\text{RefMic}}$, welche im Idealfall 1, d.h. nur in unbedeutendem Maße abweicht.

neutral ist, mit Hilfe eines adaptiven Filters, welches, wie in Abbildung 5.31 aufgezeigt, zur Systemmodellierung konfiguriert ist, abgeglichen werden.

Ferner kann das in Abbildung 5.31 aufgezeigte Verfahren auch zur Erstellung des Ausgleichsfilters, mit welchem das Messsignal modifiziert wird und somit der vorhandene Lautsprecher zu einem Referenzlautsprecher wird, verwendet werden. Dazu muss lediglich die RIR $\mathbf{h}_{\text{RefMic}}$ durch 1, also durch eine neutrale RIR und $\mathbf{h}_{\text{ArrayMic}_q}$ durch die, mit dem ursprünglichen Messsignal gemessene RIR $\mathbf{h}_{\text{RefMic}}$, ersetzt werden.

Der Block Δt in Abbildung 5.31, welcher eine Zeitverzögerung darstellt, soll verhindern, dass akausale Kompensationsfilter entstehen, deren Impulsantwort sich durch ein akustisch wahrnehmbares und somit störendes *pre-ringing*³³ auszeichnet. Gemäß [WS85] kann vermieden werden, dass $\mathbf{h}_{\text{Comp}_q}$ akausal wird, wenn die Zeitverzögerung Δt der halben Länge des für die Realisierung von $\mathbf{h}_{\text{Comp}_q}$ vorgesehenen FIR-Filters entspricht.

Mathematisch lässt sich dieses Verfahren, d.h. ein adaptives Filter im Zeitbereich, welches zur Systemmodellierung verwendet wird, wie folgt angeben ([WS85, Hay96]):

1. Filterung:

$$\begin{aligned}
 y(n) &= \mathbf{h}_{\text{Comp}_q}(n) \mathbf{x}(n), \text{ mit} \\
 \mathbf{x}(n) &= [x_0(n), x_1(n), \dots, x_{N-1}(n)]^T, \\
 \mathbf{h}_{\text{Comp}_q}(n) &= [h_{\text{Comp}_{q(0)}}(n), h_{\text{Comp}_{q(1)}}(n), \dots, h_{\text{Comp}_{q(N-1)}}(n)]^T, \text{ wobei} \\
 y(n) &= n. \text{ Sample des Ausgangssignals des adaptiven Filterkerns,} \\
 \mathbf{x}(n) &= \text{Eingangssignalvektor der Länge } N, \\
 \mathbf{h}_{\text{Comp}_q}(n) &= \text{Filterkern des } q. \text{ adaptiven Filters (FIR-Filterkoeffizienten)} \\
 &\quad \text{der Länge } N, \text{ mit } 1 \leq q \leq Q, \\
 Q &= \text{Anzahl der Mikrophone innerhalb des Mikrophonarrays,} \\
 N &= \text{Länge des FIR-Filters bzw. der Vektoren.}
 \end{aligned} \tag{5.42}$$

2. Berechnung des Fehlersignals $e(n)$:

$$\begin{aligned}
 e(n) &= d(n) - y(n), \text{ wobei} \\
 d(n) &= n. \text{ Sample des Referenzsignals (engl. } \mathbf{desired\ signal}), \\
 e(n) &= n. \text{ Sample des Fehlersignals (engl. } \mathbf{error\ signal}).
 \end{aligned} \tag{5.43}$$

3. Adaption:

³³Der Begriff *pre-ringing* stammt aus dem Englischen und bezeichnet hier das akustische Verhalten einer Impulsantwort, welche Koeffizienten besitzt, die zeitlich vor dem eigentlichen Hauptpuls erscheinen, was sich durch ein akustisch störendes *Vorklingeln* bemerkbar macht.

Adaptive Berechnung der Filterkoeffizienten des q . adaptiven Filterkerns $\mathbf{h}_{\text{Comp}_q}(n)$, nach dem NLMS³⁴-Algorithmus:

$$\begin{aligned} \mathbf{h}_{\text{Comp}_q}(n+1) &= \mathbf{h}_{\text{Comp}_q}(n) + \mu \frac{e(n) \mathbf{x}(n)}{\|\mathbf{x}(n)\|^2 + \epsilon}, \text{ wobei} \\ \epsilon &= \text{Regularisierungsparameter zur Vermeidung von Divisionen} \\ &\quad \text{durch Null,} \\ \mu &= \text{Adaptionsschrittweite.} \end{aligned} \tag{5.44}$$

5.7 Zusammenfassung

Kapitel 5.1 widmet sich der Codierung natürlicher Wellenfelder, d.h. der Messung HOA-Komponenten mit Hilfe von Mikrophonarrays. Dazu ist eine Verallgemeinerung der Fourier-Bessel-Reihe, sodass diese auch für Kugelwellen gültig ist, notwendig, auf die anfänglich eingegangen wird. Benötigt werden diese Erkenntnisse dabei insbesondere für die Bestimmung diverser, radialer Funktionen $W_m(ka)$ kugelförmiger Mikrophonarrays.

Kapitel 5.2 verschafft einen Überblick über verschiedene, geometrische Anordnungen von Mikrophonarrays und deren Verwendbarkeit für Auralisationszwecke, wobei sich kugelförmige Anordnungen als optimal herausstellten. Ferner wird hier auch eine störungsfreie, allgemeine Bestimmung der HOA-Komponenten ($B_{m,n}^\sigma$) mit Hilfe von kugelförmigen Mikrophonarrays, bis zu einer maximalen Ordnung von M , betrachtet. Zudem wird diskutiert, welche minimale Anzahl an Sensoren (Q), welche regulär auf einer Kugeloberfläche angebracht sind, für die Bestimmung des Wellenfeldes bis zu einer Ordnung von M benötigt werden.

Kapitel 5.3 widmet sich der Frage, wie man die radialen Funktionen $W_m(ka)$ für unterschiedliche Mikrophonarrays (offene und geschlossene Kugel mit omnidirektionalen und gerichteten Sensoren) mathematisch berechnen kann. Die dabei erzielten Ergebnisse werden miteinander verglichen und schließlich zusammengefasst.

Basierend auf den Erkenntnissen aus Kapitel 5.3 wurde in Kapitel 5.4 das geschlossene, kugelförmige Mikrophonarray genauer untersucht. Dabei beschränkte man sich aus praktischen Gründen auf eine maximale Anzahl omnidirektionaler Sensoren von $Q = 32$. Anschließend erfolgte eine anschauliche Interpretation der Codierung eines natürlichen Wellenfeldes mit Hilfe eines geschlossenen, kugelförmigen Mikrophonarrays. Zudem wurde eine Erklärung dafür geliefert, warum beim Entwurf der radialen *equalizing*-Funktionen $EQ_m(ka)$ eine Regularisierung notwendig ist, die sich z.B. mit Hilfe von Tikhonov-Filter realisieren lässt. Ein zentraler Punkt dieses Kapitels stellt die Bestimmung eines optimierten Regularisierungsparameters ϵ dar. Hierzu wurde auch hier, wie in vielen anderen *beamforming*-Verfahren, als Bestimmungsgröße der WNG-Wert verwendet. Eine anschließende Untersuchung verschiedener Fehlermöglichkeiten, die beim Entwurf von kugelförmigen

³⁴NLMS = engl. *Normalized Least Mean Square*

Mikrophonarrays auftreten können, bekräftigte dabei die Entscheidung den WNG-Wert als Grundlage für die Optimierung zu verwenden.

In Kapitel 5.5 wird ein neuartiger Optimierungsalgorithmus zur Bestimmung eines frequenzabhängigen Regularisierungsparameters $\epsilon(\omega)$ vorgestellt, welcher letztendlich zu einer weiteren Erfindungsmeldung [Chr11b] führte. Dabei wurden zunächst die Grundlagen zur Bestimmung der dazu nötigen Werte (*White noise gain*: $WNG_{\text{dB}}(\omega)$, Ausgangssignal des *beamformer* in Blickrichtung (θ_0, φ_0) : $d(\theta_0, \varphi_0, \omega)$, Eigenrauschen des q . Sensors: $H_q(\theta_q, \varphi_q, \omega)$, Array-Gewinn: $G(\omega)$, Richtdiagramm in Blickrichtung: $\Psi(\theta_0, \varphi_0, \omega)$, Richtdiagramm bei omnidirektionalen Schalleinfall: $\Psi(\theta, \varphi, \omega)$) vorgestellt. Ziel der Optimierung ist es, ausgehend vom Optimalfall ($\epsilon(\omega) = 0$), über der Frequenz ω den Regularisierungsparameter $\epsilon(\omega)$ so zu modifizieren (erhöhen), dass eine vorgegebene, minimale WNG-Schranke $WNG_{\text{dB}_{\text{Min}}}$ nicht unterschritten wird und zugleich am Arrayausgang in Blickrichtung ein von der Frequenz unabhängiger, konstanter Wert von 1 erreicht wird, d.h. eine frequenzunabhängige Übertragungsfunktion in Blickrichtung resultiert. Somit entsteht ein *beamformer*, welcher eine bestimmte durch die WNG-Schwelle vorgegebene Robustheit und gleichzeitig ein maximales Bündelungsmaß $G(\omega)$ aufweist. Simulationen mit unterschiedlich entworfenen, radialen *equalizing*-Funktionen $EQ_m(ka)$ verdeutlichen die Wirksamkeit des Optimierungsalgorithmus bzw. der optimierten Regularisierung. Außerdem zeigten die Ergebnisse, in welchen Umfang die Regularisierung zu einer Abnahme des Bündelungsmaßes über der Frequenz führt.

Kapitel 5.6 widmet sich praxisrelevanten Faktoren, die beim Bau eines kugelförmigen Mikrophonarrays berücksichtigt werden sollten. Hierzu zählen neben der Positionierung der Sensoren auf der Oberfläche, die idealerweise gleichförmig sein sollte, auch die Festlegung des Kugelradius a , welcher sich in unserem Fall an den Kopfdurchmesser eines erwachsenen Mannes orientiert. Außerdem werden diverse Möglichkeiten zur Erhöhung der räumlichen Abtastfrequenz aufgezeigt, zu denen neben einer Erhöhung der Anzahl der Sensoren diverse Möglichkeiten zur räumlichen Tiefpassfilterung sowie die frequenzabhängige Beschränkung der Ordnung M , wie auch die Nutzung der natürlichen Abschattung geschlossener Kugelmikrophone, zählen. Ferner wird gezeigt, dass eine der größten Fehlerquellen in der Exemplarstreuung der verwendeten Sensoren zu finden ist, welche sich jedoch durch Einführung einer geeigneten Kalibrierung reduzieren lässt. Eine mögliche Realisierung einer solchen Kalibrierung beschließt dieses Kapitel.

Kapitel 6

Verifikation unterschiedlicher, sphärischer Mikrofonarrays in der Praxis

Basierend auf den Ergebnissen des Kapitels 5 wurden drei sphärische Mikrofonarrays in der Praxis aufgebaut. Dabei handelte es sich um *zwei geschlossene Mikrofonarrays*, deren Radien a sich jeweils an der Geometrie des menschlichen Kopfes orientierten, genauer gesagt einen Wert von $a = 0.09$ [m] aufwiesen und um *ein offenes Mikrofonarray* mit einem Radius von $a = 0.0125$ [m]. Die geschlossenen Mikrofonarrays hatten dabei den in Kapitel A.5.7 näher beschriebenen *Ikosaederstumpf* als Grundkörper, welcher jeweils mit $Q = 32$ omnidirektionalen Mikrofonen an den aus Tabelle A.25(e) bekannten Positionen bestückt war. Eines der beiden geschlossenen Kugelmikrofonarrays wurde dabei als *klassisches* Mikrofonarray realisiert, bei dem sich die omnidirektionalen Mikrophone direkt an der Oberfläche der jeweiligen Positionen befanden. Beim zweiten hingegen befanden sich die Mikrophone an den Fusspunkten, der in die Oberfläche eingebrachten Dellen, was gemäß der in Kapitel 5.6.3 beschriebenen Vorgehensweise geschah. Somit eröffnete sich die Möglichkeit, einen echten Praxisvergleich dieser beiden Varianten durchzuführen und somit die Wirkungsweise dieser Art der *räumlichen Tiefpassfilterung* (Einbringung von Dellen in die Oberfläche) messtechnisch zu untersuchen.

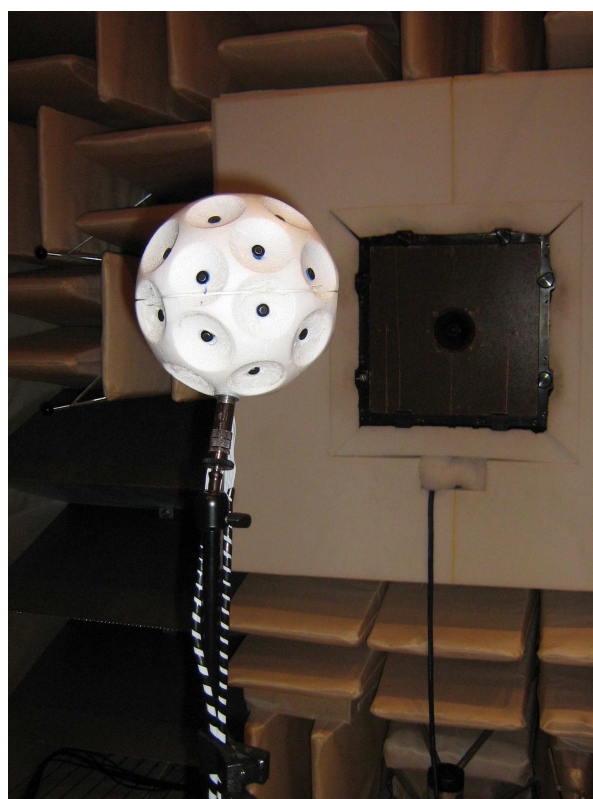
Für das offene Mikrofonarray wurde hingegen der in Kapitel A.5.2 vorgestellte *Hexaeder* als Basis verwendet, welcher mit $Q = 7$ omnidirektionalen Mikrofonen bestückt wurde. Dabei wurden 6 an den aus der Tabelle A.20(e) bekannten Flächenmittelpunkten des *Kubus* angebracht und das 7. Mikrofon im Zentrum des *platonischen Körpers*. Hintergrund dieser Anordnung war der bereits in Kapitel 5.3 diskutierte Umstand, dass offene Kugelmikrophone nur dann störungsfrei funktionieren, wenn gerichtete Mikrophone verwendet werden. Diese wiederum lassen sich gemäß Kapitel A.1 durch eine einfache Zusammenfassung von zwei omnidirektionalen Mikrofonen, d.h. durch Bildung eines differentiellen Mikrophons 1. Ordnung, auf einfache Weise realisieren. Das zentral ange-

ordnete, omnidirektionale Mikrofon dient dabei für jedes der sechs korrekt platzierten, omnidirektionalen Mikrophone als Gegenpart zur Bildung der benötigten sechs gerichteten Mikrophone 1. Ordnung.

6.1 Vergleich eines geschlossenen, kugelförmigen Mikrofonarrays mit und ohne räumlichen Tiefpassfilter



(a) Klassisches Kugelmikrofonarray.



(b) Kugelmikrofonarray mit Dellen zur räumlichen Tiefpassfilterung.

Abbildung 6.1: Darstellung der beiden geschlossenen Kugelmikrofonarrays mit (6.1(b)) und ohne räumlichen Tiefpassfilter (6.1(a)), bei der Vermessung in einer Schirmkammer.

Abbildung 6.1 stellt die beiden für die Untersuchung verwendeten, geschlossenen Kugelmikrofonarrays während der Vermessung in einem *reflexionsarmen Raum* dar. Als Grundkörper für die beiden Kugelmikrophone wurde jeweils eine zweiteilige Styroporkugel mit einem Radius von $a = 0.09$ [m] verwendet. Dabei wurde für die Auffindung bzw. für die Markierung der in Tabelle A.25(e) angegebenen $Q=32$ Sensorpositionen wie folgt

verfahren:

(a) Auffinden der Pole:

Zum Auffinden der Pole wurde am oberen Pol der Kugel (Nordpol) eine Nadel mit Faden in die Styroporkugel gesteckt und verankert, indem die Nadel im Inneren der Kugel quergelegt wurde. Anschließend wurde das Lot gefällt, indem die Kugel am Faden aufgehängt und solange gewartet wurde, bis diese auspendelte. Im ausgependelten Zustand wurde dann die Kugel in ein Stempelkissen abgelassen, wobei der dabei erzielte Abdruck den unteren Pol der Kugel, also den Südpol, kennzeichnete.

(b) Kennzeichnung des Breitengrades:

Am Äquator der Kugel, welcher durch den Spalt der beiden Kugelschalen gekennzeichnet ist, wurde nun eine maßstabsgetreue Skala, die mit Hilfe des Zeichenprogramms *Draw*[®], welches wiederum Bestandteil des freien Softwareprogramms *OpenOffice*[®] ist, mit einer Granularität von 1° erstellt und dort zur Kennzeichnung des *Breitengrades*, aufgeklebt.

(c) Kennzeichnung des Längengrades:

Zum Auffinden der Sensorpositionen ist neben dem Breitengrad noch eine Skala mit ausreichender Granularität für den Längengrad notwendig. Dies wurde mit Hilfe eines Globus, mit einem Durchmesser von $\varnothing = 18$ [cm], wie in Abbildung 6.2(a) dargestellt, bewerkstelligt.

Dabei wurde an den beiden Polen jeweils ein Loch entsprechender Größe gebohrt, so dass man die Styroporkugel anstelle des Globus in die Halterung einspannen konnte (siehe Abbildung 6.2(b)).

Für die Beschreibung wie die Dellen in die Oberfläche der Styroporkugel eingebracht wurden sei an dieser Stelle auf Kapitel 5.6.3 verwiesen.

6.1.1 Messergebnisse der geschlossenen, kugelförmigen Mikro- phonarrays mit und ohne räumlichen Tiefpassfilter

Basierend auf den Messungen, welche in der Schirmkammer auf Höhe des verwendeten Meßlautsprechers, wie in Abbildung 6.1 dargestellt, in einer Entfernung von 1 [m] durchgeführt wurden, ergaben sich die in Abbildung 6.3 dargestellten Ergebnisse.

Die amorphen Formen der in den Abbildungen 6.3(a) und 6.3(b) aufgezeigten, zylindrischen 3D-Darstellungen der Polardiagramme über der Frequenz begründet sich in der Tatsache, dass die Messungen noch nicht kalibriert wurden. Aus diesem Grund wurde zur einfacheren Beurteilung das Bündelungsmaß über der Frequenz bestimmt, welche für beide Varianten, d.h. für das geschlossene Kugelmikrophonarray mit und ohne räumlichen Tiefpassfilter in Abbildung 6.4 ersichtlich ist.



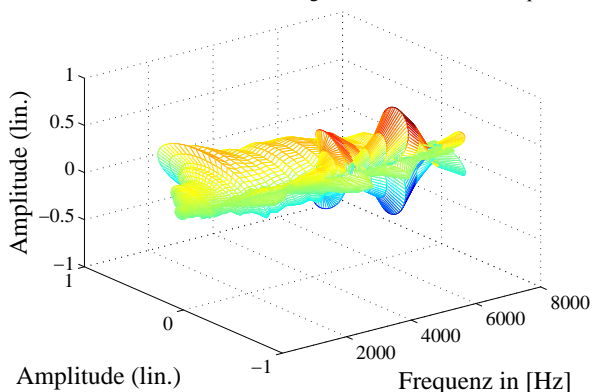
(a) Globus mit einem Durchmesser von $\varnothing = 18$ [cm], zur Kennzeichnung des Längengrades.



(b) Kugel mit Dellen eingespannt in die Halterung des Globus zur Bestimmung des Längengrades.

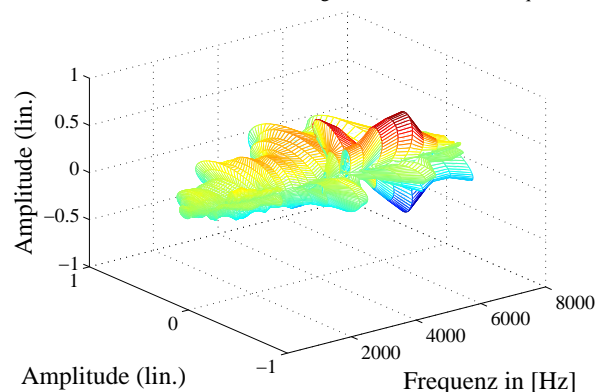
Abbildung 6.2: Globus zur Bestimmung des Längengrades.

Zylindrische 3D-Darstellung des optimal nach Tikhonov regularisierten modalen beamformer 4. Ordnung ohne Dellen über der Frequenz



(a) Messergebnis des geschlossenen Kugelmikrophonarrays ohne räumlichen Tiefpassfilter.

Zylindrische 3D-Darstellung eines optimal nach Tikhonov regularisierten modalen beamformer 4. Ordnung mit Dellen über der Frequenz



(b) Messergebnis des geschlossenen Kugelmikrophonarrays mit räumlichen Tiefpassfilter.

Abbildung 6.3: Zylindrische 3D-Darstellung der Polardiagramme des beamformer mit und ohne räumlichen Tiefpassfilter, dargestellt in einem Frequenzbereich von 20 [Hz] $\geq f \geq 8000$ [Hz].

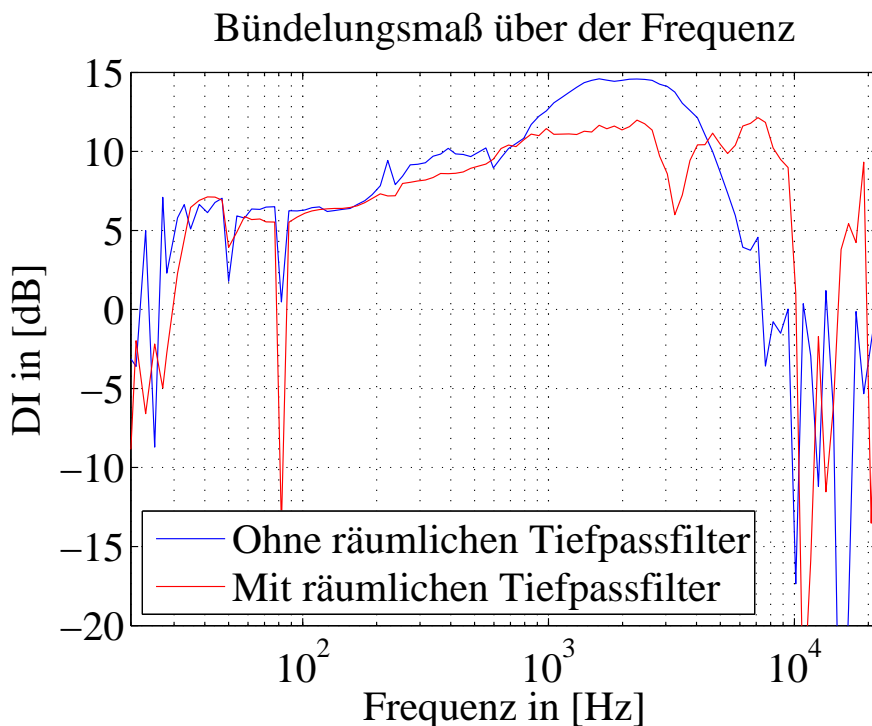


Abbildung 6.4: Bündelungsmaße über der Frequenz der vermessenen, geschlossenen Kugelmikrofonarrays mit und ohne räumlichen Tiefpassfilter.

Aufgrund der bereits oben erwähnten Tatsache, dass die zugrundeliegenden Messungen noch nicht kalibriert wurden sowie des unprofessionellen, mechanischen Aufbaus, insbesondere des Kugelmikrofonarrays mit Dellen (dieses wurde noch nicht nach der in Kapitel 5.6.3 angegebenen Methode gefertigt, sondern freihändig, manuell modifiziert), ist es schwer, aus den Polardiagrammen über der Frequenz aus Abbildungen 6.3(a) und 6.3(b) nützliche Informationen zu extrahieren. Aus diesem Grund ist das Bündelungsmaß, welches gemäß Gleichung 5.36 das Verhältnis der Energie bei Schalleinfall aus Blickrichtung zur Energie bei omnidirektionalen Schalleinfall angibt und für beide Fälle aus Abbildung 6.4 ersichtlich ist, als alternative Darstellungsform zur Beurteilung der beiden Kugelmikrophone besser geeignet. Auf diese Weise kann trotz fehlender Kalibrierung eine brauchbare Aussage aus den Messungen gewonnen werden. Wie man aus Abbildung 6.4 unschwer erkennt, lässt sich dadurch bei nahezu unveränderter Größe des Arrays, wie zuvor bereits erwähnt, die räumliche Abtastfrequenz f_{Alias} (siehe Gleichung 5.8) in etwa verdoppeln. Ferner lässt sich das geringer ausfallende, maximale Bündelungsmaß des geschlossenen Kugelmikrofonarrays sowie dessen Einbruch bei $f \approx 3.3$ [kHz] (roter Graph in Abbildung 6.4) durch die unzureichende, mechanische Präzision des Aufbaus erklären.

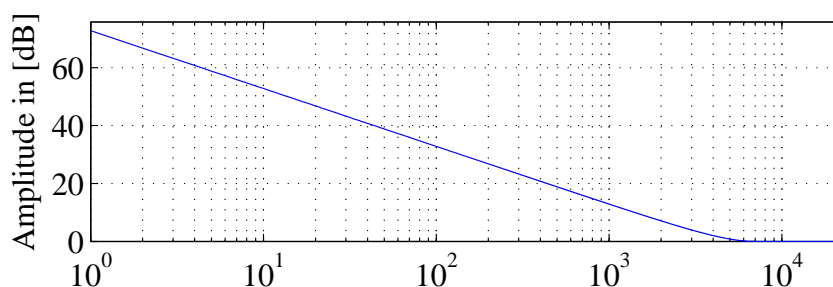
Eine andere Alternative, wie man die räumliche Abtastfrequenz f_{Alias} erhöhen kann, besteht darin, die Geometrie des Kugelmikrofonarrays zu verringern. Dadurch sinkt zwar der nutzbare untere Spektralbereich, gleichzeitig erhöht sich aber f_{Alias} aufgrund des dadurch geringer werdenden Mikrofonabstandes d_{Mic} .

Da das Kugelmikrophonarray sowohl mit Hilfe einer geschlossenen, soliden, als auch mit einer offenen Kugel realisiert werden kann, wird nachfolgend, wie zu Beginn dieses Kapitels bereits angedeutet, noch ein Beispiel aufgeführt, welches auf einem kleinen, offenen Kugelmikrophonarray basiert.

6.2 Untersuchung eines offenen, kugelförmigen Mikrophonarrays geringer Größe

Bei dem hier untersuchten offenen, kugelförmigen Mikrophonarray werden jeweils aus zwei benachbarten, omnidirektionalen Mikrofonen, gemäß der aus Kapitel A.1 vorgestellten Methode, differentielle Mikrophone 1. Ordnung berechnet. Dabei kann in unserem Fall der Abstand der benachbarten, omnidirektionalen Mikrophone jeweils mit $d_{Mic} = 0.0125$ [m] angegeben werden. Bei Verwendung einer Schallgeschwindigkeit von $c = 343$ [$\frac{m}{s}$] lässt sich das notwendige Tiefpassfilter zur Erzeugung eines frequenzunabhängigen Verhaltens bezüglich der Haupteinfallrichtung $W_i(j\omega)$ mit $i = 1$ gemäß Gleichung A.4 berechnen.

Betragsfrequenzgang des equalizing–Tiefpassfilter mit $d_{Mic}=0.0125$ [m]



Phasenfrequenzgang des equalizing–Tiefpassfilter mit $d_{Mic}=0.0125$ [m]

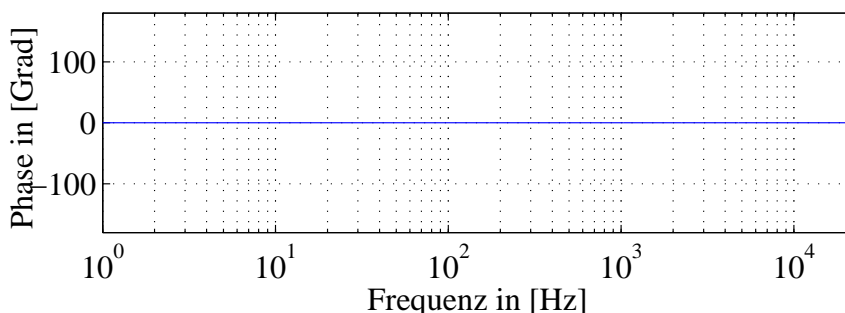


Abbildung 6.5: Bodediagramm des equalizing–Tiefpassfilter 1. Ordnung $W_1(j\omega)$, welches sich für einen Mikrophonabstand von $d_{Mic} = 0.0125$ [m] ergibt.

Abbildung 6.5 zeigt dabei den Verlauf des idealen equalizing–Tiefpassfilter 1. Ordnung $W_1(j\omega)$ auf, der sich ergibt, wenn man von Mikrofonen ohne Eigenrauschen sowie von den zuvor genannten Vorgaben ($d_{Mic} = 0.0125$ [m], $c = 343$ [$\frac{m}{s}$]) ausgeht. In der Praxis

muss jedoch auch das *equalizing*-Tiefpassfilter 1. Ordnung $W_1(j\omega)$ ab einer bestimmten unteren Frequenz in seiner maximalen Amplitude begrenzt, d.h. regularisiert werden, damit das dadurch verstärkte *white noise gain* nicht zu groß wird, sondern unterhalb einer vorgegebenen, minimalen Schranke WNG_{dBMin} bleibt.

Das *white noise gain* $WNG(j\omega)$ entspricht dem Kehrwerte der Suszeptibilität $K(j\omega)$, d.h. $WNG(j\omega) = \frac{1}{K(j\omega)}$. Da die Suszeptibilität $K(j\omega)$ für differentielle Mikrophone aus Gleichung A.8 bekannt ist, wird an dieser Stelle auf eine explizite Angabe der Berechnung des *white noise gain* $WNG(j\omega)$ für differentielle Mikrophone verzichtet.

**WNG(j ω) eines differentiellen Mikrophons 1. Ordnung,
mit $d_{\text{Mic}}=0.0125$ [m] und einem Richtfaktor von $\alpha=0.5$**

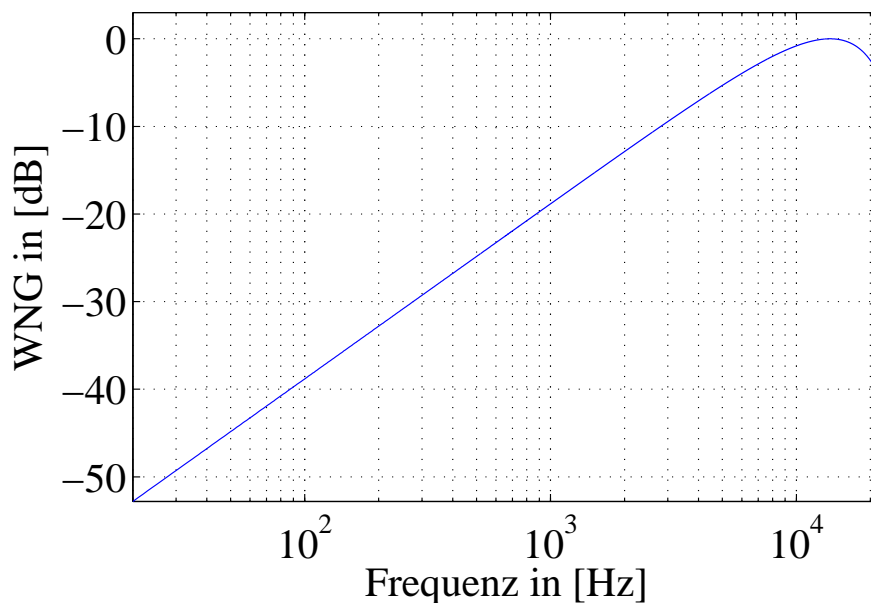
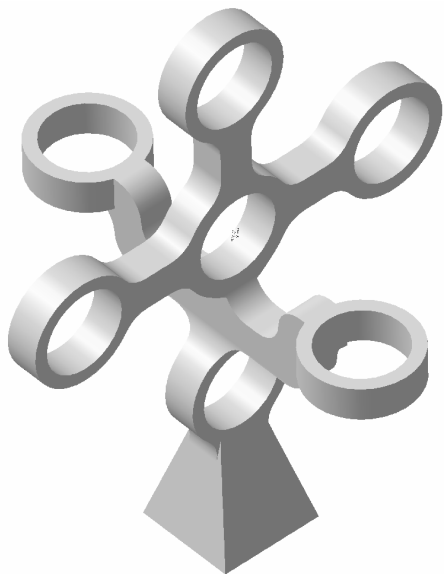


Abbildung 6.6: *White noise gain* $WNG(j\omega)$ eines differentiellen Mikrophons 1. Ordnung, mit einem Mikrophonabstand von $d_{\text{Mic}} = 0.0125$ [m] und einem Richtfaktor von $\alpha = \frac{1}{2}$.

Abbildung 6.6 veranschaulicht den WNG-Verlauf eines differentiellen Mikrophons 1. Ordnung, mit einem Mikrophonabstand von $d_{\text{Mic}} = 0.0125$ [m] und einem Richtfaktor von $\alpha = \frac{1}{2}$ über der Frequenz. Die Schranke für differentielle Mikrophone 1. Ordnung kann gemäß [Elk00] mit $WNG_{\text{dBMin}} \leq -40$ [dB] angegeben werden. Dies bedeutet, dass das hier untersuchte differentielle Mikrophon 1. Ordnung theoretisch erst ab einer Frequenz von ca. $f \approx [100, \dots, 200]$ [Hz] verwendet werden kann.

6.2.1 Beispielhafter Entwurf eines realisierbaren, robusten, offenen Kugelmikrophonarrays

Für das folgende Beispiel wird das in Abbildung 6.7(b) dargestellte Chassis, welches die Positionen der omnidirektionalen Mikrophone festlegt und einen Radius von $d_{Mic} = 0.0125$ [m] hat, zugrunde gelegt.



(a) Ideales Chassis auf dem das offene Kugelmikrophonarray, d.h. der offene modale *beamformer* beruhen sollte. (b) Real verwendetes, mit omnidirektionalen Mikrophenen bestücktes, offenes Kugelmikrophonarray.

Abbildung 6.7: Offenes Kugelmikrophonarray, bestehend aus 7 omnidirektionalen Mikrophenen.

Ein Vergleich der beiden in Abbildung 6.7 dargestellten Bilder verdeutlicht, dass das Chassis des real aufgebauten, offenen Kugelmikrophonarrays, welches in Abbildung 6.7(b) zu sehen ist, von dem als ideal angesehenen Chassis, welches aus Abbildung 6.7(a) ersichtlich ist, abweicht. Hintergrund dieser Abweichungen waren fertigungstechnische Gründe, die eine Umsetzung der notwendigen, geschwungenen Form entlang der X-Achse verhinderten (Koordinatensystem siehe Abbildung 6.8). Stattdessen konnte jede Achse ausschließlich entlang einer Geraden gefertigt werden, was ein Blick auf Abbildung 6.7(b) offenbart. Um dennoch eine belastbare Aussage über die Wirkungsweise dieser Art des Aufbaus zu erhalten, beschränkte man sich bei der Auswertung der Messergebnisse auf die Y/Z-Ebene, in welcher die nicht optimal angebrachten Mikrophone auf der X-Achse bei der Auswertung keine Rolle spielten.

Die Idee hinter dieser Anordnung ist jeweils ein außenliegendes, omnidirektionales Mikrophen mit dem sich im Mittelpunkt befindlichen, ebenfalls omnidirektionalen Mikrophen gemäß der in Abbildung 6.8 aufgezeigten Art und Weise zu einem differentiellen, also einem gerichteten Mikrophen 1. Ordnung, zu kombinieren.

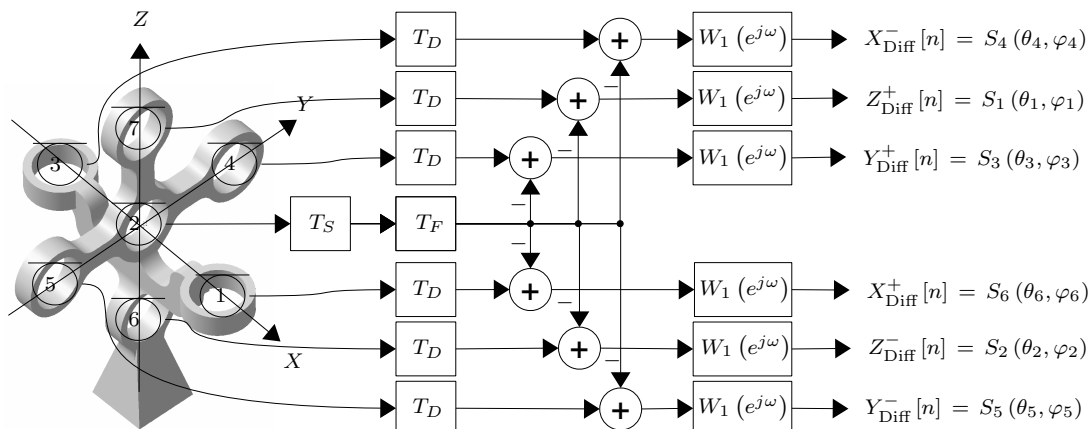


Abbildung 6.8: Erzeugung der notwendigen sechs gerichteten Mikrophone 1. Ordnung aus den sieben omnidirektionalen Mikrophone des offenen Kugelmikrophonarrays.

Wie man aus Abbildung 6.8 sieht, ist dazu lediglich die Umsetzung eines Verzögerungselements T für das in der Mitte angebrachte Mikrophone notwendig, das in Abbildung 6.8 bereits wie folgt aufgespalten wurde:

$$\begin{aligned}
 T &= T_S + T_F, \text{ mit} \\
 T_S &= \text{Verzögerungszeit in [Samples]}, \\
 T_F &= \text{Fractional delay}^1 \text{ in [Samples]}.
 \end{aligned}
 \tag{6.1}$$

Hintergrund dieser Aufspaltung ist, dass man es in der Praxis mit einem digitalen Signalprozessor zu tun hat, der mit einer bestimmten Abtastfrequenz f_A digitale Eingangssignale verarbeitet, die zuvor über einen A/D-Wandler diskretisiert wurden. Diese *Abtastwerte*, die gemein hin als *Samples* bezeichnet werden, können dann lediglich im Takt der Abtastfrequenz, also mit $T_s = \frac{1}{f_A}$ [s], verarbeitet bzw. verzögert werden. Im Allgemeinen kann nicht davon ausgegangen werden, dass die benötigte Verzögerungszeit T ein ganzzahliges Vielfaches von T_s ist, also gilt: $\frac{T}{T_s} = i \in \mathbb{N}$. Aus diesem Grund ist es ratsam, in der Praxis die benötigte Verzögerungszeit T in ein ganzzahliges Vielfaches der Abtastfrequenz und einen Rest, das sogenannte *fractional delay* T_F , das stets kleiner als ein Sample ist ($[0 < T_F < 1]$ [Sample]), aufzuspalten.

Umgesetzt wird ein solches *fractional delay* üblicherweise mit Hilfe eines FIR-Filters. Dabei wird ein *idealer Tiefpassfilter*, der auch als *idealer Interpolator* bezeichnet wird und dessen Impulsantwort durch eine SI-Funktion² gegeben ist, um das gewünschte *fractional*

¹Der Begriff *fractional Delay* stammt aus dem englischen und bedeutet *fraktionelle Verzögerung*, wobei die deutsche Übersetzung in der Fachliteratur kaum Verwendung findet.

²SI-Funktion = *Sinus Cardinalis*-Funktion ($si(x) = \frac{\sin(x)}{x}$).

delay $T_F = T - T_S = \frac{d_{Mic} f_A}{c} - \lfloor \frac{d_{Mic} f_A}{c} \rfloor$ zeitlich verschoben $\left(si(t - T_F) = \frac{\sin(t - T_F)}{t - T_F} \right)$. Anschließend wird die zeitlich verschobene Impulsantwort mit T_S abgetastet und dann mit einer Fensterfunktion (z.B. Hamming-Fenster) gewichtet, um ansonsten auftretende, störende Randeffekte (*Gibbssches Phänomen*) zu vermeiden.

Somit lassen sich die Filterkoeffizienten eines FIR-Filters, die sich als Vektor $\mathbf{h}_L = [h_0, h_1, \dots, h_{L-1}]^T$ der Länge L angeben lassen und ein *fractional delay* der Länge T_F realisieren, wie folgt berechnen:

$$\begin{aligned} h_n &= W(n) si\left(n - \frac{L}{2} - T_F\right), \text{ mit} \\ W(n) &= 0.54 - 0.46 \cos\left(2\pi \frac{n}{L}\right) \text{ (Hamming-Fenster), und} \\ n &= 0, \dots, L - 1, \text{ wobei} \\ h_n &= n. \text{ Filterkoeffizient des } fractional \text{ delay FIR-Filters,} \\ W(n) &= n. \text{ Gewichtungsfaktor der verwendeten Fensterfunktion,} \\ L &= \text{Länge des FIR-Filters,} \\ T_F &= \text{Länge des } fractional \text{ delay in [Sample].} \end{aligned} \tag{6.2}$$

Mit $d_{Mic} = 0.0125$ [m] und $c = 343$ [$\frac{m}{s}$] ergibt sich eine zu realisierende Verzögerungszeit von $T = \frac{d_{Mic} f_A}{c} = 1.6071$ [Sample]. Daraus resultiert eine ganzzahlige Verzögerungszeit von $T_S = \lfloor \frac{d_{Mic} f_A}{c} \rfloor = 1$ [Sample] sowie ein durch das FIR-Filter umzusetzendes *fractional delay* von $T_F = T - T_S = \frac{d_{Mic} f_A}{c} - \lfloor \frac{d_{Mic} f_A}{c} \rfloor = 0.6071$ [Sample]. Bei einer FIR-Filterlänge von $L = 30$ [Taps], ergeben sich für die Umsetzung des *fractional delay* folgende Ergebnisse:

- Impulsantwort:

Aus Abbildung 6.9 erkennt man, dass es sich bei den L Filterkoeffizienten des FIR-Filters, welches das gewünschte *fractional delay* umsetzen soll, um eine entsprechend verschobene, abgetastete SI-Funktion handelt. Aus Gründen der Kausalität wurde der Nullpunkt der Impulsantwort zuvor um $\frac{L}{2}$, in unserem Fall also um 15 [Sample] verzögert, d.h. nach rechts verschoben.

- Bode-Diagramm:

Wie man aus Abbildung 6.10 erkennen kann, wird das *fractional delay*-Filter, was im Idealfall ein Allpassfilter mit entsprechendem Phasenfrequenzgang ist, in der Praxis durch ein Tiefpassfilter mit korrespondierendem Phasenfrequenzgang approximiert. In unserem Fall, d.h. mit $L = 30$ [Taps] @ $f_A = 44100$ [Hz] ergibt sich dadurch eine obere Eckfrequenz des Tiefpassfilter von $f_c \approx 21$ [kHz], was in Anbetracht einer oberen akustischen Grenzfrequenz des menschlichen Gehörs von $f < 20$ [kHz] als akzeptabel erachtet werden kann.

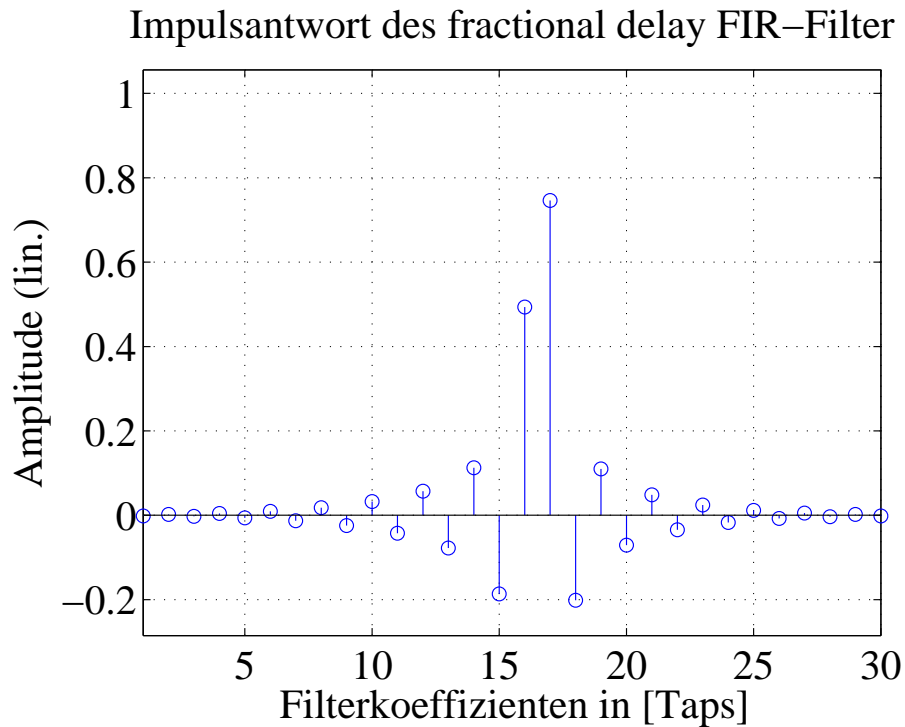


Abbildung 6.9: Impulsantwort eines FIR-Filters der Länge $L = 30$ [Taps], welches ein fractional delay von $T_F = 0.6071$ [Sample] realisieren soll.

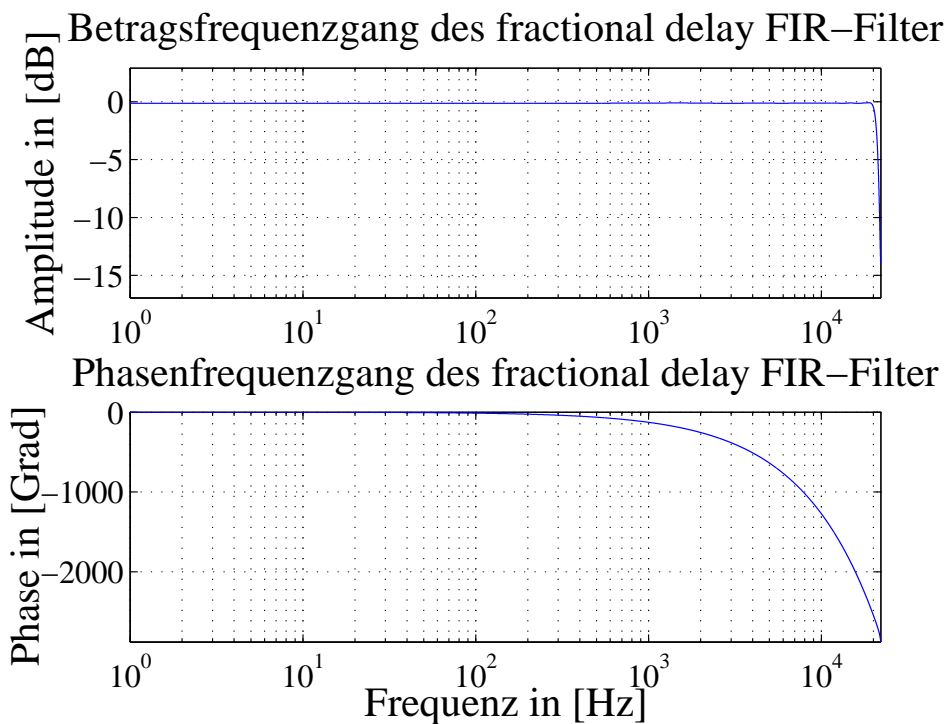


Abbildung 6.10: Bode-Diagramm eines FIR-Filters der Länge $L = 30$ [Taps], welches ein fractional delay von $T_F = 0.6071$ [Sample] realisieren soll.

- Gruppenlaufzeit:

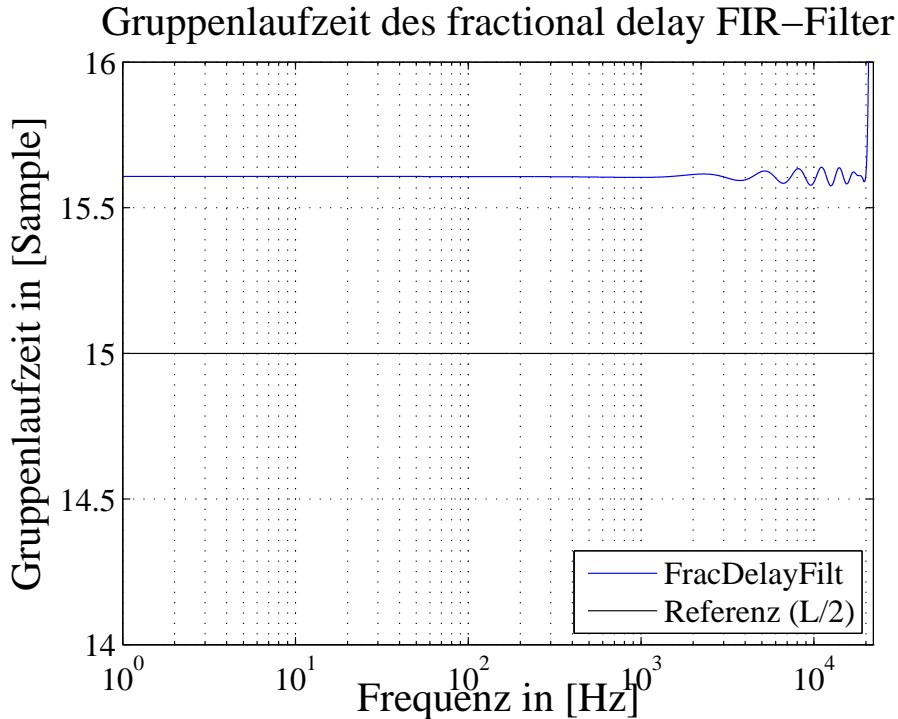


Abbildung 6.11: Gruppenlaufzeit eines FIR-Filters der Länge $L = 30$ [Taps], welches ein fractional delay von $T_F = 0.6071$ [Sample] realisieren soll.

Wie Abbildung 6.11 verdeutlicht, kann durch das verwendete FIR-Filter bis zu einer Frequenz von $f \approx 1500$ [Hz] das gewünschte *fractional delay* von $T_D + T_F = 15.6071$ [Sample] mit hoher Genauigkeit realisiert werden. Oberhalb dieser Frequenz kommt es dann zu einer gewissen Abweichung, welche sich mit $\Delta T_t \approx \pm 0.2$ [Sample] beziffern lässt. Falls diese Abweichungen als zu groß erachtet werden sollten, müsste lediglich eine größere FIR-Filterlänge L beim Entwurf verwendet werden, was dann zu einem höheren Implementierungsaufwand führt. Dadurch sinkt der verbleibende Restfehler und die Grenzfrequenz des Tiefpassfilter steigt. Prinzipiell sollte stets mit geraden Filterlängen, also $L = 2n$, $n \in \mathbb{N}$ gearbeitet werden, da sich ansonsten keine ganzzahligen Verzögerungszeiten in [Sample] für T_D ergeben, die dann ebenfalls mit Hilfe weiterer *fractional delay*-Filter aufwändig realisiert werden müssten.

Wie in Gleichung 6.2 aufgezeigt, muss das *fractional delay* FIR-Filter \mathbf{h}_L beim Entwurf um die halbe Filterlänge L , also um $T_D = \frac{L}{2}$ [Sample] verschoben werden, wodurch, wie bereits erwähnt, für dessen Kausalität gesorgt wird. Exakt um diese Zeit, also um T_D , müssen die anderen Mikrophone ebenfalls verzögert werden, damit aus diesen, zusammen mit dem so verzögerten, mittleren, omnidirektionalen Mikrophon, die gewünschten sechs

gerichteten Mikrophone 1. Ordnung entstehen können. Auf diese Weise lassen sich in der Praxis beliebige Verzögerungen T relativ aufwandsgünstig und einfach realisieren.

Die sechs auf diese Weise entstandenen, differentiellen Mikrophone 1. Ordnung befinden sich jeweils an den Enden von drei, jeweils orthogonal zueinander stehenden Achsen. Diese Positionen entsprechen somit den Flächenmittelpunkten des in Abbildung A.20(b) dargestellten Hexaeders, der wiederum zu den insgesamt 5 platonischen Körpern zählt. Somit liegt auch in diesem Fall eine reguläre Abtastung der Kugeloberfläche vor, weshalb die Orthonormalitätsbedingung aus Gleichung 5.7 zumindest weitestgehend erfüllt wird.

Insgesamt lassen sich mit Anordnung dieser $Q = 6$ regulär angebrachten, gerichteten Mikrophone 1. Ordnung sphärische Harmonische bis zu einer Ordnung von $M = \lfloor \sqrt{Q} - 1 \rfloor = 1$, bzw. die ersten $N = (M + 1)^2 = 4$ sphärischen Harmonischen einwandfrei bestimmen, was ein Blick auf die in Abbildung A.20(f) dargestellte Orthonormalitätsfehlermatrix eines Hexaeders bestätigt. Im Prinzip können die ersten $N = 4$ sphärischen Harmonischen auch fehlerfrei mit nur $Q = 4$ regulär, also auf den Flächenmittelpunkten eines Tetraeders (platonischer Körper) angebrachten Sensoren, gemessen werden. Allerdings offenbart ein Blick auf Abbildung A.20(f), dass bei Verwendung eines Hexaeders, bedingt durch die höhere Anzahl an Sensoren im Vergleich zu einem Tetraeder der Fehler bei der Messung höherer, sphärischer Harmonischer deutlich reduziert und an manchen Stellen sogar zu Null wird. So kommt es bei der Messung der sphärischen Harmonischen $N = [5, \dots, 9]$ lediglich zu Störungen durch sphärische Harmonische gleicher Ordnung, was in Abbildung A.20(f) aus dem diagonalen Verlauf in diesem Bereich zu erkennen ist.

Das so entworfene, offene Kugelmikrophonarray liefert dann, wie aus Abbildung 6.8 ersichtlich, die sechs äquivalenten Signale gerichteter Mikrophone 1. Ordnung, die dort mit $S_q(\theta_q, \varphi_q)$ und $q = [1, \dots, 6]$ bezeichnet werden. Diese Signale entsprechen nun den Eingangssignalen des aus Abbildung A.18 bekannten, modalen *beamformer*. Auf dieser Basis kann nun in Anlehnung an das vorhergegangene Kapitel 5.5 der Entwurf eines optimal regularisierten, modalen *beamformer* erfolgen, was nachfolgend beruhend auf dem hier vorgestellten Chassis aufgezeigt wird.

White noise gain des offenen Kugelmikrophonarrays, basierend auf dem verwendeten Chassis

Obwohl nun ein Vergleich mit dem zuvor behandelten, geschlossenen Kugelmikrophonarray nicht mehr offensichtlich ist, da unterschiedliche Radien a ($a = 0.09$ [m] für das geschlossene und $a = 0.0125$ [m] für das offene Kugelmikrophonarray) verwendet werden, verdeutlicht dennoch ein Vergleich von Abbildung 6.12 mit Abbildung 5.21, dass der offene Typ eines Kugelmikrophonarrays prinzipiell eine höhere Robustheit aufweist als der geschlossene. Deutlich wird dies insbesondere bei Betrachtung des spektralen Verlaufes des WNG 1. Ordnung, da hier über den gesamten Audiospektralbereich ($f \in [20, \dots, 20000]$ [Hz]) eine WNG-Schwelle von $WNG_{\text{Min}_{\text{dB}}}(j\omega) = -10$ [dB] nicht unterschritten wird und unter diesen Voraussetzungen zumindest ein *beamformer* dieser

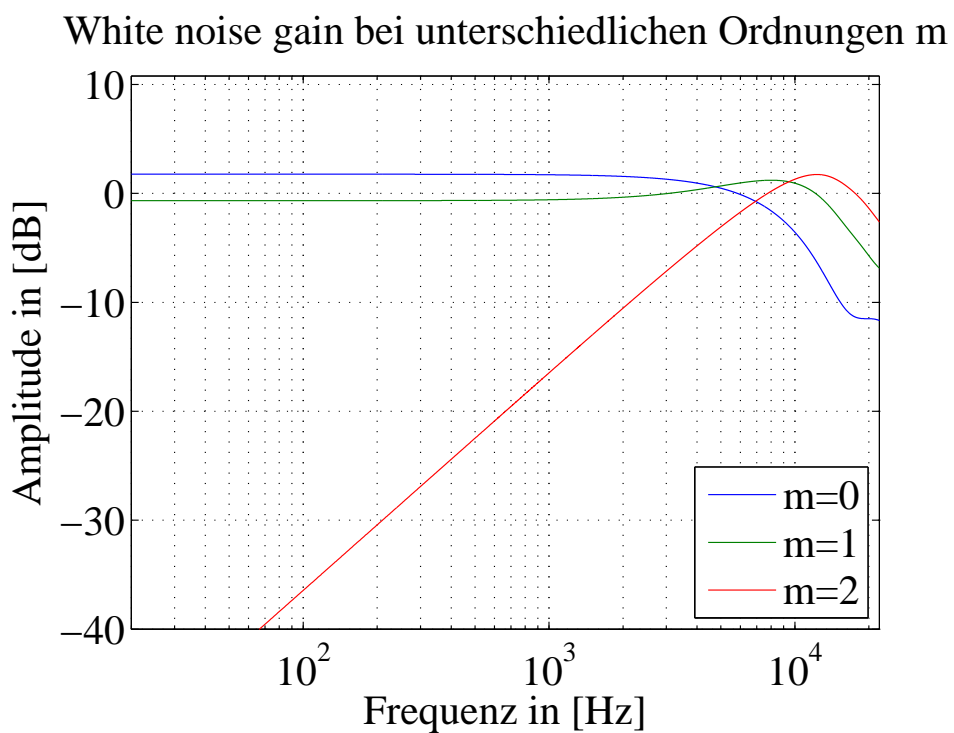


Abbildung 6.12: White noise gain unterschiedlicher Ordnungen ($m = [0, \dots, 2]$) dargestellt über der Frequenz $WNG_{dB}(j\omega)$, bei Verwendung eines idealen, offenen Kugelmikrophonarrays mit einem Radius von $a = 0.0125$ [m] und differentiellen Mikrofonen 1. Ordnung mit einem Richtfaktor von $\alpha = \frac{1}{2}$.

Ordnung breitbandig ohne Einschränkungen in der Praxis robust betrieben werden könnte.

Radiale Funktionen $W_m(ka)$ des offenen Kugelmikrophonarrays basierend auf dem verwendeten Chassis

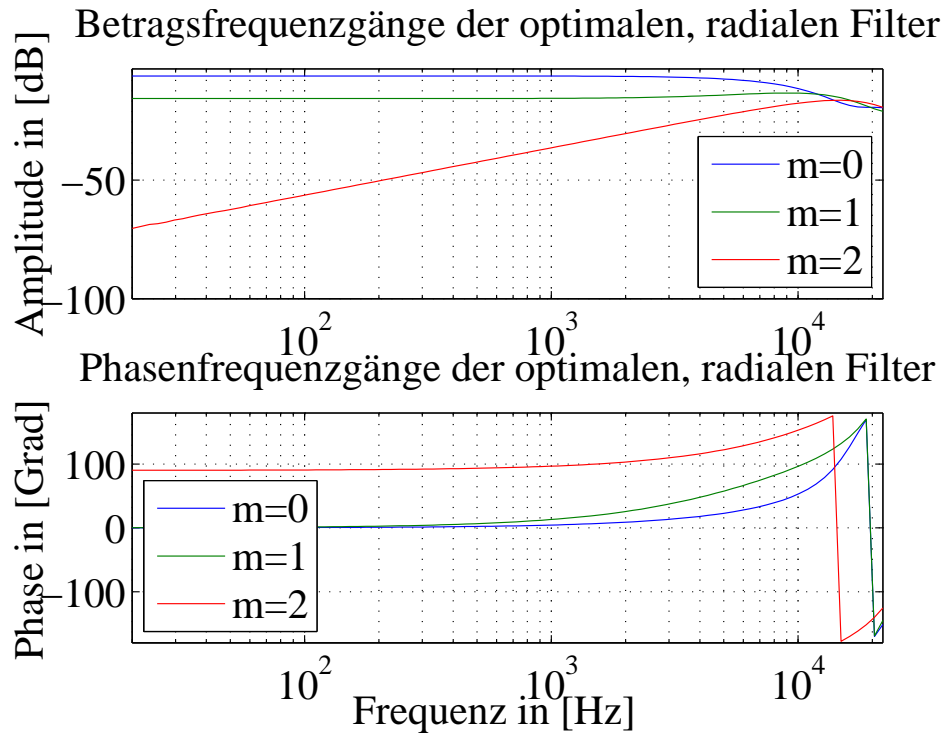


Abbildung 6.13: Radiale Funktionen $W_m(ka)$ unterschiedlicher Ordnungen ($m = [0, \dots, 2]$) dargestellt über der Frequenz bei Verwendung eines idealen, offenen Kugelmikrophonarrays mit einem Radius von $a = 0.0125$ [m] und differentiellen Mikrofonen 1. Ordnung mit einem Richtfaktor von $\alpha = \frac{1}{2}$.

Abbildung 6.13 veranschaulicht, wie sich das Wellenfeld auf einer offenen Kugeloberfläche bis zu einer Ordnung von $M = 2$ über der Frequenz zusammensetzt, wenn dabei gerichtete Mikrophone 1. Ordnung mit kardioider Richtcharakteristik ($\alpha = \frac{1}{2}$) verwendet werden, die sich an den Flächenmittelpunkten eines Hexaeders im Raum befinden und dabei ein Kugelradius von $a = 0.0125$ [m] angewendet wird.

Equalizing-Filter $EQ_m(ka)$ des offenen Kugelmikrophonarrays basierend auf dem verwendeten Chassis

Die inversen Funktionen der radialen Funktionen $W_m(ka)$, also $EQ_m(ka) = \frac{1}{W_m(ka)}$ liefern im Idealfall, d.h. wenn perfekte Bedingungen vorausgesetzt werden, direkt die benötigten equalizing-Funktionen $EQ_m(ka)$, damit der modale beamformer frequenzunabhängig die

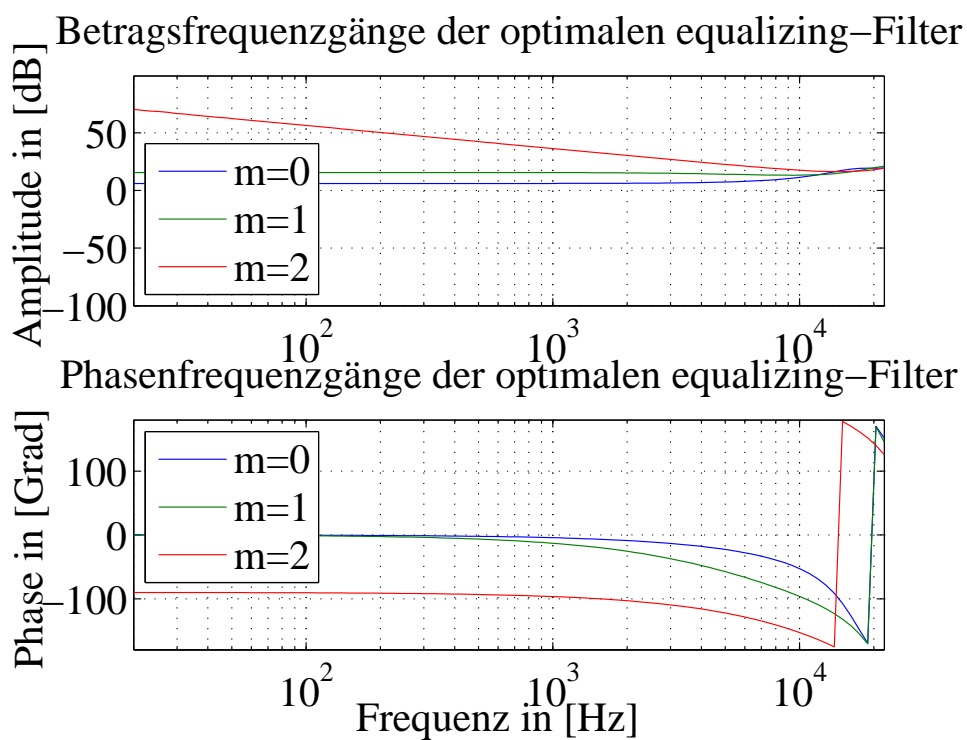


Abbildung 6.14: Equalizing-Funktionen $EQ_m(ka) = W_m^{-1}(ka)$ unterschiedlicher Ordnungen ($m = [0, \dots, 2]$) dargestellt über der Frequenz bei Verwendung eines idealen, offenen Kugelmikrophonarrays mit einem Radius von $a = 0.0125$ [m] und differentiellen Mikrofonen 1. Ordnung mit einem Richtfaktor von $\alpha = \frac{1}{2}$.

gewünschte Richtcharakteristik aufweist. Wie man aus Abbildung 6.14 erkennen kann, ist bis zur 1. Ordnung auch unter praktischen Bedingungen mit keinem Problem zu rechnen. Lediglich im unteren Spektralbereich ist bei Verwendung der 2. Ordnung eine praktische Umsetzung, aufgrund des geringen Kugelradius von $a = 0.0125$ [m] problematisch. Dies kann wieder mit einer Beschränkung der Ordnung von $m = 2$ auf $m = 1$ innerhalb dieses Frequenzbereiches, in Abhängigkeit des vorliegenden WNG (Eigenrauschen der verwendeten Mikrophone), gelöst werden. Im einfachsten Fall kann auch hier die Eckfrequenz, bei welcher man diese Beschränkung der Ordnung von $m = 2$ auf $m = 1$ durchführen muss, aus dem Schnittpunkt des Verlaufes des WNG 2. Ordnung, wie in Abbildung 6.12 dargestellt, mit einer beliebig vorgegebenen WNG-Schwelle gefunden werden. Bei einer WNG-Schwelle von $WNG_{\text{Min}_{\text{dB}}}(\omega) = -10$ [dB] ergibt sich dabei eine Grenzfrequenz von $f \approx 2.2$ [kHz]. Bessere Ergebnisse lassen sich aber auch hier erzielen, wenn das in Kapitel 5.5 vorgestellte Optimierungsverfahren angewendet wird.

Anwendung des Optimierungsverfahrens

- Regularisierungsparameter $\epsilon(j\omega)$:

Regularisierungsparameter $\epsilon(j\omega)$ eines optimalen, modalen beamformer 2. Ordnung unter Einhaltung einer minimalen WNG-Schwelle von $WNG_{\text{Min}_{\text{dB}}} = -10$ [dB]

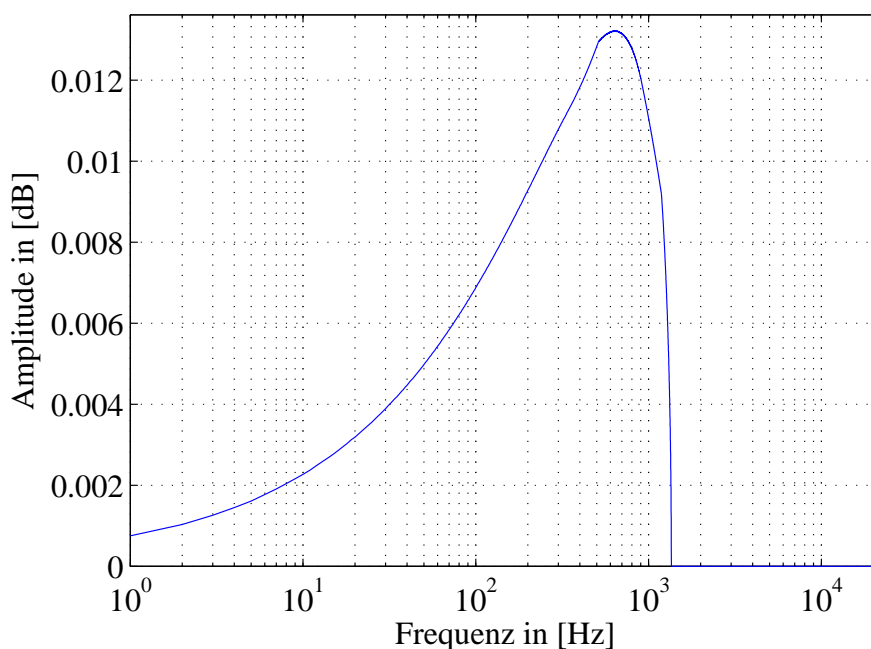


Abbildung 6.15: Regularisierungsparameter $\epsilon(j\omega)$ zur Erzeugung eines robusten, modalen beamformer 2. Ordnung, welcher auf einem offenen Kugelmikrophonarray mit einem Radius von $a = 0.0125$ [m] und differentiellen Mikrophone 1. Ordnung mit einem Richtfaktor von $\alpha = \frac{1}{2}$ basiert, wobei eine WNG-Schwelle von $WNG_{\text{Min}_{\text{dB}}} = -10$ [dB] nicht unterschritten wird.

Wird davon ausgegangen, dass beim Aufbau des offenen Kugelmikrophonarrays Mikrophone mit den gleichen (Rausch)Eigenschaften als im vorherigen Beispiel bei Verwendung eines geschlossenen Kugelmikrophonarrays verwendet werden, kann auch hier als Grenzwert die gleiche WNG-Schwelle von $WNG_{\text{Min}_{\text{dB}}} = -10$ [dB] angewendet werden. Der dabei resultierende, frequenzabhängige Regularisierungsparameter $\epsilon(j\omega)$ für den Entwurf der M Tikhonov-Filter aus Gleichung 5.28 ist in Abbildung 6.15 dargestellt.

- Regularisierungsfiler $T_m(j\omega)$:

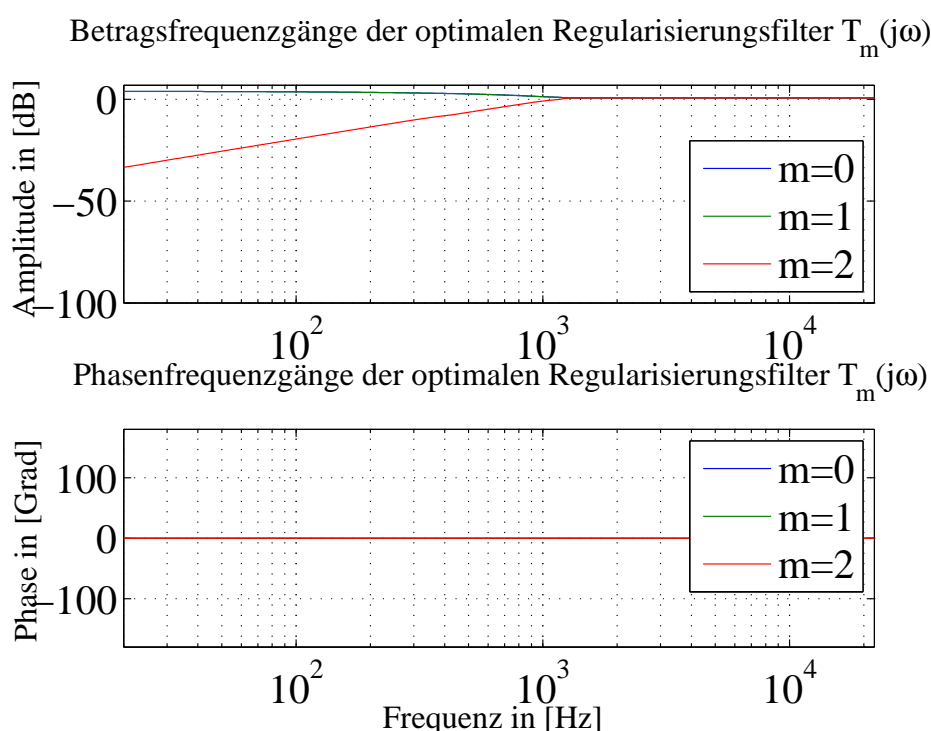


Abbildung 6.16: Regularisierungsfiler $T_m(j\omega)$ für die Ordnungen $m = [0, \dots, 2]$ zur Erzeugung eines robusten, modalen beamformer 2. Ordnung, welcher auf einem offenen Kugelmikrophonarray mit einem Radius von $a = 0.0125$ [m] und differentiellen Mikrofonen 1. Ordnung mit einem Richtfaktor von $\alpha = \frac{1}{2}$ basiert, wobei eine WNG-Schwelle von $WNG_{\text{Min}_{\text{dB}}} = -10$ [dB] nicht unterschritten wird.

Unter Verwendung der in Abbildung 6.13 dargestellten, radialen Funktionen $W_m(ka)$ sowie des in Abbildung 6.15 aufgezeigten, frequenzabhängigen Regularisierungsparameter $\epsilon(j\omega)$ ergeben sich nach Gleichung 5.28 die in Abbildung 6.16 dargestellten M Regularisierungsfiler $T_m(j\omega)$. Wie man aus Abbildung 6.16 erkennen kann, kommt es dabei zu einer Absenkung der sphärischen Harmonischen 2. Ordnung, wobei diese nun nach Anwendung der Optimierung nicht mehr wie zuvor ab einer Frequenz von $f \approx 2.2$ [kHz], sondern nun erst ab einer Frequenz von $f \approx 1.2$ [kHz] abgesenkt werden muss. Somit konnte in diesem Fall mit Hilfe der

Optimierung aus Kapitel 5.5 der spektrale Einsatzbereich des modalen *beamformer* 2. Ordnung um ca. $f \approx 1$ [kHz] zu tiefen Frequenzen hin erweitert werden.

- Equalizing-Filter $EQ_m(j\omega)$:

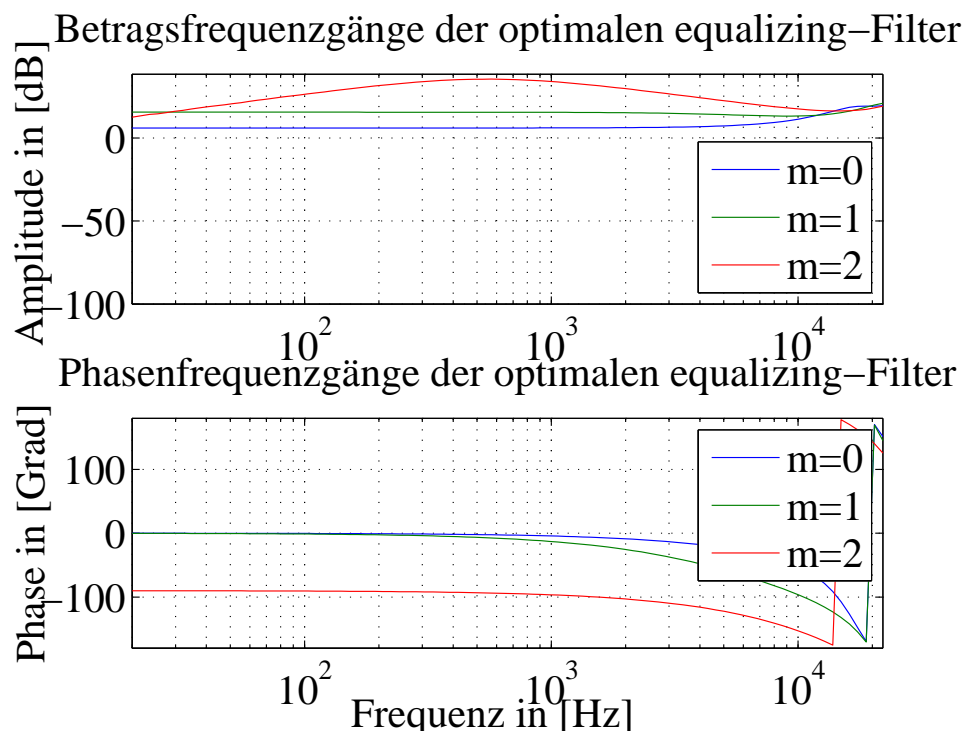


Abbildung 6.17: Regularisierte equalizing-Filter $EQ_m(j\omega)$ für die Ordnungen $m = [0, \dots, 2]$ zur Erzeugung eines robusten, modalen *beamformer* 2. Ordnung, welcher auf einem offenen Kugelmikrophonarray mit einem Radius von $a = 0.0125$ [m] und differentiellen Mikrofonen 1. Ordnung mit einem Richtfaktor von $\alpha = \frac{1}{2}$ basiert, wobei eine WNG-Schwelle von $WNG_{\text{Min}_{\text{dB}}} = -10$ [dB] nicht unterschritten wird.

Werden nun die in Abbildung 6.14 aufgezeigten, optimalen, radialen *equalizing*-Funktionen mit den in Abbildung 6.16 dargestellten, für eine WNG-Schwelle von $WNG_{\text{Min}_{\text{dB}}} = -10$ [dB] optimierten Regularisierungsfiler $T_m(j\omega)$ gewichtet, resultieren daraus die in Abbildung 6.17 dargestellten, regularisierten, radialen *equalizing*-Filter. Wenn man diese für $EQ_m(ka)$ in dem in Abbildung 5.2 dargestellten Signalfluss einsetzt, lassen sich damit auf robuste Weise Ambisonickomponenten bis zu einer maximalen Ordnung von $M = 2$ bestimmen. Eingesetzt für die modalen Gewichtungsfaktoren $C_m(j\omega)$ in Abbildung A.18 lässt sich alternativ auch ein robuster, modaler *beamformer* 2. Ordnung realisieren.

- White noise gain $WNG_{\text{dB}}(j\omega)$:

Zur Verifizierung des Optimierungsverfahrens sowie zur Darstellung der Auswirkung des selbigen auf das Bündelungsmaß, stellt Abbildung 6.18 diese im Vergleich zum

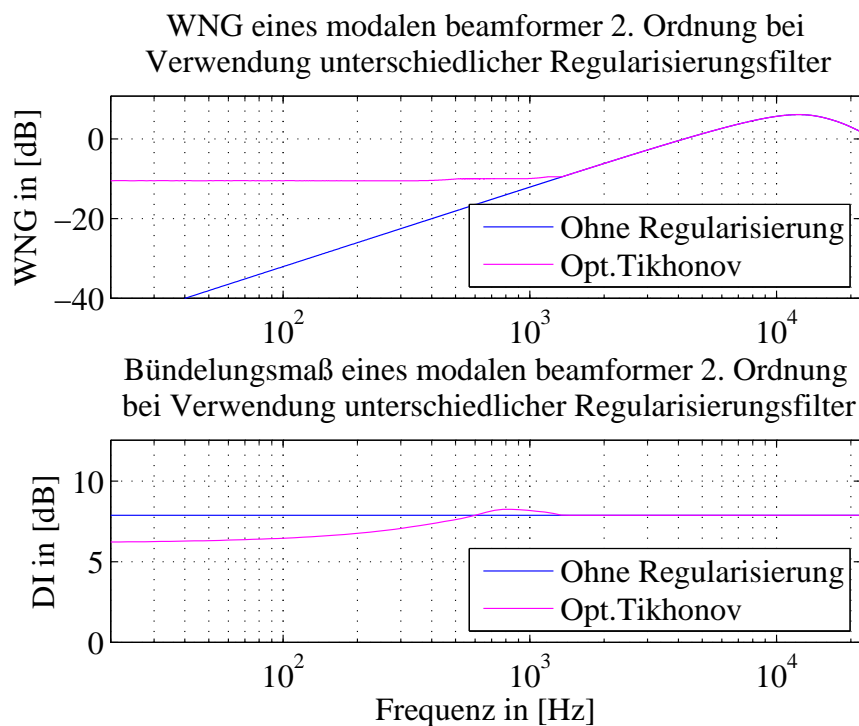


Abbildung 6.18: White noise gain $WNG_{dB(j\omega)}$ nach Anwendung der regularisierten equalizing-Filter $EQ_m(j\omega)$ für die Ordnungen $m = [0, \dots, 2]$ zur Erzeugung eines robusten, modalen beamformer 2. Ordnung, welcher auf einem offenen Kugelmikrophonarray mit einem Radius von $a = 0.0125$ [m] und differentiellen Mikrofonen 1. Ordnung mit einem Richtfaktor von $\alpha = \frac{1}{2}$ basiert, wobei eine WNG-Schwelle von $WNG_{Min_{dB}} = -10$ [dB] nicht unterschritten wird.

Idealfall dar. Auch hier kann einerseits erkannt werden, dass die WNG-Schwelle von $WNG_{\text{Min dB}} = -10 \text{ [dB]} \pm \Delta WNG_{\text{dB}} = 0.5 \text{ [dB]}$ eingehalten, d.h. nicht unterschritten wird und zudem das Bündelungsmaß $DI(j\omega)$ ab einer Frequenz von $f \approx 1.2 \text{ [kHz]}$ anfängt sich zu verändern, wobei es erst ab einer Frequenz von $f \approx 600 \text{ [Hz]}$ unterhalb des maximalen Wertes von $DI_{\text{Max}} = 10 \log_{10}(Q) = 7.78 \text{ [dB]}$ fällt. Selbst im untersten Spektralbereich fällt das Bündelungsmaß $DI(j\omega)$ nicht unterhalb eines *beamformer* 1. Ordnung, das in diesem Fall $DI(j\omega) \approx 6 \text{ [dB]}$ beträgt.

- Zylindrische 3D-Darstellung der Polardiagramme über der Frequenz:

Zylindrische 3D-Darstellung der Polardiagramme eines modalen beamformer 2. Ordnung über der Frequenz nach Anwendung eines optimal nach Tikhonov entworfenen Regularisierungsfilters

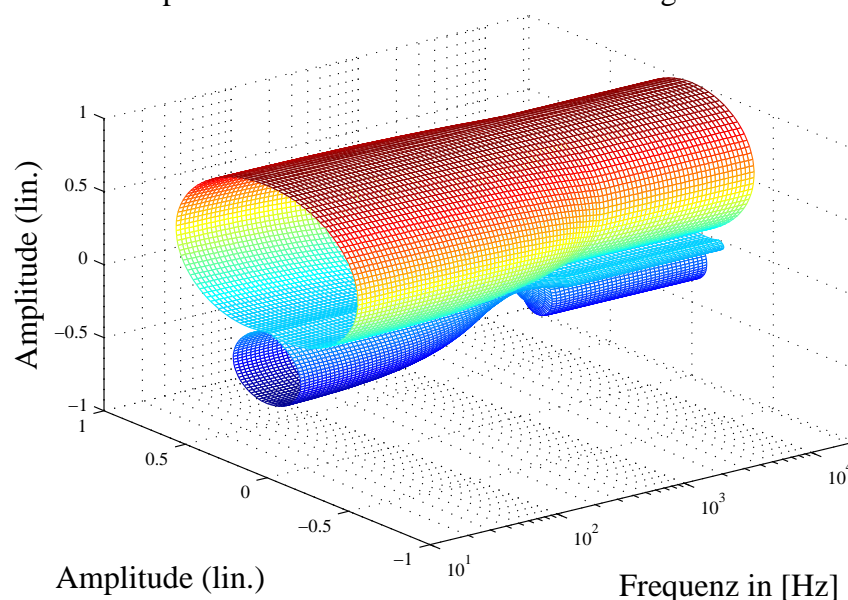


Abbildung 6.19: Resultierende Richtcharakteristik $\psi(\theta_0, \varphi_0, \omega)$ über der Frequenz nach Anwendung der regularisierten equalizing-Filter $EQ_m(j\omega)$ für die Ordnungen $m = [0, \dots, 2]$ zur Erzeugung eines robusten, modalen beamformer 2. Ordnung, welcher auf einem offenen Kugelmikrophonarray mit einem Radius von $a = 0.0125 \text{ [m]}$ und differentiellen Mikrofonen 1. Ordnung mit einem Richtfaktor von $\alpha = \frac{1}{2}$ basiert, wobei eine WNG-Schwelle von $WNG_{\text{Min dB}} = -10 \text{ [dB]}$ nicht unterschritten wird.

Abbildung 6.19 veranschaulicht dieses frequenzabhängige Verhalten des Bündelungsmaßes auf plastische Weise. Im unteren Frequenzbereich ist die Charakteristik eines gerichteten Mikrophons 1. Ordnung und im oberen Spektralbereich die eines gerichteten Mikrophons 2. Ordnung deutlich zu erkennen. Dazwischen kommt es gemäß der angewendeten Regularisierung zu einem fließenden Übergang zwischen den beiden Ordnungen. Bei $f \approx 800 \text{ [Hz]}$ kommt es dabei zu einem Einbruch der

Nebenkeule bei $\theta = 0^\circ$ bzw. $\theta = 360^\circ$ (siehe Abbildung 6.20), was den Anstieg des Bündelungsmaßes aus Abbildung 6.18 bei dieser Frequenz erklärt, wobei dort sogar der ansonsten maximale Wert von $DI_{\text{Max}} = 7.78$ [dB] überschritten wird.

- Draufsicht auf die zylindrische 3D-Darstellung der Polardiagramme über der Frequenz:

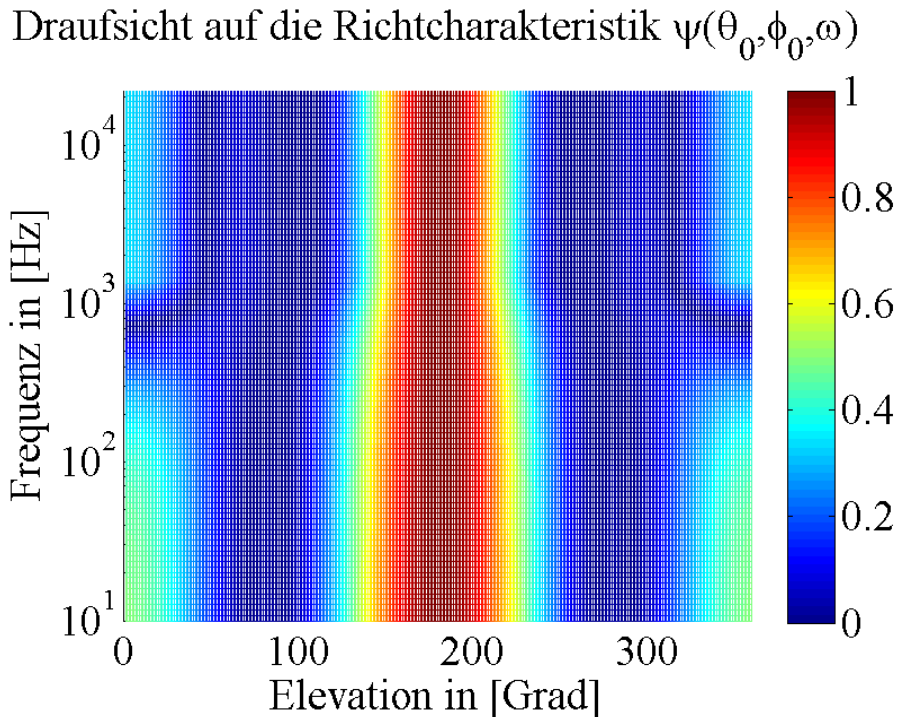


Abbildung 6.20: Draufsicht auf die resultierende Richtcharakteristik $\psi(\theta_0, \varphi_0, \omega)$ über der Frequenz nach Anwendung der regularisierten equalizing-Filter $EQ_m(j\omega)$ für die Ordnungen $m = [0, \dots, 2]$ zur Erzeugung eines robusten, modalen beamformer 2. Ordnung, welcher auf einem offenen Kugelmikrophonarray mit einem Radius von $a = 0.0125$ [m] und differentiellen Mikrofonen 1. Ordnung mit einem Richtfaktor von $\alpha = \frac{1}{2}$ basiert, wobei eine WNG-Schwelle von $WNG_{\text{Min}_{dB}} = -10$ [dB] nicht unterschritten wird.

Abbildung 6.20 zeigt die Draufsicht auf die zylindrische 3D-Darstellung. Dabei bestätigen sich die zuvor getroffenen Aussagen, dass es einerseits bei Frequenzen um $f \approx 800$ [Hz] zu einem Einbruch der Nebenkeule, welche in die zur Blickrichtung entgegengesetzte Richtung zeigt, kommt und andererseits, dass der modale Beamformer im unteren Spektralbereich die Form eines gerichteten Mikrophons 1. Ordnung und im oberen Spektralbereich die eines 2. Ordnung annimmt.

Gemäß Gleichung 5.8 kann die räumliche Aliasingfrequenz in diesem Beispiel, d.h. bei einem Mikrophonabstand von $d_{\text{Mic}} = a = 0.0125$ [m] und einer Schallgeschwindigkeit von $c = 343$ [$\frac{\text{m}}{\text{s}}$], mit $f_{\text{Alias}} = 13720$ [Hz] angegeben werden. Somit könnte eine Kombination des im ersten Beispiel behandelten modalen beamformer mit dem

bzw. mit einem ähnlich entworfenen, kleineren kugelförmigen Mikrophonarray nahezu den gesamten Audiospektralbereich ohne räumliche Aliasingfehler abdecken.

Messergebnisse des offenen, kugelförmigen Mikrophonarrays, welches auf dem gegebenen Chassis basiert

Ähnlich zu den in Abbildung 6.1 dargestellten Kugelmikrophonarrays wurde auch das in Abbildung 6.7(b) dargestellte, offene Kugelmikrophonarray in der Schirmkammer vermessen. Dabei wurden abermals sämtliche Raumimpulsantworten von dem verwendeten Messlautsprecher zu den sieben Mikrofonen aufgenommen. Diese Messungen wurden dabei viermal gemacht, wobei das Mikrophonarray jeweils um 90° in seiner horizontalen Lage gedreht wurde, sodass jeweils das Mikrophon mit der aus Abbildung 6.8 bekannten Nummerierung 1, 3, 4 und 5 am nächsten zum Messlautsprecher angeordnet war.

Zylindrische 3D-Darstellung eines optimal nach Tikhonov regularisierten, modalen beamformer 2. Ordnung über der Frequenz

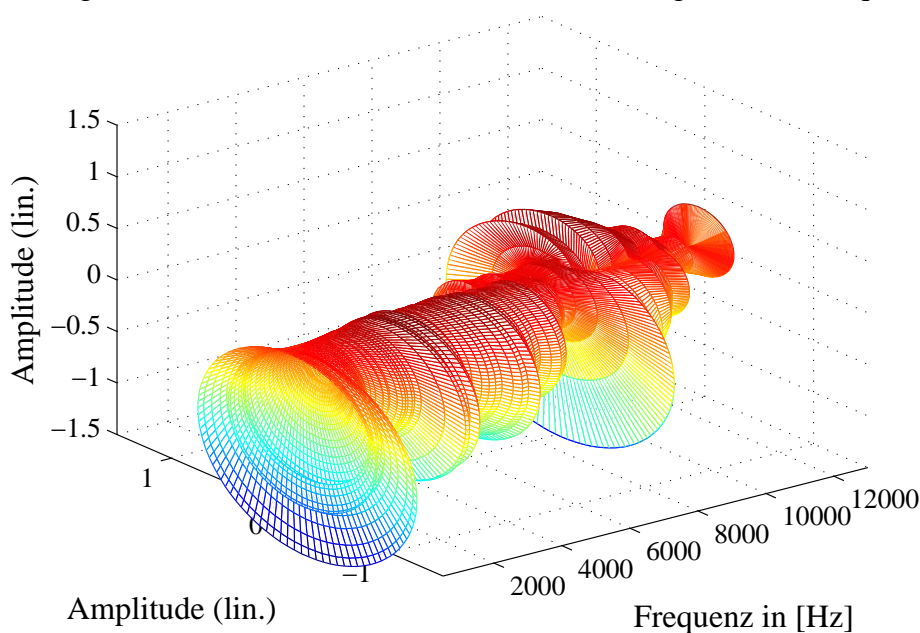


Abbildung 6.21: Messergebnisse eines robusten, modalen beamformer 2. Ordnung, welcher auf einem offenen Kugelmikrophonarray mit einem Radius von $a = 0.0125$ [m] und differentiellen Mikrofonen 1. Ordnung mit einem Richtfaktor von $\alpha = \frac{1}{2}$ basiert, dargestellt in einem Frequenzbereich von 500 [Hz] $\geq f \geq 13000$ [Hz].

Die amorphe Form der Richtdiagramme über der Frequenz aus Abbildung 6.21 begründet sich auch hier in der Tatsache, dass die Messungen noch nicht kalibriert wurden. Aus diesem Grund wird auch in diesem Fall zur leichteren Beurteilung das Bündlungsmaß $DI(j\omega)$ dieses beamformer berechnet.

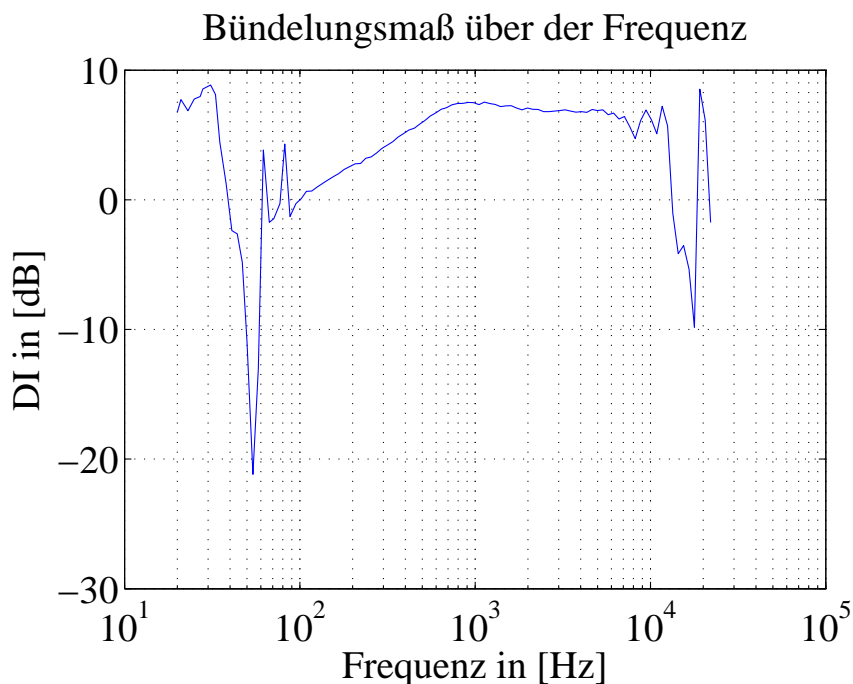


Abbildung 6.22: Bündelungsmaß über der Frequenz des vermessenen, offenen Kugelmikrophonarrays mit einem Radius von $a = 0.0125$ [m] und differentiellen Mikrofonen 1. Ordnung mit einem Richtfaktor von $\alpha = \frac{1}{2}$.

Wie man aus Abbildung 6.22 unschwer erkennen kann, ist dieses Array einerseits unterhalb von $f \approx 100$ [Hz] nicht brauchbar, da hier keine Bündelung erzielt werden kann. Andererseits zeigt Abbildung 6.22, dass es ab einer Frequenz von $f \approx 13000$ [Hz] zum räumlichen Aliasing kommt, was die Korrektheit des in Gleichung 5.8 angegebenen Zusammenhangs verdeutlicht. Somit kann festgehalten werden, dass diese Art der Erzeugung eines modalen *beamformer* in guter Übereinstimmung mit der zuvor beschriebenen Theorie steht. Der spektrale Einsatzbereich des auf die oben beschriebene Art entworfenen, modalen *beamformer* kann gemäß Abbildung 6.22 mit $f \approx [300, \dots, 13000]$ [Hz] angegeben werden, wodurch er sich sehr gut z.B. für Anwendungen beim Freisprechen eignet.

Der untere, spektrale Einsatzbereich lässt sich außerdem bei dem hier behandelten, modalen *beamformer* auf einfache Weise mit Hilfe einer rein softwaretechnischen Maßnahme erweitern. Genauer gesagt lässt sich die untere Grenzfrequenz der differentiellen Mikrophone 1. Ordnung in etwa halbieren, indem diese nicht mit dem zentral angeordneten, omnidirektionalen Mikrofon gebildet werden, sondern mit den jeweils auf der selben Achse befindlichen, in doppelter Entfernung, wie das zentral angebrachte Mikrofon, platzierten, omnidirektionalen Mikrofonen, wie in Abbildung 6.23 dargestellt.

Bei alleiniger Verwendung dieser Struktur würde sich zwar die obere Einsatzfrequenz ebenfalls in etwa halbieren, was sich jedoch dadurch verhindern lässt, indem man an geeigneter Frequenz einen Übergang zu dem aus Abbildung 6.8 bekannten Entwurf schafft. Mit Hilfe dieser Erweiterung lässt sich auf Basis des in Abbildung 6.7 darge-

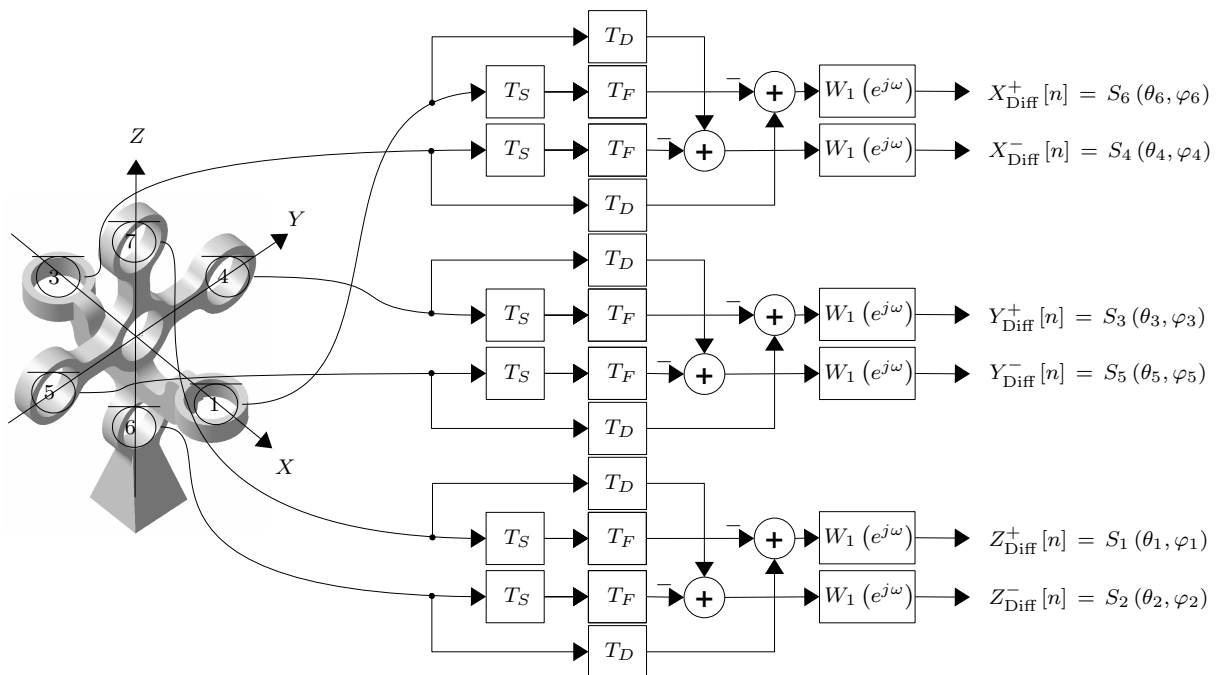


Abbildung 6.23: Erzeugung der notwendigen sechs gerichteten Mikrophone 1. Ordnung aus sechs omnidirektionalen Mikrofonen des offenen Kugelmikrophonarrays.

stellten Chassis ein modularer *beamformer* entwerfen, der in einem Frequenzbereich von $f \approx [150, \dots, 13000]$ [Hz] eingesetzt werden kann.

6.3 Zusammenfassung

In diesem Kapitel wurden drei unterschiedliche sphärische Mikrophonarrays (zwei geschlossene Ikosaederstümpfe, bestückt mit jeweils $Q = 32$ omnidirektionalen Sensoren, mit und ohne Dellen; ein offener Hexaeder mit $Q = 7$ omnidirektionalen Mikrofonen) aufgebaut und messtechnisch untersucht. Ziel dieser Untersuchungen war es, einerseits die Wirksamkeit der räumlichen Tiefpassfilterung, der in die Oberfläche der geschlossenen Kugel eingebrachten Dellen, sowie andererseits die Verwendbarkeit eines offenen, kugelförmigen Mikrophonarrays, welches ausschließlich mit und ohne zentral angebrachten, omnidirektionalen Sensoren bestückt wurde, herauszufinden. Die dabei erzielten Ergebnisse waren durchwegs positiv. So gelang es erstmals, einen messtechnischen Beweis für die Funktionalität, von in die Oberfläche eingebrachten Dellen, als einfache Form der räumlichen Tiefpassfilterung zu erbringen. Der Effekt, der dabei erzielten, räumlichen Tiefpassfilterung war, dass sich die räumliche Aliasingfrequenz nahezu verdoppelt und das Array bei gleicher Größe bis zu wesentlich höheren Frequenzen für die Analyse von Wellenfeldern eingesetzt werden kann. Außerdem konnte gezeigt werden, wie man für die Analyse von Wellenfeldern auf einfache Weise mit Hilfe einer offenen Kugel brauchbare

Messergebnisse liefern kann, ohne dabei gerichtet Mikrophone zu verwenden. Da es sich sowohl bei dem geschlossenen sphärischen Mikrophonarray mit Dellen, als auch bei dem Konzept des offenen, kugelförmigen Mikrophonarrays, zumindest nach Meinung des Autors, um eine Neuerung handelt, wurden diese jeweils als Erfindungsmeldungen eingereicht ([Chr11a, Chr11c]) bzw. in diversen Konferenzen vorgestellt ([Chr11d, Chr12d]), wobei eine Veröffentlichung, in erweiterter Form, Einzug in ein Buchkapitel fand ([Chr13]), welches demnächst erscheinen wird. Zudem erwies sich die in Kapitel 5 beschriebene Optimierung beim Entwurf der regularisierten, radialen *equalizing*-Filter $EQ_m(ka)$ in der praktischen Umsetzung als funktionell.

Kapitel 7

Zusammenfassung und Ausblick

7.1 Zusammenfassung

Im Rahmen dieser Arbeit wurden mehrere lautsprecherbasierte Verfahren zur Auralisation vorgestellt.

Inverse Filter wurden lediglich hinsichtlich einer transauralen Wiedergabe genauer betrachtet. Dieses Verfahren zielt darauf ab, möglichst unverfälscht die Schalldrücke an den Trommelfellen zu reproduzieren, wodurch die transaurale Wiedergabe prinzipiell zu der Gattung der physikalisch motivierten Auralisationsmethoden zählt. Dadurch bedingt, dass die gewünschte Wirkung nur an einem sehr kleinen Bereich im Raum erzielt werden kann, müssen nicht unerhebliche Maßnahmen getroffen werden, um einen Erfolg zu gewähren. Dazu zählt neben einer notwendigen, individuellen Vermessung der 3D-HRTFs mit ausreichender Granularität auch eine präzise Verfolgung der Kopfbewegung im Raum, was mittels eines *head tracker* erfolgt. Letzteres erfordert natürlich das Tragen eines entsprechend geeigneten Sensors, der wiederum möglichst unveränderlich am Kopf des Probanden angebracht sein muss. Trotz dieser Schwierigkeiten wird diese Methode bereits erfolgreich eingesetzt. In dieser Arbeit wurde insbesondere auf die Problematik bei der Erstellung der inversen Filter eingegangen. Diese sorgen im Idealfall dafür, dass jeweils der direkte Pfad, gegeben z.B. vom linken Lautsprecher zum linken Ohr, einer Kroneker-Delta-Funktion entspricht, was bedeutet, dass am Trommelfell ausschließlich das entsprechende Quellsignal anliegt. Zudem sorgen inverse Filter dafür, dass die jeweiligen Kreuzpfade kein Signal übertragen, d.h. eine optimale Übersprechdämpfung zwischen dem linken und rechten Kanal vorherrscht, sodass das linke Ohr ausschließlich das linke Quellsignal empfängt und das rechte Ohr nur das rechte Quellsignal wahrnimmt. Mathematisch kann dieses Problem augenscheinlich sehr einfach gelöst werden. Leider entstehen dabei inverse Filter, die sich in der Praxis nicht realisieren lassen, da sie meist eine viel zu hohe Dynamik aufweisen und zudem ein sehr starkes *pre-ringing* zeigen, was akustisch nicht toleriert werden kann. Dieses Problem lässt sich jedoch, zumindest approximativ gut lösen, indem man es in zwei Teile zerlegt. Einerseits zeigt sich, dass durch Verwendung der adjungierten Matrix

zur Matrixinvertierung sich das *Problem der Übersprechkompensation* elegant lösen lässt, was ansonsten, d.h. bei Verwendung der „klassischen“ Matrixinversion, zu inversen Filtern führt, die sich durch das bereits erwähnte starke *pre-ringing* auszeichnen und deshalb für den akustischen Einsatz ungeeignet sind. Allerdings verbleibt bei Verwendung der adjungierten Matrix für die Berechnung der inversen Filter das *Problem der Realisierung des Kompensationsfilters* $G(j\omega) = \frac{1}{\det(\mathbf{H}(j\omega))}$, welches sich aus der Inversen der Determinante der RIR-Matrix $\mathbf{H}(j\omega)$ ergibt. Neben einigen bekannten Verfahren zur Berechnung des Kompensationsfilters $G(j\omega)$, wie etwa die regularisierte Berechnung im Spektralbereich, wurde in dieser Arbeit in Kapitel 2.5.4 eine neuartige Methode vorgestellt, welche auf einer Regularisierung des nichtlinear geglätteten, minimalphasigen Anteils von $G(j\omega)$ im Spektralbereich basiert. Auf diese Weise lässt sich das Problem des *pre-ringing* auf ein Minimum reduzieren und trotzdem ein gewünschter Betragsfrequenzgang nach Anwendung der vollständigen inversen Filter erreichen. Dabei musste nach Anwendung der so berechneten inversen Filter lediglich auf eine zeitliche Komprimierung der resultierenden Impulsantworten verzichtet werden. Das bedeutet, dass nach Anwendung der so berechneten vollständigen inversen Filter für die direkten Pfade nun nicht mehr die idealen Kronecker-Delta-Funktionen resultieren, sondern exponentiell abfallende Impulsantworten, was beispielsweise Abbildung 2.14 verdeutlicht. Diese Form rührt dabei vorrangig von den CTC-Filtern, die sich aus der Berechnung der adjungierten Matrix ergeben und somit ausschließlich von den Eigenschaften des Wiedergaberaumes abhängig sind. Dadurch erklärt sich der in der Regel ähnlich gestaltete, exponentiell abfallende, zeitliche Verlauf, was ein Vergleich mit den in Abbildung 2.7 dargestellten ursprünglichen RIRs bestätigt. Auf diese Weise bleibt das natürliche Abklingverhalten des Raumes weitestgehend erhalten, was unter Berücksichtigung der akustischen Qualität der resultierenden inversen Filter als wesentlich praxistauglicher angesehen werden kann, als die zuvor behandelten, bereits bekannten Methoden.

Als nächstes wurde über die Beschreibung des Ambisonicsystems 1. Ordnung der Einstieg in die Ambisonictheorie vollzogen, welche aufgrund ihrer Komplexität in ihrer Gesamtheit den Hauptteil dieser Arbeit ausmacht. Den Grundgedanke der Ambisonictheorie beschreibt Gerzon, der getrost als Begründer dieser Theorie bezeichnet werden kann, in der Schallenergievektorthorie sowie in der Schallschnellevektorthorie. Dabei wird ein Wellenfeld dann optimal reproduziert, wenn die Längen beider Vektoren, d.h. sowohl $r_{\mathbf{v}}$ als auch $r_{\mathbf{e}}$ idealerweise eine Länge von jeweils 1 aufweisen. Ob bzw. wie gut dies erreicht werden kann, hängt dabei von mehreren Faktoren ab. So spielt neben der Anzahl sowie der Anordnung der Lautsprecher im Raum, die idealerweise regulär sein sollte, auch die Wahl des Richtfaktors q , welcher die Richtcharakteristik des virtuellen Mikrophons 1. Ordnung bestimmt, eine entscheidende Rolle. Dabei sollte der Richtfaktor frequenzabhängig ausgelegt werden, wobei er im unteren Spektralbereich den Schallschnellevektor \mathbf{v} und im oberen Spektralbereich ab einer Frequenz von $f = 700$ [Hz] den Schallenergievektor \mathbf{e} optimieren sollte. Eine Methode, wie diese optimalen Werte für den Richtcharakteristik-

parameter q bestimmt werden können, wurde ebenfalls in diesem Kapitel aufgezeigt, wobei die Übergänge von dem einen zum anderen Wert über der Frequenz mit Hilfe einfacher *shelving*-Filter 1. Ordnung realisiert werden können. Die Bestimmung der Ambisonicsignale (W, X, Y, Z) mit Hilfe eines Klangfeldmikrophons, welches ein Mikrofon mit Kugelcharakteristik und drei, jeweils orthogonal zueinander angeordneter bidirektionaler Richtcharakteristika in sich vereint, vermittelt dabei einen ersten Eindruck, wie man sich die Zusammensetzung eines Wellenfeldes anhand von sphärischen Harmonischen, deren Formen den Richtcharakteristika der Mikrophone bis zur 1. Ordnung entsprechen, vorstellen kann. Anhand eines Beispiels wurde sowohl die Codierung virtueller Quellen als auch deren Decodierung mit Hilfe eines Lautsprecherrings erörtert. Dabei stellte sich heraus, dass beide auch miteinander kombiniert werden können, was zu einer gewissen Richtcharakteristik führt, aus der man dann auf einfache Weise auf die Eigenschaften des Systems, wie z.B. auf die Lokalisationsschärfe, schließen kann.

Im darauffolgenden Kapitel wurde die Ambisonictheorie für System höherer Ordnung verallgemeinert. Der Hintergrund warum letztendlich das HOA-System für die Umsetzung einer lautsprecherbasierten Auralisierung verwendet wurde und kein anderes eventuell gängigeres Verfahren, wie etwa die *Wellenfeldsynthese* (WFS), war, dass das HOA-System in der Lage ist, an einem zentralen Punkt im Raum (*sweet spot*) ein zuvor analysiertes oder künstlich erzeugtes Wellenfeld exakt nachzubilden. Diese Eigenschaft hat zwar im Prinzip auch die WFS-Methode, allerdings nur bis zu einer wesentlich geringeren oberen Grenzfrequenz. Dafür ist die WFS-Methode in der Lage, diese Eigenschaft in einem wesentlich größeren Raumbereich umzusetzen. Da es in der Akustik, insbesondere in der Fahrzeugakustik üblich ist, lediglich in eng abgegrenzten Bereichen innerhalb des Zielraums, den sogenannten *sweet spots* einen gewünschten Klang zu erzeugen, kommen die Eigenschaften der HOA-Methode unserer Aufgabenstellung entgegen. Als Grundlage für die Verallgemeinerung des Codier- und Decodier-Prozesses diente die Approximation des Wellenfeldes mit Hilfe der Fourier-Bessel-Reihe, die je nach gewünschter bzw. erreichbarer Ordnung M abgebrochen werden kann, wodurch sich Ambisonicsysteme beliebiger Ordnung erzeugen lassen. Im Prinzip geht es bei dem Codier- und Decodier-Prozess nach dem Ambisonicverfahren darum, das vorherrschende Wellenfeld an bestimmten durch die sogenannten sphärischen Harmonischen gegebenen Orte abzutasten, wodurch quasi die Eigenwerte des Wellenfeldes extrahiert werden können (Codier-Prozess), oder mit Hilfe sphärisch bzw. kreisförmig angeordneter Quellen, eben diese (Eigen)Moden des Wellenfeldes wieder zu erzeugen (Decodier-Prozess), die dann in ihrer Überlagerung je nach Ordnung M in einem bestimmten Frequenzbereich exakt dem zuvor codierten Wellenfeld entsprechen. Für die Decodierung wird meist ein 2D-System in Form eines Lautsprecherrings, an dem die Lautsprecher in regulären Abständen angebracht sind, verwendet. Da es nicht einfach ist, ein dreidimensionales Wiedergabesystem zu erstellen, bei dem sich die verwendeten Lautsprecher an regulären Abtastpunkten auf einer Kugeloberfläche befinden und der gewünschte Abhörort dem Zentrum dieser Kugel entspricht. Insbesondere bei An-

bringung dieses 2D-Systems (regulär abgetasteter Lautsprecherring) auf Ohrhöhe konnten selbst in reflexionsbehafteten Räumen gute akustische Ergebnisse erzielt werden. Leider kann dies bei Verwendung von 2D-Systemen für die Bestimmung der Ambisonickomponenten mit Hilfe von zirkularen Mikrophonarrays so nicht behauptet werden, auch wenn nach wie vor einige Forscher damit arbeiten. Der Hauptgrund hierfür sind die störenden Einflüsse der Anteile des stets dreidimensionalen Wellenfeldes, die das Array in vertikaler Richtung durchlaufen. Dies führt zu Fehlern, die ohne entsprechende Gegenmaßnahmen im Nachhinein nicht mehr aus den Aufnahmen herausgerechnet werden können und es somit zwangsläufig zu einer Verfälschung der Codierung des Wellenfeldes kommt. Aus diesem Grund ist es insbesondere bei der Vermessung von Autoinnenräumen unablässig, auf 3D-Systeme, d.h. auf kugelförmige Mikrophonarrays bei der Codierung zurückzugreifen. Speziell in diesem Bereich, d.h. bei der Codierung bzw. Analyse von Wellenfeldern eröffnete die HOA-Theorie ein komplett neues Feld, das mittlerweile reges Interesse in der Forschung geweckt hat. Selbst die älteren WFS-Systeme bedienen sich dieser Methoden, da es für diese bis dahin nicht möglich war, Wellenfelder zu vermessen, sondern lediglich künstlich generierte Wellenfelder in der Praxis zu erzeugen.

Aus diesen Gründen war der Schwerpunkt dieser Arbeit, neben einem umfassenden Einblick in die Theorie der Codier- und Decodierung von HOA-Systemen zu geben, vor allem Methoden zu finden und zu bewerten, die eine weitere Verbesserung der Analyse natürlicher Klangfelder mit Hilfe von Mikrophonarrays erlaubt. Deshalb wurden verschiedene Mikrophonarrays vorgestellt und beurteilt, wobei sich rasch zeigte, dass ausschließlich zirkulare bzw. sphärische Mikrophonarrays imstande sind, die geforderten Aufgaben zu erfüllen. Dabei existieren bereits offene, kugelförmige Arrays sowie ein geschlossenes, kugelförmiges Array, welche man derzeit käuflich am Markt erwerben kann. Daneben wurde aber auch noch eine Sonderform betrachtet, die eine Mischung aus einem offenen Kugelmikrophonarray darstellt, das ein kleineres, geschlossenes Kugelmikrophonarray umschließt. Leider konnte dieses theoretisch bestechende Konzept in der Praxis die in ihm gesteckten Erwartungen nicht erfüllen. Ferner zeigte sich, dass offene Kugelmikrophonarrays nur dann problemlos verwendet werden können, wenn an den regulären Positionen auf der Kugeloberfläche gerichtete Mikrophone angebracht sind. Werden stattdessen omnidirektionale Mikrophone verwendet, kommt es zu Kammfiltereffekten, die zumindest an bestimmten Frequenzen keine Analyse des Wellenfeldes mehr zulassen. Da sich dieses Problem in der Praxis einigermaßen leicht beheben lässt und diese Form ansonsten gute Eigenschaften aufweist, wurde dieses Konzept letztendlich weiterverfolgt.

Vorrangig beschäftigt sich diese Arbeit jedoch mit der Untersuchung von geschlossenen Kugelmikrophonarrays und deren Eigenschaften, da diese diverse positive Eigenschaften aufweisen, die gut zu der hier gestellten Aufgabe passen. So lässt sich etwa der Abschattungseffekt des menschlichen Kopfes bereits in der Aufnahme simulieren, wenn der Grundkörper die Maße des selbigen hat, d.h. einen Radius von ca. $a = 0.09$ [m] aufweist. Wird ein geschlossenes Kugelmikrophonarray dieser Größe dann noch auf einen kopflosen

Dummy angebracht, werden zusätzlich noch Abschattungs-, Beugungs- und Reflexionseffekte des Körpers in der Aufnahme automatisch korrekt berücksichtigt. Dies ist vor allem bei Messungen von Klangfeldern in Autoinnenräumen wichtig, da es hier häufig zu Abschattungen von diversen, z.B. in der Tür verbauter Lautsprecher durch die Beine kommt. Solche Effekte könnten ansonsten ebenfalls kaum, z.B. durch eine simulatorische Nachbehandlung, praxisgerecht erfasst werden, sind jedoch für eine spätere Auralisation wichtig.

Unabhängig davon, welches Mikrophonarray verwendet wird, ist es unablässig, die für die Codierung notwendigen *equalizing*-Filter $EQ_m(ka)$ aufgrund von nicht-idealen Bedingungen, wie etwa Exemplarstreuungen der verwendeten Mikrophone, deren Eigenrauschen oder mechanische Abweichungen der Anbringungsorte, zu regularisieren. Aufgrund der Art, wie sich ein Wellenfeld auf der Oberfläche des verwendeten Kugelmikrophonarrays durch die Überlagerung sphärischer Harmonischer über der Frequenz zusammensetzt, kommt es zu sehr hohen Verstärkungen der *equalizing*-Filter $EQ_m(ka)$, speziell im unteren Frequenzbereich. Diese sind vor allem aufgrund des Eigenrauschens der verwendeten Mikrophone problematisch und müssen auf eine bestimmte Obergrenze beschränkt werden. Diese Beschränkung (Regularisierung) der *equalizing*-Filter hat zur Folge, dass das Wellenfeld im unteren Spektralbereich nicht bis zu einer maximalen Ordnung von M analysiert werden kann, sondern lediglich bis zu einer geringeren Ordnung. Um diesen negativen Effekt der notwendigen Regularisierung so gering wie möglich zu halten, wurde ein eigenes Optimierungsverfahren zur Bestimmung der regularisierten *equalizing*-Filter entwickelt, das in Kapitel 5.5 detailliert beschrieben wurde. Dieses basiert, wie viele andere aus dem Bereich des Entwurfes von *beamforming*-Filter bekannten Verfahren darauf, das WNG auf einen minimalen Wert zu begrenzen. Eine Besonderheit dieses Verfahrens ist, dass zusätzlich noch das Richtdiagramm des *beamformer* in Blickrichtung $\psi(\theta_0, \varphi_0, \omega)$ über der Frequenz ω konstant bei 0 [dB] gehalten wird. Auf diese Weise zeichnet sich das regularisierte *equalizing*-Filter $EQ_m(ka)$ durch eine minimale Beschränkung des Bündelungsmaßes hin zu tiefen Frequenzen aus, die ansonsten wesentlich höher ausfallen würde, was ein Vergleich z.B. mit dem ebenfalls in dieser Arbeit untersuchten Tikhonov-Filter verdeutlicht.

Ferner konnte gezeigt werden, dass man die maximale Ordnung M auch als Ordnung eines (modalen) *beamformer* auffassen kann, dessen Richtwirkung mit steigender Ordnung zunimmt. Die dabei maximal erreichbare Ordnung hängt abermals von dem verwendeten, kugelförmigen Mikrophonarray ab (offen oder geschlossen), den Eigenschaften der verwendeten Sensoren (Eigenrauschen usw.) sowie von der Größe des Arrays (Radius a). Davon jedoch unabhängig wird die effektiv erreichbare, maximale Ordnung stets mit sinkender Frequenz fallen. Auch in diesem Fall erwies sich, die Wahl den Radius der Kugel an die Kopfgröße des menschlichen Kopfes anzupassen als günstig, da sich letztendlich herausstellte, dass damit in etwa eine dem subjektiven Empfinden des Menschen ähnliche Lokalisationsgenauigkeit erreicht werden konnte. Das bedeutet, dass das menschliche Ge-

hör eine Lokalisationsgenauigkeit aufweist, die in ähnlichem Maße mit sinkender Frequenz abnimmt, wie das Bündelungsmaß eines modalen *beamformer*, der auf einer geschlossenen Kugel mit einem Radius basiert, der dem eines menschlichen Kopfes entspricht. Die Festlegung der Anzahl der regulär auf der Kugeloberfläche anzubringenden Mikrophone Q ist keine triviale Aufgabe. Einerseits wird hierdurch die maximale Ordnung M bestimmt, bis zu welcher das Wellenfeld analysiert werden kann - diese sollte möglichst groß ausfallen, um physikalisch einwandfreie Ergebnisse bis zu einer möglichst hohen Frequenz zu erhalten. Andererseits gibt es technische Grenzen, welche die maximale Anzahl von parallel zu verarbeiteten Mikrophenen Q begrenzt sowie physiologische Grenzen des menschlichen Gehörs. Das heißt, es gibt eine maximale Ordnungszahl M , welche die Lokalisationseigenschaften des menschlichen Gehörs vollständig abdeckt - eine weitere Steigerung von M wäre dann sinnlos, da wir Menschen dies ohnehin nicht mehr unterscheiden könnten. Wo diese physiologische Grenze letztendlich liegt, kann mit letzter Gewissheit derzeit nicht genau vorhergesagt werden. In [Far08] wird davon ausgegangen, dass bereits eine Ordnung von $M = 3$ ausreichen sollte, allerdings gibt es auch Experten, die dies bezweifeln und stattdessen eine maximale Ordnung von bis zu $M = 7$ (vgl. [FM12]) vorschlagen. Diese Arbeit beschränkt sich auf die Verwendung von $Q = 32$ omnidirektionalen Mikrophenen, die gemäß der in Kapitel A.25 aufgezeigten Weise auf einer soliden Styroporkugel mit einem Radius von $a = 0.09$ [m] angeordnet sind, womit eine Analyse des Wellenfeldes bis zu einer maximalen Ordnung von $M = 4$ möglich ist. Aufgrund der vorherrschenden geometrischen Verhältnisse und der vorliegenden maximalen Mikrophenabstände ergibt sich daraus eine minimale räumliche Aliasingfrequenz von $f_{Alias} \approx 2.58$ [kHz]. Diese räumliche Abtastfrequenz ist jedoch nicht physikalisch zwingend an die Mikrophenabstände gebunden, sondern ergibt sich aufgrund der störenden Einflüsse höherer, sphärischer Harmonischer, die mit steigender Frequenz stetig zunehmen (siehe Abbildung 5.14).

Mit Hilfe von räumlichen Tiefpassfilter ist es möglich, diesen störenden Einfluss dieser sphärischen Harmonischen höherer Ordnung zu begrenzen und dadurch den spektralen Einsatzbereich eines soliden, kugelförmigen Mikrophenarrays zu erweitern, was einen weiteren Schwerpunkt dieser Arbeit darstellt. Es wurden verschiedene Verfahren zur Steigerung der oberen Einsatzfrequenz aufgezeigt und analysiert, wobei sich vor allem die mechanische Veränderung der Kugeloberfläche zur Erhöhung der Richtwirkung der einzelnen Sensoren über der Frequenz als sehr interessante und einleuchtende Methode herauskristallisierte. Zudem existierten von dieser Methode lediglich erste Untersuchungen, die noch keine abschließende Aussage über deren Wirksamkeit zuließen, wodurch sich ein neues Forschungsfeld eröffnete, das mit dieser Arbeit behandelt wurde.

Zur Klärung der Frage, ob die mechanisch in die Oberfläche einer Kugel eingebrachten, die Richtwirkung der dort befindlichen Sensoren steigernde Modifikationen auch tatsächlich einen positiven Einfluss auf die obere Einsatzfrequenz von soliden, kugelförmigen Mikrophenarrays haben, war es unumgänglich, ein solches aufzubauen. Zur Verifikation der Wirkung dieser Form der räumlichen Tiefpassfilterung war es zudem unumgänglich,

zusätzlich ein solides, kugelförmiges Mikrofonarray gleichen Ausmaßes ohne modifizierter Oberfläche zu erstellen, welches dann als Referenz diene. Messungen dieser beiden Mikrofonarrays, die jeweils unter Verwendung des gleichen Abstandes und Lautsprecher durchgeführt wurden, ergaben dabei eindeutig, dass sich diese Form der räumlichen Tiefpassfilterung sehr gut dazu eignet, die obere Einsatzfrequenz eines solchen Mikrofonarrays zu steigern. In den hier untersuchten Beispiel konnte in etwa eine Verdoppelung der oberen Einsatzfrequenz von ursprünglich $f_{Alias} \approx 4$ [kHz] auf $f_{Alias} \approx 8$ [kHz] erreicht werden, was ein Blick auf die dabei erzielten Bündelungsmaße über der Frequenz, die aus Abbildung 6.4 ersichtlich sind, beweist. Somit konnte zum ersten Mal in einem praktischen Versuch bestätigt werden, dass diese einfache Maßnahme zur Realisierung eines räumlichen Tiefpassfilter sehr wirksam ist und aufgrund der dabei erzielten Ergebnisse empfohlen wird, diese Maßnahme beim Entwurf zukünftiger Generationen von soliden, kugelförmigen Mikrofonarrays zu berücksichtigen.

Außerdem wurde noch ein neuartiges, offenes Kugelmikrofonarray untersucht, welches auf einem Hexaeder basiert, auf dessen Flächenmittelpunkte sowie in dessen Zentrum jeweils ein omnidirektionales Mikrofon platziert wurde. Die Besonderheit dieser Anordnung ist, dass die an der Oberfläche der Kugel befindlichen omnidirektionalen Mikrophone mit Hilfe des zentral angeordneten, ebenfalls omnidirektionalen Mikrophon durch Bildung eines differentiellen Mikrophones 1. Ordnung auf einfache Weise zu einem jeweils gerichteten Mikrophon 1. Ordnung transformiert werden konnte, was bekanntermaßen eine Voraussetzung für einen störungsfreien Einsatz offener Kugelmikrofonarrays ist. Auch für dieses Array wurden mit Hilfe des in dieser Arbeit neue entwickelten Optimierungsverfahrens aus Kapitel 5.5 regularisierte *equalizing*-Filter entworfen, die auch hier ihre Wirksamkeit bestätigten. Messungen des aufgebauten Mikrofonarrays dieses Typs, welches einen Radius von $a = 0.0125$ [m] hatte, ergaben eine gute Übereinstimmung mit der Theorie, was sich etwa in der dabei erzielten, maximalen, räumlichen Aliasingfrequenz von $f_{Alias} \approx 13.7$ [kHz] widerspiegelte.

7.2 Ausblick

Ziel weiterführender Untersuchungen in diese hier propagierte Richtung sollte sein zu versuchen, eine weitere Steigerung der nutzbaren Bandbreite sphärischer Mikrofonarrays, vorrangig solider Bauart zu erreichen. Eine Möglichkeit, wie dies geschafft werden kann, ist, innerhalb der Dellen nicht ein omnidirektionales Mikrofon in den Fusspunkt, sondern neben diesen noch mehrere, omnidirektionale Mikrophone möglichst regulär verteilt auf der Oberfläche der Dellen zu platzieren, wie etwa in Abbildung 7.1 dargestellt.

Durch diese Maßnahme kann die effektive Membranfläche gesteigert und gleichzeitig das resultierende Eigenrauschen gemäß Gleichung 5.41 reduziert werden. Dabei muss sich die effektive Anzahl der zu verarbeitenden Eingangskanäle nicht erhöhen, sofern die Sensoren mit Hilfe einer einfachen, analogen Summierschaltung, wie in Abbildung 7.2 darge-

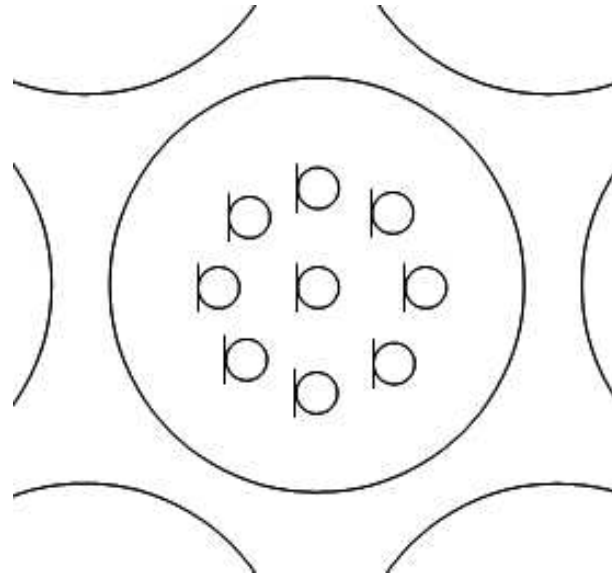


Abbildung 7.1: Exemplarische Darstellung einer Delle die mit mehreren omnidirektionalen Mikrofonen besetzt wurde.

stellt, pro Delle wieder zu einem Eingangskanal, gemäß der in Gleichung 7.1 angegebenen Weise, zusammengefasst werden.

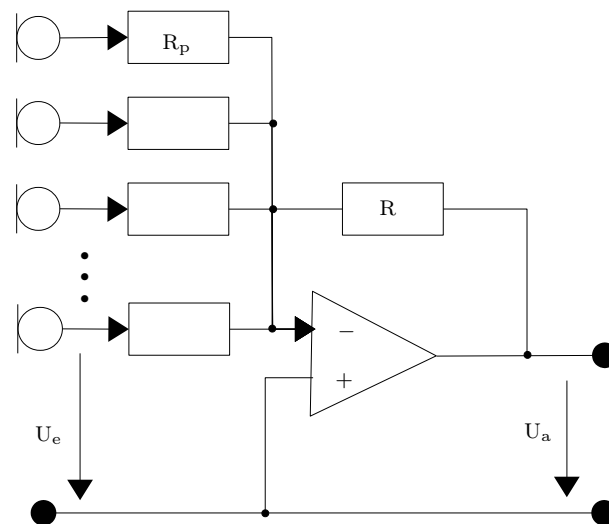


Abbildung 7.2: Prinzipieller Aufbau einer analogen Summierschaltung, mit deren Hilfe sich die Q_p omnidirektionalen Mikrophone pro Flächeneinheit (z.B. pro Delle) zusammenfassen lassen.

Die Werte der Widerstände R_p können nach [Hul04] gemäß einer der bekannten Fensterfunktionen (z.B. Hamming-Fenster) ausgelegt werden, sodass die Summation der einzelnen Sensorsignale, entsprechend gewichtet, erfolgen kann. Eine solche Fensterung stellt dabei eine weitere Möglichkeit zur Einflussnahme in die räumliche Tiefpassfilterwirkung dar, was den Entwurf des *beamformer* noch flexibler gestaltet.

$$\begin{aligned}
 U_a &= - \sum_{p=1}^{Q_p} \frac{R}{R_p} U_e, \text{ mit} \\
 U_a &= \text{Spannung des Ausgangssignals in [V]}, \\
 Q_p &= \text{Anzahl der omnidirektionalen Mikrophone pro Flächeneinheit,} \\
 R_p &= \text{Widerstand des } p. \text{ omnidirektionalen Mikrophons der zugehörigen Flächeneinheit} \\
 &\quad \text{in } [\Omega], \\
 U_e &= \text{Eingangsspannung der } Q_p \text{ Mikrophonsignale in [V].}
 \end{aligned} \tag{7.1}$$

Somit kann durch das geringere, effektive Rauschen die WNG-Schwelle beim Entwurf der regularisierten *equalizing*-Filter erniedrigt werden, was dazu führt, dass eine bessere Richtwirkung, d.h. ein höheres Bündelungsmaß auch bei niedrigeren Frequenzen erreicht werden kann. Falls eine weitere Steigerung des tieffrequenten Bündelungsmaßes jedoch als nicht notwendig erachtet werden sollte, wenn diese etwa die physiologisch bedingten Eigenschaften des menschlichen Gehörs übersteigen sollte, könnte aufgrund der nun niedrigeren WNG-Schwelle auch die Größe der Kugel verringert werden, was wiederum zu einer Erhöhung der oberen Einsatzfrequenz führen würde.

Unabhängig davon verstärkt sich aufgrund der vergrößerten, effektiven Membranfläche die räumliche Tiefpassfilterwirkung, was letztendlich zu einer weiteren Erhöhung der oberen Einsatzfrequenz führt, ohne dabei die effektive Anzahl, der zu verarbeitenden Kanälen erhöhen zu müssen.

Das am Ende des Kapitels 6.2 angegebene Konzept zur Modifikation eines offenen Kugelmikrophonarrays, basierend auf einem Hexaeder als Grundkörper, lässt sich wie nachfolgen ausgeführt erweitern, sodass letztendlich ein kompaktes, offenes Kugelmikrophonarray entsteht, auf dessen Basis ein modaler *beamformer* 2. Ordnung erzeugt werden kann, der im kompletten Audiofrequenzbereich Gültigkeit besitzt. Dies kann geschehen, indem die Abstände der innenliegenden Mikrophone z.B. auf $d_{\text{Mic}} \approx 0.85$ [cm] verringert werden, wodurch sich die obere Eckfrequenz dann auf $f \approx 20$ [kHz] steigert und zudem Mikrophone mit einem Abstand zum Zentrum von $d \approx 10$ [cm] angebracht werden, was einen effektiven Mikrophonabstand von $d_{\text{Mic}} \approx 20$ [cm] entspricht, um eine untere Einsatzfrequenz von $f \approx 20$ [Hz] zu erreichen. Auf diese Weise lässt sich, wie in Abbildung 7.3 dargestellt, auf einfache Art mit Hilfe von $Q = 13$ omnidirektionalen Mikrophenen ein modaler *beamformer* 2. Ordnung entwerfen, der im gesamten Audiofrequenzbereich eingesetzt werden kann.

Ferner erscheint es dringend notwendig weitere Untersuchungen bezüglich der minimal benötigten Ordnung M für die Nachbildung eines Wellenfeldes mit Hilfe von sphärischen Harmonischen zu unternehmen, welche die physiologischen Grenzen des menschlichen Gehörs erreicht bzw. übertrifft. Diese Aufgabe dürfte sich jedoch als ziemlich schwierig erweisen, da unser Gehör je nach verwendetem Testsignal und Darbietungs-

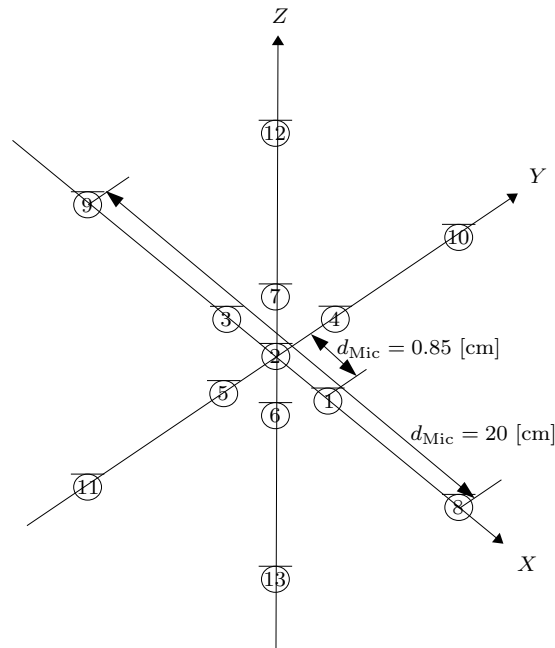


Abbildung 7.3: Beispielhafte Anordnung eines offenen Kugelmikrophonarrays, basierend auf einem Hexaeder als Grundkörper, zur Erzeugung eines modalen beamformer 2. Ordnung, bestehend aus $Q = 13$ omnidirektionalen Mikrofonen, welcher im gesamten Audiofrequenzbereich eingesetzt werden kann.

ebene (horizontale- oder vertikale Ebene) unterschiedliche Lokalisationsgenauigkeiten aufweist (Für Klicklaute, die besonders leicht bzw. scharf lokalisiert werden können gilt dabei für den Azimut: $\pm 3.6^\circ @ \theta = 0^\circ, \dots, \pm 10^\circ @ \theta = \pm 90^\circ$ und für die Elevation: $\pm 9^\circ @ \varphi = 0^\circ, \dots, \pm 22^\circ @ \varphi = 90^\circ$). Geht man unabhängig von der Darbietungsebene von einer maximalen Lokalisationsgenauigkeit von $\pm 10^\circ$ aus, was für „normale Signale“ (z.B. Musik) ausreichen sollte, würde es genügen eine Lokalisationsunschärfe von $\alpha_E \leq 20^\circ$ zu erreichen. Dies kann für 2D-Systeme, wie aus Tabelle 4.3 zu entnehmen, bereits für Ordnungen von $M \geq 4$ erreicht werden. Für 3D-Systeme ergibt sich hingegen, gemäß den Ausführungen aus Kapitel 4.2, eine maximale Lokalisationsunschärfe für eine Ordnung von $M = 4$ bei Verwendung des $\text{Max } r_e$ -Decoders von $\alpha_E \approx 25^\circ$. In diesem Fall würde eine Lokalisationsschärfe von $\alpha_E \leq 20^\circ$ erst ab einer Ordnung von $M \geq 6$ erreicht werden. Zudem stehen, wie in Kapitel 7.1 aufgezeigt, Aussagen unterschiedlicher Forscher im Raum, die behaupten, dass Ordnungen von $M \geq 3$ ([Far08]) bzw. von $M \geq 7$ ([FM12]) nötig seien, um dieses Ziel zu erreichen.

In diesem Zusammenhang wurden etwa in [FZS08] auch Experimente bezüglich der Lokalisationsgenauigkeit bei Verwendung unterschiedlicher Decodervarianten und Benutzung eines 2D-Lautsprecherrings mit 12 Lautsprechern und unterschiedlichen Ordnungen $M = [1, 3, 5]$ durchgeführt, wobei der $\text{Max } r_e$ -Decoder die beste Performance gegenüber dem Basic- und dem In-Phase-Decoder aufwies.

Sobald ein für die Vermessung des Wellenfeldes in Kraftfahrzeugen geeignetes Mikro-

phonarray gefunden ist, sollte dieses auch eingesetzt und die dabei erzielten Ergebnisse sowohl subjektiv als auch objektiv verifiziert werden. Sofern die dabei erzielten Ergebnisse qualitativ gut und reproduzierbar erscheinen, kann das hier vorgestellte Codier- und Decodierverfahren z.B. für die Erstellung einer virtuellen akustischen Umgebung verwendet werden. Dabei wird etwa ein Autoinnenraum (Quellraum) vermessen, d.h. codiert, um anschließend dessen Wellenfeld in einem anderen Raum (Zielraum), z.B. in einem akustischen Abhörraum wieder rekonstruieren, d.h. decodieren zu können. Dieser Vorgang wird als *akustische Dokumentation* bzw. als *Auralisation* bezeichnet.

Sobald eine Auralisation sowohl objektiv als auch subjektiv erfolgreich umgesetzt werden kann, ergeben sich weitere, interessante Möglichkeiten. Eine davon kann als *virtuelles Tuning* bezeichnet werden, wobei im Zielraum (z.B. Abhörraum, Studio) der Quellraum (z.B. Fahrzeuginnenraum) virtuell akustisch eingestellt werden kann, ohne das dabei der Quellraum physikalisch vorhanden sein muss. Auf diese Weise lassen sich die Zeiten an denen ein teurer und rarer Prototyp eines Fahrzeugs für eine akustische Abstimmung vor Ort sein muss, auf ein Minimum, d.h. lediglich auf die Zeit bis der Quellraum auf allen gewünschten Positionen im Raum (z.B. allen Sitzplätzen) codiert ist, reduzieren. Neben einer Reduzierung der Bereitstellungszeit von Prototypen lässt sich somit auch die Zeit für das akustische Tuning erhöhen, was letztendlich zu einer gesteigerten akustischen Qualität führt.

Eine logische Weiterentwicklung ist die Erstellung eines Tools zum automatischen, akustischen *equalizing*. Das Einzige was dabei benötigt wird, ist die Bereitstellung der Daten eines codierten *Referenzraums*. Dazu könnte etwa eine für gut befundene akustische Umgebung (Quellraum) dienen, die vermessen und anschließend codiert wird, um anschließend den Zielraum mit Hilfe der vorhandene Möglichkeiten zum akustischen *equalizing* so zu modifizieren, dass dessen Codierung so exakt wie möglich der des Referenzraums entspricht. Auf diese Weise ist es möglich, mit den vorhandenen Mitteln (Anzahl der vorhandenen Kanälen, Lautsprecher usw.) das Wellenfeld und somit die Akustik eines beliebigen Zielraums (z.B. Innenraum eines Fahrzeugs) an ein beliebiges Wellenfeld, welches beispielsweise ebenfalls vom Fahrzeughersteller vorgegeben werden kann und als Referenz für die Decodierung dient, so gut wie möglich anzugleichen. Somit könnte in Zukunft neben anderen Dingen, wie dem Klang des Motors oder dem Design, auch die Akustik eines Fahrzeugs, eine dem Markennamen entsprechenden und vom jeweiligen Kunden definierbaren Charakter aufweisen und somit den Wiedererkennungswert eines Produkts steigern. Die Realisierung einer marken- bzw. produktspezifischen Akustik wäre demnach mit Hilfe dieser Technik möglich.

Neben diesen Einsatzgebieten bietet vor allem die Verwendung von modalen *beamformer* weitere interessante Möglichkeiten. So bilden diese etwa die Basis für einen erfolgreichen Einsatz einer *akustischen Kamera* (siehe Abbildung 5.5). Aber auch ohne Überlagerung mit einem optischen Bild sind modale *beamformer* nützlich, um etwa Klappern im Fahrzeuginnenraum während der Fahrt bzw. beim Betrieb der Audioanlage aufzuspüren

oder herauszufinden aus welcher Raumrichtung, zu welchen Zeiten, verursacht von einem oder mehreren Lautsprechern Reflexionen generiert werden, die wiederum Aufschluss über die Qualität des Lautsprechereinbaus bzw. der Lautsprecheranordnung liefern. Da ein solches Wissen essentiell für die erreichbare akustische Qualität ist, können Möglichkeit solche Defekte messtechnisch erfassen zu können, nicht hoch genug eingeschätzt werden. Um jedoch das volle Potential dieser Möglichkeiten auszuschöpfen, fehlt es derzeit noch an geeigneten, leicht verständlichen und bedienbaren Tools für die Analyse - in diese Richtung eröffnet sich ein weiteres, großes Feld für eine zukünftige Forschung.

Darüber hinaus sind modale *beamformer* sehr flexibel. Es lassen sich *beamformer* mit beliebiger Richtcharakteristik in beliebiger Anzahl rein softwaretechnisch erzeugen, die sich zudem ebenfalls nur mit Hilfe der Signalverarbeitung in Echtzeit oder zu einem späteren Zeitpunkt in jede beliebige dreidimensionale Raumrichtung steuern lassen.

Dieser Ansatz hat nicht nur für die Gestaltung von Aufnahmesystemen, d.h. von Kugelmikrophonarrays Gültigkeit, sondern kann auch für Wiedergabesysteme angewendet werden. Dazu ist anstelle eines Kugelmikrophonarrays ein Kugellautsprecherarray notwendig, bei dem die einzelnen, im Idealfall omnidirektionalen Lautsprecher, ebenfalls auf einer (soliden) Kugeloberfläche möglichst regulär angeordnet sein müssen, sodass auch hier die Orthonormalitätsbedingung aus Gleichung 5.7 möglichst exakt erfüllt wird. Zu diesem Zweck wurden zwei Kugellautsprecherarrays aufgebaut, die in Abbildung 7.4 dargestellt sind, wobei hier der aus Kapitel A.5.5 bekannte Ikosaeder (platonischer Körper) als Grundkörper verwendet wurde, also jedes Array aus $Q = 20$ Einzellautsprecher mit einem Durchmesser $\varnothing_{\text{Lautsprecher}}$ von je $\varnothing_{\text{Lautsprecher}} = 10$ [cm] bestand.

Die Befestigungskörbe der Lautsprecher erhöhten deren effektiven Durchmesser $\varnothing_{\text{LautsprecherEffektiv}}$ auf $\varnothing_{\text{LautsprecherEffektiv}} = 11.3$ [cm], was unter Berücksichtigung einer für den Bau notwendigen Reserve von ≈ 0.7 [cm] zu einer Kantenlänge $a_{\Delta} = (r_{\text{LautsprecherEffektiv}} + \text{Reserve}) \frac{6}{\sqrt{3}}$ der gleichseitigen Dreiecke, die als Grundfläche des Ikosaeders dienen, von $a_{\Delta} = 20.8$ [cm] führte. Schließlich wurden gleichseitige Dreiecke Δ mit einer Kantenlänge von $a_{\Delta} = 21$ [cm] verwendet, die jeweils einen *Gehrungswinkel* \angle_{Gehrung} von $\angle_{\text{Gehrung}} = 21^{\circ}$ aufweisen mussten, damit zwei zusammengesetzte Flächen einen für Ikosaeder typischen *Flächenwinkel*¹ $\angle_{\text{Fläche}}$ von $\angle_{\text{Fläche}} = 138^{\circ}$ aufwiesen. Somit ergab sich ein Ikosaeder mit einem Volumen $V_{\text{Ikosaeder}}$ von $V_{\text{Ikosaeder}} = \frac{5}{12} (3 + \sqrt{2}) a_{\Delta}^3 = 0.0249$ [m³] @ $a_{\Delta} = 21$ [cm] sowie einem äußeren Durchmesser $\varnothing_{\text{IkosaederAußen}}$ von $\varnothing_{\text{IkosaederAußen}} = \frac{\sqrt{10+2\sqrt{5}}}{2} a_{\Delta} \approx 40$ [cm] @ $a_{\Delta} = 21$ [cm]. Der spektrale Einsatzbereich des hier vorgestellten Kugellautsprecherarrays kann zur Zeit nur geschätzt werden, da diese noch nicht vermessen wurden. Jedoch dürfte sich dieser zwischen $f \approx [100, \dots, 1340]$ [Hz] befinden, wobei sich die untere Grenze durch die Lautsprecher bestimmt und die oberen Grenzfrequenz aus der Geometrie des Kugellautsprechers, genauer gesagt aus dem Abstand der einzelnen Lautsprecher, resultiert. Der Abstand entspricht dabei der Seitenlänge der Grundfläche des dem Ikosaeder eingeschrie-

¹Einen von zwei Flächen gebildeten Winkel bezeichnet man als *Flächenwinkel*.



Abbildung 7.4: Darstellung der beiden fertiggestellten Kugellautsprecherarrays, die jeweils auf einen Ikosaeder als Grundkörper basieren und einen Außendurchmesser von jeweils $\varnothing_{\text{Ikosaeder}_{\text{Außen}}} \approx 40$ [cm] aufweisen.

benen Dualkörpers (Dodekaeders) a_{Pentagon} , wobei es sich hierbei, wie aus Kapitel A.5.5 bekannt, um ein Pentagon, also um ein regelmäßiges Fünfeck handelt. Der Durchmesser des Inkreises eines Ikosaeder $\varnothing_{\text{Ikosaeder}_{\text{Innen}}}$ berechnet sich dabei nach [ZPS08] wie folgt: $\varnothing_{\text{Ikosaeder}_{\text{Innen}}} = \varnothing_{\text{Ikosaeder}_{\text{Außen}}} \sqrt{\frac{7+3\sqrt{5}}{3(5+\sqrt{5})}} = 31.8$ [cm]. Aus diesem ergibt sich gemäß [ZPS08] die Seitenlänge des Pentagons und somit der Abstand der einzelnen Lautsprecher zu: $a_{\text{Pentagon}} = \frac{2\varnothing_{\text{Ikosaeder}_{\text{Innen}}}}{(1+\sqrt{5})\sqrt{3}} = 12.8$ [cm]. Aus diesem ergibt sich mit Hilfe der Gleichung 5.8 eine obere Grenzfrequenz von $f = 1340$ [Hz].

Näheres bezüglich Kugellautsprecherarrays kann der interessierte Leser z.B. in [ZN07, ZPS, ZPS08, Zot09, Zot12] finden.

Anhang A

Anhang

A.1 Differentielle Mikrophone

Differentielle Mikrophone gehören zu der Gattung der *Richtmikrophone*. *Gradientenmikrophone*, wie sie auch genannt werden, besitzen bei kleinen Abmessungen eine gute Richtwirkung.

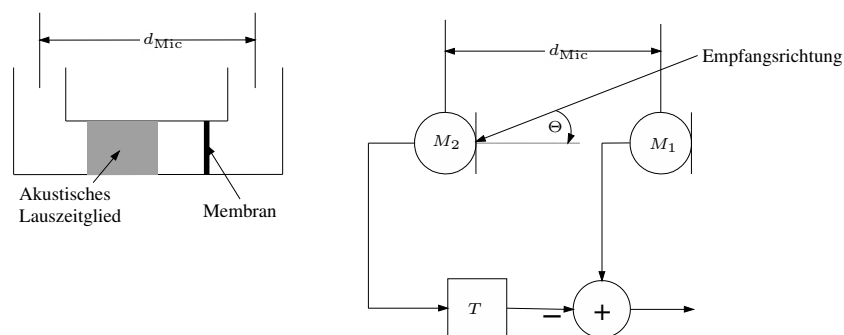


Abbildung A.1: Differentielle Mikrophone 1. Ordnung (Links: Prinzipieller Aufbau; Rechts: Ersatzschaltung).

Das linke Teilbild von Abbildung A.1 zeigt den prinzipiellen Aufbau eines Gradientenmikrophons erster Ordnung sowie im rechten Teilbild eine äquivalente Ersatzschaltung, basierend auf zwei omnidirektionalen Mikrophen, deren Ausgangssignale in geeigneter Weise kombiniert werden. Bei dem Gradientenmikrophon ist die Membran des elektroakustischen Wändlers derart angeordnet, dass der Schall die Membran von beiden Seiten erreicht und somit eine Membranauslenkung nur durch eine *Schalldruckdifferenz* hervorgerufen wird. Das reine Gradientenmikrophon erster Ordnung, bei dem die Membran exakt in der Mitte platziert ist, ohne dass sich ein akustisches Laufzeitglied dazwischen befindet, stellt einen *Dipol* dar, also ein Mikrophon mit Achtercharakteristik. Mit Hilfe des akustischen Laufzeitglieds T kann die Charakteristik z.B. so verändert werden, dass das resultierende Mikrophon nur noch eine einzige *Hauptkeule* aufweist, was im Allgemeinen

als *Richtmikrofon* bezeichnet wird. Bei der im rechten Teilbild von Abbildung A.1 dargestellten Anordnung weist die Hauptkeule in die Richtung $\Theta = 0^\circ$, d.h. nach rechts. Das akustische Laufzeitglied kann im einfachsten Fall, wie im linken Teil von Abbildung A.1 dargestellt, durch eine asymmetrische Anordnung der Membran zwischen den beiden Schalleintrittsöffnungen und/oder durch den Einsatz eines Mediums mit einer niedrigeren Schallgeschwindigkeit realisiert werden. Beschreibt man die mechanische Anordnung, wie im rechten Teil von Abbildung A.1 dargestellt, durch ein signaltheoretisches Blockschaltbild, so sind die beiden Schalleintrittsöffnungen durch zwei omnidirektionale Mikrophone und die auf die Schalldruckdifferenz reagierende Membran durch eine Subtraktion der Mikrophonsignale zu ersetzen, wobei das Signal des aus Sicht der Nutzschaquelle hinteren Mikrophons M_2 entsprechend dem akustischen Laufzeitglied verzögert wird. Soll das Gradientenmikrofon für die Haupteinfallrichtung ein frequenzunabhängiges Verhalten aufweisen, ist die Ersatzschaltung um ein Tiefpassfilter $H_{TP}(f)$ mit einer zur Frequenz reziproken Übertragungsfunktion $H_{TP}(f) \approx f^{-1}$ zu erweitern, um die zu tiefen Frequenzen hin zunehmende Empfindlichkeit des *Druckgradientenempfängers* zu kompensieren. Dieser Tiefpass kann bei einem mechanisch realisierten Richtmikrofon (linkes Teilbild von Abbildung A.1) auch elektromechanisch durch Einsatz eines entsprechenden Wandlers oder mit Hilfe akustischer Filter erfolgen.

Die Wirkungsweise des Richtmikrophons kann anschaulich am Beispiel der *kardioiden* Charakteristik erläutert werden. Diese Richtcharakteristik zeichnet sich dadurch aus, dass die rückwärtige Einfallrichtung ($\Theta = 180^\circ$) vollständig unterdrückt wird. Um eine kardioide Charakteristik zu erreichen, ist die Verzögerung des akustischen Laufzeitglieds T derart zu wählen, dass sie die Zeitdifferenz zwischen den beiden Schalleintrittsöffnungen bei Schalleinfall aus Richtung $\Theta = 180^\circ$ entspricht, also:

$$\begin{aligned} T &= \frac{d_{\text{Mic}}}{c}, \text{ mit} \\ d_{\text{Mic}} &= \text{Mikrofonabstand in [m]}, \\ c &= \text{Schallgeschwindigkeit (Bei } \vartheta = 25^\circ \text{ ist } c = 346 \left[\frac{\text{m}}{\text{s}} \right]). \end{aligned} \quad (\text{A.1})$$

Mit dieser Wahl der Zeitkonstanten sind bei Schalleinfall aus der rückwärtigen Richtung die beiden in den Subtraktionspunkt einfließenden Signale um die gleiche Zeit verzögert, sodass die Subtraktion im Idealfall zu einer perfekten Auslöschung führt und somit kein Ausgangssignal generiert wird. Andere Richtcharakteristika ergeben sich durch die Wahl einer geringeren Verzögerung, sodass die Nullstelle für eine andere Einfallrichtung (beispielsweise $\Theta = \pm 109^\circ$ im Fall der *hyperkardioiden* Charakteristik) auftritt. Die Verzögerung T darf also lediglich aus einem Bereich von: $0 \leq T \leq \frac{d_{\text{Mic}}}{c}$ gewählt werden, wobei bei $T = 0$ eine Dipol- und bei $T = \frac{d_{\text{Mic}}}{c}$, wie bereits erörtert, eine kardioide Charakteristik entsteht. Durch geschickte Verknüpfung mehrerer Gradientenmikrophone erster Ordnung können auf relativ einfache Weise auch differentielle Mikrophone höherer Ordnung realisiert werden.

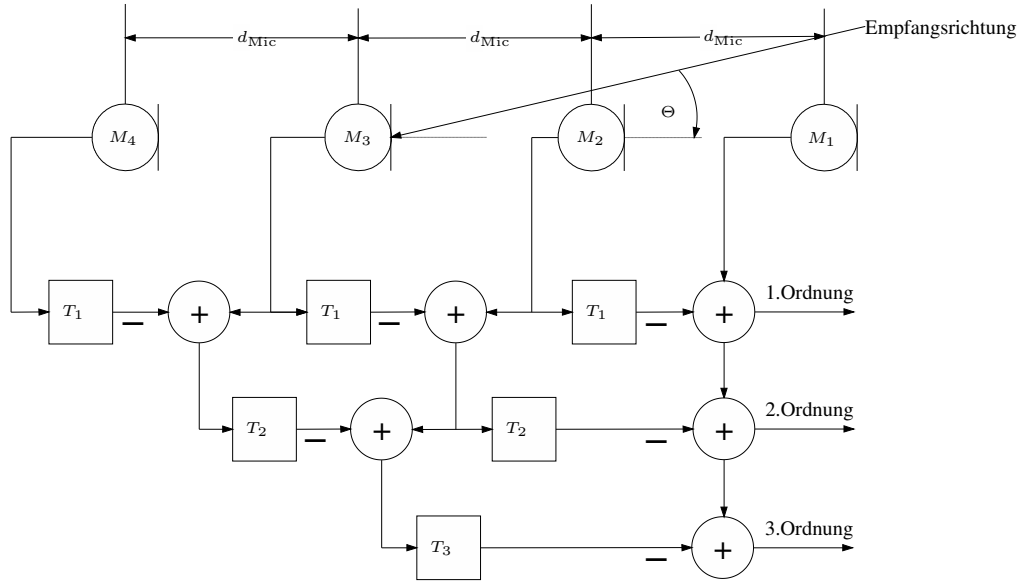


Abbildung A.2: Differentielles Mikrofonarray 3. Ordnung, zusammengesetzt aus Gradientenmikrophone 1. Ordnung.

A.1.1 Berechnung der Verzögerungszeiten T_i

Die Verzögerungszeiten T_i sind prinzipiell nur vom Mikrophonabstand d_{Mic} sowie von der gewünschten Richtcharakteristik abhängig. Damit die Richtcharakteristik von den Verzögerungszeiten T_i entkoppelt werden kann, werden neue Parameter α_i eingeführt, mit deren Hilfe sich die Verzögerungszeiten T_i folgendermaßen berechnen:

$$T_i = \frac{d_{Mic}}{c} \left(\frac{\alpha_i}{1 - \alpha_i} \right), \text{ mit}$$

$$\alpha_i = \text{Parameter zur Einstellung der Richtcharakteristik}$$

eines differentiellen Mikrophons i . Ordnung. (A.2)

Durch Variation der Parameter α_i können also beliebige Richtcharakteristika eingestellt werden, wobei in der Praxis üblicherweise nur an bestimmten Charakteristika ein Interesse besteht. Die Rede ist dabei vor allem von den Richtcharakteristika: *Dipol*, *Kardioide* (Nierencharakteristik), *Hyperkardioide* (Hypernierencharakteristik) und *Superkardioide* (Supernierencharakteristik). Dabei weist die bidirektionale Charakteristik (Dipol), zumindest beim Design erster- und zweiter Ordnung den geringsten Öffnungswinkel auf. Die kardioide Charakteristik zeichnet sich, wie bereits erläutert dadurch aus, dass sie eine (doppelte) Nullstelle bei $\Theta = 180^\circ$ hat und dementsprechend Störer, die sich hinter dem Array befinden, perfekt unterdrückt. Die hyperkardioide Charakteristik maximiert den *Array-Gewinn* und besitzt für Ordnungen größer zwei die geringste 3[dB]-Bandbreite. Der Vorteil der superkardioiden Charakteristik besteht darin, dass sie über das optimale *Vor/Rück-Verhältnis* verfügt. Die Richtcharakteristikparameter α_i können nicht ohne

weiteres berechnet werden, sondern werden in der Regel je nach Ordnungszahl i empirisch ermittelt. Tabelle A.1 listet für alle oben genannten Richtcharakteristika bis zur dritten Ordnung sämtliche Parameterwerte α_i zusammen mit den daraus resultierenden Arrayparameterwerten auf.

Richtcharakteristik	Richtcharakteristikparameter α_i	Maximaler Array-Gewinn $G_{\text{Max}}(\omega)$ in [dB]	Maximales Vor/Rück-Verhältnis $FBR_{\text{Max}}(\omega)$ in [dB] ¹	Minimale 3 [dB]-Bandbreite in [Grad]	Räumliche Nullstelle(n) $\Theta_{-\infty}$ in [Grad]
---------------------	---	---	---	--------------------------------------	--

Design 1. Ordnung ($i = 1$)

Dipol	0	4.8	0	90	90
Kardioide	0.5	4.8	8.5	131	180
Hyperkardioide	0.25	6	8.5	105	109
Superkardioide	0.366	5.7	11.4	115	125

Design 2. Ordnung ($i = 2$)

Dipol	[0, 0]	7	0	65	90
Kardioide	[0.5, 0.5]	7	14.9	94	180
Hyperkardioide	[0.41, -0.41]	9.5	8.5	66	[73, 134]
Superkardioide	[0.2, 0.45]	8.3	24	80	[104, 144]

Design 3. Ordnung ($i = 3$)

Dipol	[0, 0, 0]	8.5	0	54	90
Kardioide	[0.5, 0.5, 0.5]	8.5	21	78	180
Hyperkardioide	[0.45, 0.15, -1.35]	12	11.2	48	[55, 100, 145]
Superkardioide	[0.113, 0.473, 0.346]	9.9	37.7	66	[97, 122, 153]

Tabelle A.1: Parameter für den Entwurf differentieller Mikrophone 1., 2. und 3. Ordnung gemäß [Elk00].

A.1.2 Berechnung der Arrayparameter differentieller Mikrophone

In der Tabelle A.1 werden die Auswirkungen der Richtcharakteristikparameter α_i auf unterschiedliche Arrayparameter aufgezeigt, ohne bis dato genauer darauf eingegangen zu sein, was hinter diesen Begriffen steckt bzw. wie diese berechnet werden. Aus diesem Grund wird nachfolgend aufgezeigt, wie sich die wichtigsten Arrayparameter differentieller Mikrophone berechnen.

¹FBR = engl. *Front-to-Back Ratio* (Vor/Rück-Verhältnis)

Berechnung des Richtdiagramms $\Psi(\omega, \Theta)$

Eine anschauliche Größe zur Beschreibung des Richtverhaltens eines Arrays stellt das *Richtdiagramm* $\Psi(\omega, \Theta)$ dar, wobei dieses die Empfindlichkeit des Arrays in Abhängigkeit von der Einfallsrichtung Θ bei verschiedenen Frequenzen ω angibt.

$$\begin{aligned} \Psi(\omega, \Theta) &= \prod_{i=1}^M W_i(\omega) j e^{-j\pi f T_i} \sin \left(\left(\frac{\pi f d_{\text{Mic}}}{c} \right) \left(\frac{T_i c}{d_{\text{Mic}}} + \cos(\Theta) \right) \right), \text{ mit} \\ i &= 1, \dots, M, \text{ wobei} \\ \Psi(\omega, \Theta) &= \text{Richtdiagramm,} \\ W_i(\omega) &= i. \text{ equalizing-Tiefpassfilter,} \\ T_i &= i. \text{ Verzögerungselement in [s],} \\ f &= \text{Frequenz in [Hz],} \\ d_{\text{Mic}} &= \text{Mikrophonabstand in [m],} \\ M &= \text{Anzahl der Mikrophone,} \\ \Theta &= \text{Empfangs- bzw. Einfallsrichtung in [Rad],} \\ c &= \text{Schallgeschwindigkeit in } \left[\frac{\text{m}}{\text{s}} \right]. \end{aligned} \tag{A.3}$$

Berechnung der *equalizing*-Tiefpassfilter $W_i(\omega)$

$$\begin{aligned} W_i(\omega) &= \begin{cases} \frac{1}{\sin\left(\frac{\pi f}{2f_{c_i}}\right)} & \text{wenn } 0 < f \leq f_{c_i} \\ 1 & \text{sonst} \end{cases}, \text{ mit} \\ f_{c_i} &= \frac{1}{2 \left(\frac{d_{\text{Mic}}}{c} + T_i \right)}, \\ i &= 1, \dots, M, \text{ wobei} \\ f_{c_i} &= \text{Grenzfrequenz des } i. \text{ equalizing-Filter in [Hz].} \end{aligned} \tag{A.4}$$

Oberhalb von f_c macht sich mehr und mehr eine Nullstelle in der Hauptempfangsrichtung Θ_0 bemerkbar, wobei diese bei Frequenzen, für die gilt:

$$\begin{aligned} f &= \Lambda \frac{c}{d_{\text{Mic}} (1 - \cos(\Theta_{-\infty}))}, \text{ mit} \\ \Lambda &\in \mathbb{R}^+, \\ \Theta_0 &= \text{Hauptempfangsrichtung in [Grad],} \\ \Theta_{-\infty} &= \text{Nullstelle der Richtcharakteristik in [Grad],} \\ \mathbb{R}^+ &= \text{Menge der positiven, reellen Zahlen ohne der Null,} \end{aligned} \tag{A.5}$$

exakt mit Θ_0 zusammenfällt und somit einen Empfang aus der Hauptempfangsrichtung unmöglich macht. Bei der hyperkardioiden Richtcharakteristik eines differentiellen Mikrofonarrays erster Ordnung beträgt dieser $\Theta_{-\infty} \pm 109^\circ$.

Berechnung des Array-Gewinns $G(\omega)$

Der Gewinn eines Arrays ist ein kompakteres Maß zur Beschreibung der Richtwirkung, das im Gegensatz zum Richtdiagramm nicht vom Einfallswinkel Θ abhängt. Der Gewinn ist definiert als die Empfindlichkeit des Arrays in Haupteinfallrichtung Θ_0 bezogen auf die Empfindlichkeit bei omnidirektionalem, d.h. diffusem Schalleinfall.

$$\begin{aligned} G(\omega) &= M \frac{|\Psi(\omega, \Theta_0)|}{\frac{1}{2\pi} \int_{\Theta=0}^{\Theta=2\pi} |\Psi(\omega, \Theta)| d\Theta}, \text{ wobei} \\ G(\omega) &= \text{Array Gewinn,} \\ m &= \text{Anzahl der Mikrophone.} \end{aligned} \tag{A.6}$$

Berechnung des Vor/Rück-Verhältnis $FBR(\omega)$

Das Vor/Rück-Verhältnis ($FBR(\omega)$) ist ein Maß dafür, um wie viel stärker Signale aus der vorderen Empfangsrichtung, also der Halbkugel bzw. Halbebene in Blickrichtung, im Verhältnis zur hinteren Empfangsrichtung empfangen werden.

$$\begin{aligned} FBR(\omega) &= \frac{\int_{\Theta=-\frac{\pi}{2}}^{\Theta=\frac{\pi}{2}} |\Psi(\omega, \Theta)|^2 d\Theta}{\int_{\Theta=\frac{\pi}{2}}^{\Theta=\frac{3\pi}{2}} |\Psi(\omega, \Theta)|^2 d\Theta}, \text{ wobei} \\ FBR(\omega) &= \text{Vor/Rück-Verhältnis über der Frequenz (front-to-back ratio)} \end{aligned} \tag{A.7}$$

Berechnung des Suszeptibilität $K(\omega)$

Die letzte Kenngröße, die sogenannte *Suszeptibilität* ($K(\omega)$), beschreibt die Fähigkeit des Arrays, räumlich unkorreliertes Rauschen, wie z.B. das Eigenrauschen der Mikrophone zu verstärken. Werte der Suszeptibilität $K(\omega) < 1$ repräsentieren dabei eine Unterdrückung von unkorrelierten Rauschen, wogegen Werte von $K(\omega) > 1$ eine Verstärkung des selbigen darstellen.

$$\begin{aligned} K(\omega) &= \frac{M+1}{\left[\prod_{i=1}^M 2 \sin\left(\pi f \left(\frac{d_{\text{Mic}}}{c} + T_i\right)\right) \right]^2}, \text{ oder wenn } \left(\omega \frac{d_{\text{Mic}}}{c}\right) \ll 1 \Rightarrow \\ K(\omega) &\approx \frac{M+1}{\left[\prod_{i=1}^M 2\pi f \left(\frac{d_{\text{Mic}}}{c} + T_i\right) \right]^2}, \text{ mit} \\ i &= 1, \dots, M, \\ K(\omega) &= \text{Suszeptibilität über der Frequenz.} \end{aligned} \tag{A.8}$$

A.1.3 Modelle nach Gerzon zur Vorhersage der subjektiv empfundenen Horizontallokalisation

Zur Vorhersage der empfundenen Horizontallokalisation entwickelte Gerzon zwei Theorien, nämlich die:

- *Schallschnellevektorthorie* sowie die
- *Schallenergievektorthorie*.

Der Unterschied der beiden Theorien ist in ihrem spektralen Einsatzgebiet zu finden. Die Schallschnellevektorthorie liefert für Frequenzen unterhalb von $f < 700$ [Hz] die besten Ergebnisse bei der Vorhersage der empfundenen Horizontallokalisation, wobei ihre Ergebnisse durchaus bis zu einer maximalen Frequenz von $f < 2000$ [Hz] verwendbar wären. Wie ihr spektrales Einsatzgebiet vermuten lässt, werden hier interaurale Laufzeitunterschiede, die in diesem Frequenzbereich die Horizontallokalisation dominieren, verwendet. Gemäß der Schallschnellevektorthorie kann diese aus dem Verhältnis des Schallschnellektors \mathbf{v} und des Schalldrucks p am Abhörort geschätzt werden. Der zweidimensionale Schallschnellevektor $\mathbf{v} = (v_x, v_y)$ kann für eine beliebige Anzahl von Lautsprechern wie folgt geschätzt werden:

$$\begin{aligned}
 v_x &= \sum_{i=1}^L G_i \cdot \cos(\Theta_i), \\
 v_y &= \sum_{i=1}^L G_i \cdot \sin(\Theta_i), \\
 \mathbf{v} &= (v_x, v_y), \\
 p &= \sum_{i=1}^L G_i, \\
 L &= \text{Anzahl der Lautsprecher}, \\
 G_i &= \text{Verstärkungsfaktor des } i. \text{ Lautsprechers}, \\
 \Theta_i &= \text{Horizontaler Aufstellwinkel des } i. \text{ Lautsprechers, gemäß Abbildung A.3}, \\
 p &= \text{Gesamter Schalldruckpegel an der zentralen Abhörposition.} \tag{A.9}
 \end{aligned}$$

Die *empfundene Horizontallokalisation* $\tilde{\Theta}_{\mathbf{v}}$ in kartesischen Koordinaten ergibt sich demnach aus:

$$\begin{aligned}
 (\tilde{\Theta}_{v_x}, \tilde{\Theta}_{v_y}) &= \left(\frac{\mathbf{v}}{p} \right) = \left(\frac{v_x}{p}, \frac{v_y}{p} \right), \text{ mit} \\
 (\tilde{\Theta}_{v_x}, \tilde{\Theta}_{v_y}) &= \text{Schallschnellevektorbasierte Horizontallokalisation in kartesischen} \\
 &\quad \text{Koordinaten} \tag{A.10}
 \end{aligned}$$

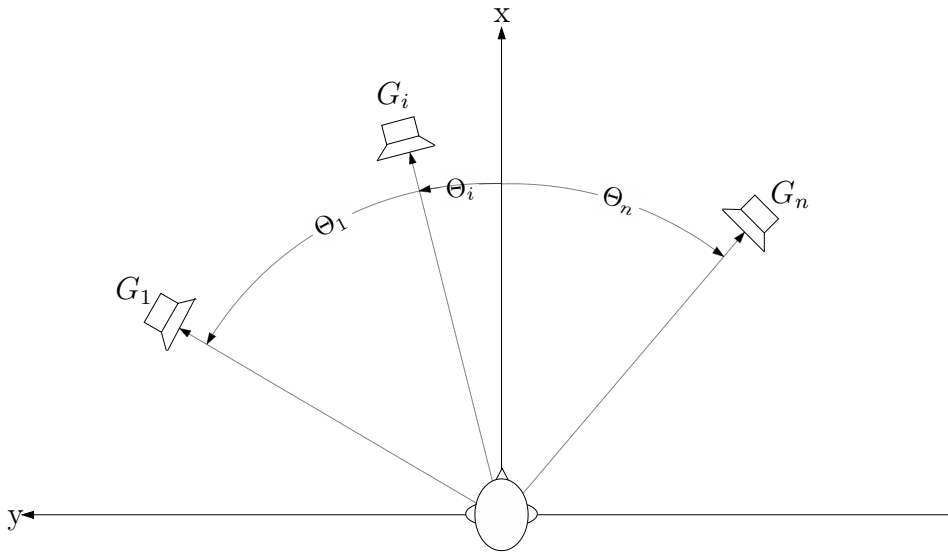


Abbildung A.3: Koordinatensystem nach Gerzon [Ger92b].

Eine anschließende Wandlung der kartesischen Koordinaten in Polarkoordinaten liefert einerseits die gesuchte Approximation der empfundenen Horizontallokalisation $\tilde{\Theta}_{\mathbf{v}}$, d.h. den Winkel des Schallschnellektors \mathbf{v} und andererseits dessen Länge $r_{\mathbf{v}}$.

$$\begin{aligned} \tilde{\Theta}_{\mathbf{v}} &= \arctan\left(\frac{\tilde{\Theta}_{v_y}}{\tilde{\Theta}_{v_x}}\right) \\ r_{\mathbf{v}} &= \sqrt{\tilde{\Theta}_{v_x}^2 + \tilde{\Theta}_{v_y}^2}, \text{ mit} \\ \tilde{\Theta}_{\mathbf{v}} &= \text{Winkel des Schallschnellektors } \mathbf{v} \\ r_{\mathbf{v}} &= \text{Länge des Schallschnellektors } \mathbf{v} \end{aligned} \tag{A.11}$$

Die Länge $r_{\mathbf{v}}$ des Schallschnellektors kann dabei maximal einen Wert von 1 annehmen, wobei dieser eine kompakte Horizontallokalisation kennzeichnet. Ein kleiner werdender Wert $r_{\mathbf{v}}$ bezeichnet hingegen eine zunehmend diffuser werdende und somit schwerer zuordenbare Horizontallokalisation. Bei höheren Frequenzen gibt Gerzon einen Frequenzbereich von $f \in [700, \dots, 5000]$ [Hz] an. Hier liefert die Schallenergievektorthorie Werte, die besser mit den Ergebnissen der Hörversuche zur Horizontallokalisation übereinstimmen. Dabei dominieren die interauralen Pegelunterschiede, die sich gemäß [Ger92b] aus dem Verhältnis der *Schallintensität*, gegeben durch den *Schallintensitätsvektor* \mathbf{e} und der gesamten Energie am Abhörort E schätzen lassen. Der zweidimensionale Schallintensitätsvektor $\mathbf{e} = (e_x, e_y)$ kann für eine beliebige Anzahl von Lautsprechern wie folgt geschätzt

werden:

$$\begin{aligned}
 e_x &= \sum_{i=1}^L |G_i|^2 \cdot \cos(\Theta_i), \\
 e_y &= \sum_{i=1}^L |G_i|^2 \cdot \sin(\Theta_i), \\
 \mathbf{e} &= (e_x, e_y), \\
 E &= \sum_{i=1}^L |G_i|^2, \text{ mit} \\
 E &= \text{Gesamte Schallenergie an der zentralen Abhörposition.} \quad (\text{A.12})
 \end{aligned}$$

Die empfundene Horizontallokalisation $\tilde{\Theta}_{\mathbf{e}}$ in kartesischen Koordinaten ergibt sich demnach aus:

$$\begin{aligned}
 (\tilde{\Theta}_{e_x}, \tilde{\Theta}_{e_y}) &= \left(\frac{\mathbf{e}}{E} \right) = \left(\frac{e_x}{E}, \frac{e_y}{E} \right) \text{ mit} \\
 (\tilde{\Theta}_{e_x}, \tilde{\Theta}_{e_y}) &= \text{Schallintensitätsvektorbasierte Horizontallokalisation in kartesischen} \\
 &\quad \text{Koordinaten.} \quad (\text{A.13})
 \end{aligned}$$

Eine anschließende Wandlung der kartesischen Koordinaten in Polarkoordinaten liefert auch hier zum einen die Approximation der empfundenen Horizontallokalisation $\tilde{\Theta}_{\mathbf{e}}$, die nun aber den Winkel des Schallenergievektors \mathbf{e} angibt und zum anderen dessen Länge $r_{\mathbf{e}}$ so zu interpretieren ist wie der Wert von $r_{\mathbf{v}}$:

$$\begin{aligned}
 \tilde{\Theta}_{\mathbf{e}} &= \arctan \left(\frac{\tilde{\Theta}_{e_y}}{\tilde{\Theta}_{e_x}} \right) \\
 r_{\mathbf{e}} &= \sqrt{\tilde{\Theta}_{e_x}^2 + \tilde{\Theta}_{e_y}^2}, \text{ mit} \\
 \tilde{\Theta}_{\mathbf{e}} &= \text{Winkel des Schallenergievektors } \mathbf{e}, \\
 r_{\mathbf{e}} &= \text{Länge des Schallenergievektors } \mathbf{e}. \quad (\text{A.14})
 \end{aligned}$$

Wendet man die beiden Theorien nach Gerzon gemäß der Gleichungen A.9 - A.14 zur Schätzung der empfundenen Horizontallokalisation, am Beispiel einer klassischen Stereoaufstellung² an, ergeben sich für die beiden Theorien, die in Abbildung A.4 dargestellten Verläufe.

Dabei wurden die Verstärkungsfaktoren für den linken und rechten Lautsprecher in einem Bereich von $[-30, \dots, 30]$ [dB] in 1 [dB]-Schritten variiert. Wie man aus Abbildung A.4 erkennen kann, kommt es dabei zum Teil zu erheblichen Abweichungen in den

²Man spricht von einer klassischen Stereoaufstellung, wenn zwei Lautsprecher in einem Winkel von $\Theta = \pm 30^\circ$ aufgestellt sind.

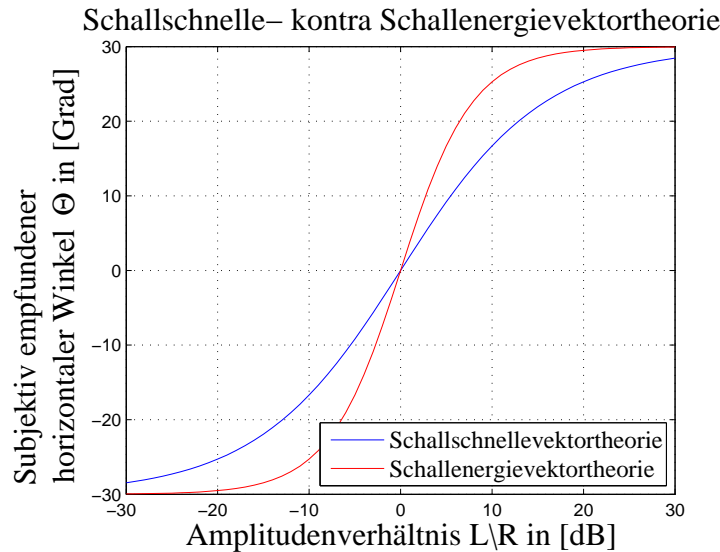


Abbildung A.4: Geschätzte Horizontallokalisation bei Verwendung der Schallschnelle- und der Schallenergievektortheorie nach [Ger92b].

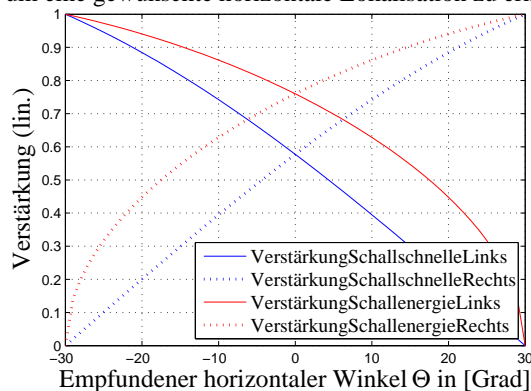
geschätzten Horizontallokalisationen der beiden Theorien. Wird gemäß der Schallschnellevektortheorie bei einem Amplitudenverhältnis von $\frac{L}{R} = 10$ [dB] ein Wert von $\tilde{\Theta}_v = 25.28^\circ$ vorhergesagt, so reduziert sich dieser Wert bei Anwendung der Schallenergievektortheorie bei gleichen Voraussetzungen auf $\tilde{\Theta}_e = 16.7^\circ$.

In der Praxis ist man üblicherweise eher daran interessiert, welche Verstärkungsfaktoren G_i für die jeweiligen Ausgangskanäle benötigt werden, um für eine gegebene Lautsprecherkonstellation eine gezielte Horizontallokalisation zu erreichen. Durch Umformung der Gleichungen A.9 - A.14 lassen sich diese für die beiden Theorien entsprechend berechnen. Für den hier betrachteten, einfachen Stereofall ergeben sich die in Abbildung A.5 dargestellten Verläufe.

Aus Abbildung A.5(a) wird ersichtlich, dass die linearen Ausgangsverstärkungswerte G_i , die nach der Schallschnellevektortheorie berechnet wurden, in einem nahezu linearem Verhältnis zur vorhergesagten Horizontallokalisation stehen. Dagegen verdeutlicht Abbildung A.5(b), dass bei energetischer Betrachtung der Ausgangsverstärkungsfaktoren, welche nach der Schallenergievektortheorie berechnet wurden, diese in einem direkt proportionalen Verhältnis zur empfundenen Horizontallokalisation stehen.

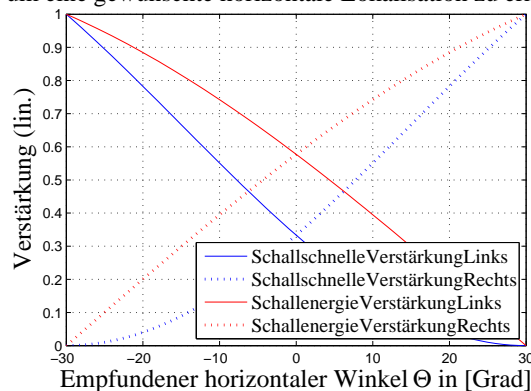
Unabhängig von den vorherrschenden Unterschieden, werden durch die beiden Theorien nach Gerzon Verstärkungsfaktoren für die jeweiligen Ausgangskanäle G_i berechnet, die üblicherweise breitbandig also frequenzunabhängig eingesetzt werden. Aus diesem Grunde überrascht es nicht, dass in praktischen Anwendungen meist die Schallenergievektortheorie verwendet wird, da diese auf der Nutzung der interauralen Pegeldifferenzen basiert. Diese leidet im Gegensatz zur ITD unter keiner Dispersion und ist deshalb für eine frequenzunabhängige Beeinflussung der Horizontallokalisation geeignet.

Verstärkungen für den linken und rechten Lautsprecher um eine gewünschte horizontale Lokalisation zu erhalten



(a) Ausgangsverstärkungswerte G_i für beide Theorien

Energiewerte für den linken und rechten Lautsprecher um eine gewünschte horizontale Lokalisation zu erhalten



(b) Ausgangseergiewerte $|G_i|^2$ für beide Theorien

Abbildung A.5: Lineare und quadratische Ausgangsverstärkungsfaktoren G_i , berechnet nach beiden Theorien

Ziel dieser beiden Theorien ist, die Gewichtungsfaktoren G_i so zu bestimmen, dass einerseits die Winkel der beiden Vektoren $\tilde{\Theta}_v$ und $\tilde{\Theta}_e$ exakt mit dem Azimut der realen bzw. virtuellen Schallquelle Θ_{SRC} übereinstimmt, wodurch die Horizontallokalisation erhalten bleibt und andererseits die Längen der beiden Vektoren r_v und r_e idealerweise zu 1 werden, wodurch eine maximale *Lokalisationsschärfe*³ gewährleistet wird.

³Mit dem Begriff *Lokalisationsschärfe* wird die räumliche Ausdehnung um die Hauptempfangsrichtung Θ einer Schallquelle bezeichnet. Ist die Lokalisationsschärfe dabei hoch wird die Lokalisation als scharf, exakt, präzise oder fokussiert bezeichnet, andernfalls mit unscharf, breit oder „breiig“.

A.2 Matrizen

In der vorliegenden Arbeit werden unterschiedlichste Matrizen verwendet, die in diesem Kapitel näher beschrieben werden.

A.2.1 Berechnung der inversen Matrix mit Hilfe der adjungierten Matrix

1. Berechnung der inversen Matrix:

Nach dem *laplaceschen Entwicklungssatz* gilt:

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \text{adj}(\mathbf{A}), \text{ wobei}$$

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} \\ a_{21} & a_{22} & \dots & a_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} \end{pmatrix}, \text{ mit}$$

$$\begin{aligned} \text{adj}(\mathbf{A}) &= \text{Adjungierte der Matrix } \mathbf{A}, \\ \det(\mathbf{A}) &= \text{Determinante der Matrix } \mathbf{A}. \end{aligned} \tag{A.15}$$

2. Berechnung der adjungierten Matrix $\text{adj}(\mathbf{A})$:

(a) Berechnung der *Kofaktoren* c_{ij} :

$$\begin{aligned} c_{ij} &= (-1)^{i+j} n_{ij}, \text{ mit} \\ c_{ij} &= \text{Kofaktor der Matrix } \mathbf{A} \text{ in der } i. \text{ Zeile und } j. \text{ Spalte,} \\ n_{ij} &= \text{Minor von } \mathbf{A} \text{ in der } i. \text{ Zeile und } j. \text{ Spalte.} \end{aligned} \tag{A.16}$$

(b) Berechnung des *Minors* n_{ij} :

Der Minor n_{ij} des Elementes $a_{ij} \in \mathbf{A}$ ergibt sich aus der Determinanten der zugehörigen Untermatrix \mathbf{A}_{ij} von \mathbf{A} .

$$\begin{aligned} n_{ij} &= \det(\mathbf{A}_{ij}), \text{ mit} \\ \mathbf{A}_{ij} &= \text{Untermatrix } \mathbf{A}_{ij} \text{ der Matrix } \mathbf{A}, \\ \det(\mathbf{A}_{ij}) &= \text{Determinante der Untermatrix } \mathbf{A}_{ij}. \end{aligned} \tag{A.17}$$

(c) Berechnung der *Untermatrix* \mathbf{A}_{ij} :

Die Untermatrix \mathbf{A}_{ij} ergibt sich, wenn man die i . Zeile und j . Spalte von \mathbf{A}

entfernt, z.B.

$$\mathbf{A}_{11} = \begin{pmatrix} a_{22} & a_{23} & \dots & a_{2j} \\ a_{32} & a_{33} & \dots & a_{3j} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i2} & a_{i3} & \dots & a_{ij} \end{pmatrix}. \quad (\text{A.18})$$

(d) Zusammenstellung der *Matrix der Kofaktoren*⁴ \mathbf{C} :

$$\mathbf{C} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1j} \\ c_{21} & c_{22} & \dots & c_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ c_{i1} & c_{i2} & \dots & c_{ij} \end{pmatrix}. \quad (\text{A.19})$$

(e) Berechnung der adjungierten Matrix $\text{adj}(\mathbf{A})$:

$$\text{adj}(\mathbf{A}) = \begin{cases} (\mathbf{C})^T & , \text{ wenn } \mathbf{A} \in \mathbb{R} \\ (\mathbf{C})^H & , \text{ wenn } \mathbf{A} \in \mathbb{C} \end{cases} \quad (\text{A.20})$$

A.2.2 Hermitische Matrix

Man spricht von einer *hermitischen Matrix* bzw. von einer *selbstadjungierten Matrix*, wenn gilt:

$$\begin{aligned} \mathbf{A} &= \mathbf{A}^H, \text{ bzw.} \\ \mathbf{A}^T &= \mathbf{A}^*, \text{ oder elementweise} \\ a_{ij} &= a_{ji}^*. \end{aligned} \quad (\text{A.21})$$

Eigenschaften einer hermitischen Matrix:

- Alle Elemente auf der Hauptdiagonalen sind reell, d.h. $a_{ij} = \Re\{a_{ij}\} \forall i = j$, wohingegen die restlichen Elemente komplex sein können.
- Sie besitzt einen symmetrischen Realteil, d.h. $\Re\{a_{ij}\} = \Re\{a_{ji}\}$ sowie einen schiefsymmetrischen Imaginärteil, für den gilt: $\Im\{a_{ij}\} = -\Im\{a_{ji}\}$.
- Ist \mathbf{A} reell, spricht man von einer *symmetrischen Matrix* für die gilt: $\mathbf{A} = \mathbf{A}^T$.
- Hermitische Matrizen sind *normal*, d.h. sie erfüllen folgende Eigenschaft: $\mathbf{A}^H \mathbf{A} = \mathbf{A} \mathbf{A}^H$.

⁴Die Matrix der Kofaktoren wird auch als *komplementäre Matrix* bezeichnet.

- Sie sind *diagonalisierbar*, d.h. die *Ähnlichkeitsbedingung*, gegeben durch:

$$\begin{aligned}\text{diag}\{\mathbf{A}\} &= \mathbf{S}^{-1} \mathbf{A} \mathbf{S}, \text{ mit} \\ \mathbf{S} &= \{E(\lambda_1), \dots, E(\lambda_i)\}, \text{ wobei} \\ E(\lambda_n) &= \text{Eigenräume zu allen Eigenwerten } \lambda_n, \quad (\text{A.22})\end{aligned}$$

wird stets erfüllt.

- Sie besitzen rein reelle Eigenwerte.
- Ihre Determinante, d.h. $\det(\mathbf{A})$ ist rein reell.

A.3 Besselfunktionen

In dieser Arbeit werden zur Bestimmung unterschiedlicher Funktionen, welche radiale Abhängigkeiten von Wellenfeldern um einen bestimmten Punkt im Raum beschreiben, wobei es sich meist um das Zentrum des Koordinatensystems handelt, unterschiedliche, sphärische Besselfunktionen verwendet. Alle dabei angewendeten Funktionen lassen sich im Prinzip auf die Besselfunktion 1. Art, m . Ordnung $J_m(z)$ zurückführen. Diese stellt, neben der Besselfunktion 2. Art, m . Ordnung, die auch als *Neumannfunktion* $Y_m(z)$ bzw. der Besselfunktion 3. Art, m . Ordnung, die auch als *Hankelfunktion* von der wiederum zwei Arten nämlich die Hankelfunktion 1. Art, m . Ordnung $H_m^{(1)}(z)$ sowie die Hankelfunktion 2. Art, m . Ordnung $H_m^{(2)}(z)$ existieren, bezeichnet werden, Lösungen der *besselschen Differentialgleichung* dar, bei der es sich um eine gewöhnliche Differentialgleichung 2. Ordnung handelt und die wie folgt angegeben werden kann:

$$z^2 \frac{\partial^2 y}{\partial z^2} + z \frac{\partial y}{\partial z} + (z^2 - m^2) y = 0. \quad (\text{A.23})$$

Eine geschlossene Lösung dieser Differentialgleichung 2. Ordnung kann z.B. mit Hilfe der Besselfunktion 1. Art, m . Ordnung wie folgt angegeben werden [BS00]:

$$\begin{aligned} J_m(z) &= \left(\frac{z}{2}\right)^m \sum_{k=0}^{\infty} \frac{(-1)^k}{\Gamma(k+1) \Gamma(m+k+1)} \left(\frac{z}{2}\right)^{2k}, \quad m \in \mathbb{R}_0^+, \text{ mit} \\ \Gamma(x) &= \int_0^{\infty} t^{(x-1)} e^{-t} dt, \quad (x > 0), \\ \mathbb{R}_0^+ &= \text{Menge der positiven, reellen Zahlen inklusive der Null,} \end{aligned} \quad (\text{A.24})$$

wobei für $m \in \mathbb{Z}$ folgende Beziehungen gelten:

$$\begin{aligned} J_{-m}(z) &= J_m(-z) = (-1)^m J_m(z), \text{ wobei} \\ \Gamma(x) &= \text{Gammafunktion, definiert für Argumente } x > 0, \\ \mathbb{Z} &= \text{Menge der ganzen Zahlen.} \end{aligned} \quad (\text{A.25})$$

Als Lösung zweiter Art für die besselsche Differentialgleichung nach Gleichung A.23 kann die Neumannfunktion $Y_m(z)$ angegeben werden, die sich wie folgt mit Hilfe der Besselfunktion 1. Art darstellen lässt:

$$Y_m(z) = \frac{J_m(z) \cos(m\pi) - J_{-m}(z)}{\sin(m\pi)}, \quad (\text{A.26})$$

Wobei auch hier wie schon für $J_m(z)$ folgende Zusammenhänge gelten:

$$Y_{-m}(z) = Y_m(-z) = (-1)^m Y_m(z). \quad (\text{A.27})$$

Als Lösungen der dritten Art können die Hankelfunktionen 1. und 2. Art angegeben werden, die sich wie folgt aus den Besselfunktionen 1. und 2. Art berechnen lassen:

$$\begin{aligned} H_m^{(1)}(z) &= J_m(z) + j Y_m(z), \text{ und} \\ H_m^{(2)}(z) &= J_m(z) - j Y_m(z), \end{aligned} \quad (\text{A.28})$$

wobei hier folgende Beziehungen gelten:

$$\begin{aligned} H_{-m}^{(1)}(z) &= H_m^{(1)}(-z) = e^{jm\pi} H_m^{(1)}(z), \text{ und} \\ H_{-m}^{(2)}(z) &= H_m^{(2)}(-z) = e^{-jm\pi} H_m^{(2)}(z). \end{aligned} \quad (\text{A.29})$$

Zur besseren Vorstellung veranschaulicht Abbildung A.6 die Verläufe der Besselfunktionen 1., 2. und 3. Art für Ordnungen von $m = [0, \dots, 5]$.

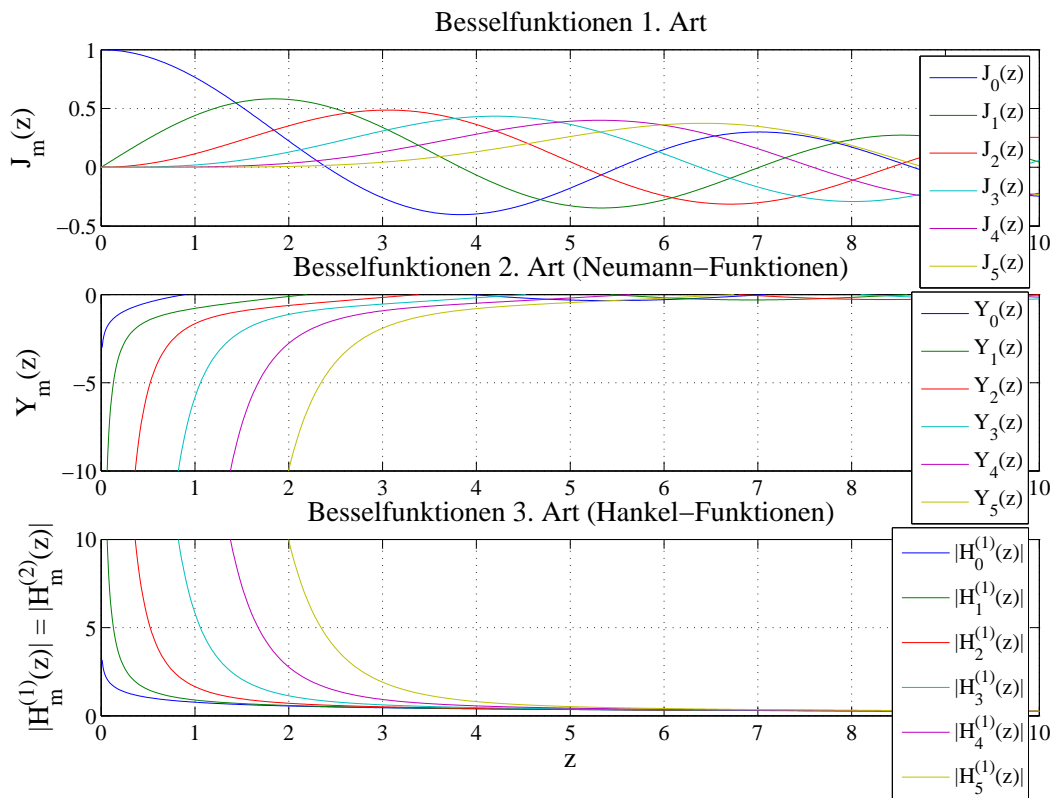


Abbildung A.6: Besselfunktionen der 1., 2. und 3. Art für die Ordnungen $m = [0, \dots, 5]$.

Wie aus Gleichung A.28 ersichtlich, unterscheiden sich die Hankelfunktionen 1. und 2. Art lediglich im Vorzeichen ihres Imaginärteils. Da ebenfalls ihre Realteile gleich sind und jeweils dem Verlauf der Besselfunktion 1. Art $J_m(z)$ entsprechen, macht eine Darstellung sowohl des Real- als auch des Imaginärteils keinen Sinn. Aus diesem Grund wird stattdessen der Betrag der Hankelfunktion graphisch dargestellt, der, wie leicht nachvollzogen

werden kann, für beide Arten der Hankelfunktion gleich ist, d.h. $\left|H_m^{(1)}(z)\right| = \left|H_m^{(2)}(z)\right|$. Deshalb findet in Abbildung A.6 keine Unterscheidung zwischen der Hankelfunktion 1. und 2. Art statt.

Abschließend werden noch einige Beziehungswiederholungen⁵ angegeben, die für alle oben vorgestellten Besselfunktionen gelten, weshalb anstelle der oben verwendeten Namen $J_m(z)$, $Y_m(z)$, $H_m^{(1)}(z)$ und $H_m^{(2)}(z)$ als Platzhalter $K_m(z)$ verwendet wird.

$$\begin{aligned}\frac{2m}{z} K_m(z) &= K_{(m-1)}(z) + K_{(m+1)}(z), \\ 2 K'_m(z) &= K_{(m-1)}(z) - K_{(m+1)}(z), \\ K'_m(z) &= K_{(m-1)}(z) - \frac{m}{z} K_m(z), \\ K'_m(z) &= -K_{(m+1)}(z) + \frac{m}{z} K_m(z).\end{aligned}\tag{A.30}$$

A.3.1 Sphärische Besselfunktionen

Die sphärischen Besselfunktionen, die in Analogie zu den zuvor vorgestellten Besselfunktionen zu ihrer Unterscheidung nun mit Kleinbuchstaben bezeichnet werden, stellen Lösungen für die folgende Differentialgleichung dar:

$$z^2 \frac{\partial y^2}{\partial z^2} + 2z \frac{\partial y}{\partial z} + [z^2 - m(m+1)] y = 0.\tag{A.31}$$

Die sphärischen Besselfunktionen lassen sich, wie schon die Besselfunktionen höherer Art, d.h. $Y_m(z)$, $H_m^{(1)}(z)$ sowie $H_m^{(2)}(z)$, mit Hilfe von Besselfunktionen 1. Art $J_m(z)$ wie folgt darstellen:

$$\begin{aligned}j_m(z) &= \sqrt{\frac{\pi}{2z}} J_{(m+\frac{1}{2})}(z), \\ y_m(z) &= \sqrt{\frac{\pi}{2z}} Y_{(m+\frac{1}{2})}(z), \\ h_m^{(1)}(z) &= j_m(z) + j y_m(z) = \sqrt{\frac{\pi}{2z}} \left[J_{(m+\frac{1}{2})}(z) + j Y_{(m+\frac{1}{2})}(z) \right], \\ h_m^{(2)}(z) &= j_m(z) - j y_m(z) = \sqrt{\frac{\pi}{2z}} \left[J_{(m+\frac{1}{2})}(z) - j Y_{(m+\frac{1}{2})}(z) \right].\end{aligned}\tag{A.32}$$

Im Gegensatz zu den Besselfunktionen aus Kapitel A.3 lassen sich sämtliche sphärischen Besselfunktionen auch mit Hilfe von trigonometrischen Funktionen als Alternative zu

⁵Anstelle des deutschen Begriffs „Beziehungswiederholungen“ wird in der Literatur auch oft dessen englisches Pendant *recurrence relations* verwendet.

Gleichung A.32 wie folgt berechnen:

$$\begin{aligned}
 j_m(z) &= (-z)^m \left(\frac{1}{z} \frac{d}{dz} \right)^m \left(\frac{\sin(z)}{z} \right), \\
 y_m(z) &= -(-z)^m \left(\frac{1}{z} \frac{d}{dz} \right)^m \left(\frac{\cos(z)}{z} \right), \\
 h_m^{(1)}(z) &= (-z)^m \left(\frac{1}{z} \frac{d}{dz} \right)^m \left(\frac{e^{jz}}{jz} \right), \\
 h_m^{(2)}(z) &= (-z)^m \left(\frac{1}{z} \frac{d}{dz} \right)^m \left(\frac{e^{-jz}}{-jz} \right).
 \end{aligned} \tag{A.33}$$

Auch für die sphärischen Besselfunktionen lassen sich einige Beziehungswiederholungen angeben, die auch hier für alle sphärischen Besselfunktionen und deren Linearkombinationen gleichermaßen gelten. Auch hier wird wieder der Einfachheit halber ein Platzhalter für die die sphärischen Besselfunktionen $j_m(z)$, $y_m(z)$, $h_m^{(1)}(z)$ und $h_m^{(2)}(z)$ verwendet, welcher mit $k_m(z)$ bezeichnet wird.

$$\begin{aligned}
 \frac{2m+1}{z} k_m(z) &= k_{(m-1)}(z) + k_{(m+1)}(z), \\
 k'_m(z) &= k_{(m-1)}(z) - \frac{m+1}{z} k_m(z),
 \end{aligned} \tag{A.34}$$

wobei durch Einsetzen der oberen in die untere Formel der Gleichung A.34 sich die Ableitung $k'_m(z)$ auch wie folgt berechnen lässt:

$$k'_m(z) = \frac{m}{z} k_m(z) - k_{(m+1)}(z). \tag{A.35}$$

Abschließend veranschaulicht Abbildung A.7 die sphärischen Besselfunktionen 1., 2. und 3. Art für die Ordnungen $m = [0, \dots, 5]$, wobei für die Darstellung der sphärischen Hankelfunktionen das Gleiche gilt wie bei der Darstellung der Hankelfunktionen in Abbildung A.6, d.h. es wird lediglich der Betrag dargestellt, da auch hier gemäß Gleichung A.32 gilt: $\left| h_m^{(1)}(z) \right| = \left| h_m^{(2)}(z) \right|$.

Für die Simulationen, wobei dafür in dieser Arbeit das Programm *MATLAB*[®] der Firma *Mathworks* verwendet wurde, waren insbesondere die Beziehungen aus Formel A.32 sowie die Beziehungswiederholungen der Formeln A.34 und A.35 von Bedeutung, da *MATLAB*[®] zwar imstande ist sämtliche Besselfunktionen, d.h. $J_m(z)$, $Y_m(z)$, $H_m^{(1)}(z)$ und $H_m^{(2)}(z)$ zu berechnen, jedoch für die Berechnung ihrer sphärischen Pendanten also für die Bestimmung von $j_m(z)$, $y_m(z)$, $h_m^{(1)}(z)$ und $h_m^{(2)}(z)$ und deren Ableitungen keine Funktionen anbietet.

Somit lassen sich die sphärischen Besselfunktionen in *MATLAB*[®] bis zur Ordnung M , d.h. in *MATLAB*[®] Konvention für $m = [1, \dots, M+1]$, für das Argument z , mit $z = kr$, wie folgt berechnen:

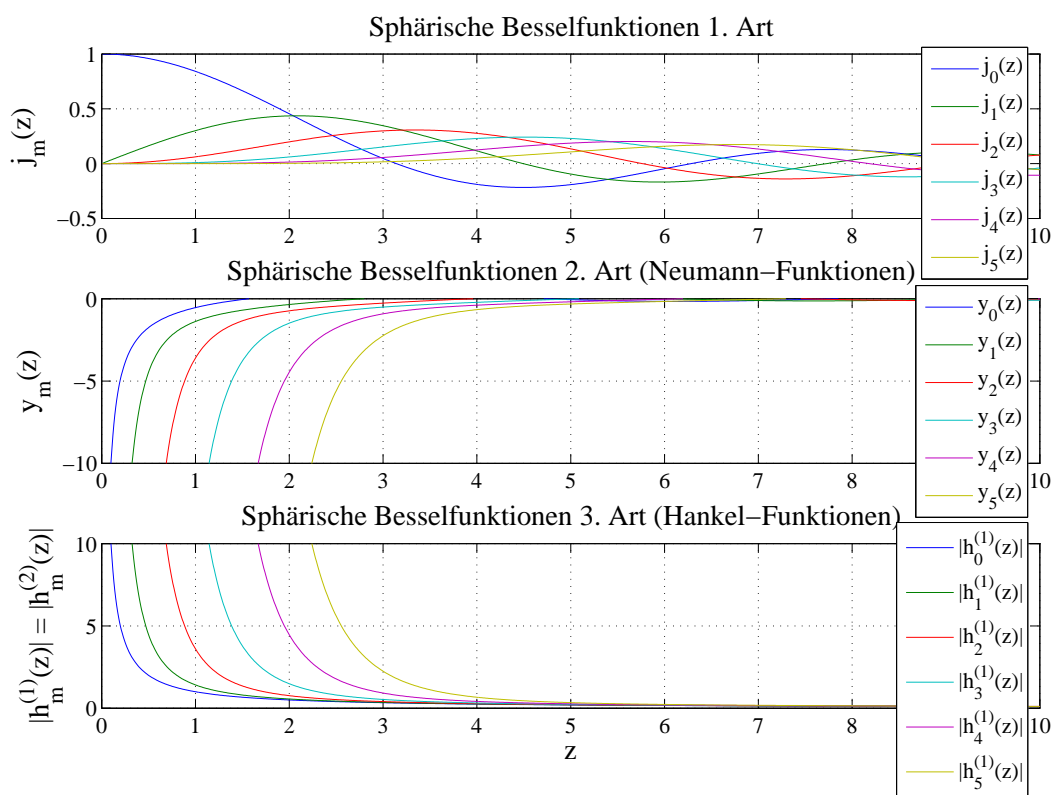


Abbildung A.7: Sphärische Besselfunktionen der 1., 2. und 3. Art für die Ordnungen $m = [0, \dots, 5]$.

1. Sphärische Besselfunktion 1. Art $j_m(z)$:

$$j(m,z) = \text{sqrt}(\text{pi.}/(2*z)) .* \text{besselj}((m-1)+1/2,z);$$

2. Sphärische Besselfunktion 2. Art $y_m(z)$ (=Neumannfunktion):

$$y(m,z) = \text{sqrt}(\text{pi.}/(2*z)) .* \text{bessely}((m-1)+1/2,z);$$

3. Sphärische Hankelfunktion 1. Art $h_m^{(1)}(z)$:

$$h1(m,z) = \text{sqrt}(\text{pi.}/(2*z)) .* (\text{besselj}((m-1)+1/2,z) + \text{sqrt}(-1)*\text{bessely}((m-1)+1/2,z));$$

4. Sphärische Hankelfunktion 2. Art $h_m^{(2)}(z)$:

$$h2(m,z) = \text{sqrt}(\text{pi.}/(2*z)).* (\text{besselj}((m-1)+1/2,z) - \text{sqrt}(-1)*\text{bessely}((m-1)+1/2,z));$$

sowie deren Ableitungen

1. Ableitung der sphärischen Besselfunktion 1. Art $j'_m(z)$:

$$dj(m,z) = \text{sqrt}(\text{pi.}/(2*z)) .* (((m-1)./z).*\text{besselj}((m-1)+1/2,1,z) - \text{besselj}((m-1)+1/2+1,1,z));$$

2. Ableitung der sphärische Besselfunktion 2. Art $y'_m(z)$ (=Ableitung der Neumannfunktion):

$$dy(m,z) = \text{sqrt}(\text{pi.}/(2*z)) .* (((m-1)./z).*\text{bessely}((m-1)+1/2,1,z) - \text{bessely}((m-1)+1/2+1,1,z));$$

3. Ableitung der sphärische Hankelfunktion 1. Art $h_m^{(1)}(z)$:

$$dh1(m,z) = \text{sqrt}(\text{pi.}/(2*z)) .* (((m-1)./z).*\text{besselh}((m-1)+1/2,1,z) + \text{besselh}((m-1)+1/2+1,1,z));$$

4. Ableitung der sphärische Hankelfunktion 2. Art $h_m^{(2)}(z)$:

$$dh2(m,z) = \text{sqrt}(\text{pi.}/(2*z)) .* (((m-1)./z).*\text{besselh}((m-1)+1/2,2,z) - \text{besselh}((m-1)+1/2+1,2,z));$$

A.4 Entwurf eines modalen *beamformer*

In diesem Kapitel soll folgende Frage beantwortet werden: „Was versteht man unter einem modalen *beamformer* und wie kann ein solcher entworfen werden?“.

Der erste Teil der Frage lässt sich dabei auf recht anschauliche Weise beantworten. Ein modaler *beamformer* realisiert eine Wunschrictcharakteristik $\Psi(\Theta, \varphi)$ in dem diverse sphärische Harmonische $Y_{m,n}^\sigma(\Theta, \varphi)$, d.h. gewisse Raummoden, womit sich auch der Name „modaler *beamformer*“ erklärt, mit zugehörigen von der Raumrichtung unabhängigen Gewichtungsfaktoren $C_{m,n}^\sigma$ skaliert und anschließend aufsummiert werden. Sofern ein kugelförmiges Mikrofonarray, auf dessen Oberfläche z.B. in regelmäßiger Anordnung Mikrophone angebracht sind, verwendet wird, lassen sich die hierfür benötigten Raummoden auf sehr einfache Weise generieren. Unter diesen Voraussetzungen lassen sich die für den Entwurf benötigten Raummoden aus den Mikrophonsignalen des Kugelmikrophonarrays gewinnen, in dem für jede Raummode $Y_{m,n}^\sigma(\Theta, \varphi)$ sämtliche Mikrophone speziell gewichtet und zusammengefasst werden. Die nötigen Gewichte gewinnt man dabei durch Abtastung der in ihrer Form bekannten zugehörigen, sphärischen Harmonischen an den Orten, an denen sich die Mikrophone des Kugelmikrophonarrays befinden.

Mathematisch lässt sich dieser Prozess wie folgt zusammenfassen:

$$\Psi(\Theta, \varphi) = \sum_{m=0}^M \sum_{0 \leq n \leq m, \sigma = \pm 1} C_{m,n}^\sigma Y_{m,n}^\sigma(\Theta, \varphi), \text{ wobei}$$

$$C_{m,n}^\sigma = \text{Frequenzunabhängige Gewichtungsfaktoren zur Erzeugung eines modalen } \textit{beamformer}, \quad (\text{A.36})$$

bzw. in Matrixschreibweise:

$$\Psi = \mathbf{C} \mathbf{Y}. \quad (\text{A.37})$$

Zur Beantwortung des zweiten Teils der anfänglichen Fragestellung würde es im Prinzip genügen zu erwähnen, dass sich die modalen Gewichtungsfaktoren \mathbf{C} gemäß Gleichung A.37 wie folgt bestimmen lassen:

$$\mathbf{C} = \mathbf{Y}^+ \Psi. \quad (\text{A.38})$$

Nichtsdestotrotz wird nachfolgend schrittweise erklärt, wie sich der Entwurf eines modalen *beamformer* interpretieren lässt.

1. Definition der Wunschrictcharakteristik bzw. der Zielfunktion $\Psi(\Theta, \varphi)$.

Eine einfache und flexible Form stellt in diesem Zusammenhang die Verwendung der kardioiden Richtcharakteristik dar, da diese im Gegensatz zu anderen Richtcharakteristika, wie aus Tabelle A.1 ersichtlich, einen Richtcharakteristikparameter α_m ,

mit $m = [1, \dots, M]$ hat, welcher unabhängig von der gewünschten Ordnung m ist und stets einen Wert von $\alpha_m = 0.5$ aufweist.

Somit lässt sich die Zielfunktion $\Psi(\Theta, \varphi)$ bei Verwendung der kardioiden Richtcharakteristik auf einfache Weise wie folgt berechnen ([CCFS10]):

$$\begin{aligned}\Psi(\Theta, \varphi) &= [\Psi_{\text{Cardioid}}(\Theta, \varphi)]^m, \text{ mit} \\ \Psi_{\text{Cardioid}}(\Theta, \varphi) &= 0.5 + 0.5 \cos(\Theta) \cos(\varphi), \text{ wobei} \\ \Psi_{\text{Cardioid}}(\Theta, \varphi) &= \text{Richtcharakteristik eines Kardioidmikrophons} \\ &\quad \text{1. Ordnung.}\end{aligned}\tag{A.39}$$

2. Auswahl der sphärischen Harmonischen $Y_{m,n}^\sigma(\Theta, \varphi)$, welche für den Entwurf des modalen *beamformer* verwendet werden sollen. Je nach gewählter Ordnung M stehen hierbei insgesamt maximal $N = (M + 1)^2$ sphärische Harmonische zur Verfügung.

Hinweis:

Der Einfachheit halber sowie aus Effizienzgründen beschränken sich an dieser Stelle einige Autoren (siehe hierzu beispielsweise: [EKM03, ME04, Teu05]) auf *beamformer* die in Z -Richtung zeigen. Hintergrund hierfür ist, dass wenn $\varphi = \pm 90^\circ$, wie aus Tabelle 4.1 entnommen werden kann, lediglich die reellen ($\sigma = 1$) M Moden des Grades $n = 0$, d.h. die in der mittleren Spalte in Abbildung 4.4 dargestellten, sphärischen Harmonischen $Y_{m,0}^1(\Theta, \varphi)$ für den Entwurf verwendet werden. Dies entspricht der geringst möglichen Anzahl und spiegelt somit die effizienteste Realisierung eines modalen *beamformer* wider.

3. Geht man von idealen Verhältnissen aus, sind sowohl die Wunschrichtcharakteristik $\Psi(\Theta, \varphi)$ als auch die Richtcharakteristika der sphärischen Harmonischen $Y_{m,n}^\sigma(\Theta, \varphi)$ frequenzunabhängig. Somit genügt es, die Gewichtungsfaktoren $C_{m,n}^\sigma$ zu finden, deren Linearkombination mit den korrespondierenden, sphärischen Harmonischen möglichst exakt die Zielfunktion approximiert, was bereits durch Gleichung A.36 verdeutlicht wurde.

Beschränkt man sich hierbei auf eine Blickrichtung in Z -Richtung, vereinfacht sich Gleichung A.36 zu:

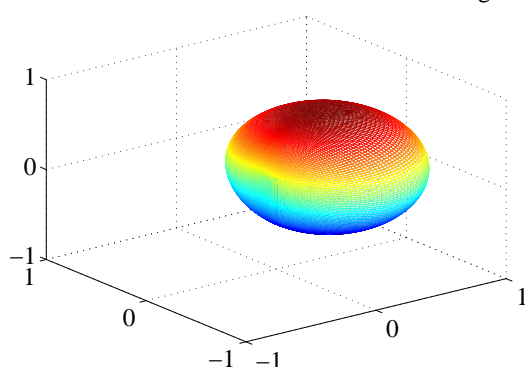
$$\Psi(\Theta, \varphi) = \sum_{m=0}^M C_m Y_m(\Theta, \varphi).\tag{A.40}$$

4. Unter Berücksichtigung der Rotationssymmetrie kann das Optimierungsproblem zum Auffinden der optimalen Gewichtungsfaktoren $C_{m,n}^\sigma$ vom drei- in den zwei-dimensionalen Bereich verlagert werden. Aufgrund der Achsensymmetrie der Funktionen ist es ferner ausreichend, wenn lediglich ein Winkelbereich von 180° bei der Optimierung berücksichtigt wird.

Im Falle der Ausrichtung in Z -Richtung bedeutet das, dass für die Bestimmung der Gewichtungsfaktoren C_m die Auswertung des Polardiagramms bei $\Theta = 0^\circ$ in einem Winkelbereich von $\varphi = -90^\circ, \dots, 90^\circ$ ausreicht.

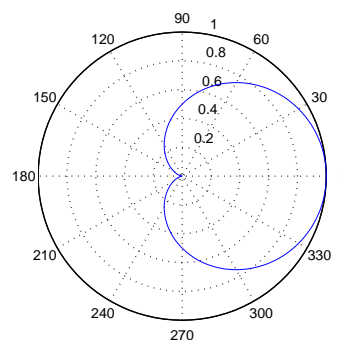
Zum einfacheren Verständnis wird das oben beschriebene schrittweise Verfahren zunächst an einem einfachen Beispiel verdeutlicht. Dabei wird ein modaler *beamformer* der Ordnung $m = 1$ entworfen, welcher in Z -Richtung zeigt und dabei die Richtcharakteristik eines Kardiodmikrophons 1. Ordnung aufweist.

3D-Darstellung eines beamformer mit einer kardioden Richtcharakteristik 1. Ordnung



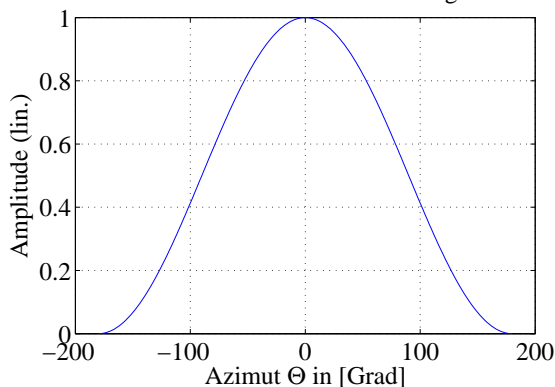
(a) 3D-Darstellung

Polardiagramm eines beamformer mit einer kardioden Richtcharakteristik 1. Ordnung



(b) Polardiagramm

Darstellung des Richtdiagramms eines beamformer mit einer kardioden Charakteristik 1. Ordnung als Funktion



(c) Polardiagramm im kartesischen Koordinatensystem

Abbildung A.8: Graphische Darstellungen der Wunschrictcharakteristik (kardioiden Richtcharakteristik 1. Ordnung).

Wie Abbildung A.8(a) verdeutlicht, handelt es sich im Normalfall bei $\Psi(\Theta, \varphi)$ um eine rotationssymmetrische Funktion, dessen Projektion auf die X/Y -Ebene in das in Abbildung A.8(b) aufgezeigte Polardiagramm resultiert. Dies veranschaulicht die in Punkt 4 angesprochene Verlagerung vom drei- in den zweidimensionalen Bereich.

Betrachtet man sich die in Abbildung A.8(c) dargestellte, zweidimensionale Projektion des Richtdiagramms im kartesischen Koordinatensystem, fällt auf, dass dieses achsensymmetrisch ist und somit eindeutig in einem reduzierten Winkelbereich von hier $\Theta = [-180^\circ \dots 0^\circ]$ beschrieben werden kann. Das entspricht der in Punkt 4 erwähnten Zielfunktion.

Ferner ist zu erkennen, dass es bei der Auswahl der Wunschrictcharakteristik $\Psi(\Theta, \varphi)$ prinzipiell egal ist, welche Orientierung, d.h. welche Ausrichtung, diese hat, sofern man sich auf den redundanzfreien Teil der Zielfunktion beschränkt. In obigem Beispiel soll, wie eingangs erwähnt, ein in die Z -Richtung deutender modaler *beamformer* entworfen werden, wobei als Referenz eine Wunschrictcharakteristik $\Psi(\Theta, \varphi)$ verwendet wird, welche in X -Richtung zeigt. Dies gelingt, indem einerseits der Übergang vom drei- in den zweidimensionalen Bereich durch einen Schnitt in der X/Y -Ebene erfolgt und man sich andererseits auf einen Winkelbereich des Azimut Θ von $\Theta = [-180^\circ \dots 0^\circ]$ beschränkt, um die redundanzfreie Zielfunktion zu generieren.

In Abbildung A.9 werden die zur Approximation der Zielfunktion ausgewählten Moden auf die gleiche Weise wie in Abbildung A.8 aufgezeigt. Diese deuten nun, wie gefordert, in Z -Richtung, sodass zur Generierung der redundanzfreien, modalen Funktionsverläufe zum einen ein Schnitt in der X/Z -Ebene für den Übergang vom drei- in den zweidimensionalen Bereich nötig ist und man sich zum anderen, wie eingangs erwähnt, nun auf einen Winkelbereich der Elevation von $\varphi = [-90^\circ \dots 90^\circ]$ beschränken darf.

Nachdem man die in Abbildung A.9(c) dargestellten Funktionsverläufe auf den minimal zulässigen, redundanzfreien Winkelbereich reduziert hat, sind alle Startbedingungen für die Optimierung erfüllt, was durch die in Abbildung A.10 aufgezeigten Funktionsverläufe visualisiert wird.

Ziel der Optimierung ist es, Gewichtungsfaktoren C_m zu finden, sodass die Überlagerung der gewichteten, modalen Funktionsverläufe möglichst exakt mit der Zielfunktion übereinstimmt.

Berechnet man die optimalen Gewichtungsfaktoren C_m nach Gleichung A.38, ergeben sich die in Abbildung A.11 dargestellten Verläufe. Wie man sieht ist es in dem hier gewählten Beispiel möglich, durch eine Summierung der skalierten, sphärischen Harmonischen 0. Ordnung $Y_{0,0}^1(\Theta, \varphi)$ mit einem Gewichtungsfaktor von $C_0 = 0.5$ und der skalierten, sphärischen Harmonischen 1. Ordnung $Y_{1,0}^1(\Theta, \varphi)$ mit einem Gewichtungsfaktor von $C_1 = 0.289$, die Wunschrictcharakteristik $\Psi(\Theta, \varphi)$, d.h. eine kardioide Richtcharakteristik 1. Ordnung, exakt nachzubilden.

Führt man obiges Beispiel auch für weitere Ordnungen M durch, erhält man die in Tabelle A.2 für $m = [0, \dots, 4]$ aufgelisteten Gewichtungsfaktoren C_m .

Berechnet man ferner für die Ordnungen $m = [0, \dots, 4]$ die in Abbildung 5.21 dargestellten, spektralen Verläufe des *white noise gain* $WNG_{\text{dB}}(\omega)$ und setzt dabei eine Schwelle von $WNG_{\text{dBMin}} = -10$ [dB] an, ergeben sich bei Verwendung eines soliden Kugelmikrophonarrays mit Radius $a = 0.09$ [m], auf dessen Oberfläche $Q = 32$ regulär an-

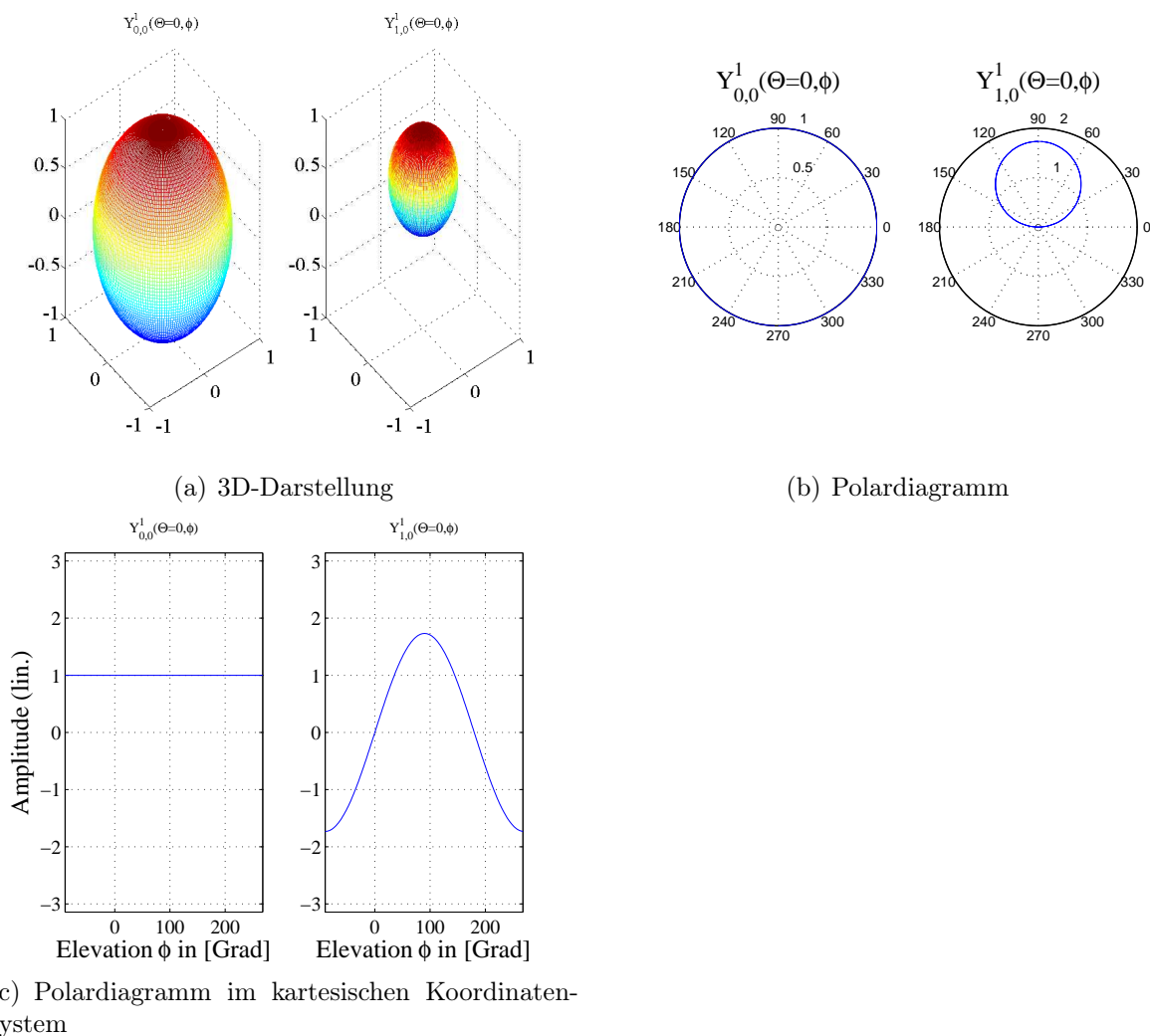


Abbildung A.9: Graphische Darstellungen der Moden mit deren Hilfe die Wunschrictcharakteristik (kardioide Richtcharakteristik 1. Ordnung) approximiert werden soll.

Ordnung	C_0	C_1	C_2	C_3	C_4	Grenzfrequenz f_c in [Hz] bei der gilt: $WNG = -10$ [dB]
0	1	0	0	0	0	0
1	0.5	0.2887	0	0	0	30
2	0.333	0.2887	0.0745	0	0	214
3	0.25	0.2598	0.1118	0.0189	0	504
4	0.2	0.2309	0.1278	0.0378	0.0048	852

Tabelle A.2: Gewichtungskoeffizienten C_m für einen modalen *beamformer* m . Ordnung, welcher in Z -Richtung ausgerichtet ist und eine nach Gleichung A.39 berechnete, kardioide Wunschrictcharakteristik m . Ordnung perfekt nachbildet.

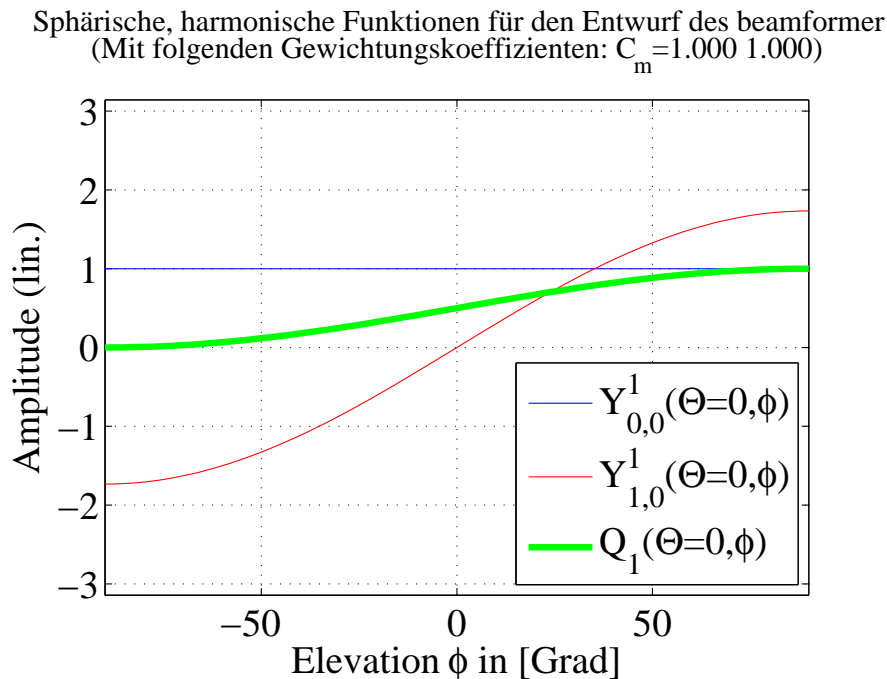


Abbildung A.10: Darstellungen der Funktionsverläufe der Wunschrictcharakteristik sowie der Richtcharakteristika der Moden, welches durch eine entsprechende Gewichtung durch die Faktoren C_m die Zielfunktion approximieren sollen.

gebrachte, omnidirektionale Mikrophone installiert sind, die in Tabelle A.2 angegebenen Übergangsfrequenzen oberhalb derer gilt: $WNG_{\text{dB}}(\omega) \geq WNG_{\text{dBMin}}$.

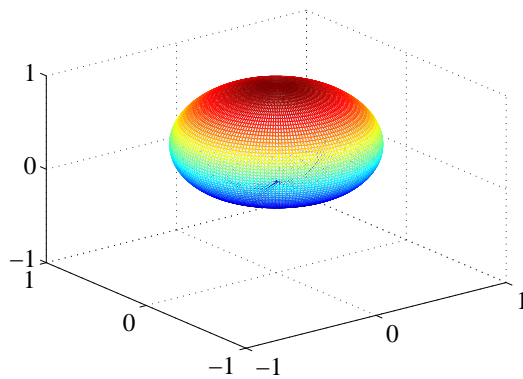
Soll nun die vorgegebene WNG-Schwelle von $WNG_{\text{dBMin}} = -10$ [dB] für einen modalen beamformer 4. Ordnung über der Frequenz eingehalten werden, müssen sich auch die Gewichtungsfaktoren, wie in Tabelle A.2 angegeben, entsprechend mit der Frequenz ändern. Abbildung A.12 stellt dabei die prinzipiellen Verläufe der nun frequenzabhängigen Gewichtungsfaktoren $C_m(\omega)$ graphisch dar.

A.4.1 Praktische Anwendungen modaler beamformer

Wie eingangs in diesem Kapitel gezeigt, handelt es sich bei einem modalen beamformer um eine sehr flexible und einfache Möglichkeit, unterschiedliche Richtcharakteristika zu realisieren. Da es sich dabei im Prinzip lediglich um eine Approximation einer beliebigen Zielfunktion mit Hilfe von vorgegebenen, festen Grundfunktionen gegeben durch die ausgewählten, sphärischen Harmonischen handelt, was durchaus mit einer Spline- oder Polynominterpolation verglichen werden kann, ist man beim Entwurf nicht auf bekannte Richtcharakteristika beschränkt, sondern kann mit dieser Methode beamformer mit beliebiger Richtcharakteristik entwerfen.

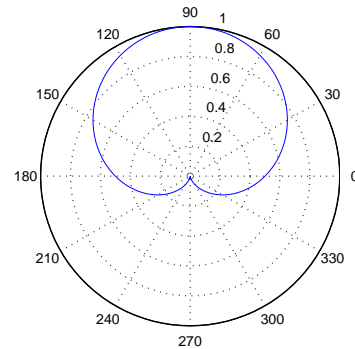
Als praktische Anwendung resultiert daraus, dass ein modaler beamformer als vielseitiger Ersatz für Mikrophone mit unterschiedlichsten Richtcharakteristika, wie sie etwa

3D-Darstellung eines modalen beamformer welcher eine kardioidie Richtcharakteristik 1. Ordnung approximiert



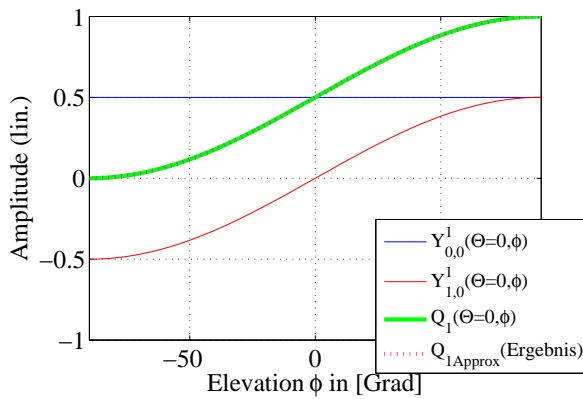
(a) 3D-Darstellung

Polardiagramm eines modalen beamformer welcher eine kardioidie Richtcharakteristik 1. Ordnung approximiert



(b) Polardiagramm

Sphärische, harmonische Funktionen für den Entwurf des gewünschten beamformer (Mit folgenden Gewichtungskoeffizienten: $C_m = [0.500 \ 0.289]$)



(c) Polardiagramm im kartesischen Koordinatensystem

Abbildung A.11: Graphische Darstellungen der mit den Gewichtungsfaktoren $C_m = [0.25, 0.433]$ skalierten Moden, welche die Wunschrictcharakteristik (kardioidie Richtcharakteristik 1. Ordnung) approximieren.

Optimale, modale Gewichtungskoeffizienten C_m , welche perfekt eine kardioidie Richtcharakteristik mit unterschiedlichen Ordnungen m über der Frequenz so approximiert, dass ein $\text{WNG}_{\text{dB}} = -10[\text{dB}]$ erreicht wird

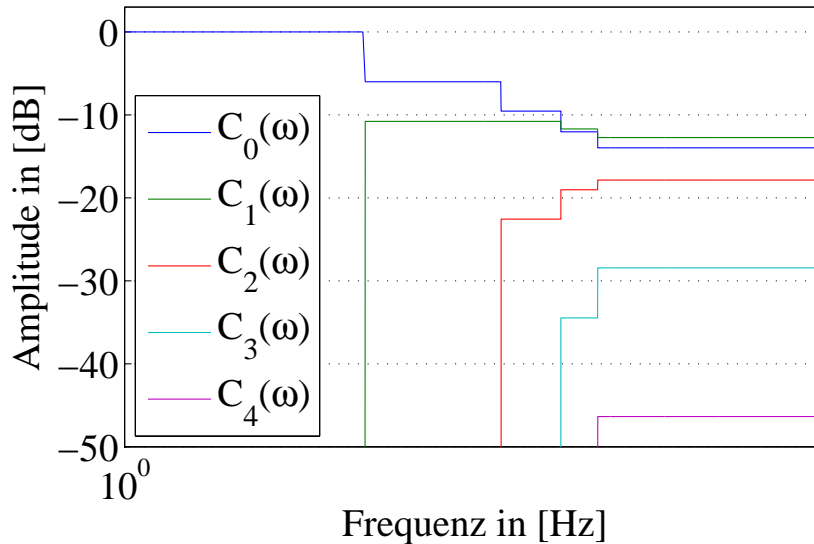


Abbildung A.12: Darstellungen der modalen Gewichtungskoeffizienten C_m über der Frequenz, welche perfekt kardioidie Richtcharakteristika der Ordnungen $m = [0, \dots, 4]$ nachbilden.

in der Aufnahmetechnik eingesetzt werden, agieren kann. Bei 5.1⁶-Aufnahmen werden in Anlehnung an die bei 5.1-Systemen genormte Lautsprecheraufstellung drei Kardioidmikrophone, wobei zwei nach hinten und eines zentral nach vorne zeigt und zwei Hyperkardioidmikrophone, bei denen eines nach links und das andere nach rechts zeigt, verwendet. Für das von der Firma *Auro3D* stammende 9.1-Format, bei dem das 5.1-System noch um vier zusätzliche Lautsprecher, welche in der Höhe, genauer gesagt oberhalb der beiden vorderen linken (FLH⁷) und rechten (FRH⁸) sowie oberhalb der beiden hinteren linken (RLH⁹) und rechten Lautsprecher (RRH¹⁰) angebracht sind, um den akustischen Effekt von oben, d.h. aus der dritten Dimension abzubilden, erweitert wurde, werden für die entsprechende Aufnahmetechnik ebenfalls vier zusätzliche nach oben zeigende, Richtmikrophone mit ausgeprägtem Richtverhalten benötigt, die ebenfalls problemlos mit einem modalen *beamformer* per Software abgedeckt werden können.

⁶Der Begriff 5.1 bezeichnet in der Akustik ein System, welches sich aus vier breitbandigen Lautsprechern, die wie folgt im Raum angeordnet sind: vorne links (FL = engl. **F**ront **L**eft), zentral vorne (C = engl. **C**enter), vorne rechts (FR = engl. **F**ront **R**ight), hinten links (RL = engl. **R**ear **L**eft), hinten rechts (RR = engl. **R**ear **R**ight) und einem schmalbandigen Tiefton- bzw. Bass-Lautsprecher, der als Subwoofer oder kurz unter der Bezeichnung *Sub* bekannt ist, zusammensetzt.

⁷FLH = engl. **F**ront **L**eft **H**igh

⁸FRH = engl. **F**ront **R**ight **H**igh

⁹RLH = engl. **R**ear **L**eft **H**igh

¹⁰RRH = engl. **R**ear **R**ight **H**igh

Unter diesem Gesichtspunkt betrachtet, kann nachvollzogen werden, dass es mit Hilfe modaler *beamformer* auch möglich ist, zumindest approximativ, Mikrophone mit extremer Richtcharakteristik der Ordnung M_Q , die auch als *shotgun*-Mikrophone bezeichnet werden, mit sphärischen Harmonischen der Ordnung M , wobei gilt: $M_Q > M$, nachzubilden. In [CCFS10] wurde basierend auf einem kugelförmigen Mikrophonarray eine Möglichkeit aufgezeigt, wie beliebige Richtcharakteristika realisiert werden können. Allerdings basierte der dort vorgestellte Ansatz auf einem klassischen *filter-and-sum beamformer*. Dadurch konnten nach einer einmaligen jedoch aufwändigen Vermessung sehr gute Ergebnisse erzielt werden. Jedoch besitzt diese Methode nicht die, einem modalen *beamformer* innewohnende Fähigkeit, dass die Blickrichtung der durch ihm realisierten Richtkeulen ohne viel Aufwand durch eine einfache Änderung der Gewichtung dynamisch, offline oder online über den gesamten 3D-Bereich gesteuert werden kann. Diese Eigenschaft konnte lediglich durch Interpolation vorab berechneter *beamforming*-Filter erreicht werden, was wiederum den Speicheraufwand extrem erhöhte.

Im abschließenden Beispiel wird auf den Entwurf eines *shotgun*-Mikrophons, welches die Richtcharakteristik eines Kardioidmikrophons 9. Ordnung besitzt und mit Hilfe eines in Z -Richtung ausgerichteten, modalen *beamformer* 4. Ordnung approximiert werden soll, aufgezeigt.

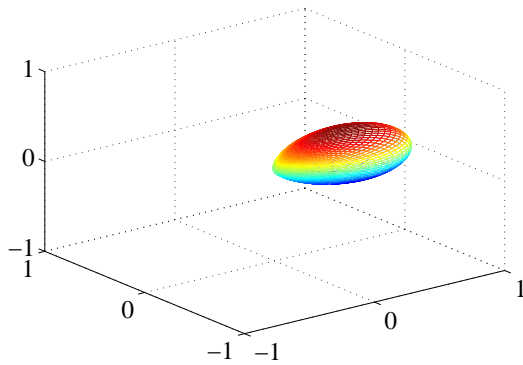
Bei exakter Betrachtung von Abbildung A.16(c) erkennt man, dass die Approximation der Zielfunktion (kardioidale Richtcharakteristik 9. Ordnung) durch die hier verwendeten 5 Grundfunktionen, gegeben durch die reellen ($\sigma = 1$), sphärischen Harmonischen 0. Grades, $Y_{m,0}^1(\Theta, \varphi)$ der Ordnungen $m = [0, \dots, 4]$, zwar nicht perfekt, aber dennoch als sehr gut bezeichnet werden kann.

A.4.2 Rotation des modalen *beamformer*

Soll ein modaler *beamformer* in eine andere Richtung als etwa in die zuvor erwähnte am effizientest zu realisierenden Z -Richtung zeigen, kann dies auf mehrere Arten geschehen. Entweder man entwirft den modalen *beamformer*, wie in den allgemeingültigen Gleichungen A.37, A.37 und A.38 beschrieben, was jedoch einen erhöhten Berechnungsaufwand in sich birgt oder man rotiert einen bereits in eine bestimmte Richtung (z.B. Z -Richtung) zeigenden modalen *beamformer* in einem nachfolgenden separaten Signalverarbeitungsschritt. Letztere Möglichkeit bietet dabei mehrere Vorteile. Einerseits kann der zugrundeliegende, modale *beamformer* aufwandsgünstig entworfen werden, in dem man ihn z.B. in Z -Richtung zeigend entwirft, andererseits kann durch eine separate Rotation eine Neuberechnung der Gewichtungsfaktoren C_m bzw. $C_m(\omega)$ entfallen. Lediglich eine Neuberechnung der für die Rotation notwendigen Gewichte, was z.B. durch die in der Tabelle 4.1 angegebenen Gleichungen auf sehr einfache Weise geschehen kann, wird benötigt, was ohne Probleme in Echtzeit erfolgen kann. Somit entsteht ein sehr flexibler *beamformer*, dessen Richtkeule in Echtzeit in jede, dreidimensionale Richtung geschwenkt werden kann, was vorteilhaft, z.B. für Liveaufnahmen von Theateraufführungen oder all-

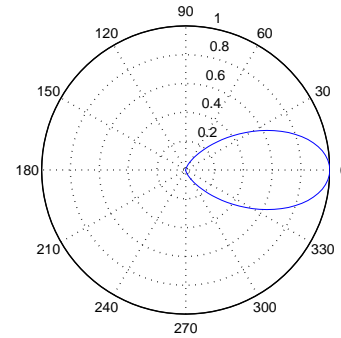
Darstellung der Wunschrictcharakteristik:

3D-Darstellung eines beamformer mit einer kardioiden Richtcharakteristik 9. Ordnung



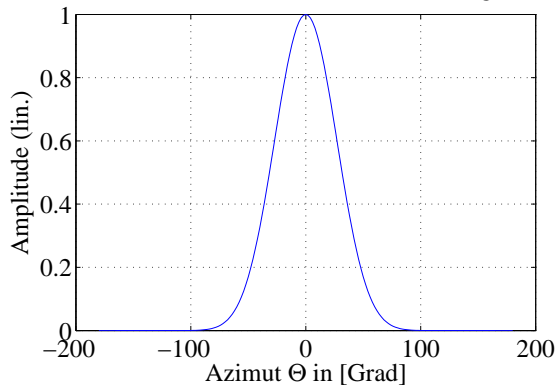
(a) 3D-Darstellung

Polardiagramm eines beamformer mit einer kardioiden Richtcharakteristik 9. Ordnung



(b) Polardiagramm

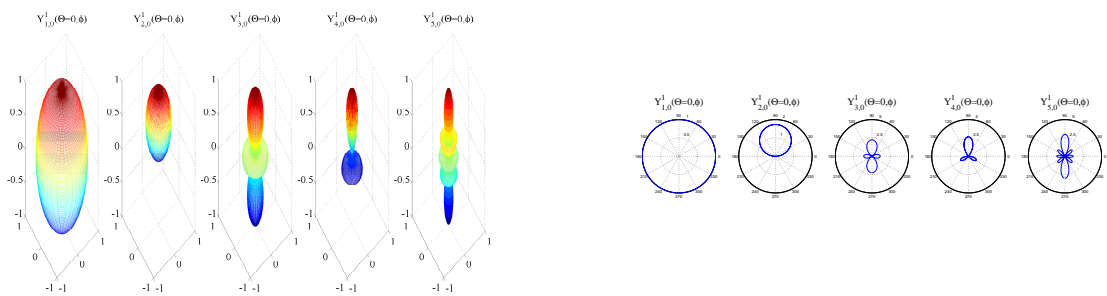
Richtdiagramm eines beamformer mit kardioider Charakteristik 9. Ordnung



(c) Polardiagramm im kartesischen Koordinatensystem

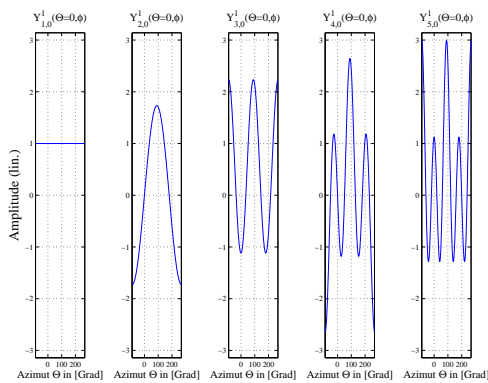
Abbildung A.13: Graphische Darstellungen der Wunschrictcharakteristik (kardioider Richtcharakteristik 9. Ordnung).

Darstellung der für die Approximation der Wunschrictcharakteristik, zur Verfügung stehenden Moden:



(a) 3D-Darstellung

(b) Polardiagramm



(c) Polardiagramm im kartesischen Koordinatensystem

Abbildung A.14: Graphische Darstellungen der Moden mit deren Hilfe die Wunschrictcharakteristik (kardioide Richtcharakteristik 9. Ordnung) approximiert werden soll.

Darstellung der Startsituation vor Berechnung der modalen Gewichtungsfaktoren C_m :

Sphärische, harmonische Funktionen für den Entwurf eines beamformer
(Mit folgenden Gewichtungskoeffizienten: $C_m = 1.0 \ 1.0 \ 1.0 \ 1.0 \ 1.0$)

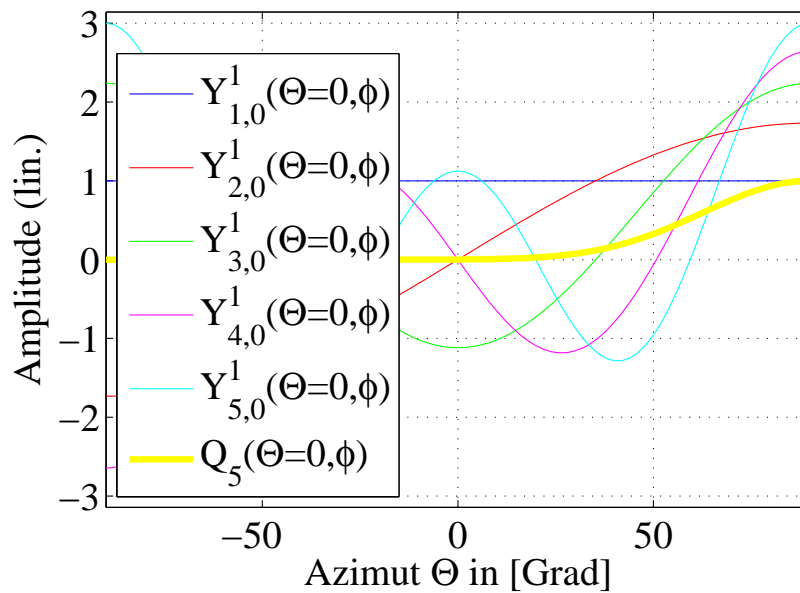
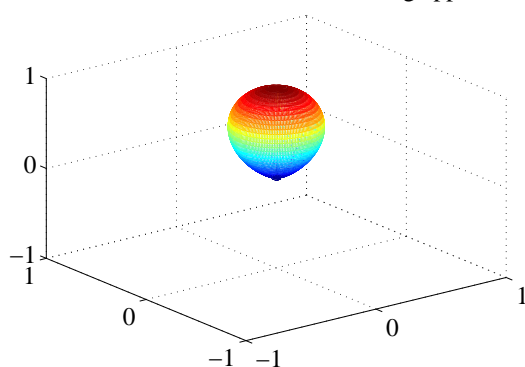


Abbildung A.15: Darstellungen der Funktionsverläufe der Wunschrictcharakteristik sowie der Richtcharakteristika der Moden, welches durch eine entsprechende Gewichtung durch die Faktoren C_m , die Zielfunktion approximieren sollen.

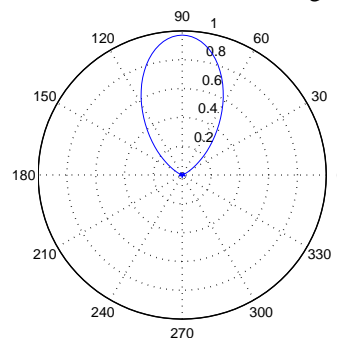
Darstellung des Resultats, d.h. des modalen *beamformer* 4. Ordnung, welche die Wunschrichtcharakteristik möglichst exakt approximiert:

3D-Darstellung eines modalen *beamformer*, welcher eine kardioidale Richtcharakteristik 9. Ordnung approximiert



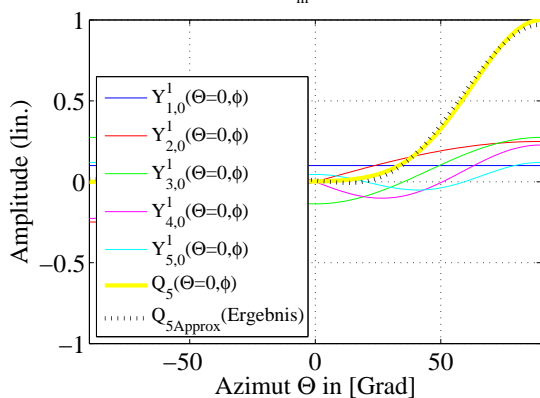
(a) 3D-Darstellung

Polardiagramm eines modalen *beamformer* welcher eine kardioidale Richtcharakteristik 9. Ordnung approximiert



(b) Polardiagramm

Sphärische, harmonische Funktionen für den Entwurf des *beamformer*
(Mit folgenden Gewichten: $C_m = [0.100 \ 0.144 \ 0.123 \ 0.086 \ 0.040]$)



(c) Polardiagramm im kartesischen Koordinatensystem

Abbildung A.16: Graphische Darstellungen der mit den Gewichtungsfaktoren $C_m = [0.100, 0.144, 0.123, 0.086, 0.040]$ skalierten Moden, welche die Wunschrichtcharakteristik (kardioidale Richtcharakteristik 9. Ordnung) approximieren.

gemein bei sich dynamisch ändernden, akustischen Verhältnisse (z.B. sich auf der Bühne bewegendende Schauspieler), ist.

Mathematisch lässt sich allgemein die Berechnung einer gewünschten Richtcharakteristik $\Psi(\Theta_{Des}, \varphi_{Des})$, inklusive Rotation wie folgt zusammenfassen:

$$\Psi(\Theta_{Des}, \varphi_{Des}) = \sum_{m=0}^M \sum_{0 \leq n \leq m, \sigma = \pm 1} C_{m,n}^{\sigma} Y_{m,n}^{\sigma}(\Theta, \varphi) Y_{m,n}^{\sigma}(\Theta_{Des}, \varphi_{Des}), \text{ wobei} \quad (\text{A.41})$$

$$Y_{m,n}^{\sigma}(\Theta_{Des}, \varphi_{Des}) = \text{Modale Gewichtungskoeffizienten, welche multipliziert mit den sphärischen Harmonischen } Y_{m,n}^{\sigma}(\Theta_{Des}, \varphi_{Des}) \text{ in Summe } (M+1) \text{ Moden ergibt, die in die gewünschte Richtung gedreht wurden.}$$

Für den Sonderfall, dass der *beamformer* in Z -Richtung entworfen wurde, d.h. das lediglich $(M+1)$ modale Gewichtungskoeffizienten C_m existieren, vereinfacht sich Gleichung A.42 zu:

$$\Psi(\Theta_{Des}, \varphi_{Des}) = \sum_{m=0}^M C_m \sum_{0 \leq n \leq m, \sigma = \pm 1} Y_{m,n}^{\sigma}(\Theta, \varphi) Y_{m,n}^{\sigma}(\Theta_{Des}, \varphi_{Des}). \quad (\text{A.42})$$

Anhand der Gleichung A.42 lässt sich die Rotation wie folgt interpretieren: Während beim Entwurf eines modalen *beamformer* in Z -Richtung, lediglich $(M+1)$ sphärische Harmonische beteiligt sind, muss dieser Zustand mit Hilfe der Rotation erst geschaffen werden. Dabei gilt es lediglich $(M+1)$ sphärische Harmonische, also je eine Mode pro Ordnung m , mit $m = [0, \dots, M]$, was bekanntermaßen der minimalen Anzahl entspricht, zu bearbeiten, was den Aufwand gering hält. Durch die Rotation, d.h. durch die Gewichtungskoeffizienten $Y_{m,n}^{\sigma}(\Theta_{Des}, \varphi_{Des})$ werden die $(2m+1)$ Moden $Y_{m,n}^{\sigma}(\Theta, \varphi)$, die pro Ordnung m existieren, so gewichtet und anschließend zusammengefasst, dass einerseits nur noch $(M+1)$ sphärische Harmonische $Y_m(\Theta_{Des}, \varphi_{Des})$ übrig bleiben. Diese werden bedingt durch Anwendung der Rotationskoeffizienten $Y_{m,n}^{\sigma}(\Theta_{Des}, \varphi_{Des})$ so in ihrer Lage verändert, dass eine anschließende Linearkombination mit den modalen Gewichtungsfaktoren C_m letztendlich zu dem modalen *beamformer* $\Psi(\Theta_{Des}, \varphi_{Des})$ führt. Dieser zeigt dann in die gewünschte Richtung $(\Theta_{Des}, \varphi_{Des})$.

Am einfachsten lässt sich dieser Vorgang nachvollziehen, wenn man sich die 3 (komplexen) Moden 1. Ordnung vorstellt, wobei es sich bekanntermaßen um drei, jeweils senkrecht zu einander stehende Achtercharakteristika handelt. Durch eine gewichtete Zusammenfassung dieser drei Richtcharakteristika lässt sich eine Richtkeule erzeugen, die in jede gewünschte Raumrichtung zeigen kann. Die dabei angesprochenen komplexen Achtercharakteristika entsprechen dabei den zugrunde liegenden, $(m+1) = 3$ komplexen ($\sigma = \pm 1$) sphärischen Harmonischen $Y_{1,n}^{\sigma}(\Theta, \varphi)$, die erwähnten Gewichte den Rotationskoeffizienten $Y_{1,n}^{\sigma}(\Theta_{Des}, \varphi_{Des})$ und die resultierende Richtkeule der auf die gewünschte Blickrichtung ausgerichteten, sphärischen Harmonischen $Y_1(\Theta_{Des}, \varphi_{Des})$.

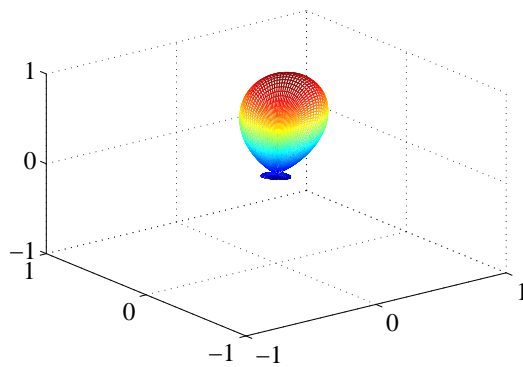
Erst in einem zweiten Schritt werden die so auf die gewünschte Blickrichtung $(\Theta_{\text{Des}}, \varphi_{\text{Des}})$ ausgerichteten, $(M + 1)$ sphärischen Harmonischen $Y_m(\Theta_{\text{Des}}, \varphi_{\text{Des}})$ mit Hilfe der modalen Gewichtungskoeffizienten C_m wie folgt zusammengefasst:

$$\Psi(\Theta_{\text{Des}}, \varphi_{\text{Des}}) = \mathbf{Y}(\Theta_{\text{Des}}, \varphi_{\text{Des}}) \mathbf{C}, \quad (\text{A.43})$$

und ergeben dann die gewünschte Richtcharakteristik $\Psi(\Theta_{\text{Des}}, \varphi_{\text{Des}})$.

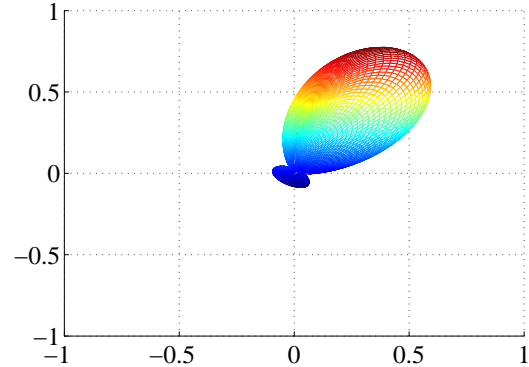
Abbildung A.17 stellt den resultierenden, modalen *beamformer* 4. Ordnung, welcher so rotiert wurde, dass er in die Richtung $(\Theta_{\text{Des}} = 45^\circ, \varphi_{\text{Des}} = 45^\circ)$ zeigt, dar.

3D-Darstellung eines modalen beamformer, welcher eine kardioidale Richtcharakteristik 9. Ordnung approximiert



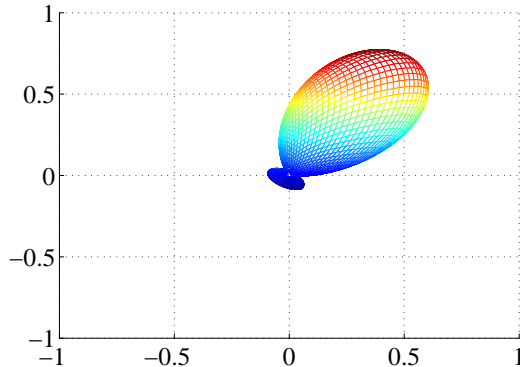
(a) 3D-Darstellung

3D-Darstellung eines modalen beamformer, welcher eine kardioidale Richtcharakteristik 9. Ordnung approximiert



(b) 3D-Darstellung - Schnitt in der X/Z-Ebene

3D-Darstellung eines modalen beamformer, welche eine kardioidale Richtcharakteristik 9. Ordnung approximiert



(c) 3D-Darstellung - Schnitt in der Y/Z-Ebene

Abbildung A.17: Graphische Darstellungen der mit den Gewichtungsfaktoren $C_m = [0.100, 0.144, 0.123, 0.086, 0.040]$ skalierten Moden, welche die Wunschrictcharakteristik (kardioidale Richtcharakteristik 9. Ordnung) approximieren, so rotiert, dass der resultierende, modale beamformer nach $\Theta = 45^\circ, \varphi = 45^\circ$ zeigt.

Abbildung A.18 fasst den Entwurf modaler *beamformer* zusammen, wobei erwähnt werden muss, dass es sich dabei um ein lineares System handelt und es somit gestattet

ist, einige Blöcke im Ablauf zu vertauschen. So kann es etwa vorteilhaft sein, den Block in dem die modale Gewichtung stattfindet mit dem Block, in dem die Rotation realisiert wird, zu vertauschen. Dies darf selbstverständlich erst dann erfolgen, wenn zuvor eine entsprechende Erweiterung der modalen Gewichte von $(M + 1)$ auf $N = (M + 1)^2$ vollzogen wurde. Eine solche Vertauschung ist speziell dann sinnvoll, wenn eine Richtcharakteristik in mehrere von einander unabhängige Raumrichtungen deuten soll, dann muss lediglich pro Richtkeule die Rotation neu berechnet werden.

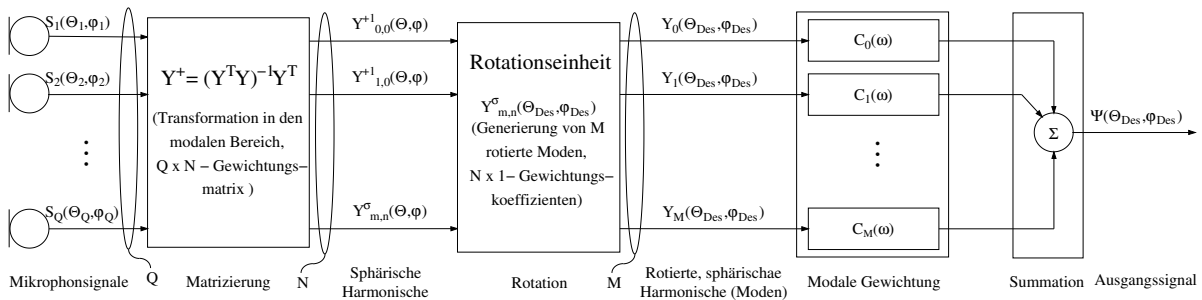


Abbildung A.18: Signalflussdiagramm, welches die notwendigen Schritte beim Entwurf eines modalen beamformers zusammenfasst.

Sind hingegen mehrere voneinander unabhängige Richtcharakteristika zu erzeugen, die zudem in unterschiedliche Richtungen zeigen sollen, dann wäre die in Abbildung A.18 dargestellte Variante zu bevorzugen, da diese weniger Signalpfade und somit einen geringeren Signalverarbeitungsaufwand beinhaltet. Pro Richtcharakteristik würden dann eine separate modale Gewichtung- und Rotationsseinheit benötigt werden.

Der Übergang vom Zeit in den räumlichen bzw. modalen Bereich wird durch eine einfache $Q \times N$ Gewichtungsmatrix, die lediglich frequenzunabhängige Verstärkungsfaktoren enthält, realisiert und verursacht somit keinen großen Aufwand. Selbiges gilt, wie im obigen Beispiel gezeigt wurde, in ähnlicher Weise auch für die Rotation. Es werden lediglich frequenzunabhängige Gewichte auf alle $(M + 1)^2$ Signale der sphärischen Harmonischen angewandt und anschließend die $(2m + 1)$ gewichteten Signale pro Ordnung m aufsummiert. Theoretisch verhält es sich bei den modalen Gewichten ebenso, allerdings muss hier der Realität Tribut gezollt werden, d.h. die modalen Gewichtungskoeffizienten $C_m(\omega)$ müssen frequenzabhängig ausgelegt werden, damit das WNG keine zu kleinen Werte annehmen kann. Dieser Verlauf kann dabei wie in Abbildung A.12 dargestellt aussehen, sollte jedoch, um das Maximale aus dem echten System herauszuholen, in Anlehnung an dem in Kapitel 5.5 aufgezeigten Optimierungsalgorithmus entworfen werden. Abbildung 5.23(b) zeigt dabei beispielhaft auf, wie ein praktischer Verlauf der Gewichtungskoeffizienten aussehen könnte. Wie man sieht, handelt es sich dabei um Filter, die einen sehr „handsamen“ Betragsfrequenzgang aufweisen, welcher sich im Prinzip auch durch ein aufwandsgünstiges IIR¹¹-Filter nachbilden lässt.

¹¹IIR = engl. *Infinite Impulse Response* = Rekursives Filter mit unendlicher Impulsantwort

A.5 Reguläre und quasi-reguläre Körper

In [Raf05, BDM06] wurde aufgezeigt, dass es günstig ist, Sensoren an regulären bzw. quasi-regulären Positionen auf einer Kugeloberfläche zu positionieren. Zum einen reduziert sich dadurch das räumliche Aliasing und zum anderen lässt sich dadurch die Anzahl der verwendeten Sensoren reduzieren, was sowohl im Sinne der Qualität, als auch bezüglich der Wirtschaftlichkeit sinnvoll ist. Einen detailreichen Überblick über reguläre und quasi-reguläre Körper sowie Verfahren zum Entwurf von beliebigen *Polyedern* können in [Hol06] gefunden werden. Wie sich beweisen lässt, existieren lediglich fünf reguläre Körper, die auch als *platonische Körper* bekannt sind: der *Tetraeder* (Vierflächner), der *Hexaeder* (Sechseckflächner), der *Oktaeder* (Achtflächner), der *Dodekaeder* (Zwölfflächner) sowie der *Ikosaeder* (Zwanzigflächner), welche in Kapitel 4.2.4 bereits kurz vorgestellt wurden und nachfolgend im Detail behandelt werden. Bevor jedoch mit der Behandlung unterschiedlicher Polyeder begonnen wird, gilt es zunächst, ein paar Begriffe zu erklären, die nachfolgend des öfteren verwendet werden.

Platonischer Körper:

Regulärer Polyeder. Von einem solchen wird gesprochen, wenn alle Oberflächen aus demselben regelmäßigen Vieleck bestehen und in jeder Ecke gleich viele dieser Vielecke zusammenstoßen.

Dualkörper:

Jeder platonische und *archimedische Körper* besitzt einen sogenannten *Dualkörper*, wobei jeder platonische Dualkörper wieder einen platonischen Körper ergibt und die archimedischen Dualkörper die sogenannten *catalanischen Körper* ergeben. Allgemein ergeben sich Dualkörper, indem die Mittelpunkte der Flächen des Ausgangskörpers miteinander verbunden werden. Dabei kennzeichnen die Flächenmittelpunkte des Ausgangskörpers bzw. die Eckpunkte seines zugehörigen Dualkörpers die Lage der Sensoren auf der Kugeloberfläche. Die Anzahl der Flächen des Ausgangskörpers entspricht der Anzahl der Ecken des Dualkörpers und umgekehrt, wobei die Anzahl der Kanten unverändert bleibt. Bildet man Dualkörper der platonischen Körper gelten folgende, in Tabelle A.3 zusammengefasste, Sonderfälle:

Platonischer Körper	Dualkörper
Tetraeder	Tetraeder (Selbstdualität)
Hexaeder	Oktaeder
Oktaeder	Hexaeder
Dodekaeder	Ikosaeder
Ikosaeder	Dodekaeder

Tabelle A.3: Dualkörper der platonischen Körper

Eulerscher Polyedersatz:

Dieser beschreibt eine fundamentale Eigenschaft von *konvexen* Polyedern die sich wie folgt formulieren lässt:

$$\boxed{\text{Anzahl der Flächen} + \text{Anzahl der Ecken} - \text{Anzahl der Kanten} = 2}.$$

Archimedischer Körper:

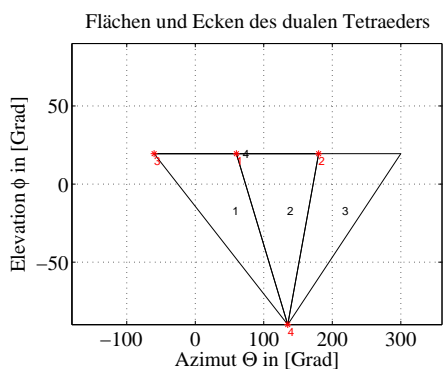
Diese werden auch als *halb-reguläre* bzw. auch als *quasi-reguläre* Polyeder bezeichnet. Sie entstehen beispielsweise, wenn von einem platonischen Körper die Ecken so abgeschnitten werden, dass dadurch alle Kanten gleich lang werden. Alternativ lassen sich diese auch durch einen Schnitt eines platonischen Körpers mit seinem, passend vergrößerten *Dualkörper* konstruieren.

A.5.1 Der Tetraeder

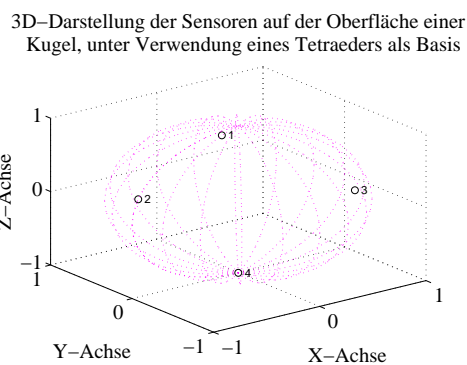


(a) 3D-Darstellung des Grundkörpers (Draufsicht)

(b) 3D-Darstellung des Dualkörpers (Draufsicht)



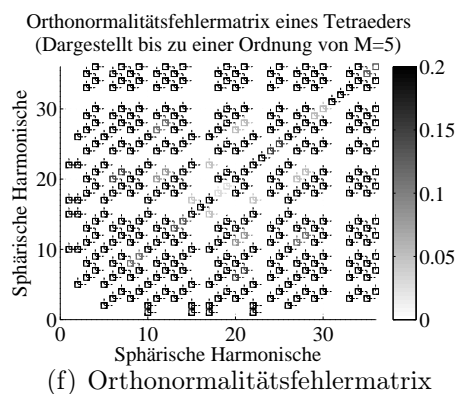
(c) 2D-Darstellung der Flächen und Ecken des Dualkörpers



(d) Positionierung der Sensoren auf der Kugeloberfläche

Azimut in [Grad]	Elevation in [Grad]	Radius (linear)
60.0000	19.4712	1.0000
180.0000	19.4712	1.0000
-60.0000	19.4712	1.0000
135.0000	-90.0000	1.0000

(e) Sensorpositionen



(f) Orthonormalitätsfehlermatrix

Abbildung A.19: 3D-Darstellung des Tetraeders, seines dualen Körpers, dessen 2D-Darstellung der Flächen und Ecken, der korrespondierenden Sensorpositionen auf der Kugeloberfläche sowie der daraus resultierenden Orthonormalitätsfehlermatrix.

Abbildung A.19(a) beinhaltet die Draufsicht des Tetraeders in seiner dreidimensionalen Form.

Wendet man die zuvor beschriebene Konstruktion des Dualkörpers an, d.h. verbind-

det man die Flächenmittelpunkte des Tetraeders miteinander, so ergibt sich wieder ein Tetraeder, der nun jedoch, wie die Draufsicht auf den dreidimensionalen Dualkörper aus Abbildung A.19(b) verdeutlicht, nach unten zeigt und horizontal um 180° gedreht ist.

Rollt man den dualen Körper auf einer Ebene ab, ergibt sich die in Abbildung A.19(c) aufgezeigte zweidimensionale Darstellung, in welcher sowohl die Flächen, als auch die Eckpunkte aufgetragen und durchnummeriert sind. Dabei kennzeichnen die schwarzen Nummern die Flächen und die roten mit einem Stern versehenen Zahlen die Eckpunkte. Darüber hinaus werden für die Darstellung von Flächen, die über die Grenzen des horizontalen Winkelbereichs, welcher eine Größe von $\Theta = [-180^\circ, \dots, 180^\circ]$ hat, hinausragen, nach rechts, d.h. zu Azimutwinkeln von maximal $\Theta = 360^\circ$ erweitert, wodurch sich auch der Darstellungsbereich dieser Abbildung von $\Theta = [-180^\circ, \dots, 360^\circ]$ und $\varphi = [-90^\circ, \dots, 90^\circ]$ erklärt. Um Verwirrungen zu vermeiden, werden für die erweiterte Darstellung der Flächen die Nummern der duplizierten Eckpunkte nicht noch einmal angegeben. Die Fläche die in Abbildung A.19(c) mit der schwarzen Nummer 4 bezeichnet wird, kennzeichnet die in dieser Darstellung nicht darstellbare Bodenfläche, welche durch die roten Eckpunkte 4, 2 und 3 aufgespannt wird.

Bekanntermaßen kennzeichnen die Eckpunkte des dualen Körpers die Lagen der Sensoren, welche sich projiziert auf eine Kugeloberfläche, wie in Abbildung A.19(d) dargestellt, ergeben. Diese entsprechen exakt den Mikrofonpositionen eines Klangfeldmikrophons, deren Vertreter der Firma „SoundField“ beispielhaft in Darstellung 4.1 abgebildet ist. Allerdings gibt es hier einen kleinen aber entscheidenden Unterschied, den es zu beachten gilt. Während bei den kommerziell erhältlichen Klangfeldmikrophonen Mikrophone mit kardioider Charakteristik, eingesetzt werden, wird in dieser Untersuchung stets von Mikrofonen mit Kugelcharakteristik ausgegangen. Da diese logischerweise anders kombiniert werden müssen als Mikrophone mit kardioider Charakteristik um letztendlich die gesuchten Ambisonicsignale (W , X , Y und Z) zu erhalten, unterscheiden sich die hier ermittelten Mikrofonpositionen von den Koordinaten kommerziell erhältlicher Klangfeldmikrophone.

Berechnet man gemäß Gleichung 5.4 bzw. Gleichung 5.5 die in Abbildung 5.2 und Abbildung A.18 dargestellte, pseudoinverse Matrix $\mathbf{Y}_{N \times Q}^+$, welche die Transformation vom Zeit- in den Wellenfeldbereich vollzieht, erkennt man, dass sich hier die Ambisonicsignale W , X , Y und Z wie folgt aus den gemessenen Eingangssignalen ergeben:

$$\begin{pmatrix} W(j\omega) \\ Z(j\omega) \\ X(j\omega) \\ Y(j\omega) \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\sqrt{3} \\ \sqrt{\frac{2}{3}} & -\sqrt{\frac{8}{3}} & \sqrt{\frac{2}{3}} & 0 \\ \sqrt{2} & 0 & -\sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} S_1(j\omega) \\ S_2(j\omega) \\ S_3(j\omega) \\ S_4(j\omega) \end{pmatrix}. \quad (\text{A.44})$$

Diese unterscheiden sich im Vergleich z.B. zu dem in Abbildung 4.1 dargestellten Klangfeldmikrophon, welches bekanntermaßen aus vier Mikrofonen mit kardioider Charakte-

ristik besteht und wie in Tabelle A.4 auf der Kugeloberfläche angeordnet sind.

Azimut in [Grad]	Elevation in [Grad]	Radius (linear)
45.0	35.3	1.0
135.0	-35.3	1.0
-45.0	-35.3	1.0
-135.0	35.3	1.0

Tabelle A.4: Positionen der Kardiodmikrophone des Klangfeldmikrophons aus Abbildung 4.1

Aus diesen lassen sich die Ambisonicsignale W , X , Y und Z wie folgt bestimmen:

$$\begin{pmatrix} W(j\omega) \\ Z(j\omega) \\ X(j\omega) \\ Y(j\omega) \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} S_1(j\omega) \\ S_2(j\omega) \\ S_3(j\omega) \\ S_4(j\omega) \end{pmatrix}. \quad (\text{A.45})$$

Abbildung A.19 beinhaltet zudem neben den Kugelkoordinaten der Sensorpositionen, welche in tabellarischer Form im Teilbild A.19(e) aufgeführt sind und in Übereinstimmung mit der in Abbildung A.19(d) aufgezeigten, dreidimensionalen Sensoranordnung gebracht werden kann, wenn diese von unten her betrachtet wird auch die sich daraus ergebende *Orthonormalitätsfehlermatrix* \mathbf{E} . Letztere enthält die aus der Lage und Anzahl der Sensoren resultierenden Abweichungen von der Orthonormalität, die stets bis zu einer Ordnung von $M = 5$, bzw. für die ersten $N = (M + 1)^2 = 36$ sphärischen Harmonischen dargestellt werden, um einen vergleichbaren Wert für die hier untersuchten Polyeder zu gewinnen. Die Einträge dieser Fehlermatrix ergeben sich aus der Differenz der aus Gleichung 5.7 bekannten Orthonormalitätsmatrix und einer Einheitsmatrix \mathbf{I} gleicher Größe. Da die Orthonormalitätsmatrix im Idealfall zu einer Einheitsmatrix wird, sollte die Orthonormalitätsfehlermatrix optimalerweise nur Nullen enthalten. Mathematisch lässt sich die Berechnung der Orthonormalitätsfehlermatrix \mathbf{E} , bei welcher es sich um eine $N \times N$ -Matrix handelt, wie folgt ausdrücken:

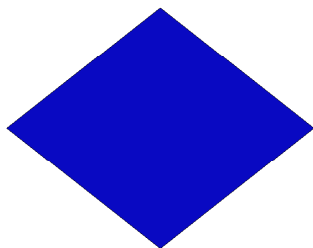
$$\begin{aligned} \mathbf{E}_{N \times N} &= \mathbf{I}_{N \times N} - \frac{1}{Q} Y_{m,n}^\sigma(\Theta_q, \varphi_q)^T Y_{m',n'}^{\sigma'}(\Theta_q, \varphi_q), \text{ mit} \\ &0 \leq m, m' \leq M, 0 \leq n, n' \leq M, \sigma, \sigma' = \pm 1, \text{ und} \\ N &= (M + 1)^2, \text{ wobei} \\ Q &= \text{Anzahl der Arraymikrophone,} \\ M &= \text{Ordnung der sphärischen Harmonischen,} \\ N &= \text{Anzahl der sphärischen Harmonischen,} \\ q &= \text{Mikrofonindex mit } q = [1, \dots, Q]. \end{aligned} \quad (\text{A.46})$$

Ferner erkennt man aus Abbildung A.19(f), dass sich bei Verwendung von $Q = 4$

Mikrofonen, welche gemäß der in Tabelle A.19(e) angegebenen Positionen angeordnet sind die ersten $N = 4$ sphärische Harmonischen fehlerfrei bestimmen lassen. Höhere sphärische Harmonische können, wie aus Abbildung A.19(f) ersichtlich, nicht bzw. nur mit sehr großen Fehlern gemessen werden. Aus diesem Grund wird eine auf einem Tetraeder basierte Mikrofonanordnung auch nur bis zu einer maximalen Ordnung von $M = \sqrt{N} - 1 = \sqrt{4} - 1 = 1$ verwendet, d.h. es werden lediglich die W , X , Y und Z Komponenten der sphärischen Harmonischen (Ambisonicsignale) aus den vier Mikrophonsignalen extrahiert und verwendet.

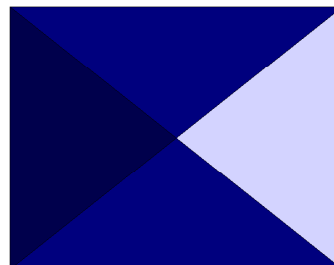
A.5.2 Der Hexaeder

Hexaeder



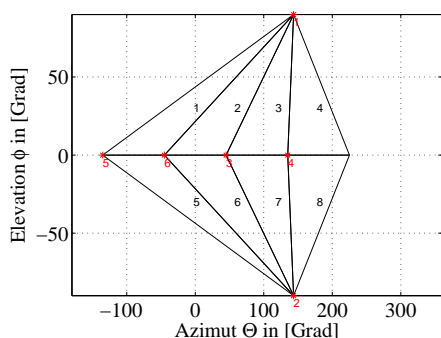
(a) 3D-Darstellung des Grundkörpers (Draufsicht)

Dualkörper des Hexaeders



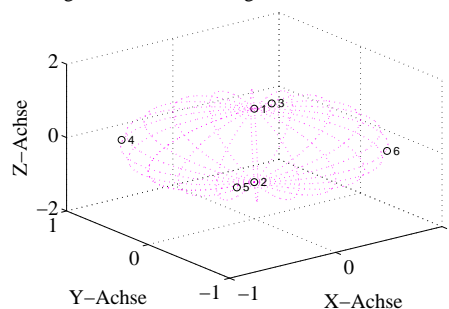
(b) 3D-Darstellung des Dualkörpers (Draufsicht)

Flächen und Ecken des dualen Hexaeders



(c) 2D-Darstellung der Flächen und Ecken des Dualkörpers

3D-Darstellung der Sensoren auf der Oberfläche einer Kugel, unter Verwendung eines Hexaeders als Basis

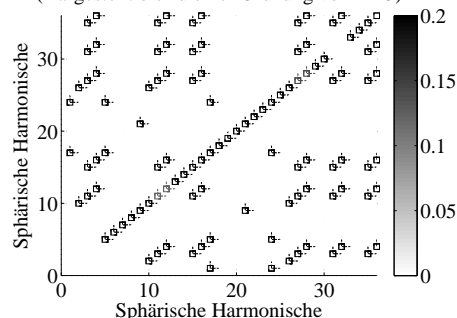


(d) Positionierung der Sensoren auf der Kugeloberfläche

Azimut in [Grad]	Elevation in [Grad]	Radius (linear)
143.4909	90.0000	1.0000
143.4909	-90.0000	1.0000
45.0000	0.0000	1.0000
135.0000	0.0000	1.0000
-135.0000	0.0000	1.0000
-45.0000	0.0000	1.0000

(e) Sensorpositionen

Orthonormalitätsfehlermatrix eines Hexaeders (Dargestellt bis zu einer Ordnung von $M=5$)



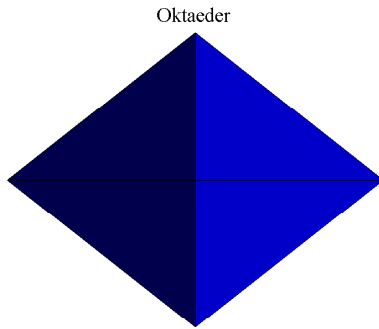
(f) Orthonormalitätsfehlermatrix

Abbildung A.20: 3D-Darstellung des Hexaeders, seines dualen Körpers, dessen 2D-Darstellung der Flächen und Ecken, der korrespondierenden Sensorpositionen auf der Kugeloberfläche sowie der daraus resultierenden Orthonormalitätsfehlermatrix.

Beim Dualkörper des in Teilbild A.20(a) dargestellten Hexaeders handelt es sich, wie aus Abbildung A.20(b) ersichtlich, um einen Oktaeder, dessen Eckpunkte abermals die örtlichen Lagen der Sensoren kennzeichnen. Dabei befinden sich je zwei Sensoren an den

äußeren Enden der jeweils senkrecht aufeinander stehenden X , Y und Z -Achsen, was Abbildung A.20(d) verdeutlicht. Mit den so angeordneten 6 Sensoren können dabei, wie mit den 4 Sensoren des Tetraeders, lediglich die ersten $N = 4$ sphärischen Harmonischen ohne Fehler messtechnisch bestimmt werden, was ein Blick auf Abbildung A.20(f) verdeutlicht. Ein Vergleich der Orthonormalitätsfehlermatrix des Tetraeders (Abbildung A.19(f)) mit der des Hexaeders (Abbildung A.20(f)) offenbart nun, dass bedingt durch die höhere Anzahl der Sensoren der Fehler bei der Messung höherer sphärischer Harmonischer deutlich reduziert und an manchen Stellen sogar zu Null wird. So kommt es bei der Messung der sphärischen Harmonischen $N = [5, \dots, 9]$ lediglich zu Störungen durch sphärische Harmonische gleicher Ordnung, was in Abbildung A.19(f) aus dem diagonalen Verlauf in diesem Bereich zu erkennen ist.

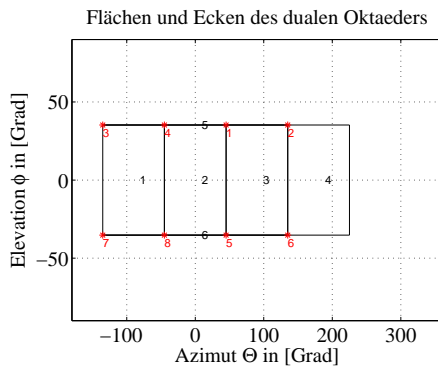
A.5.3 Der Oktaeder



(a) 3D-Darstellung des Grundkörpers (Draufsicht)

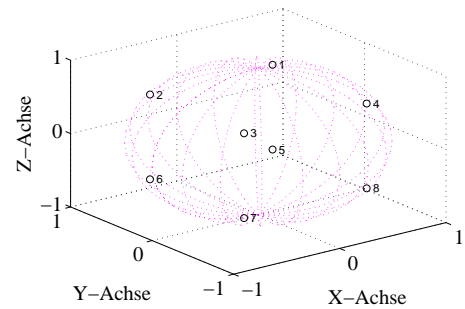


(b) 3D-Darstellung des Dualkörpers (Draufsicht)



(c) 2D-Darstellung der Flächen und Ecken des Dualkörpers

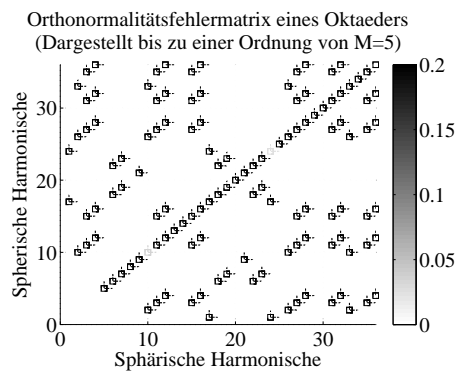
3D-Darstellung der Sensoren auf der Oberfläche einer Kugel, unter Verwendung eines Oktaeders als Basis



(d) Positionierung der Sensoren auf der Kugeloberfläche

Azimut in [Grad]	Elevation in [Grad]	Radius (linear)
45.0000	35.2644	1.0000
135.0000	35.2644	1.0000
-135.0000	35.2644	1.0000
-45.0000	35.2644	1.0000
45.0000	-35.2644	1.0000
135.0000	-35.2644	1.0000
-135.0000	-35.2644	1.0000
-45.0000	-35.2644	1.0000

(e) Sensorpositionen



(f) Orthonormalitätsfehlermatrix

Abbildung A.21: 3D-Darstellung des Oktaeders, seines dualen Körpers, dessen 2D-Darstellung der Flächen und Ecken, der korrespondierenden Sensorpositionen auf der Kugeloberfläche sowie der daraus resultierenden Orthonormalitätsfehlermatrix.

Beim Oktaeder sind die Sensoren wie dessen Dualkörper, welcher in Abbildung A.21(b) dargestellt ist, verdeutlicht an den Ecken eines Hexaeders (Würfel) angeordnet, was auch aus der in Abbildung A.21(d) aufgezeigten, dreidimensionalen Positionierung auf einer

Kugeloberfläche ersichtlich ist. Ein Vergleich der Orthonormalitätsfehlermatrix des Hexaeders aus Abbildung A.20(f) mit der des Oktaeders aus Abbildung A.21(f) lässt keinen nennenswerten Unterschied erkennen, wodurch dessen Einsatz zur messtechnischen Bestimmung von sphärischen Harmonischen fragwürdig erscheint, da sich kein Gewinn durch die Steigerung der Anzahl der Sensoren von $Q = 6$ (Hexaeder) auf $Q = 8$ (Oktaeder) erkennen lässt. Man ist also selbst bei Verwendung von $Q = 8$ regulär angeordneten Sensoren nicht in der Lage, mehr als $N = 4$ sphärische Harmonische fehlerfrei zu bestimmen, was bekanntermaßen bereits mit $Q = 4$ regulär angeordneten Sensoren (Tetraeder) möglich ist.

A.5.4 Der Dodekaeder

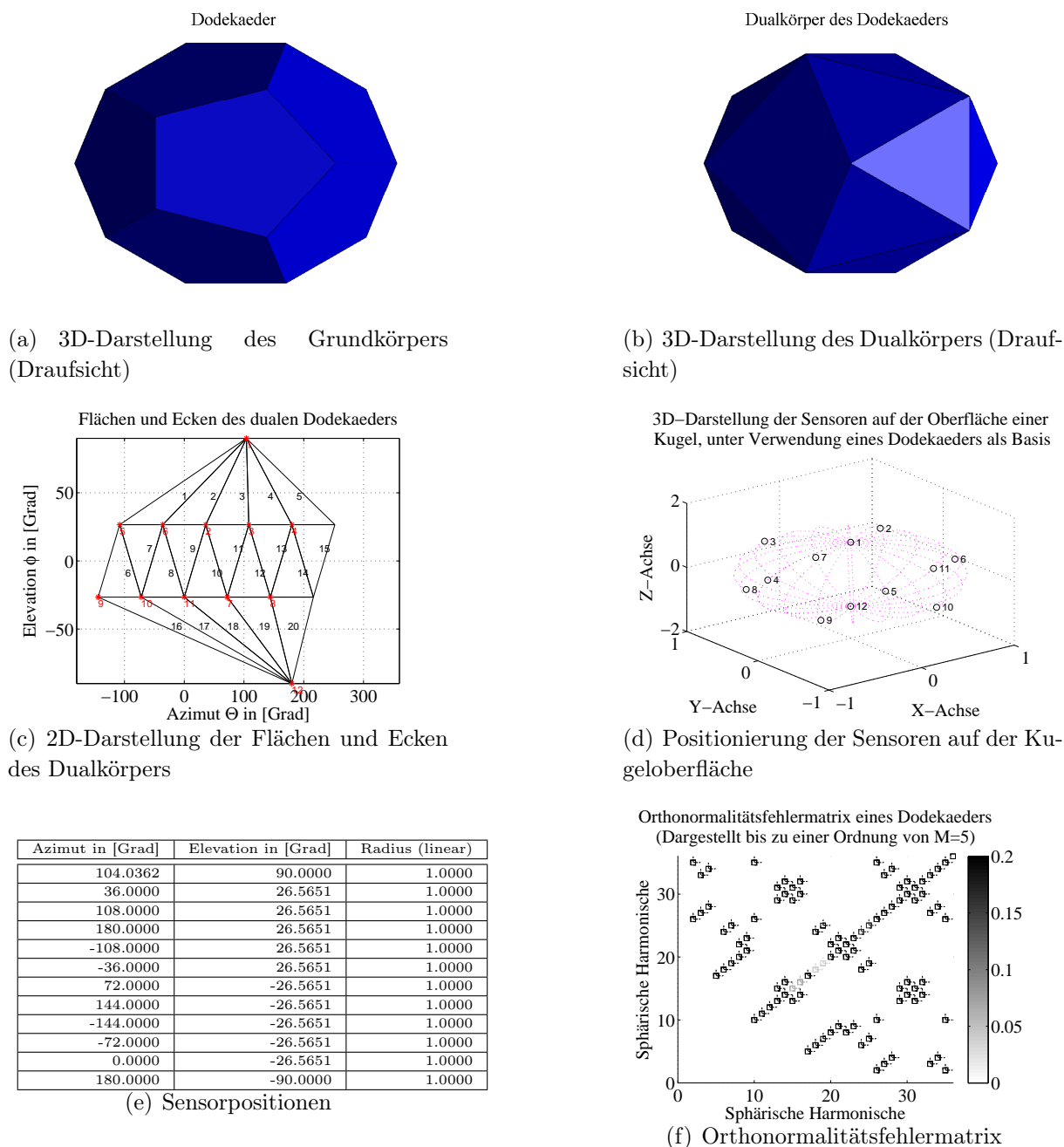
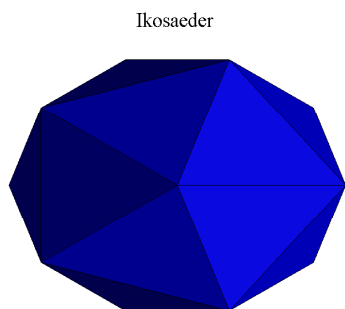


Abbildung A.22: 3D-Darstellung des Dodekaeder, seines dualen Körpers, dessen 2D-Darstellung der Flächen und Ecken, der korrespondierenden Sensorpositionen auf der Kugeloberfläche sowie der daraus resultierenden Orthonormalitätsfehlermatrix.

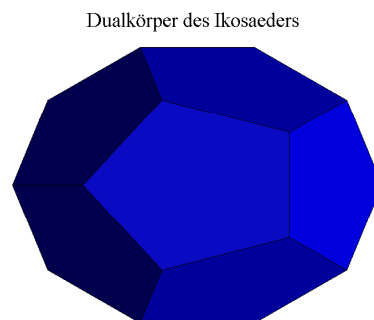
Mit dem in Teilbild A.22(a) dargestellten Dodekaeder ist man in der Lage, mit Hilfe seiner $Q = 12$ Sensoren, die an den Ecken eines Icosaeders sitzen, was Abbildung A.22(b) verdeutlicht, sphärische Harmonische bis zu einer Ordnung von $M = \lfloor \sqrt{Q} - 1 \rfloor = 2$ bzw.

der ersten $N = (M + 1)^2 = 9$ sphärischen Harmonischen messtechnisch, fehlerfrei zu erfassen, was ein Blick auf Abbildung A.22(f) beweist.

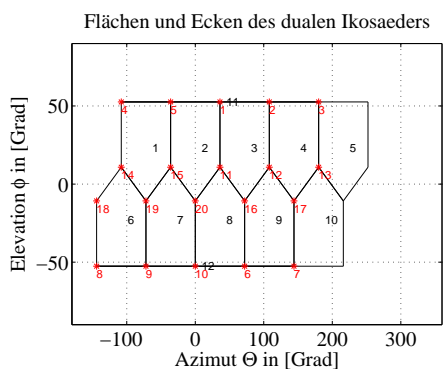
A.5.5 Der Ikosaeder



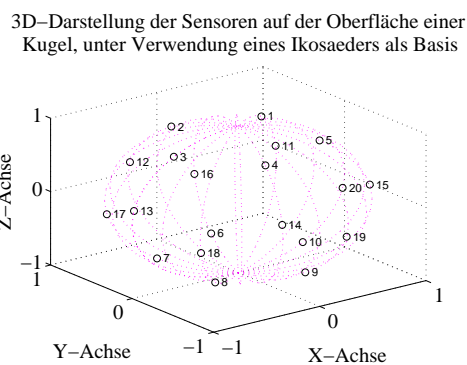
(a) 3D-Darstellung des Grundkörpers (Draufsicht)



(b) 3D-Darstellung des Dualkörpers (Draufsicht)



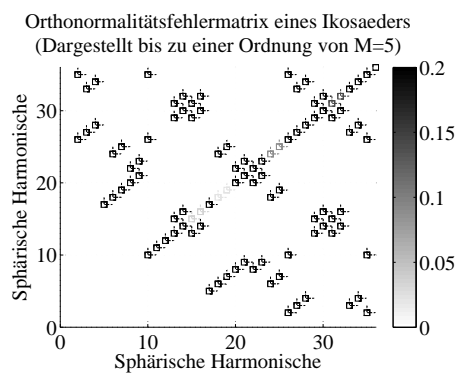
(c) 2D-Darstellung der Flächen und Ecken des Dualkörpers



(d) Positionierung der Sensoren auf der Kugeloberfläche

Azimut in [Grad]	Elevation in [Grad]	Radius (linear)
36.0000	52.6226	1.0000
108.0000	52.6226	1.0000
180.0000	52.6226	1.0000
-108.0000	52.6226	1.0000
-36.0000	52.6226	1.0000
72.0000	-52.6226	1.0000
144.0000	-52.6226	1.0000
-144.0000	-52.6226	1.0000
-72.0000	-52.6226	1.0000
0.0000	-52.6226	1.0000
36.0000	10.8123	1.0000
108.0000	10.8123	1.0000
180.0000	10.8123	1.0000
-108.0000	10.8123	1.0000
-36.0000	10.8123	1.0000
72.0000	-10.8123	1.0000
144.0000	-10.8123	1.0000
-144.0000	-10.8123	1.0000
-72.0000	-10.8123	1.0000
0.0000	-10.8123	1.0000

(e) Sensorpositionen



(f) Orthonormalitätsfehlermatrix

Abbildung A.23: 3D-Darstellung des Ikosaeders, seines dualen Körpers, dessen 2D-Darstellung der Flächen und Ecken, der korrespondierenden Sensorpositionen auf der Kugeloberfläche sowie der daraus resultierenden Orthonormalitätsfehlermatrix.

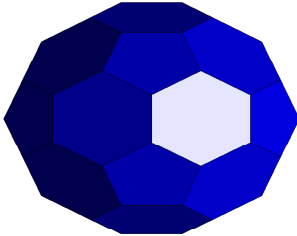
Wie bereits beim Vergleich des Hexaeders mit dem Oktaeder kommt es auch beim Vergleich der Orthonormalitätsfehlermatrix des Dodekaeders (Abbildung A.22(f)) mit der des Ikosaeders (Abbildung A.23(f)) zu keinen nennenswerten Unterschied, obwohl Letztgenannter einerseits 8 Sensoren mehr besitzt und andererseits mit diesem theoretisch sphärische Harmonische bis zu einer Ordnung von $M = \lfloor \sqrt{20} - 1 \rfloor = 3$ bzw. die ersten $N = 16$ sphärischen Harmonischen fehlerfrei bestimmt werden können sollten, was offensichtlich jedoch nicht der Fall ist.

→ *Eine regulär abgetastete Kugeloberfläche gegeben z.B. durch einen platonischen Körper bedeutet nicht zwangsweise, dass die daraus resultierende Sensoranordnung im Sinne der Orthonormalitätsbedingung optimal ist.*

Allerdings resultieren Sensoren, die regulär auf einer Kugeloberfläche angebracht sind, in Ambisonicsysteme, die meist sehr gute, jedoch nicht unbedingt optimale Ergebnisse liefern. Wie gut dabei eine spezifische Sensoranordnung für ein Ambisonicsystem tatsächlich ist, muss letztendlich z.B. durch Aufnahme der Orthonormalitätsfehlermatrix individuell von Fall zu Fall ermittelt werden.

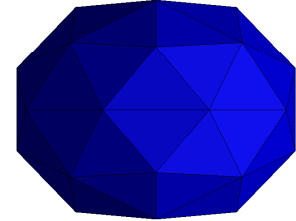
A.5.6 Der Buckyball

Buckyball (Ikosaederstumpf)



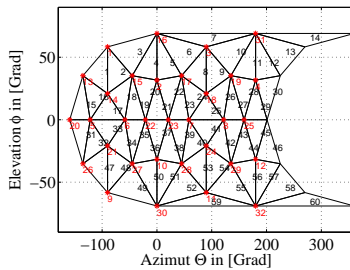
(a) 3D-Darstellung des Grundkörpers (Draufsicht)

Dualkörper des Buckyballs (Pentakisdodekaeder)



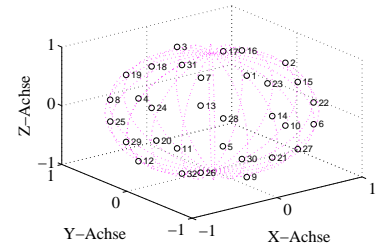
(b) 3D-Darstellung des Dualkörpers (Draufsicht)

Flächen und Ecken eines dualen Buckyballs



(c) 2D-Darstellung der Flächen und Ecken des Dualkörpers

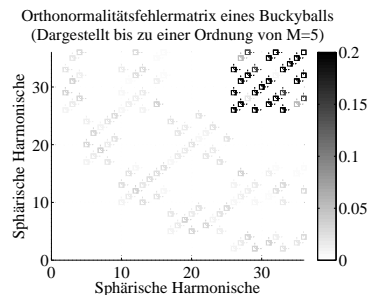
3D-Darstellung der Sensoren auf der Oberfläche einer Kugel, unter Verwendung eines Buckyballs als Basis



(d) Positionierung der Sensoren auf der Kugeloberfläche

Azimut in [Grad]	Elevation in [Grad]	Radius (linear)
-90.0000	58.2825	1.0164
0.0000	31.7175	1.0164
90.0000	58.2825	1.0164
180.0000	31.7175	1.0164
-121.7175	0.0000	1.0164
-58.2825	0.0000	1.0164
58.2825	0.0000	1.0164
121.7175	0.0000	1.0164
-90.0000	-58.2825	1.0164
0.0000	-31.7175	1.0164
90.0000	-58.2825	1.0164
180.0000	-31.7175	1.0164
-135.0000	35.2644	0.9901
-90.0000	20.9052	0.9901
-45.0000	35.2644	0.9901
0.0000	69.0948	0.9901
45.0000	35.2644	0.9901
90.0000	20.9052	0.9901
135.0000	35.2644	0.9901
-159.0948	0.0000	0.9901
-90.0000	-20.9052	0.9901
-20.9052	0.0000	0.9901
20.9052	0.0000	0.9901
90.0000	-20.9052	0.9901
159.0948	0.0000	0.9901
-135.0000	-35.2644	0.9901
-45.0000	-35.2644	0.9901
45.0000	-35.2644	0.9901
135.0000	-35.2644	0.9901
0.0000	-69.0948	0.9901
180.0000	69.0948	0.9901
180.0000	-69.0948	0.9901

(e) Sensorpositionen



(f) Orthonormalitätsfehlermatrix

Abbildung A.24: 3D-Darstellung des Buckyball, seines dualen Körpers, dessen 2D-Darstellung der Flächen und Ecken, der korrespondierenden Sensorpositionen auf der Kugeloberfläche sowie der daraus resultierenden Orthonormalitätsfehlermatrix.

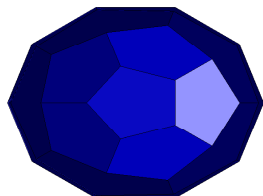
Bei dem nachfolgenden Körper handelt es sich um einen Vertreter der insgesamt 13 archimedischen Körper von denen sich 5, wie eingangs erwähnt, durch eine Modifizierung, z.B. durch eine Abstumpfung der 5 platonischen Körper ergeben. Der dabei wohl interessanteste Vertreter für unsere Zwecke ist der in Abbildung A.24(a) dargestellte, sogenannte *Buckyball*, der diesen Namen trägt, weil er die Form eines Fußballs hat. Dabei handelt es sich um einen *abgestumpften Ikosaeder*, von denen es, wie wir später noch sehen werden, mehrere Vertreter gibt. Abgestumpfte Ikosaeder, wie der hier vorgestellte Buckyball, setzen sich allgemein aus einem Dodekaeder und einem Ikosaeder zusammen. Sie bestehen somit aus 12 *Pentagons* (Fünfecke) und 20 *Hexagons* (Sechsecke), also insgesamt aus 32 Flächen. Ferner weisen diese neben den 32 Flächen noch 60 Ecken und 90 Kanten auf, sodass der *eulersche Polyedersatz* (siehe Kapitel A.5) durch diesen Körper erfüllt wird. Der in Abbildung A.24(b) dargestellte, duale Körper des Buckyball, dessen Ecken bekanntermaßen die Positionen der 32 Sensoren kennzeichnen, ist in der Literatur als *Pentakisdodekaeder* bekannt und gehört wie jeder andere duale, archimedische Körper zu den *catalanischen Körpern*.

Wie man anhand der Radien, die aus der im Teilbild A.24(e) aufgezeigten Tabelle ersichtlich sind, erkennen kann, befinden sich die Sensoren nicht mehr exakt auf der Kugeloberfläche, da es zur Ausbildung von zwei unterschiedlichen Radien kommt, je nachdem ob die Positionen vom zugrunde liegenden Dodekaeder, von welchem die größeren Radien herrühren oder vom Ikosaeder, welcher der Lieferant der kleineren Radien ist, stammen. Die letztendlich verwendeten Radien wurden dabei so berechnet, dass deren Mittelwert eine Länge von 1 aufweist.

Verwendet man die durch die Ecken des Pentakisdodekaeders gegebenen in Tabelle A.24(e) aufgelisteten Sensorpositionen für die Messung der sphärischen Harmonischen, resultiert die in Abbildung A.24(f) aufgezeigte Orthonormalitätsfehlermatrix. In Anlehnung an den Dodekaeder bzw. den Ikosaeder liefert dieser ebenfalls nur bis zu einer Ordnung von $M = 2$ perfekte, d.h. fehlerfreie Ergebnisse, jedoch kann er bis zu einer Ordnung von $M = \lfloor \sqrt{32} - 1 \rfloor = 4$ bzw. für die Bestimmung der ersten $N = 25$ sphärischen Harmonischen verwendet werden, da hier der maximale Fehler lediglich bei 3.5855% liegt, was als akzeptabler Wert angesehen werden kann, insbesondere da die Exemplarstreuung der verwendeten Sensoren (Mikrophone) normalerweise diesen Wert übersteigt.

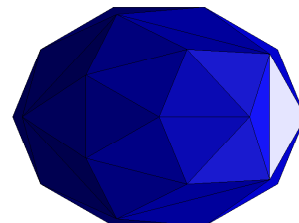
A.5.7 Der verwendete Ikosaederstumpf

Angewendeter Ikosaederstumpf



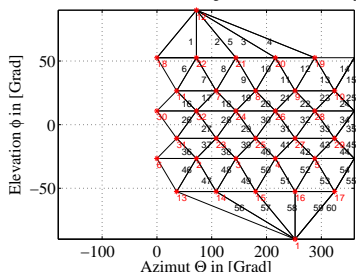
(a) 3D-Darstellung des Grundkörpers (Draufsicht)

Angewendeter Dualkörper des Ikosaederstumpfes



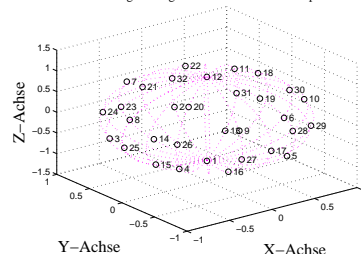
(b) 3D-Darstellung des Dualkörpers (Draufsicht)

Flächen und Ecken des dualen, angewendeten Ikosaederstumpfes



(c) 2D-Darstellung der Flächen und Ecken des Dualkörpers

3D-Darstellung der Sensoren auf der Oberfläche einer Kugel, unter Verwendung des angewendeten Ikosaederstumpfes als Basis

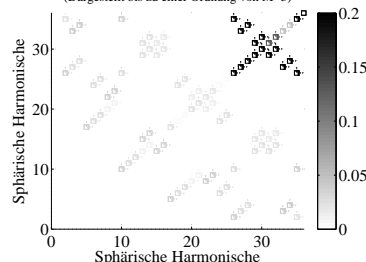


(d) Positionierung der Sensoren auf der Kugeloberfläche

Azimut in [Grad]	Elevation in [Grad]	Radius (linear)
252.0000	-90.0000	1.0160
72.0000	-26.6000	1.0160
144.0000	-26.6000	1.0160
216.0000	-26.6000	1.0160
288.0000	-26.6000	1.0160
0.0000	-26.6000	1.0160
108.0000	26.6000	1.0160
180.0000	26.6000	1.0160
252.0000	26.6000	1.0160
324.0000	26.6000	1.0160
36.0000	26.6000	1.0160
72.0000	90.0000	1.0160
36.0000	-52.6000	0.9900
108.0000	-52.6000	0.9900
180.0000	-52.6000	0.9900
252.0000	-52.6000	0.9900
324.0000	-52.6000	0.9900
0.0000	52.6000	0.9900
288.0000	52.6000	0.9900
216.0000	52.6000	0.9900
144.0000	52.6000	0.9900
72.0000	52.6000	0.9900
108.0000	-10.8000	0.9900
144.0000	10.8000	0.9900
180.0000	-10.8000	0.9900
216.0000	10.8000	0.9900
252.0000	-10.8000	0.9900
288.0000	10.8000	0.9900
324.0000	-10.8000	0.9900
0.0000	10.8000	0.9900
36.0000	-10.8000	0.9900
72.0000	10.8000	0.9900

(e) Sensorpositionen

Orthonormalitätsfehlermatrix des angewendeten Ikosaederstumpfes (Dargestellt bis zu einer Ordnung von $M=5$)



(f) Orthonormalitätsfehlermatrix

Abbildung A.25: 3D-Darstellung des verwendeten Ikosaederstumpfes, seines dualen Körpers, dessen 2D-Darstellung der Flächen und Ecken, der korrespondierenden Sensorpositionen auf der Kugeloberfläche sowie der daraus resultierenden Orthonormalitätsfehlermatrix.

Beim letzten hier behandelten Körper handelt es sich um eine weitere Variante eines abgestumpften Ikosaeders. Die letztendlich für die Positionierung der Mikrophone auf der Kugeloberfläche verwendeten, in Tabelle A.25(e) aufgelisteten Koordinaten stammen aus dem in Abbildung A.25(b) dargestellten Dualkörper des in Abbildung A.25 aufgezeigten Ikosaederstumpfes, die auch in [ME04] zu finden sind. Der Unterschied zum Buckyball ist, dass sich bei diesem Körper jeweils 6 der insgesamt 12 Pentagons an den beiden Polen des Körpers befinden und nicht wie beim Buckyball regelmäßig auf der Kugeloberfläche verteilt sind, wodurch dieser Körper eine etwas höhere Sensordichte um den Äquator hat. Dies ist in der Praxis positiv zu werten, da üblicherweise dort, d.h. in der horizontalen Ebene in welcher sich die Ohren befinden, das Wellenfeld am exaktesten gemessen bzw. nachgebildet werden muss. Betrachtet man sich die in Abbildung A.25(f) aufgezeigte Orthonormalitätsfehlermatrix und vergleicht diese mit der Orthonormalitätsfehlermatrix des Buckyballs aus Abbildung A.24(f), fällt auf, dass sich diese ziemlich stark ähneln. So ist der verwendete Ikosaederstumpf ebenfalls in der Lage, sphärische Harmonische bis zu einer maximalen Ordnung von $M = 4$ messtechnisch zu erfassen, wobei nun ein maximaler Fehler von 3.6716% auftritt, wodurch er sich nur unwesentlich von der Qualität des Buckyballs in diesem Bereich unterscheidet. Erwähnenswert ist zudem, dass der verwendete Ikosaederstumpf nun nicht mehr in der Lage ist, sphärische Harmonische bis zu einer Ordnung von $M = 2$ fehlerfrei zu bestimmen, sondern es nun auch in diesem Bereich zu einem gewissen Fehler kommt, der jedoch lediglich 0.1089% beträgt.

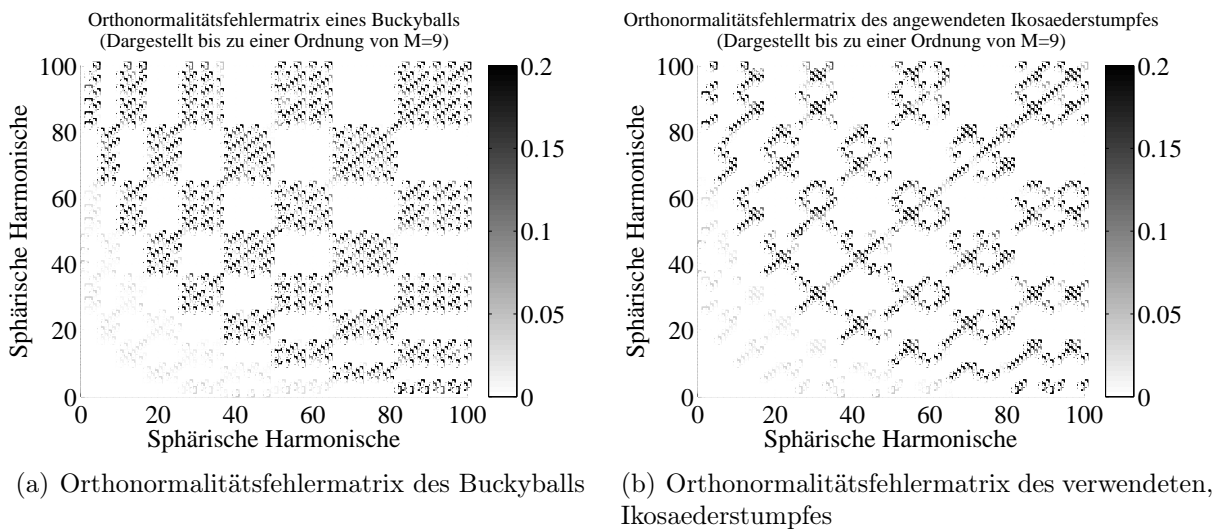


Abbildung A.26: Orthonormalitätsfehlermatrix des Buckyballs sowie des verwendeten Ikosaederstumpfes, dargestellt bis zu einer maximalen Ordnung von $M = 9$

Ferner fällt auf, dass sich die beiden Ikosaederstümpfe vorrangig außerhalb ihres vorgesehenen Einsatzbereichs, d.h. im Bereich oberhalb von $N > 25$ unterscheiden, was durch einen Vergleich der beiden in Abbildung A.26 aufgezeigten Orthonormalitätsfehlermatrizen des Buckyballs und des verwendeten Ikosaederstumpfes, die hier bis zu einer Maxi-

malen Ordnung von $M = 9$ dargestellt wurden, ersichtlich ist.

Somit wird deutlich, dass sich sowohl hinsichtlich seiner erhöhten Sensordichte um den Äquator als auch bezüglich seiner geringeren Fehleranfälligkeit durch höhere, sphärische Harmonische, der in Kapitel A.5.7 vorgestellte, abgestumpfte Ikosaeder sehr gut für die messtechnische Erfassung sphärischer Harmonischer bis zu einer Ordnung von $M = 4$ eignet.

Literaturverzeichnis

- [ALS05] I. Assenmacher, T. Lentz und J. Sokoll, "Performance of spatial audio using dynamic cross-talk cancellation," in *Audio Engineering Society Convention 119*, 10 2005. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=13296>
- [BA05] T. Betlehem und D. Abhayapala, T., "Theory and design of sound field reproduction in reverberant rooms," *JASA*, Bd. 117, Nr. 4, S. 2100–2111, 2005.
- [Ban10] B. Bank, "Audio equalization with fixed-pole parallel filters: An efficient alternative to complex smoothing," in *Audio Engineering Society Convention 128*, 5 2010. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=15262>
- [BDM06] S. Bertet, J. Daniel und S. Moreau, "3D sound field recording with higher order ambisonics - objective measurements and validation of spherical microphone," *Audio Engineering Society Convention 120*, 5 2006. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=13661>
- [BK99] R. K. Bock und W. Krischer, "The data analysis brief book - pseudoinverse," 3 1999. [Online]. Available: <http://rkb.home.cern.ch/rkb/AN16pp/node220.html#SECTION00022000000000000000000000000000.html>
- [Bmw07] BmwPress, "Anordnung der Lautsprecher im Fahrzeug," 2007. [Online]. Available: <http://www.bmwarchiv.de/foto/312-40-anordnung-der-lautsprecher-im-fahrzeug.html>
- [BS00] I. N. Bronstein und K. A. Semendjajew, *Taschenbuch der Mathematik*. Harri Deutsch, Frankfurt, Germany, 2000.
- [CB89] D. H. Cooper und J. L. Bauck, "Prospects for transaural recording," *J. Audio Eng. Soc*, Bd. 37, Nr. 1/2, S. 3–19, 1989. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=6108>

- [CCFS10] A. Capra, L. Chiesi, A. Farina und L. Scopece, “A spherical microphone array for synthesizing virtual directive microphones in live broadcasting and in post production,” in *Audio Engineering Society Conference: 40th International Conference: Spatial Audio: Sense the Sound of Space*, 10 2010. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=15577>
- [Chr11a] M. Christoph, “Design of a modal beamformer, based on a microphone array where 1st order differential microphones are placed on an open sphere,” Invention Disclosure Record (IDR), 10 2011.
- [Chr11b] —, “Design of regularized, radial equalizing filter, necessary for the design of a robust, modal beamformer with enhanced noise, respectively directivity properties,” Invention Disclosure Record (IDR), 10 2011.
- [Chr11c] —, “Methods for the reduction of spatial aliasing for modal beamformer,” Invention Disclosure Record (IDR), 10 2011.
- [Chr11d] —, “Room in a room - a neglected concept for auralization,” in *5th Biennial Workshop on DSP for In-Vehicle Systems, Kiel, Germany, 2011*, 9 2011.
- [Chr12a] —, “Concepts for the creation of individual sound zones in automobiles,” Invention Disclosure Record (IDR), 4 2012.
- [Chr12b] —, “Creation of individual sound zones by utilization of directional loudspeakers,” Invention Disclosure Record (IDR), 4 2012.
- [Chr12c] —, “Inverse filter design,” Invention Disclosure Record (IDR), 4 2012.
- [Chr12d] —, “A modal beamforming concept for the auralization of car interiors,” in *6th Annual Workshop on Acoustics, Deggendorf, Germany, December, 2012*, 12 2012.
- [Chr13] —, *Smart Mobile In-Vehicle Systems Next Generation Advancements*. Springer, Berlin/Heidelberg, 2013, Kap. Room in a Room: A Neglected Concept for Auralization, S. 181–235.
- [CS08] M. Christoph und L. Scholz, “Akustisch korrekter Entwurf der unteren Cross-Over-Filter,” Internal document, 1 2008.
- [CS10] —, “Untersuchung der Einflussnahme des menschlichen Körpers auf die Lokalisierung,” Internal document, 6 2010.
- [Dan00] J. Daniel, “Acoustic field representation, application to the transmission and the reproduction of complex sound environments in a multimedia context (english translation),” Dissertation, Université Paris 6, 9 2000.

- [Online]. Available: http://gyronymo.free.fr/audio3D/download_Thesis_PwPt.html#Thesis_download
- [Dan03] —, “Spatial Sound Encoding Including Near Field Effect: Introducing Distance Coding Filters and a Viable, New Ambisonic Format,” *Audio Engineering Society Conference: 23rd International Conference: Signal Processing in Audio Recording and Reproduction*, 5 2003. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=12321>
- [DE08] J. Daniel und N. Epain, “Improving Spherical Microphone Arrays,” in *Audio Engineering Society Convention 124*, 5 2008. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=14609>
- [DMN03] J. Daniel, S. Moreau und R. Nicol, “Further Investigations of High-Order Ambisonics and Wavefield Synthesis for Holophonic Sound Imaging,” *Audio Engineering Society Convention 114*, 3 2003. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=12567>
- [Dör98] M. Dörbecker, “Mehrkanalige Signalverarbeitung zur Verbesserung akustisch gestörter Sprachsignale am Beispiel elektronischer Hörhilfen,” Dissertation, IND, RWTH Aachen, 9 1998.
- [DRP98] J. Daniel, J.-B. Rault und J.-D. Polack, “Ambisonics Encoding of Other Audio Formats for Multiple Listening Conditions,” *Audio Engineering Society Convention 105*, 9 1998. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=8385>
- [EKM03] G. W. Elko, R. A. Kubli und J. M. Meyer, “Audio system based on at least second-order eigenbeams,” Patent application: US 0147539A1, 2003.
- [Elk00] G. W. Elko, *Acoustic Signal Processing for Telecommunication*, Reihe The Kluwer International Series in Engineering and Computer Science. Kluwer Academic Publishers, 2000, Kap. Superdirectional Microphone Arrays, S. 181–235.
- [Ell01] S. Elliott, *Signal processing for active control*. Academic Press, 2001.
- [EM08] G. W. Elko und J. Meyer, “Handling spatial aliasing in spherical array applications,” in *Hands-Free Speech Communication and Microphone Arrays, 2008. HSCMA 2008*, 5 2008, S. 1–4.
- [Far08] A. Farina, “Impulse response measurements by exponential sine sweeps,” *Audio Precision Workshop: Casa dellaMusica, Parma, Italy*, 10 2008.

- [FATG01] A. Farina, E. Armelloni, A. Torger und R. Glasgal, "Ambiophonic principles for the recording and reproduction of surround sound for music," *Audio Engineering Society Conference: 19th International Conference*, 6 2001.
- [FBA00] A. Farina, A. Bellini und E. Armelloni, "Implementation of cross-talk canceling filters with warped structures - subjective evaluation of the loudspeaker reproduction of stereo recordings," in *Proc. of SHARC2000*, Boston, (USA), 9 2000.
- [FM12] S. E. Favrot und M. Marschall, *Metrics for performance assessment of mixed-order Ambisonics spherical microphone arrays*, 2012.
- [FU97] A. Farina und E. Ugolotti, "Spatial equalisation of sound systems in cars by digital inverse filtering," in *Proc. of Reproduced Sound 13, Windermere (GB)*, 1997, S. 23–26.
- [FU98] —, "Spatial equalization of sound systems in cars," in *Audio Engineering Society Conference: 15th International Conference*, 10 1998. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=8099>
- [FU99] —, "Subjective comparison between stereo dipole and 3d ambisonics surround systems for automotive applications," in *16th AES Conference*, Bd. 1214. Rovaniemi, (Finland): Audio Engineering Society (AES), 1999, S. 1–12.
- [FZS08] M. Frank, F. Zotter und A. Sontacchi, "Localization experiments using different 2D ambisonics decoders," 2008.
- [Gar00] J. Garas, *Adaptive 3D Sound Systems*. Kluwer Academic Publishers, 2000.
- [Ger74] M. A. Gerzon, "Sound reproduction systems," Patent application: US 1494751, 1974.
- [Ger92a] —, "General metatheory of auditory localisation," in *Audio Engineering Society Convention 92*, 3 1992. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=6827>
- [Ger92b] —, "Optimum reproduction matrices for multispeaker stereo," *Journal of the Audio Engineering Society (JAES)*, Bd. 40, S. 571–589, 1992.
- [GL96] G. H. Golub und C. F. Loan, *Matrix Computations*. The Johns Hopkins University Press, Baltimore/London, 1996.
- [Gri09] D. Griesinger, "Frequency Response Adaptation in Binaural Hearing," in *Audio Engineering Society Convention 126*, 5 2009. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=14964>

- [Hay96] S. O. Haykin, *Adaptive Filter Theory (3rd Edition)*, 3rd Ausg. Prentice Hall, 1996.
- [Hes93] W. Hess, *Digitale Filter: Eine Einfuehrung*, Reihe Teubner-Studienbücher. Teubner, 1993. [Online]. Available: <http://books.google.de/books?id=LJwkAAAACAAJ>
- [HKP⁺99] U. Horbach, A. Karamustafaoglu, R. Pellegrini, P. Mackensen und G. Theile, "Design and Applications of a Data-Based Auralization System for Surround Sound," in *Audio Engineering Society Convention 106*, 5 1999. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=8204>
- [Hol06] F. Hollerweger, "Periphonic sound spatialization in multi-user virtual environments," Master's thesis, Technische Universität Graz, Center for Research in Electronic Art Technology at the University of California Santa Barbara, 3 2006.
- [Hul04] E. Hulsebosm, "Auralization using wave field synthesis," Dissertation, Technische Universiteit Delft, Niederlande, 10 2004.
- [IIE] IIEF, "Consulting zu wissenschaftlich-technischen vorhaben." [Online]. Available: <http://www.iief.de/sites/lconsult.html>
- [JMKM11] J. O. Jungmann, R. Mazur, M. Kallinger und A. Mertins, "Robust combined crosstalk cancellation and listening-room compensation," in *Proc. IE-EE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2011)*, New Paltz, New York, USA, 10 2011, S. 9–12.
- [Kim03] S.-M. Kim, "Authentic reproduction of stereo sound - a wiener filter approach," in *Audio Engineering Society Convention 115*, 10 2003. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=12462>
- [KNHOB98] O. Kirkeby, P. A. Nelson, H. Hamada und F. Orduna-Bustamante, "Fast deconvolution of multichannel systems using regularization," *Speech and Audio Processing, IEEE Transactions on*, Bd. 6, Nr. 2, S. 189–194, 3 1998.
- [KP06] M. Karjalainen und T. Paatero, "Equalization of audio systems using kautz filters with log-like frequency resolution," in *Audio Engineering Society Convention 120*, 5 2006. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=13571>
- [LD07] Z. Li und R. Duraiswami, "Flexible and optimal design of spherical microphone arrays for beamforming," *Audio, Speech, and Language Processing, IEEE Transactions on*, Bd. 15, Nr. 2, S. 702–714, 2 2007.

- [LP99] J. Lekkala und M. Paaajanen, “Emfi-new electret material for sensors and actuators,” in *Electrets, 1999. ISE 10. Proceedings. 10th International Symposium on*, 1999, S. 743–746.
- [LPY08] J. Lee, Y.-C. Park und D.-H. Youn, “Robust crosstalk cancellation based on energy-based control,” in *Audio Engineering Society Conference: 34th International Conference: New Trends in Audio for Mobile and Handheld Devices*, 8 2008. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=14423>
- [MA03] J. Meyer und T. Agnello, “Spherical microphone array for spatial sound recording,” in *Audio Engineering Society Convention 115*, 10 2003. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=12353>
- [Mal92] D. Malham, “Experience with large area 3-d ambisonic sound systems,” *Proceedings of the Institute of Acoustics*, 1992.
- [MD04] S. Moreau und J. Daniel, “Study of higher order ambisonic microphone,” *7ème Congrès Français d’Acoustique Joint congress CFADAGA04*, 2004.
- [ME02] J. Meyer und G. Elko, “A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield,” in *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*, Bd. 2, 5 2002, S. II–1781 –II–1784.
- [ME04] J. Meyer und G. W. Elko, *Audio Signal Processing for Next-Generation Multimedia Communication Systems*. Kluwer, 1 2004, Kap. Spherical Microphone Arrays for 3D Sound Recording, S. 67–89.
- [ME08] J. Meyer und G. Elko, “Spherical harmonic modal beamforming for an augmented circular microphone array,” in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, 4 2008, S. 5280–5283.
- [Mey01] J. Meyer, “Beamforming for a circular microphone array mounted on spherically shaped objects,” *Journal of the Acoustical Society of America*, Bd. 109, S. 185–193, 1 2001. [Online]. Available: <http://tubiblio.ulb.tu-darmstadt.de/17344/>
- [MFV11] B. Masiero, J. Fels und M. Vorländer, “Review of the crosstalk cancellation filter technique,” in *Proc. of ICOSA 2011*, 11 2011.
- [Mös10] M. Möser, *Acoustic Signal Processing for Telecommunication*. Springer, Berlin/Heidelberg, 2010.

- [Mou94] J. N. Mourjopoulos, "Digital equalization of room acoustics," *J. Audio Eng. Soc.*, Bd. 42, Nr. 11, S. 884–900, 1994. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=6921>
- [NA79] S. T. Neely und J. B. Allen, "Invertibility of a room impulse response," *Journal of the Acoustical Society of America*, Bd. 66, Nr. 1, S. 165–169, 1979.
- [Neu] M. Neukom, "Ambisonics equivalent panning AEP." [Online]. Available: <http://www.icst.net>
- [Neu07] —, "Ambisonic panning," *Audio Engineering Society Convention 123*, 10 2007. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=14354>
- [Ple09] P. Plessas, "Rigid sphere microphone arrays for spatial recording and holography," Master Thesis, Graz University of Technology, Graz, Austria, 11 2009.
- [PMV11] S. Pelzer, B. Masiero und M. Vorländer, "3d reproduction of room acoustics using a hybrid system of combined crosstalk cancellation and ambisonics playback," in *Proc. of the International Conference on Spatial Audio*, Detmold, 2011.
- [Pol96] M. A. Poletti, "The Design of Encoding Functions for Stereophonic and Polyphonic Sound Systems," *J. Audio Eng. Soc.*, Bd. 44, Nr. 11, S. 948–963, 11 1996. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=7879>
- [Pol05a] —, "Effect of Noise and Transducer Variability on the Performance of Circular Microphone Arrays," *J. Audio Eng. Soc.*, Bd. 53, Nr. 5, S. 371–384, 2005. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=13417>
- [Pol05b] —, "Three-Dimensional Surround Sound Systems Based on Spherical Harmonics," *J. Audio Eng. Soc.*, Bd. 53, Nr. 11, S. 1004–1025, 2005. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=13396>
- [PR10] Y. L. Parodi und P. Rubak, "Objective evaluation of the sweet spot size in spatial sound reproduction using elevated loudspeakers." *J Acoust Soc Am*, Bd. 128, Nr. 3, S. 1045–55, 2010. [Online]. Available: <http://www.biomedsearch.com/nih/Objective-evaluation-sweet-spot-size/20815441.html>
- [Raf05] B. Rafaely, "Analysis and design of spherical microphone arrays," *Speech and Audio Processing, IEEE Transactions on*, Bd. 13, Nr. 1, S. 135–143, 1 2005.

- [RWB07] B. Rafaely, B. Weiss und E. Bachmat, "Spatial aliasing in spherical microphone arrays," *Signal Processing, IEEE Transactions on*, Bd. 55, Nr. 3, S. 1003–1010, 3 2007.
- [SH06] G. Schmidt und T. Haulick, "Signal processing for in-car communication systems," *Signal Processing*, Bd. 86, Nr. 6, S. 1307 – 1326, 2006. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0165168405003300>
- [Sof02] O. Soffke, "Simulation, Bewertung und Entwicklung von Algorithmen zur raumentzerrten Wiedergabe in Fahrgastzellen," Master Thesis, Technische Universität Darmstadt, Institut für Datentechnik, Fachgebiet Rechnersysteme, Darmstadt, Germany, 7 2002.
- [Son03] A. Sontacchi, "Dreidimensionale Schallfeldreproduktion für Lautsprecher- und Kopfhöreranwendungen," Dissertation, Technische Universität Graz, 4 2003.
- [Son06] —, "Kugelflächenfunktionen Ambisonic Höherer Ordnung & Near-Field-Compensated Higher Order Ambisonic," Technische Universität Graz, Tech. Rep., 2006.
- [Sou12] Soundfield, "Benefits of a soundfield system," 2012. [Online]. Available: <http://www.soundfield.com/soundfield/soundfield.php>
- [SYS09] H. Sun, S. Yan und P. U. Svensson, "Robust spherical microphone array beamforming with multi-beam-multi-null steering, and sidelobe control," in *Applications of Signal Processing to Audio and Acoustics, 2009. WASPAA '09. IEEE Workshop on*, 10 2009, S. 113–116.
- [SYS11] H. Sun, S. Yan und U. P. Svensson, "Robust minimum sidelobe beamforming for spherical microphone arrays," *Audio, Speech, and Language Processing, IEEE Transactions on*, Bd. 19, Nr. 4, S. 1045–1051, 5 2011.
- [Teu05] H. Teutsch, "Wavefield decomposition using microphone arrays and its application to acoustic scene analysis," Dissertation, Friedrich-Alexander-Universität Erlangen-Nürnberg, 2005.
- [Too08] F. E. Tool, *Sound Reproduction: Loudspeakers and Rooms*. Elsevier, Oxford, UK, 2008.
- [TTN01] T. Takeuchi, M. Teschl und P. Nelson, "Subjective evaluation of the optimal source distribution system for virtual acoustic imaging," in *Audio Engineering Society Conference: 19th International Conference: Surround Sound - Techniques, Technology, and Perception*, 6 2001. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=10065>

- [Vai93] P. Vaidyanathan, *Multirate systems and filter banks*, Reihe Prentice-Hall signal processing series. Prentice Hall, 1993. [Online]. Available: <http://books.google.lu/books?id=pAsfAQAAIAAJ>
- [Wig04] B. Wiggins, “An investigation into the real-time manipulation and control of three-dimensional sound fields,” Dissertation, University of Derby, 2004.
- [WS85] B. Widrow und S. D. Stearns, *Adaptive signal processing*, Reihe Prentice-Hall signal processing series. Prentice-Hall, 1985. [Online]. Available: <http://books.google.ie/books?id=X74QAQAAMAAJ>
- [YSM⁺11] S. Yan, H. Sun, X. Ma, U. P. Svensson und C. Hou, “Time-domain implementation of broadband beamformer in spherical harmonics domain,” *Audio, Speech, and Language Processing, IEEE Transactions on*, Bd. 19, Nr. 5, S. 1221–1230, 7 2011.
- [YSS⁺11] S. Yan, H. Sun, P. U. Svensson, X. Ma und J. M. Hovem, “Optimal modal beamforming for spherical microphone arrays,” *IEEE Transactions on Audio, Speech & Language Processing*, Bd. 19, Nr. 2, S. 361–371, 2011.
- [ZN07] F. Zotter und M. Noisterning, “Near- and farfield beamforming using spherical loudspeaker arrays,” in *Proceedings of the 3rd Congress of the Alps Adria Acoustics Association, Graz, Austria, September 27-28, 2007*. Graz, Austria: DEGA, 2007.
- [Zot09] F. Zotter, “Analysis and Synthesis of Sound-Radiation with Spherical Arrays,” Dissertation, Institute of Electronic Music and Acoustics, University of Music and Performing Arts (IEM), Graz, Austria, 9 2009.
- [Zot12] —, “Holografie für Musikinstrumente,” in *Proc. DAGA 2012*. Darmstadt: DEGA, 2012.
- [ZPS] F. Zotter, H. Pomberger und A. Schmeder.
- [ZPS08] —, “Efficient directivity pattern control for spherical loudspeaker arrays,” *J Acoust Soc Am*, Bd. 123, Nr. 5, S. 3643, 2008.
- [ZR07] L. Zhiyun und D. Ramani, “Flexible and optimal design of spherical microphone arrays for beamforming,” *IEEE Transactions on Audio, Speech & Language Processing*, Bd. 15, Nr. 2, S. 702–714, 2007.

Index

- Fractional delay*, 163
- Modelling delay*, 5
- Room Impulse Response (RIR)*, 5
- Temporal diffusion*, 10
- Beamformer*
 - Filter-and-sum*, 223
 - Abweichungen in der Übertragungsfunktion der Sensoren, 121
 - Array-Gewinn, 128
 - Bündelungsmaß, 128
 - Modal, 119, 215
 - Positionierungsfehler der Sensoren, 121
 - Regularisierung, 121
 - Suszeptibilität, 125
- Binaural Room Scanning-Verfahren (BRS)*, 1
- Crosstalk cancellation (CTC)*, 5
- Head tracker*, 33
- Higher Order Ambisonic-Verfahren (HOA)*, 1
- Interaural Level Difference (ILD)*, 10, 67
- Interaural Time Difference (ITD)*, 10
- Least Mean Square (LMS)*, 85
- Least Square (LS)*, 121
- Localization blurr*, 83
- Motion tracker*, 34
- Multiple Error Least Mean Square (MELMS)*, 5
- Optimal source distribution (OSD)*, 9
- Signal to Noise Ratio (SNR)*, 125
- Sweet spot*, 9
- Virtual Studio Technology (VST)*, 109
- Abtastwert, 163
- Adaptives Filter, 5
- Akustische Dokumentation, 191
- Allgemeine Fourier-Bessel-Reihe, 93
- Ambisonic, 37
 - Max r_v -Decoder, 74
 - Higher Order Ambisonic (HOA)*, 54
 - Re-encoding-Matrix*, 69
 - Basic-Decoder*, 74
 - Distance coding*, 89
 - Distance coding-Filter*, 89
 - Matching condition*, 69
 - Near Field Compensated - Higher Order Ambisonic (NFC-HOA)*, 90
 - Panning-Funktion*, 74
 - Re-encoding-Vorschrift*, 69
 - Truncation error*, 80
 - Windows-applied-Decoder*, 83
- Bass-Boost-Effekt, 66
- Codierprozess, 62
- Decodervarianten, 74
- Decodierprozess, 62
- Decodiervorschrift, 69
- Fensterbasierter Decoderentwurf, 51
- In-Phase-Decoder, 50, 74
- Klangfeldanalyse, 62
- Kompensation des Nahfeldeffektes, 67
- Nahfeldeffekt, 66
- Nahfeldmodellierung, 66
- Orthonormalitätsbedingung, 100
- Räumliche Stabilität, 77
- Radiale Funktion, 98
- Richtwirkung, 43
- Super-Minimal-Decoder, 83
- Ambisonic-Verfahren 1. Ordnung, 1

- Ambisonickanäle, 63
- Ambisonicsignale, 63
- Auralisation, 191
- Auralisierung, 57
- B-Format-Signale, 38
- Bühnenabstand, 64
- Bodediagramme der optimierten *equalizing*-Filter $EQ_m(\omega)$, 138
- Bodediagramme der optimierten Regularisierungsfiler $T_m(\omega)$, 138
- Breitengrad, 157
- Cepstralbereich, 17
- Cepstrum
 - Reell, 17
- CTC-Filter, 6
- Dispersion, 204
- Ebene Welle, 37
- Eigen-*beamformer*, 120
- Eigenbeams, 119
- Eigenrauschen der Sensoren, 121
- Eigenwerte des Wellenfeldes, 119
- Equalizing, 5
- Eulersche Gleichung, 105
- Fernfeldbedingung, 64
- Filter
 - Shelving*, 47
 - Biquadratisch, 30
 - Biquads, 30
 - Idealer Interpolator, 163
 - Idealer Tiefpass, 163
 - Kuhschwanz, 47
 - Linearphasig, 29
 - Nullphasig, 138
 - Optimal, 15
 - Wiener, 15
- Filterentwurf
 - Frequency sampling method*, 29
 - Frequenzabtastverfahren, 29
- Flächenwinkel, 192
- Fourier-Bessel-Reihe, 55, 66
- Fraktionelle Verzögerung, 163
- Gehrungswinkel, 192
- Gibbssches Phänomen, 164
- Glättung
 - Nichtlinear, 21
- Greensche Funktion, 65
- Hankelfunktion, 209
- Hauptkeule, 195
- Helmholtz-Gleichung, 55
- HRTF, 34
- Interferenzbreite, 79
- Interpolation
 - Polynom, 220
 - Spline, 220
- inverse Filter, 5
- Inverse Filter-Verfahren, 1
- Jacobimatrix, 74
- Kammfiltereffekt, 10
- Kirchhoff-Helmholtz-Integral, 96
- Klangfeldsynthese, 62
- Kopfbezogene Übertragungsfunktion (HRTF), 34
- Kronecker-Delta-Funktion, 25
- Kugelflächenfunktion, 37
- Kugelwelle, 65
- Längengrad, 157
- Laplacescher Entwicklungssatz, 206
- Latenz, 66
- Lautsprecher
 - 2-Wege-System, 19
 - Koax, 19
 - Koaxial, 18
- Legendrefunktion, 59
 - Schmidt-semi-normalisiert, 59
- Lokalisationsschärfe, 50, 54, 205

- Lokalisationsunschärfe, 82
- Matrix
- Ähnlichkeitsbedingung, 208
 - Adjungiert, 206
 - Diagonalisierbar, 208
 - Hermitisch, 207
 - Inverse, 206
 - Kofaktor, 206
 - Kofaktoren, 207
 - Komplementär, 207
 - Minor, 206
 - Moore-Penrose-Inverse, 70
 - Normal, 207
 - Pseudoinverse, 8, 70
 - Selbstadjungiert, 207
 - Symmetrisch, 207
 - Untermatrix, 206
- Mikrophon
- Öffnungswinkel, 43
 - Soundfield*, 53
 - Endfire*-Ausrichtung, 95
 - Patch*, 145
 - Shotgun*, 223
 - Broadside*-Ausrichtung, 95
 - Lineares Array, 95
 - Zirkulares Array, 95
 - Achtcharakteristik, 38
 - Array-Gewinn, 197
 - Bidirektional, 38
 - Differentiell, 38, 54
 - Dipol, 38, 197
 - Druckgradientenempfänger, 196
 - Eigenrauschen, 115
 - Elektret, 146
 - Geschlossene Kugel, 101
 - Gradientenmikrophon, 195
 - Hyperkardioide, 196
 - Hypernierencharakteristik, 197
 - Kardioide, 196
 - Klangfeld, 53
 - Kreuzförmiges Array, 95
 - Kugelcharakteristik, 39
 - Kugelförmiges Array, 97
 - Nierencharakteristik, 197
 - Offene Kugel, 100
 - Onmidirektional, 39
 - Referenz, 150
 - Richtdiagramm, 199
 - Richtmikrophon, 195
 - Schalldruck, 39
 - Sphärisches Array, 97
 - Superkardioide, 197
 - Supernierencharakteristik, 197
 - Virtuell, 45
 - Vor/Rück-Verhältnis, 197
- Minimalphasiger Anteil, 16, 27
- Moden des Wellenfeldes, 119
- Moore-Penrose-Bedingungen, 70
- Nahfeldkompensation, 3
- Neumannfunktion, 209
- Optimalen Lautsprecherverteilung, 9
- Personal Computer (PC), 109
- Platonische Körper, 71
- Polyeder
- Archimedische Körper, 231, 232
 - Buckyball, 246
 - Catalanische Körper, 231, 246
 - Dualkörper, 231
 - Eulerscher Polyedersatz, 231
 - Hexaeder, 155
 - Ikosaederstumpf, 155
 - Kubus, 155
 - Platonische Körper, 231
 - Quasi-Regulär, 232
 - Regulär, 231
- Punktquelle, 66
- Quantisierungsfehler, 30
- Räumliche Tiefpassfilterung, 145

- Radialfunktion, 56
- Raumentzerrung, 5
- Referenzlautsprecher, 151
- Referenzraum, 191
- Reflexionsarmer Raum, 156
- Regelmäßiges Polygon, 71
- Reguläre Polyeder, 71
- Regularisierung, 11
- Richtfaktor
 - Frequenzabhängig, 50

- Sample, 163
- Schallenergievektorthorie, 201
- Schallintensität, 202
- Schallintensitätsvektor, 202
- Schallschnellevektorthorie, 201
- Spektrale Faktorisierung, 16
- Sphärische Besselfunktion 1. Art, 56
- Sphärische Harmonische, 37
- Streufeld, 105
- Suszeptibilität, 200
- Sweet spot, 2, 50

- Tikhonov-Filter, 122
- Transaural-Stereo, 6
- Transaurale Wiedergabe, 6
- Transformationsmatrix
 - Raum-Zeit- zu Wellenfeldbereich, 120

- Virtueller Raum, 16
- Virtuelles Tuning, 191

- Wellenfeld, 50
- Wellenfeldsynthese (WFS), 95
- Wellengleichung, 55
- Wronskian-Beziehung, 107

Bisher sind in der Reihe

Arbeiten über digitale Signalverarbeitung und Systemtheorie

folgende Bände erschienen:

- Nr. 1 Markus Christoph
 Untersuchung verschiedener Verfahren zur messtechnischen Bestimmung und Nach-
 bildung der Akustik insbesondere von Fahrzeugaudiosystemen, 2013