Technische Universität Berlin
Fakultät I - Geistes- und Bildungswissenschaften
Institut für Sprache und Kommunikation
Fachgebiet Audiokommunikation

Zur Rekonstruktion gerichteter Schallquellen mittels Monopolsynthese

Masterarbeit für die Prüfung zum Master of Science
im Studiengang Audiokommunikation und -technologie

vorgelegt von: Johannes Redlich
[redacted]

1. Gutachter: Prof. Dr. Stefan Weinzierl
FG Audiokommunikation
Technische Universität Berlin

2. Gutachter: Jakob Bergner
Fraunhofer IDMT Ilmenau

eingereicht am: 2. November 2017

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt gegenüber der Fakultät I der Technischen Universität Berlin, dass die vorliegende, dieser Erklärung angefügte Arbeit selbstständig und nur unter Zuhilfenahme der im Literaturverzeichnis genannten Quellen und Hilfsmittel angefertigt wurde. Alle Stellen der Arbeit, die anderen Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind kenntlich gemacht. Ich reiche die Arbeit erstmals als Prüfungsleistung ein. Ich versichere, dass diese Arbeit oder wesentliche Teile dieser Arbeit nicht bereits dem Leistungserwerb in einer anderen Lehrveranstaltung zugrunde lagen.

Titel der schriftlichen Arbeit

Zur Rekonstruktion gerichteter Schallquellen mittels Monopolsynthese

Verfasser

Redlich, Johannes [redacted]

Betreuender Dozent

Prof. Dr. Stefan Weinzierl

Mit meiner Unterschrift bestätige ich, dass ich über fachübliche Zitierregeln unterrichtet worden bin und verstanden habe. Die im betroffenen Fachgebiet üblichen Zitiervorschriften sind eingehalten worden.

Eine Überprüfung der Arbeit auf Plagiate mithilfe elektronischer Hilfsmittel darf vorgenommen werden.

Ort,Datum Unterschrift

Zusammenfassung

Diese Arbeit beschäftigt sich mit der Rekonstruktion von gerichteten Schallquellen mittels Monopolsynthese. Das Verfahren der Monopolsynthese beschreibt die Berechnung eines Schallfeldes aus der Überlagerung von Monopolen, welche wiederum ungerichtete Schallquellen sind. Das Interesse liegt dabei in der Wiedergabe von gerichteten Schallquellen mit 3D-Audio-Systemen. Dazu muss das abgegebene Schallfeld zunächst virtuell berechnet werden. Diese Berechnung ist am einfachsten durch die Überlagerung von Monopolen möglich.

Um nun eine gerichtete Schallquelle korrekt wiederzugeben, muss eine Anordnung und Ansteuerung von Monopolen gefunden werden, durch welche das Schallfeld der gerichteten Schallquelle am besten rekonstruiert werden kann. Die bisher bekannten Verfahren ermöglichen eine Rekonstruktion der Abstrahlcharakteristik der gerichteten Schallquelle jedoch nur für eine Frequenz und sind eher für simulatorische Zwecke zu verwenden. Mit diesen Verfahren ist eine Wiedergabe von gerichteten Schallquellen mit einer bestimmten Bandbreite in der Regel nicht möglich.

Daher werden Verfahren gesucht, welche automatisch eine optimierte Anordnung und Ansteuerung der Monopole ermitteln. Zu den untersuchten Verfahren, um die Anordnung der Monopole zu finden, zählen ein in Zusammenarbeit mit dem Fraunhofer Institut für Digitale Medientechnologie entwickelter generischer Ansatz der Anordnung von Monopolen auf Spiralen und ein Verfahren nach Giron (1996), welches erweitert wurde, um die Anordnung der Monopole für einen möglichst großen Frequenzbereich zu berechnen. Zur Ermittlung der Ansteuerung der Monopole werden die Tikhonov-Regularisierung und die LSQR-Regularisierung verwendet. Die untersuchten Verfahren werden anschließend in einem qualitativen, psychoakustischen Hörtest evaluiert.

Die Arbeit bietet somit Ansätze, um gerichtete Schallquellen wiederzugeben. Die untersuchten Verfahren müssen jedoch noch weiter entwickelt werden, um eine bessere Plausibilität zu erreichen und Abweichungen der rekonstruierten Abstrahlcharakteristik zur originalen Abstrahlcharakteristik zu minimieren.

Inhaltsverzeichnis

1. Einleitung und Fragestellung **1**

1.1. Stand der Forschung 4

1.2. Vorarbeiten 4

2. Grundlagen **5**

2.1. Konventionen 5

2.2. Monopolsynthese 7

2.3. Sphärische Harmonische Zerlegung 9

2.4. Inkorrekt gestellte Probleme und Regularisierungsmethoden 11

2.5. Prinzipieller Ablauf der Rekonstruktion von gerichteten Schallquellen mittels Monopolsynthese 11

2.6. Auswahl der Verfahren 13

3. Methoden zur Anordnung der Monopole **14**

3.1. Spiralanordnung der Monopole 15

3.2. Monopoldekomposition nach Giron 17

4. Ermittlung der Gewichte der Monopole **20**

4.1. Mit Tikhonov-Regularisierung 21

4.2. Mit LSQR-Regularisierung 22

4.3. Regularisierungsparameter 22

5. Implementierung der Verfahren mit MatLab **24**

5.1. Einlesen der originalen Abstrahlcharakteristik 24

5.2. Spiralanordnung der Monopole 25

5.3. Monopoldekomposition nach Giron 26

5.4. Regularisierungstoolbox von Hansen (2007) 30

6. Hörversuch **38**

6.1. Design 39

6.2. Aufbau 41

6.3. Restriktionen des Wiedergabesystems 44
6.4. Versuchsdurchführung 45
6.5. Ergebnisse 45
6.6. Auswertung und Verbesserungsvorschläge 47

7. Einschätzung der verwendeten Verfahren 48
7.1. Spiralanordnung der Monopole 48
7.2. Monopoldekomposition nach Giron 49
7.3. Tikhonov-Regularisierung mit Regularisierungsparameter nach L-Kurven Verfahren 49
7.4. LSQR-Regularisierung 49

8. Fazit und Ausblick 50

Abbildungsverzeichnis 52

Tabellenverzeichnis 52

Abkürzungsverzeichnis 54

Glossar 54

Normen und Standards 55

Literatur 55

Anhang A. Verwendete MatLab Funktionen vom Fraunhofer IDMT 58

Anhang B. Ermittelung der lokalen Maxima 59

Anhang C. Monopole zusammenfassen 61

Anhang D. Versuchsanleitung 62

1. Einleitung und Fragestellung

Bei der Entwicklung von Geräten und Maschinen ist nicht zuletzt auch der abgestrahlte Schall ein wichtiges Bewertungskriterium der Qualität des Gerätes oder der Maschine. Dabei spielt nicht nur die Lautstärke, sondern auch die Schallemission in bestimmte Richtungen eine wichtige Rolle. Der abgestrahlte Schall wird in der Regel an einem Prototyp oder an einem Modell gemessen und beurteilt. Durch eine Beurteilung des abgestrahlten Schalls schon während der Konstruktionsphase, könnten Kosten und Ressourcen eingespart werden.

Der abgestrahlte Schall ist normalerweise, wie bei den meisten realen Schallquellen, gerichtet. Das bedeutet, dass in einer gleichbleibenden Entfernung zur Schallquelle, aber in unterschiedlichen Richtungen, eine unterschiedliche Lautstärke und eine Veränderung des Klanges wahrgenommen wird. Diese wahrgenommenen Veränderungen entstehen durch einen unterschiedlichen frequenz- und zeitabhängigen Druck an den Abtastpositionen (z.B. Ohren, Mikrofone). Dieser Druck wird an den jeweiligen Abtastpositionen auf einer Kugeloberfläche gemessen. Der Druck an allen Abtastpositionen wird als Abstrahlcharakteristik bezeichnet und gibt die Richtwirkung der Schallquelle wieder. Gleichzeitig repräsentiert die Abstrahlcharakteristik das Schallfeld der Schallquelle an den Abtastpositionen. In anderen aktuellen Forschungsarbeiten wird an Verfahren gearbeitet, welche diese Abstrahlcharakteristik auch aus einer rechnergestützten Konstruktionszeichnung (3D-CAD-Zeichnung) der Geräte oder Maschinen mit einem dazugehörigem physikalischem Verhaltensmodel berechnen können. Eine dieser Forschungsarbeiten ist das Projekt AVP^3 des Fraunhofer Instituts für Digitale Medientechnologie (IDMT) (Bergner et al. (2016), Münch et al. (2017)).

Die gemessene oder berechnete Abstrahlcharakteristik soll nun so wiedergegeben werden, als wenn man die reale Schallquelle direkt hören würde. Das Fraunhofer IDMT forscht dazu zurzeit an der Umsetzung einer perzeptiv plausiblen Wiedergabe. Solch eine perzeptiv plausible Wiedergabe ermöglichen indirekt 3D-Audio-Systeme. Im Allgemeinen wird mit 3D-Audio-Systemen ein virtuell berechnetes Schallfeld wiedergegeben. Ein virtuelles Schallfeld kann durch die Überlagerung von ungerichteten Schallquellen (Monopolen) berechnet werden. Die Brechnung eines Schallfeldes

aus der Überlagerung von Monopolen wird Monopolsynthese genannt und erfolgt mit Monopolen, da diese mathematisch relativ einfach zu beschreiben sind. Das berechnete virtuelle Schallfeld kann dann mit Kopfhörern oder Lautsprechern durch verschiedenen Verfahren hörbar gemacht werden. Zu diesen Verfahren zählen unter anderen die Wellenfeld Synthese, Higher Order Ambisonics und die Binauralsynthese Slavik und Weinzierl (2008).

Da wie beschrieben die Abstrahlcharakteristik nicht direkt als Schallfeld synthetisiert und mit einem 3D Audio-System wiedergegeben werden kann, muss diese auf Monopole abgebildet werden. Die Zerlegung der Abstrahlcharakteristik in Monopole wird Monopolzerlegung (Monopoldekomposition) genannt und ist bisher nur für einzelne Frequenzen beschrieben (siehe Abschnitt 1.1). Da aber eine perzeptiv plausible Wiedergabe angestrebt wird und reale Schallquellen Schall über ein breites Frequenzspektrum abstrahlen, muss eine Anordnung und Ansteuerung der Monopole gefunden werden, welche für ein breites Frequenzspektrum funktioniert. Das ist die Aufgabe dieser Arbeit. Die ermittelten Monopole, mit ihren Positionen und Ansteuerungen, sollen zu dem mit einem existierenden 3D-Audio-System wiedergegeben werden können. Die bekannten 3D-Audio-Systeme unterliegen technischen und physikalischen Grenzen, welche bei der Brechung der Monopole beachtet werden müssen. Auf diese Grenzen wird später an den entsprechenden Stellen eingegangen.

Mit den in dieser Arbeit untersuchten und entwickelten Verfahren soll es möglich sein, eine ursprüngliche Abstrahlcharakteristik mit großer Bandbreite in Monopole zu zerlegen und diese mittels Monopolsynthese zu rekonstruieren. Das rekonstruierte Schallfeld soll dann mit einem existierenden 3D-Audiosystem wiedergegeben werden. Dabei geht es nicht primär um die physikalisch korrekte Reproduktion der Abstrahlcharakteristik, sondern vielmehr um die perzeptiv plausibelste Lösung.

Zukünftig könnten, nach erfolgreicher Entwicklung aller benötigten Verfahren und Komponenten, Geräte und Maschinen virtuell begutachtet werden (Abbildung 1.0.1). Auch Künstler könnten gerichtete Schallquellen, wie es zum Beispiel reale Instrumente sind, in ihren 3D-Audiokompositionen verwenden.

Abbildung 1.0.1: Mögliches Szenario der virtuellen Begutachtung einer Maschine mit Virtual Reality Brille und 3D-Audio-System, ©Fraunhofer IDMT

Die in dieser Arbeit untersuchten und entwickelten Verfahren werden mit MatLab von MathWorks® implementiert und berechnet. Die dafür verwendeten Schallquellen, wurden in einem reflektionsarmen Raum (RAR) aufgenommen. Die Verwendung von Quellen in reflektierenden Umgebungen oder von simulierten Quellen aus 3D-CAD-Zeichnungen wird in dieser Arbeit nicht betrachtet. Die Verwendung von Quellaufnahmen aus dem RAR mit relativ einfachen Abstrahlcharakteristiken ermöglicht eine gute Kontrolle über die Eigenschaften der Stimuli, welche zur Evaluierung der Verfahren zur Anordnung und Ansteuerung der Monopole verwendet werden. Zudem werden die verwendeten Quellen als lineare, zeitinvariante (LTI-)Systeme betrachtet.

Das Vorgehen zur Anordnung und Ansteuerung der Monopole, welches den plausibelsten Höreindruck erzeugt, soll durch einen psychoakustischen Hörversuch herausgefunden werden.

Durch die Ergebnisse der Berechnungen und des Hörversuchs können erste Aussagen über die verwendeten Verfahren und ihre Parameter im Bezug auf die Wiedergabe in einem existierenden 3D-Audiosystem gemacht werden.

1.1. Stand der Forschung

Die Wiedergabe von Schallquellen mit 3D-Audio-Systemen (z.B. Wellenfeldsynthese (WFS), Higher Order Ambisonics (HOA)) ist in der Literatur gut beschrieben (Slavik und Weinzierl (2008)). Häufig werden als Schallquellen Monopole eingesetzt, da diese relativ gut implementierbar und analytisch gut beschreibbar sind.

Als Ansatz, um ein Schallfeld aus einer Anordnung von Monopolen zu berechnen, ist die Monopolsynthese ein gut dokumentiertes Verfahren (Escolano et al. (2007); Müller-Trapet und Vorländer (2014); Ochmann (1990, 1995); Opdam et al. (2016)).

Zur Zerlegung der Abstrahlcharakteristik gibt es bisher einige Ansätze welche aber nur für einzelne Frequenzen verwendet werden.

In Ochmann (1990) findet man eine Beschreibung zur Anordnung der Ersatzquellen für die Erzeugung einer gewünschten Abstrahlcharakteristik, welche vorsieht, die Abstrahlgeometrie in möglichst kugelähnliche Substrukturen zu unterteilen. In den Schwerpunkt jeder Substruktur wird der Quellort $\mathbf{x}_\mathrm{i}$ gelegt.

In Opdam et al. (2016) ist die Anordnung in Form von archimedischen Spiralen beschrieben. Dabei soll durch eine logarithmische Anordnung der Monopole auf den Spiralen eine gleichmäßige Abstrahlung erzielt werden.

In Giron (1996) wird ein Verfahren zum Finden von Positionen und Amplituden von Monopolen beschrieben. Dieses Verfahren wird verfolgt, um einen Algorithmus zu entwickeln, welcher die Positionen und Amplituden der Monopole automatisch ermittelt.

Es gibt aber noch kein bekanntes automatisches Verfahren, welches ermittelt, wie Monopole angeordnet und angesteuert werden sollen, um einen möglichst plausiblen Höreindruck gegenüber der originalen Aufnahme der Quelle zu erhalten.

1.2. Vorarbeiten

Während der Recherche zur Masterarbeit ist ein Paper entstanden, welches die zugrunde liegende Problematik beschreibt (Bergner et al. (2017)).

2. Grundlagen

Zur Zerlegung von Abstrahlcharakteristiken in Monopole sowie zur Monopolsynthese sind einige mathematische und physikalische Grundlagen und Konventionen notwendig, welche in diesem Kapitel vorgestellt werden.

2.1. Konventionen

Für diese Arbeit werden folgende Konventionen festgelegt, welche in der Regel mit üblichen Konventionen der Mathematik oder der Akustik übereinstimmen.

Vektoren werden klein, gerade und fett geschrieben und Matrizen groß, gerade und fett. Die Indizes der Vektor- oder Matrix-Elemente werden klein und gerade geschrieben.

Vektor $\mathbf{v} = (v_1 \; v_2 \; ... \; v_\mathrm{i})$

Matrix $\mathbf{A} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{\mathrm{i}1} \\ a_{12} & a_{22} & \cdots & a_{\mathrm{i}2} \\ \vdots & \ddots & \vdots & \\ a_{1\mathrm{j}} & a_{2\mathrm{j}} & \cdots & a_{\mathrm{ij}} \end{pmatrix}$

Dabei gelten alle Rechenregeln für Vektoren und Matrizen im dreidimensionalen Raum (Papula (2009a), S.72 ff.). Dezimalbrüche werden mit einem Punkt „ . “ als Trennzeichen geschrieben, da in Vektoren auch das Komma „ , “ als Trennzeichen der Vektorelemente verwendet wird.

Um Positionen im Raum zu beschreiben, gibt es verschiedene Koordinatensysteme. Für diese Arbeit wird sowohl das kartesische als auch das sphärische Koordinatensystem verwendet. Eine Position kann als Vektor in kartesischen Koordinaten mit $\mathbf{x} = (x, y, z)^T$ beschrieben werden.

Die Umrechnung von kartesischen in sphärische Koordinaten wird für diese Arbeit

wie folgt definiert.

$$\mathbf{x} = (r, \varphi, \vartheta)$$

$$r = \sqrt{x^2 + y^2 + z^2} \qquad r \in [0, \infty) \qquad (2.1.1)$$

$$\varphi = arctan\left(\frac{y}{x}\right) \qquad \varphi \in [0, 2\pi) \qquad (2.1.2)$$

$$\vartheta = arccos\left(\frac{z}{r}\right) \qquad \vartheta \in [0, \pi] \qquad (2.1.3)$$

Die Winkel φ und ϑ sind im Bogenmaß angegeben und können in Winkel im Gradmaß wie folgt umgerechnet werden. ϕ sei ein Winkel im Bogenmaß und α der selbe Winkel im Gradmaß, so besteht die Beziehung $\alpha = \phi \cdot \frac{180°}{pi}$. Für ϕ kann φ oder ϑ im Bogenmaß eingesetzt werden. In dieser Arbeit werden Winkel im Bogenmaß zur Berechnung benutzt. Winkel im Gradmaß dienen in dieser Arbeit dem besseren Verständnis der Daten und Grafiken.

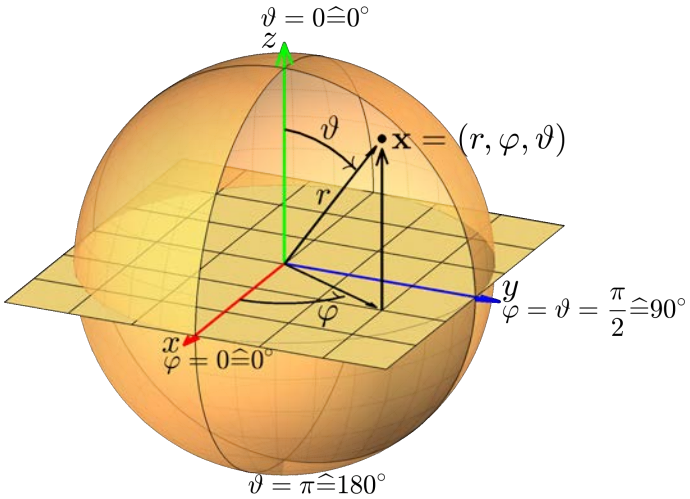

Abbildung 2.1.1: Koordinatensystem

Für einige Berechnung werden komplexen Zahlen in der Form $A = x + yi$ verwendet. Dabei ist i die imaginäre Einheit und ist mit $i = \sqrt{-1}$ definiert.

Die Kreisfrequenz ω der Frequenz f wird mit $\omega = 2\pi \cdot f$ berechnet. Als Schallgeschwindigkeit wird die Schallgeschwindigkeit in Luft bei 20°C Raumtemperatur mit $c = 343\frac{\mathrm{m}}{\mathrm{s}}$ festgelegt. Damit lässt sich die Wellenzahl mit $k = \frac{\omega}{c}$ berechnen.

Ein nicht periodisches Signal kann als gewichtete Überlagerung von Sinus- und Kosinusschwingungen aller Frequenzen betrachtet werden. Gleichung 2.1.4 beschreibt

diesen Zusammenhang und wird Fourier-Transformation $\mathcal{F}\{f(t)\} = F(\omega)$ genannt (Papula (2009b), S.539 ff.; Williams (1999)).

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}\mathrm{d}\omega \tag{2.1.4}$$

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{i\omega t}\mathrm{d}t \tag{2.1.5}$$

Um ein zeit- und wertdiskretes (digitales) Signal zu berechnen wird eine diskrete Fourier-Transformation durchgeführt, welche in der Fast-Fourier-Transformation verwendet wird ($\mathcal{FFT}$).

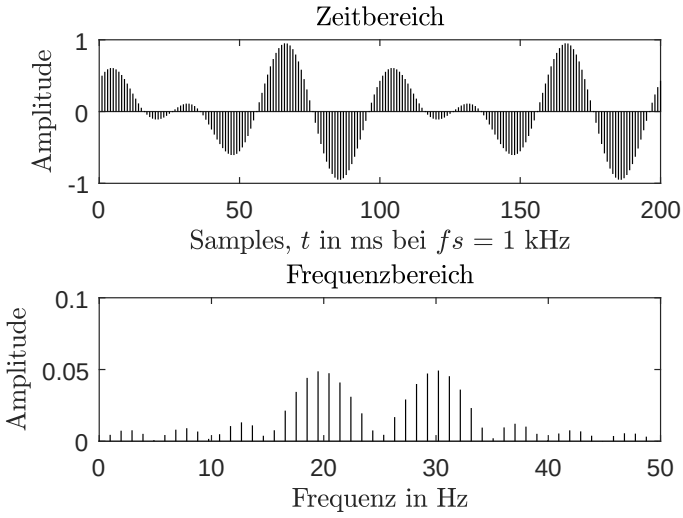

Abbildung 2.1.2: Fast-Fourier-Transformation mit 1024 Samples eines Signals bestehend aus der Überlagerung einer Sinusschwingung mit 20 Hz und einer Kosinusschwingung mit 30 Hz und einer Abtastrate fs von 1 kHz, im Frequenzbereich werden die Frequenzbins im relevanten Frequenzintervall dargestellt

2.2. Monopolsynthese

Die Monopolsynthese ist die Berechnung eines Schallfeldes aus der Überlagerung von gewichteten Monopolen. Die inverse Operation dazu ist die Monopolzerlegung (Monopoldekomposition), durch welche die Positionen und Gewichte der Monopole berechnet werden. Um das Schallfeld, das durch die gewichteten Monopole erzeugt wird, als Schalldruck an bestimmten Abtastposition im Raum abzubilden, wird eine

Transfermatrix zwischen den Monopolpositionen und den Abtastpositionen aufgestellt. Diese Transfermatrix ist keine unitäre Abbildung. Eine unitäre Abbildung, das heißt eine verlustfreie Operationskette von der Monopolzerlegung bis zur Monopolsynthese, ist nur bei idealen Problemstellungen möglich, die eine invertierbare Transfermatrix enthalten (siehe Abschnitt 2.4).

Durch die Überlagerung der gewichteten Monopole, bestehend aus den Quellgewichten $\mathbf{d} = (d_1 \ ... \ d_\mathrm{i})$ und den zugehörigen Monopolpositionen $\mathbf{x}_\mathrm{i}$, entsteht ein Schallfeld $\mathbf{p} = (p_1 \ ... \ p_\mathrm{j})$ an den Abtastpositionen $\mathbf{r}_\mathrm{j}$. Die Übertragung des Schalls von den Monopolen zu den Abtastpositionen wird durch die Matrix der Freifeldübertragungsfunktionen $\mathbf{G}$ beschrieben.

$$\mathbf{p} = \mathbf{G} \cdot \mathbf{d}. \tag{2.2.1}$$

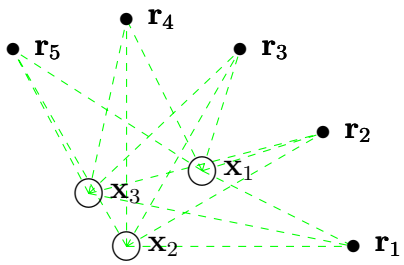

Abbildung 2.2.1: Prinzipielle Anordnung zur Monopolsynthese, $\mathbf{x}_\mathrm{i}$ - Monopolpositionen, $\mathbf{r}_\mathrm{j}$ - Abtastpositionen, (- -) - Übertragungspfade $G_{\mathrm{i,j}}$

Die Freifeldübertragungsfunktion ist durch die dreidimensionale Green'sche Funktion für Punktquellen definiert. Die $\mathbf{G}$-Matrix enthält die Green'schen Funktionen aller Übertragungspfade. Jede Green'sche Funktion $G_{\mathrm{i,j}}$ der $\mathbf{G}$-Matrix beschreibt die Schallübertragung zwischen einer Monopolposition $\mathbf{x}_\mathrm{i}$ und einer Abtastposition $\mathbf{r}_\mathrm{j}$.

$$\mathbf{G} = \begin{pmatrix} \frac{e^{-ikr_{1,1}}}{4\pi r_{1,1}} & \cdots & \frac{e^{-ikr_{\mathrm{i},1}}}{4\pi r_{\mathrm{i},1}} \\ \vdots & \ddots & \vdots \\ \frac{e^{-ikr_{1,\mathrm{j}}}}{4\pi r_{1,\mathrm{j}}} & \cdots & \frac{e^{-ikr_{\mathrm{i},\mathrm{j}}}}{4\pi r_{\mathrm{i},\mathrm{j}}} \end{pmatrix} \tag{2.2.2}$$

2.3. Sphärische Harmonische Zerlegung

Der abgestrahlte Schalldruck $p(r, \varphi, \vartheta)$ einer Schallquelle im Freifeld, gemessen auf einer Kugeloberfläche mit dem Radius r, kann vollständig durch das Produkt aus den Koeffizienten $\widetilde{A}_{\mathrm{p,q}}$, den zugehörigen Raummoden $Y_{\mathrm{p,q}}(\varphi, \vartheta)$ und der Hankelfunktion erster Art $h_{\mathrm{p}}^{(1)}(k \cdot r)$ beschrieben werden. Da bei einer Schallquelle der abgestrahlte Schall außerhalb der Quelle im Freifeld betrachtet wird, wird diese Problematik auch äußeres Problem genannt, welches durch die sphärische Hankelfunktion erster Art $h_{\mathrm{p}}^{(1)}(k \cdot r)$ beschrieben wird. Für diese Arbeit ist die Schreibweise $h_{\mathrm{p}}(kr)$ gleichbedeutend mit $h_{\mathrm{p}}^{(1)}(k \cdot r)$. Der Schalldruck außerhalb einer Schallquelle lässt sich somit mit der Gleichung (2.3.1) bestimmen (Williams (1999), Spherical Waves S.183 ff.).

$$p(r, \varphi, \vartheta) = \sum_{\mathrm{p}=0}^{\infty} \sum_{\mathrm{q}=-\mathrm{p}}^{\mathrm{p}} \widetilde{A}_{\mathrm{p,q}} \cdot h_{\mathrm{p}}(kr) \cdot Y_{\mathrm{p,q}}(\varphi, \vartheta) \tag{2.3.1}$$

Für die Ordnung p und den Grad q gilt bei der sphärisch harmonischen Transformation und ihrer Inversen immer p $\geq$ 0 und -p $\leq$ q $\leq$ p.

Bei der sphärischen harmonischen Transformation $\mathcal{SHT}$ wird der Schalldruck auf einer Kugeloberfläche in seine sphärischen harmonischen Bestandteile zerlegt. Auf diese Weise entsteht eine sphärisches Wellenspektrum $\widetilde{\mathbf{A}}$, ähnlich dem Frequenzspektrum bei der Fourier-Transformation $\mathcal{F}\{f(t)\}$. Das sphärische Wellenspektrum besteht aus den Koeffizienten $\widetilde{A}_{\mathrm{p,q}}$. Als Basisfunktionen dienen dafür die sphärischen Harmonischen $Y_{\mathrm{p,q}}(\varphi, \vartheta)$.

$$Y_{\mathrm{p,q}}(\varphi, \vartheta) = \sqrt{\frac{(2\mathrm{p}+1)}{4\pi} \frac{(\mathrm{p}-\mathrm{q})!}{(\mathrm{p}+\mathrm{q})!}} \cdot P_{\mathrm{p}}^{\mathrm{q}}(cos\vartheta) \cdot e^{i\mathrm{q}\varphi} \tag{2.3.2}$$

Wobei $P_{\mathrm{p}}^{\mathrm{q}}(cos\vartheta)$ für die zugeordneten Legendre-Polynome steht.

$$P_{\mathrm{p}}^{\mathrm{q}}(cos\vartheta) = (-1)^{\mathrm{q}}(1-x^2)^{\frac{\mathrm{q}}{2}} \frac{\mathrm{d}^{\mathrm{q}}}{\mathrm{d}(cos\vartheta)^{\mathrm{q}}} P_{\mathrm{p}}(cos\vartheta) \tag{2.3.3}$$

Für negative Grade q gilt mit den komplex konjugierten $Y_{\mathrm{p,q}}^{*}(\varphi, \vartheta)$:

$$Y_{\mathrm{p},-\mathrm{q}}(\varphi, \vartheta) = (-1)^{\mathrm{q}} \cdot Y_{\mathrm{p,q}}^{*}(\varphi, \vartheta) \tag{2.3.4}$$

Die sphärische harmonische Transformation $\mathcal{SHT}$ wird dann durch folgende Gleichung beschrieben:

$$\widetilde{A}_{\mathrm{p,q}} = \frac{1}{h_{\mathrm{p}}(kr)} \cdot \int_0^{2\pi} \int_0^{\pi} p(r, \varphi, \vartheta) \cdot Y_{\mathrm{p,q}}^{*}(\varphi, \vartheta) \cdot sin\vartheta \cdot \mathrm{d}\vartheta \cdot \mathrm{d}\varphi \tag{2.3.5}$$

Eine vollständige Rekonstruktion des Schallfeldes an jedem Ort außerhalb der Schallquelle ist durch die inverse sphärische harmonische Transformation $\mathcal{ISHT}$ (Gleichung (2.3.1)) möglich.

Für die computergestützte Berechnung müssen die Daten diskretisiert werden. Das heißt es wird ein Raster für die Messpunkte auf der Kugeloberfläche gewählt (N Elevationsschritte und M Azimutschritte) und die sphärischen Harmonischen auf eine maximale Ordnung p begrenzt. Dadurch kann nun eine numerische Lösung der $\mathcal{SHT}$ berechnet werden, welche hier sphärische harmonische Zerlegung (diskrete sphärische harmonische Transformation $\mathcal{DSHT}$) genannt wird. Durch numerische Integration von Gleichung (2.3.5) über $\mathrm{d}\vartheta = \pi/(N-1)$ und $\mathrm{d}\varphi = 2\pi/M$ erhält man (Giron (1996)):

$$\widetilde{A}_{\mathrm{p,q}} = \frac{C_{\mathrm{p,q}}}{h_{\mathrm{p}}(kr)} \cdot \frac{2\pi}{M} \sum_{n=0}^{N-1} \alpha_n \cdot P_{\mathrm{p}}^{\mathrm{q}}(cos\vartheta_n) \cdot \sum_{m=0}^{M-1} p(r, \varphi_n, \vartheta_m) \cdot e^{i\mathrm{q}\varphi_m} \tag{2.3.6}$$

$$\text{mit} \qquad C_{\mathrm{p,q}} = \sqrt{\frac{(2\mathrm{p}+1)}{4\pi} \frac{(\mathrm{p}-\mathrm{q})!}{(\mathrm{p}+\mathrm{q})!}}\ , \qquad \vartheta_n = n \cdot \mathrm{d}\vartheta, \quad 0 \leq n \leq N-1,$$

$$\varphi_m = m \cdot \mathrm{d}\varphi, \ 0 \leq m \leq M-1$$

$$\text{und} \qquad \alpha_n = \begin{cases} \left| \int\limits_{\vartheta_n - \frac{\mathrm{d}\vartheta}{2}}^{\vartheta_n + \frac{\mathrm{d}\vartheta}{2}} sin(\vartheta) \cdot \mathrm{d}\vartheta \right| \\ \qquad = \left| cos(\vartheta_n + \frac{\mathrm{d}\vartheta}{2}) - cos(\vartheta_n - \frac{\mathrm{d}\vartheta}{2}) \right|\ , & 0 \leq n \leq N-1 \\ 1 - cos(\frac{\mathrm{d}\vartheta}{2})\ , & n = 0,\ N-1 \end{cases}$$

Die maximale Ordnung p_{max} wird dabei durch die spatiale Auflösung der originalen Abstrahlcharakteristik bestimmt. Es sind mindestens doppelt so viele Azimutringe M nötig wie die maximalen Ordnung, damit das abgetastete Signal korrekt rekonstruiert werden kann. Das bedeutet im Umkehrschluss, dass die Ordnung, bis zu der die $\mathcal{DSHT}$ mindestens berechnet werden muss, der Hälfte der Anzahl der Azimutringe entspricht. Bei Abtastung mit gleich großen Azimut- und Elevationswinkeln ($\mathrm{d}\varphi = \mathrm{d}\vartheta$), kann die maximale Anzahl der Ordnung auch durch die Anzahl der Elevationsringe ($\mathrm{p}_{max} = N-1$ bei $\vartheta = \pi$ bis $-\pi$) bestimmt werden (Giron (1996), S.68 ff.).

$$\mathrm{p}_{max} = \frac{M}{2} = N - 1 \tag{2.3.7}$$

Durch die maximale Ordnung der $\mathcal{DSHT}$ kann auch die maximale korrekt rekonstruierbare Frequenz ermittelt werden. Dazu wird die Gleichung $\mathrm{p}_{max} = k \cdot r_s$ aus Giron (1996), S.68 nach der Frequenz der Wellenzahl umgestellt.

$$f_{max} = \frac{\mathrm{p}_{max} \cdot c}{2\pi \cdot r_s} \tag{2.3.8}$$

Dabei gibt r_s den Radius der maximalen mechanischen Quellausdehnung an.

Für die $\mathcal{ISHT}$ ist keine numerische Integration notwendig. Sie kann direkt mit dem Ergebnis der $\mathcal{DSHT}$ verwendet werden.

2.4. Inkorrekt gestellte Probleme und Regularisierungsmethoden

Wann ist ein Problem korrekt gestellt und wann ist es inkorrekt gestellt?
„Ein Problem heißt korrekt gestellt, wenn es eine eindeutige Lösung besitzt und diese bestimmt werden kann, ohne dass zusätzliche Informationen oder Annahmen jedweder Art einbezogen werden müssen. Anderenfalls liegt ein inkorrekt gestelltes Problem vor“. (Kaschenz (2006))

Für die Berechnung der Monopolsynthese muss der inverse Übertragungspfad gefunden werden. Da die Anzahl der Monopole und die der Empfängerpunkte normalerweise nicht gleich ist, handelt es sich um ein unter- oder überbestimmtes System und es kann der inverse Übertragungspfad nicht einfach durch $\mathbf{d} = \mathbf{G}^{-1} \cdot \mathbf{p}$ gefunden werden. Damit dieses inkorrekt gestellte Problem jedoch trotzdem gelöst werden kann, bedient man sich einer näherungsweisen Lösung im Sinne einer pseudoinversen Matrix oder einer Regularisierung.

2.5. Prinzipieller Ablauf der Rekonstruktion von gerichteten Schallquellen mittels Monopolsynthese

Die Rekonstruktion von Abstrahlcharakteristiken durch Monopolquellen wird im Folgenden erläutert (nach Escolano et al. (2007); Müller-Trapet und Vorländer (2014); Ochmann (1990, 1995); Opdam et al. (2016)).

1. Ermitteln der originalen Abstrahlcharakteristik an j Abtastpositionen

Die originale Abstrahlcharakteristik kann entweder durch den abgestrahlten Schalldruck einer Schallquelle an j Mikrofonpositionen in einem reflektionsarmen Raum (RAR) gemessen werden oder anhand von Simulationsmodellen berechnet werden. Die Anordnung der Abtastpositionen $\mathbf{x}_\mathrm{j} = \mathbf{r}_\mathrm{j}$ ist dafür sphärisch um die Quelle angeordnet.

2. Festlegen von Anzahl i und Position $\mathbf{x}_\mathrm{i}$ der Monopole

Ein geeignetes Verfahren dafür zu finden ist Hauptbestandteil dieser Arbeit. Im Abschnitt 1.1 sind ein paar Ansätze aufgeführt, welche die Anordnung der Monopole für eine Frequenz ermöglichen.

3. Bestimmen der frequenzabhängigen Gewichte für die Monopole

Sobald die Anzahl und Positionen der Monopole gewählt wurden, können die Übertragungspfade zwischen den Monopolpositionen $\mathbf{x}_\mathrm{i}$ und den Abtastpositionen $\mathbf{x}_\mathrm{j}$ durch jeweils eine Green'sche Funktion im Freifeld beschrieben werden. Dieser Übertragungspfad kann als Matrixdarstellung formuliert werden.

$$\mathbf{p} = \mathbf{G} \cdot \mathbf{d} \tag{2.5.1}$$

Die $G_{\mathrm{i,j}}$-Matrix enthält die Green'schen Funktionen

$$\mathbf{G} = \begin{pmatrix} \frac{e^{-jkr_{1,1}}}{4\pi r_{1,1}} & \cdots & \frac{e^{-jkr_{\mathrm{i},1}}}{4\pi r_{\mathrm{i},1}} \\ \vdots & \ddots & \vdots \\ \frac{e^{-jkr_{1,\mathrm{j}}}}{4\pi r_{1,\mathrm{j}}} & \cdots & \frac{e^{-jkr_{\mathrm{i},\mathrm{j}}}}{4\pi r_{\mathrm{i},\mathrm{j}}} \end{pmatrix}. \tag{2.5.2}$$

Die Gewichtungsfaktoren d_i der Monopole können durch Minimierung des linearen Problems bestimmt werden.

$$\min_{\mathbf{d}} = \|\mathbf{G} \cdot \mathbf{d} - \mathbf{p}\|_2 \tag{2.5.3}$$

Durch Benutzung einer Einzelwertzerlegung kann eine Pseudoinverse der $\mathbf{G}$-Matrix $\mathbf{G}^+$ erzeugt werden. Mit einer normalisierten Methode der kleinsten Fehlerquadrate kann das lineare Problem mit der folgenden Matrixgleichung gelöst werden, wodurch die Gewichtungsfaktoren gefunden werden:

$$\mathbf{d} = \mathbf{G}^+ \mathbf{p}. \tag{2.5.4}$$

Durch die Anordnung der Monopole und die Ermittlung der zugehörigen Gewichte, kann auf diese Weise die komplexe Abstrahlcharakteristik einer Schallquelle für eine Frequenz rekonstruiert werden.

4. Rekonstruktion durch Überlagerung der gewichteten Monopole

Rein simulativ und theoretisch lässt sich das Schallfeld aus der Überlagerung der Monopole mit der Gleichung (2.5.1) berechnen und die ursprüngliche Abstrahlcharakteristik an den Abtastpositionen $\mathbf{x}_j$ nahezu korrekt rekonstruieren. In dieser Arbeit sollen die Schallquelle aber mit realen existierenden Wiedergabesystemen hörbar gemacht werden. Die Berechnung des Schallfeldes geschieht prinzipiell auch mit der Gleichung (2.5.1). Die realen Wiedergabesysteme unterliegen aber physikalischen Grenzen, wodurch ein Kompromiss zwischen der Genauigkeit des rekonstruierten Schallfeldes und der Wiedergabe in den physikalischen Grenzen des Systems gefunden werden muss. Auch die Anzahl der Monopole in realen Wiedergabesystemen ist auf Grund der Rechenleistung begrenzt, wodurch weitere Ungenauigkeiten in der Rekonstruktion auftreten können. Außerdem geht von realen Schallquellen in der Regel ein breitbandiges Signal aus, womit der Anspruch formuliert wird eine Monopolanordnung für ein breites Frequenzspektrum zu finden.

2.6. Auswahl der Verfahren

Aus den in Abschnitt 1.1 vorgestellten Verfahren, wurde das Verfahren nach Giron ausgewählt. Dieses Verfahren wird in der zugehörigen Dissertation (Giron (1996)) ausführlich beschrieben und für die numerische Berechnung vorbereitet. Daher kann es direkt implementiert werden. Es wird als eine gute und sinnvolle Grundlage zur Berechnung von Monopolen über einen großen Frequenzbereich eingeschätzt.

Den Denkanstoß zur Anordnung der Monopole auf Spiralen gab das Paper von Escolano et al. (2007). Jedoch wurde die Anordnung durch eigene Überlegungen in Zusammenarbeit mit dem Fraunhofer IDMT erarbeitet.

Bei den weiteren Verfahren konnte kein Ansatz zum automatischen finden der Monopolpositionen erkannt werden.

Für die Rekonstruktion und somit auch das Ermitteln der Ansteuerungsfunktionen wurden diverse Methoden zur Bildung der Pseudoinversen gefunden (z.B. Moor-Penrose Pseudoinverse, mit Least Square Regularisierung (LSQR), mit Tikhonov Regularisierung). Für die Regularisierungsverfahren sind die entsprechenden Regularisierungsparameter von entscheidender Bedeutung und es gibt auch dafür verschiedene Möglichkeiten, diese zu ermitteln. (Hansen (1994, 2007); Kaschenz (2006)) Die Auswahl der Regularisierungsmethoden beschränkt sich auf zwei gebräuchliche Methoden, welche eine gute Lösung des inkorrekt gestellten Problems ermöglichen sollen.

Ausgewählte Verfahren:

- Spiralanordnung
- Monopoldekomposition nach Giron

Rekonstruktion mit unterschiedlichen Regularisierungsverfahren:

- Tikhonov-Regularisierung mit L-Kurven Verfahren zum Finden des Regularisierungsparameters
- LSQR iterativ

3. Methoden zur Anordnung der Monopole

Für die Monopolsynthese wird die Anordnung der Monopole gesucht, mit der die originale Abstrahlcharakteristik mit der geringsten Abweichung rekonstruiert werden kann. Diese Anordnung muss näherungsweise bestimmt werden. In der Regel verändert sich die Abstrahlcharakteristik einer Quelle über die Frequenz. Eine Anforderung an die gefundene Anordnung der Monopole ist, dass sie für eine möglichst große Bandbreite funktioniert und somit einheitliche Positionen für den gesamten Frequenzbereich gefunden werden sollen.

Man kann die Anordnung der Monopole sowohl anhand der originalen Abstrahlcharakteristik bestimmen, als auch einen generischen Ansatz wählen. Ein generischer Ansatz ist allein vom benötigten Frequenzbereich abhängig und lässt sich verhältnismäßig einfach berechnen. Im folgendem werden zwei Möglichkeiten zur Ermittlung der Anordnung der Monopole beschrieben.

3.1. Spiralanordnung der Monopole

Als generischer Ansatz wurde eine als pragmatisch und plausibel erscheinende Anordnung der Monopole gewählt, welche auf der Anordnung auf archimedischen Spiralen basiert. Diese Anordnung ist für alle Abstrahlcharakteristiken in einem bestimmten Frequenzbereich konstant und unabhängig von der zu synthetisierenden frequenzabhängigen Abstrahlcharakteristik.

Die archimedische Spirale ist dadurch definiert, dass der Radius r gleich dem Produkt aus Drehwinkel φ und einem Faktor $b \in \mathbb{R}^+$ ist:

$$r = b \cdot \varphi \tag{3.1.1}$$

Da ein Signal mit hoher Frequenz f eine kleine Wellenlänge λ ($\lambda = c/f$) und ein Signal mit niedriger Frequenz eine große Wellenlänge hat, ergeben sich je nach Frequenz unterschiedliche Abstände d zwischen den Monopolen, damit sich die Signale konstruktiv überlagern können. Der Abstand der Monopole ist also so zu wählen, dass sich der abgestrahlte Schall entsprechend dem Superpositionsprinzip durch Überlagerung und Auslöschung (Interferenz) so ergänzt, dass bei entsprechender Ansteuerung das gewünschte Schallfeld entsteht. Dabei wird für die hier verwendete Anordnung ein Abstand von $\lambda/2$ gewählt.

Der Frequenzbereich, in dem die Abstrahlcharakteristik rekonstruiert werden kann, wird durch die maximale Frequenz f_{max} und die minimale Frequenz f_{min} begrenzt. Die maximale Frequenz ergibt sich durch die Aliasing-Frequenz der ordnungsbegrenzten $\mathcal{SHT}$ mit der Gleichung (2.3.8) oder durch die maximale Frequenz des Testsignals mit der die Abstrahlcharakteristik vermessen wurde, wenn diese kleiner ist. Die minimale Frequenz wird durch die maximale Quellausdehnung d_s bestimmt ($d_{max} = r_s = d_s/2$). Durch die Grenzen des zu rekonstruierenden Frequenzbereichs sind der minimale und der maximale Abstand bestimmt.

$$\begin{aligned} d_{min} &= \frac{1}{2} c/f_{max} \\ d_{max} &= \frac{1}{2} c/f_{min} \end{aligned} \tag{3.1.2}$$

Um eine möglichst gleich verteilte Anordnung der Monopole zu erhalten, wurde die Positionierung entlang von Spiralen gewählt, wobei der Abstand der Monopole den

Radius der Spirale bestimmt ($r = d$). Da die Anordnung für einen großen Frequenzbereich funktionieren soll, ist es sinnvoll, den Abstand der Monopole exponentiell wachsen zu lassen. Das wird erreicht, indem der Radius r der Spirale exponentiell zum Drehwinkel φ wächst.

$$r = b \cdot e^{(a \cdot \varphi)} \tag{3.1.3}$$

Der gewünschten Winkelbereich, welcher r_{min} bis r_{max} entspricht, wird durch das Umstellen von Gleichung (3.1.3) nach φ gefunden:

$$\varphi = log(\frac{r}{b})/a \tag{3.1.4}$$

Dabei wird der exponentiell wachsende Abstand wie folgt bestimmt:

$$r_{1...K} = 10^{(r_{log,1...K})} \qquad ; K = Anzahl\ Monopole \tag{3.1.5}$$

mit $$r_{log,1...K} = log_{10}(d_{min}) + (r_{logStep} \cdot (0...K-1))$$

und $$r_{logStep} = \frac{log_{10}(d_{max}) - log_{10}(d_{min})}{(K-1)}$$

Diese Verteilung ist gleichzeitig auch äquiangular (siehe Abb. 3.1.1).

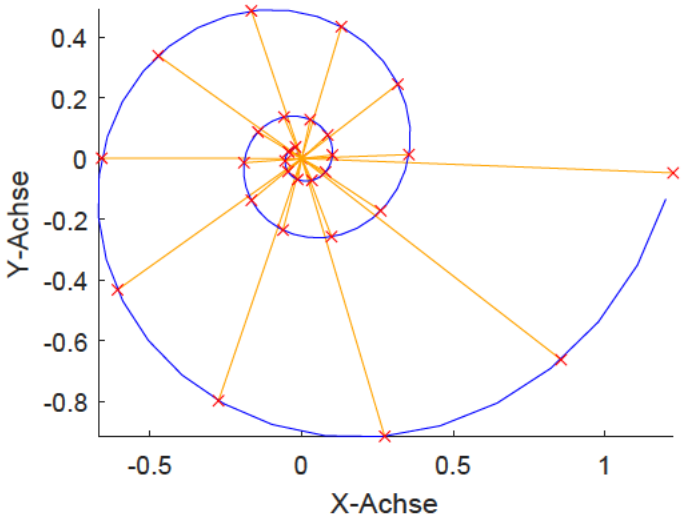

Abbildung 3.1.1: Spirale (–) mit exponentiell wachsendem Radius (–) und Abstand zwischen den Monopolen (×); $f_{min} = 140, f_{max} = 2000, K = 28, a = 0.2, b = 0.1$

Um eine dreidimensionale Abstrahlcharakteristik zu rekonstruieren wird in der Regel auch eine dreidimensionale Anordnung von Monopolen benötigt. Dazu wird

der Spirale eine dem Radius entsprechende Z-Komponente hinzugefügt($z = r$). Die Spirale wird dupliziert und orthogonal zueinander in allen drei Raumrichtungen gesetzt, beginnend von einem gemeinsamen Ursprungspunkt (siehe Abb. 3.1.2).

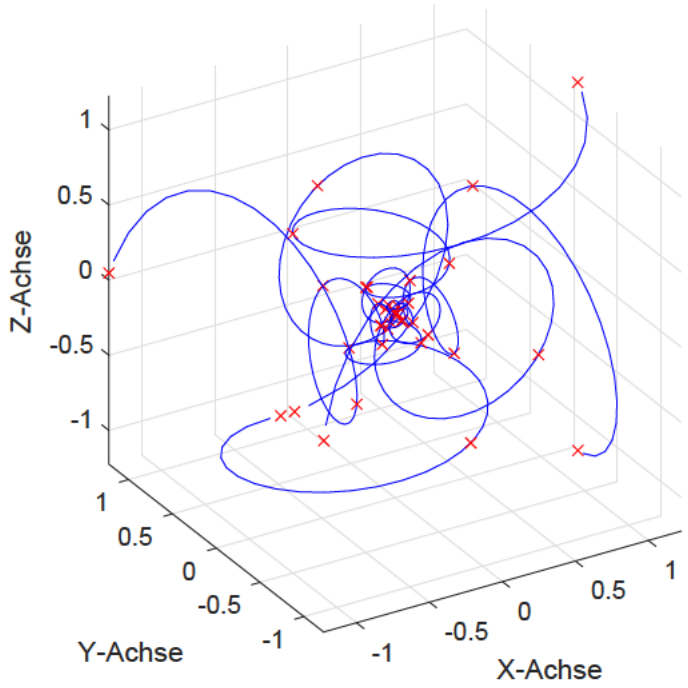

Abbildung 3.1.2: Spiralen in drei Dimensionen

Daraus ergibt sich die Anzahl der Monopole in der dreidimensionalen Anordnung nach Gleichung (3.1.6).

$$L = 6 \cdot K + 1 \tag{3.1.6}$$

Durch den logarithmisch wachsenden Abstand und die dreidimensionale Anordnung sollte eine Rekonstruktion von Abstrahlcharakteristiken sowohl kleiner als auch großer Frequenzen möglich sein.

3.2. Monopoldekomposition nach Giron

Im Zuge der Monopoldekomposition wird die Anzahl und Position der Monopole bestimmt, welche notwendig sind, um die Freifeldübertragungsfunktion zwischen den Monopol- und Abtastpositionen zu ermitteln (Müller-Trapet und Vorländer (2014); Opdam et al. (2016)). In der folgenden Abbildung 3.2.1 werden alle nötigen

Schritte aufgeführt, um die Anordnung von Monopolen zu finden, mit der die originale Abstrahlcharakteristik möglichst korrekt im Fernfeld rekonstruiert werden kann (nach Giron (1996)).

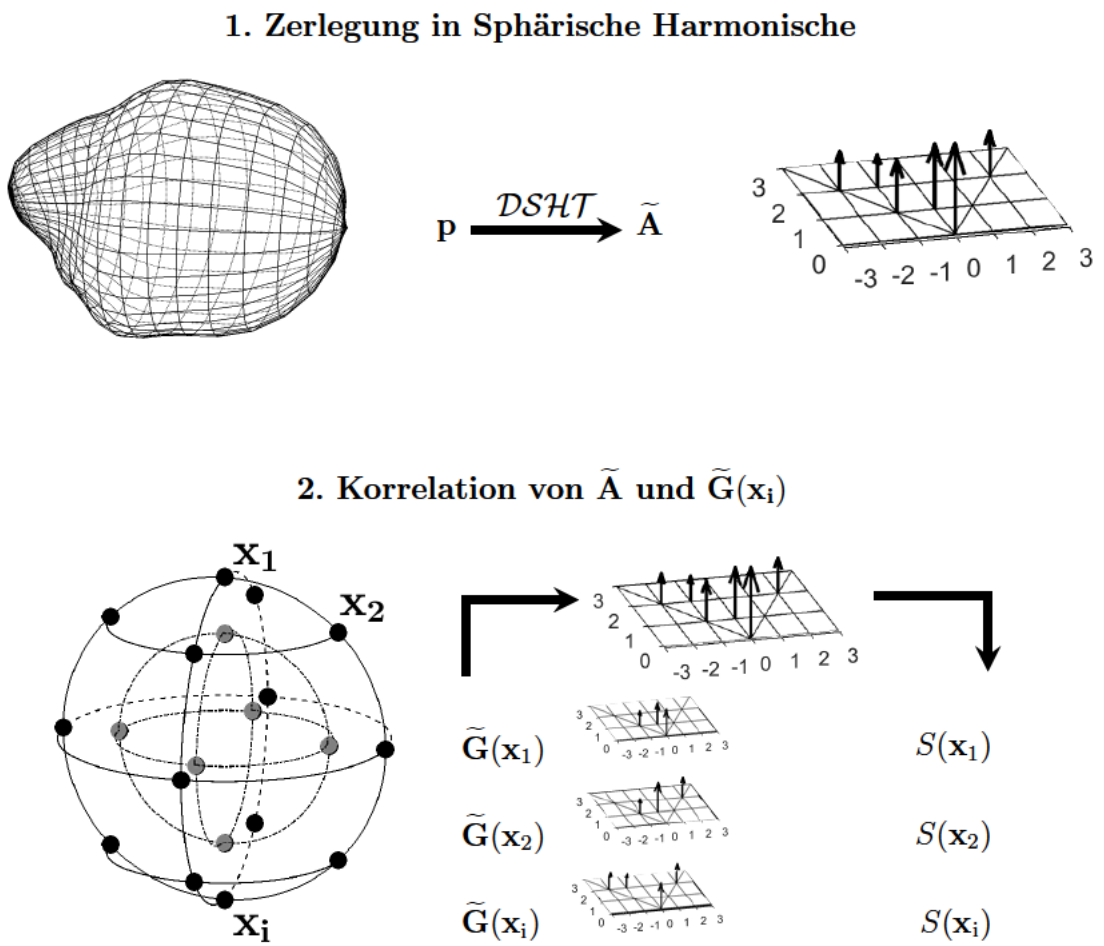

Abbildung 3.2.1: Zusammenfassung der Schritte zum Finden der Monopole

Als Grundlage zur Berechnung der Monopolpositionen dient die originale Abstrahlcharakteristik der Quelle, welche durch den Druck $p(\mathbf{x})$ an den Punkten $\mathbf{x}$ auf einer

Kugeloberfläche beschrieben ist. Durch die Berechnung der $\mathcal{DSHT}$ von $\mathbf{p}$ erhält man die Matrix $\widetilde{\mathbf{A}}$, welche die Amplituden der sphärischen Harmonischen beinhaltet.

Als mögliche Positionen der Monopole für die Rekonstruktion der originalen Abstrahlcharakteristik werden Positionen entlang eines Rasters auf Sphären mit unterschiedlichem Radius gewählt. Das Schallfeld der Monopole an diesen Positionen wird in sphärische Harmonische zerlegt $\widetilde{\mathbf{G}}(\mathbf{x}_\mathrm{i})$. Diese lassen sich direkt durch Green'sche Funktionen im Freifeld bestimmen:

$$\widetilde{G}_{\mathrm{p,q}}(\mathbf{x}_\mathrm{i}) = i \cdot k \cdot Y_{\mathrm{p,q}}(\varphi_\mathrm{i}, \vartheta_\mathrm{i}) \cdot j_m(k \cdot r_\mathrm{i}) \tag{3.2.1}$$

Dabei ist j_m die sphärische Besselfunktion.

Die Bestimmung der bestmöglichen Monopolpositionen wird durch die Korrelation von $\widetilde{\mathbf{G}}(\mathbf{x}_\mathrm{i})$ und $\widetilde{\mathbf{A}}$ ermöglicht, wobei die Korrelationsmatrix $\mathbf{S}(\mathbf{x})$ gebildet wird.

$$S(\mathbf{x}_\mathrm{i}) = \frac{\sum\limits_{\mathrm{p}=0}^{\mathrm{p}_{max}} \sum\limits_{\mathrm{q}=-\mathrm{p}}^{\mathrm{p}} \widetilde{A}_{\mathrm{p,q}} \cdot \widetilde{G}^*_{\mathrm{p,q}}(\mathbf{x}_\mathrm{i})}{\sum\limits_{\mathrm{p}=0}^{\mathrm{p}_{max}} \sum\limits_{\mathrm{q}=-\mathrm{p}}^{\mathrm{p}} \widetilde{G}_{\mathrm{p,q}}(\mathbf{x}_\mathrm{i}) \cdot \widetilde{G}^*_{\mathrm{p,q}}(\mathbf{x}_\mathrm{i})} \tag{3.2.2}$$

Durch das Ermitteln der Maxima von $\mathbf{S}(\mathbf{x})$ werden die Monopolpositionen $\mathbf{x}_\mathrm{i}$ mit der größten Übereinstimmung herausgefunden. Die Ermittlung der Maxima wird in Anhang B erläutert.

Monopolpositionen für ein breites Frequenzspektrum

Die Berechnung der Monopolpositionen nach Giron ist nur für eine Frequenz vorgesehen, jedoch ist die Rekonstruktion der Quelle für ein breites Frequenzspektrum notwendig, um eine plausible Wiedergabe zu ermöglichen. Um diese Anordnung zu erhalten, werden die Monopolpositionen für jede Frequenz einzeln bis zur maximalen Frequenz mit dem vorher beschriebenen Verfahren ermittelt. Aus den ermittelten Positionen werden die Positionen mit dem höchsten Korrelationsgrad ausgewählt und Positionen, die in einem bestimmten Toleranzbereich übereinstimmen, werden zusammengefasst. Auf diese Weise sollte es möglich sein, eine Anordnung zu finden, welche zur Rekonstruktion der Abstrahlcharakteristik am besten passt. Damit die Monopolpositionen für jede Frequenz einzeln berechnet werden können, werden

die gemessenen Impulsantworten der Quelle mittels $\mathcal{FFT}$ in den Frequenzbereich transformiert. Es werden nun jeweils die Werte eines Frequenzbins an allen Mikrofonpositionen verwendet, was gleichzeitig die originale Abstrahlcharakteristik bei einer Frequenz ist.

Die gefundenen Monopolpositionen aller Frequenzen $\mathbf{x}_\mathrm{i}(f)$ werden nach ihrem Korrelationsgrad $S(\mathbf{x}_\mathrm{i}, f)$ sortiert und die Positionen mit der größten Korrelation verwendet. Der Toleranzbereich für die Zusammenfassung von nahe beieinander liegenden Monopolen ist $\pm 10\%$ des Radius des aktuellen Monopols.

4. Ermittlung der Gewichte der Monopole

Für die Rekonstruktion der Abstrahlcharakteristik müssen die Gewichte der Monopole ermittelt werden. Dazu ist die eingangs erwähnte Invertierung der $\mathbf{G}$-Matrix notwendig. Da in der Regel weniger Monopole als Empfängerpunkte für die Rekonstruktion zur Verfügung stehen, stellt die $\mathbf{G}^{-1}$-Matrix ein unterbestimmtes System dar, welches nicht direkt lösbar ist. Solch ein unterbestimmtes System kann mit einer pseudoinversen $\mathbf{G}^{+}$-Matrix gelöst werden. Zudem enthalten die originale Abstrahlcharakteristik und der Übertragungspfad ab und zu auch Ungenauigkeiten oder Störungen, welche die Berechnung der Gewichte zu einem inkorrekt gestellten Problem machen. Durch eine Regularisierung der Lösung kann eine näherungsweise Lösung gefunden werden.

Die Erzeugung einer pseudoinversen Matrix $\mathbf{A}^{+}$, also die Invertierung einer beliebigen rechteckigen Matrix $\mathbf{A}$, kann durch die Singulärwertzerlegung erfolgen (Deuflhard und Hohmann (2008), S. 147).

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T = \sum_{\mathrm{i}=1}^{n} \mathbf{u}_\mathrm{i}\sigma_\mathrm{i}\mathbf{v}_\mathrm{i}^T \tag{4.0.1}$$

$$\mathbf{A}^{+} = \mathbf{V}\boldsymbol{\Sigma}^{+}\mathbf{U}^T \tag{4.0.2}$$

mit $$\boldsymbol{\Sigma}^{+} = \mathrm{diag}(\sigma_1^{-1}, ..., \sigma_\mathrm{i}^{-1}), \sigma_1 \geq \cdots \geq \sigma_\mathrm{i} > 0$$

Die rechteckige Matrix $\mathbf{A}$ kann somit auch durch $\mathbf{A}^T\mathbf{A} = \mathbf{V}\boldsymbol{\Sigma}^2\mathbf{V}^T$ und $\mathbf{A}\mathbf{A}^T = \mathbf{U}\boldsymbol{\Sigma}^2\mathbf{U}^T$ symmetriert werden.

Die Pseudoinverse wird auch Moore-Penrose-Inverse genannt und ergibt eine Lösung, bei der die Norm des Fehlers minimal ist. Eine solche Lösung wird als Lösung eines Least-Square Problems bezeichnet beziehungsweise Methode der kleinsten Quadrate.

$$\min_{\mathbf{d}} = \|\mathbf{G} \cdot \mathbf{d} - \mathbf{p}\|_2 \tag{4.0.3}$$

Für die Rekonstruktion von Abstrahlcharakteristiken kann das jedoch zu großen Unterschieden der Amplituden der Monopole führen. Dieser Unterschied kann dann so groß sein, dass die berechneten Pegel in einem realen Wiedergabesystem aus physikalischen und technischen Gründen nicht wiedergegeben werden können. Daher muss eine Lösung gefunden werden, die eine noch nahezu korrekte Rekonstruktion zulässt, bei der die wiedergegebenen Pegel aber nicht zu groß werden. Solche Lösungen können durch Regularisierungsverfahren ermittelt werden.

4.1. Mit Tikhonov-Regularisierung

Das Verfahren nach A.N. Tikhonov ermöglicht eine näherungsweise Lösung inkorrekt gestellter Probleme und ist robust gegenüber Störungen (Hansen (2007)).

$$\min_{\mathbf{d}} = (\|\mathbf{G} \cdot \mathbf{d} - \mathbf{p}\|_2^2 + \lambda^2 \|\mathbf{L}(\mathbf{d} - \mathbf{d}_{init})\|_2^2) \tag{4.1.1}$$

Die Lösung wird dabei durch Randbedingungen gesteuert. Zu den Randbedingungen gehören die Matrix $\mathbf{L}$, eine initiale Schätzung $\mathbf{d}_{init}$ der Lösung und der Regularisierungsparameter λ. Die Matrix $\mathbf{L}$ enthält entweder die Einheitsmatrix $\mathbf{I}_\mathrm{j}$ oder die Schätzung der Ableitungsoperatoren. Wenn die Lösung schon grob bekannt ist, kann diese als Anfangsparameter $\mathbf{d}_{init}$ mit angegeben werden, wodurch der Fehler der Lösung weiter minimiert werden kann. Als wichtigste Randbedingung, um die Regularisierung zu steuern, gilt der Regularisierungsparameter λ. Ein kleines λ bewirkt dabei eine genaue Rekonstruktion bei großen Unterschieden in dem Lösungsvektor $\mathbf{d}_\mathrm{i}$ und ein großes λ bewirkt das Gegenteil.

Bei der Anwendung der Tikhonov-Regularisierung zur Ermittlung der Ansteuerungsfunktionen der Monopole ist $\mathbf{L} = \mathbf{I}_\mathrm{j}$ und $\mathbf{d}_{init} = 0$, wodurch sich die Gleichung

4.1.1 zu

$$\min_{\mathbf{d}} = (\|\mathbf{G} \cdot \mathbf{d} - \mathbf{p}\|_2^2 + \lambda^2 \|\mathbf{d}\|_2^2) \tag{4.1.2}$$

bzw.

$$\mathbf{d} = (\mathbf{G}^T\mathbf{G} + \lambda \mathbf{I}_\mathrm{j})^{-1} \cdot \mathbf{G}^T\mathbf{p} \tag{4.1.3}$$

vereinfacht.

Es ist wichtig, λ so zu wählen, dass die Pegelunterschiede in $\mathbf{d}$ nicht zu groß werden und wiederum die Abweichungen der mit $\mathbf{d}$ berechneten Abstrahlcharakteristik zur originalen Abstrahlcharakteristik in einem akzeptablen Rahmen bleiben.

Auf die Wahl des Regularisierungsparameters wird in Abschnitt 4.3 eingegangen.

4.2. Mit LSQR-Regularisierung

LSQR ist ein iterativer **L**east **S**quare Algorithmus, welcher sich der **QR**-Zerlegung bedient (Paige und Saunders (1982)). Der LSQR-Algorithmus verzichtet auf die Verwendung der Kreuzproduktmatrix $\mathbf{A}^T\mathbf{A}$, wodurch der Genauigkeitsverlust durch den Gebrauch von $\mathbf{A}^T\mathbf{A}$ vermieden wird. Die QR-Zerlegung der Matrix $\mathbf{G} = \mathbf{Q} \cdot \mathbf{R}$ wird mit Hilfe der Lanczos-Golub-Kahan (LGK) Bidiagonalisierung berechnet. Die Anzahl der Iterationen der LSQR bestimmt die Anzahl der Schritte der LGK Bidiagonalisierung. Durch die Anzahl an Iterationen wird die Stärke der Regularisierung bestimmt, weshalb sie die Rolle des Regularisierungsparameters annimmt.

4.3. Regularisierungsparameter

Regularisierungsparameter dienen der Steuerung von Regularisierungsverfahren und bestimmen entscheidend die Lösung des inkorrekt gestellten Problems. Das inkorrekt gestellte Problem wird im allgemeinen in folgender Form beschrieben:

$$\min_{\mathbf{d}} = \|\mathbf{G} \cdot \mathbf{d} - \mathbf{p}\|_2$$

Durch die Regularisierungsparameter kann die Regularisierung so gesteuert werden, dass nicht ausschließlich eine Minimierung der Norm des Fehlers ($\|\mathbf{G} \cdot \mathbf{d} - \mathbf{p}\|_2$) erfolgt, sondern eine gute Balance zwischen der Norm des Fehlers und der Norm des

Lösungsvektors ($\|\mathbf{d}\|_2$) gefunden wird. Zur Ermittlung des optimierten Regularisierungsparameters gibt es mehrere Verfahren, welche unter anderem das Diskrepanzprinzip, die generalisierte Kreuzvalidierung (GCV) und das L-Kurven-Verfahren sind. Das L-Kurven-Verfahren ermittelt den Regularisierungsparameter nur aus den Daten des Problems ohne weitere Randbedingungen. Daher wird dieses Verfahren in dieser Arbeit zur Bestimmung des Regularisierungsparameters ausgewählt. Für die anderen Verfahren sind entweder zusätzliche Randbedingungen nötig, welche zum Teil nicht bekannt sind oder sie sind nicht robust gegenüber Störungen. Die Regularisierungsparameter aus den anderen Verfahren bedeuten zu dem nicht unbedingt eine bessere Lösung.

L-Kurve

Für das L-Kurven-Verfahren werden die Norm des Fehlers und die Norm des Lösungsvektors gegeneinander in einem doppelt logarithmischen Diagramm aufgetragen.

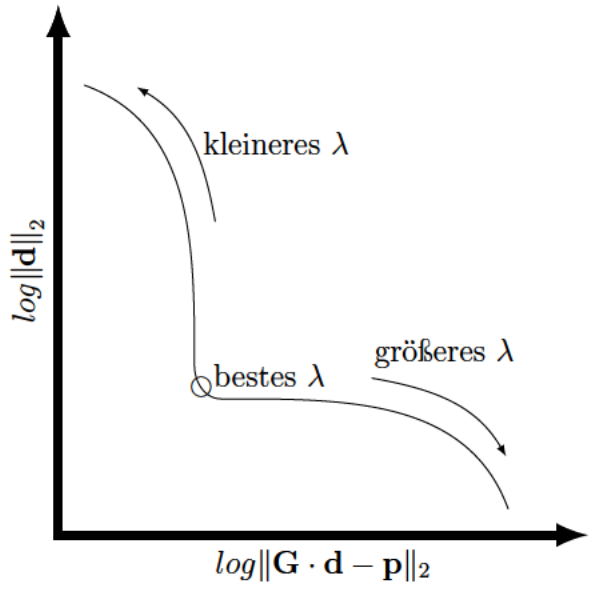

Abbildung 4.3.1: L-Kurve

Das L-Kurven-Verfahren ermöglicht die Ermittlung eines optimierten Regularisierungsparameters λ durch die Berechnung des Punktes, an dem $\|\mathbf{d}\|_2$ und $\|\mathbf{G}\cdot\mathbf{d}-\mathbf{p}\|_2$ jeweils gering sind, ohne dass die jeweils andere Norm zu groß ist. Dieser Punkt liegt

im charakteristischen Knick der L-Kurve.

5. Implementierung der Verfahren mit MatLab

Um mit den gewählten Verfahren zu rechnen, wurden diese mit MatLab von MathWorks® in der Version R2016b implementiert. Das Programm wurde auf Grund vieler schon bestehender Funktionen und Toolboxen und seiner weiten Verbreitung und Bekanntheit gewählt. Alle verwendeten Skripte sind auf der beigefügten CD zu finden und werden, wenn nötig, auszugsweise im Text wiedergegeben.

In Anhang A sind Funktionen aufgelistet, die vom Fraunhofer IDMT entwickelt und für die Arbeit zur Verfügung gestellt wurden.

Das Hauptskript `ReconBalloon_RealMeasurement.m` ist gegliedert in:

- Einlesen der originalen Abstrahlcharakteristik
- Berechnung der Monopolpositionen
- Ermittlung der Gewichte der Monopole
- Rekonstruktion der Abstrahlcharakteristik zur Ermittlung des Fehlers und zur Probe.

Zur besseren Verständlichkeit wird anhand eines Beispiels die Berechnung der Monopolpositionen und der Gewichte gezeigt. Im Beispiel wurde ein bereits vorhandener Datensatz aus der Vermessung des Kling&Freitag-Lautsprechers CA106 verwendet. Die Vermessung erfolgte mit Monkey Forest von Four Audio und wurde im equiangularen Raster von 10° -Schritten verwendet.

5.1. Einlesen der originalen Abstrahlcharakteristik

Die originale Abstrahlcharakteristik wird im Hauptskript durch die Funktion `readBalloonWAV` eingelesen. Die originalen Daten müssen dazu als Impulsantworten für jeden Abtastpunkt, in Form von wav-Dateien vorhanden sein.

```
[outMtxd,gridMtxd,fsd] = readBalloonWAV(InPath)
```

In der Ausgabe `outMtxd` werden alle Richtcharakteristiken im Frequenzbereich gespeichert. `gridMtxd` enthält das Raster, mit dem die Abstrahlcharakteristik ab-

getastet wurde und `fsd` ist die Abtastfrequenz. Damit lassen sich die Frequenzen der Frequenzbins berechnen.

$$f_i = \frac{(bin_i - 1) \cdot fsd}{2 \cdot n_{bins}}, \qquad n_{bins} = \text{Anzahl bins}, \; i = 1, 2, ..., n_{bins} \tag{5.1.1}$$

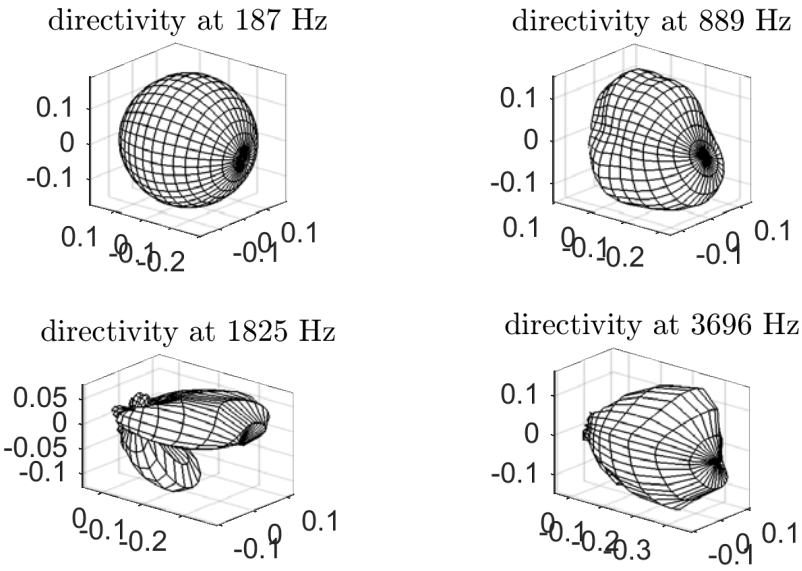

Abbildung 5.1.1: Originale Abstrahlcharakteristik als Betrag der komplexen Werte, gemessen auf einer Sphäre mit $r = 4.3$ m

5.2. Spiralanordnung der Monopole

Mit der Funktion `getMonopoleDistribution` werden die Positionen der Monopole in der Spiralanordnung berechnet. Die Funktion ist eine direkte Implementierung des in Abschnitt 3.1 beschriebenen Verfahrens und wurde in Zusammenarbeit mit dem Fraunhofer IDMT entwickelt. Die Eingangsparameter sind die Ordnung, die Frequenzlimits, die räumliche Dimension, der Typ und ein Logikschalter für die graphische Ausgabe. Die Ordnung `N` bestimmt die Anzahl `M` der Monopole nach Gleichung (3.1.6). Die Grenzfrequenzen lassen sich durch die maximal und minimal zu berechnenden Frequenzen bestimmen. Die maximale Frequenz ist durch die Abtastung der originalen Abstrahlcharakteristik mit Gleichung (2.3.8) gegeben. Sie kann jedoch auch durch die maximal im Signal vorkommende Frequenz bestimmt sein, wenn diese geringer als die aus dem Abtasttheorem resultierende Frequenz ist.

Die räumliche Dimension gibt an, ob die Anordnung für einen dreidimensionalen oder zweidimensionalen Raum erzeugt wird. Durch den Typ wird die Anordnung auf Spiralen oder zufällig im Raum gewählt. Bisher ist nur eine Anordnung auf Spiralen in der Funktion implementiert. Um durch die Funktion eine graphische Darstellung der Anordnung zu erhalten, muss der Logikschalter `verbose` auf `true` gesetzt werden. Die erzeugten Positionsdaten der Monopole $\mathbf{x}_\mathrm{i}$ können nun genutzt werden, um die Ansteuerungsfunktionen $\mathbf{d}$ mit den Regularisierungsverfahren zu ermitteln.

5.3. Monopoldekomposition nach Giron

Das Verfahren zur Zerlegung der Abstrahlcharakteristik in Monopole nach Giron wurde in der Funktion `PosFind_FX` zusammengefasst. Die Funktion untergliedert sich wiederum in die $\mathcal{DSHT}$, die Berechnug der Korrelationsmatrix $\mathbf{S}$ und die Ermittlung der lokalen Extrema von $\mathbf{S}$. Die lokalen Extrema sind gleichzeitig die wahrscheinlichsten Monopolpositionen, mit denen die Abstrahlcharakteristik rekonstruiert werden kann.

Durch die Funktion `DSHT` wird das sphärische Amplitudenspektrum $\widetilde{\mathbf{A}}$ der originalen Abstrahlcharakteristik berechnet. Die Ordnung p der $\mathcal{DSHT}$ wurde mit p = 12 bestimmt. Diese Ordnung wurde unter der Annahme berechnet, dass die Abstrahlcharakteristik mit einer equiangularen Abtastung aufgenommen wurde. Das entspricht aber nicht der Wahrheit, da die Azimut- und Elevationswinkel zwischen den Abtastpunkten der verwendeten Abstrahlcharakteristik gleich groß sind. Die equiangulare Abtastung ist jedoch mit gleich vielen Abtastpunkten in der Elevation als auch im Azimut definiert , wodurch die Elevationswinkel halb so groß sind wie die Azimutwinkel (Driscoll und Healy, Jr. (1994)). Die Ordnung der $\mathcal{DSHT}$ müsste daher mit der Gleichung (2.3.7) berechnet werden und würde p = 18 ergeben. Die Berechnung wurde jedoch, auf Grund der falschen Annahme, mit p = 12 durchgeführt, da zum Zeitpunkt der Berechnungen der Sachverhalt noch nicht klar war. Als maximale Ausdehnung der Quelle wird die Diagonale des Lautsprechers CA106 genommen ($r_s = \frac{\sqrt{0.195^2+0.365^2+0.206^2}\mathrm{m}}{2} = 0.23$ m). Daraus ergibt sich die maximal rekonstruierbare Frequenz nach Gleichung (2.3.8) mit $f_{max} = 2.8$ kHz. Durch die

maximale Ausdehnung der Quelle und der Annahme, dass die Monopole nur innerhalb der Quelle auftreten können, lässt sich die minimale rekonstruierbare Frequenz nach Gleichung (3.1.2) zu $f_{min} = 373$ Hz bestimmen.

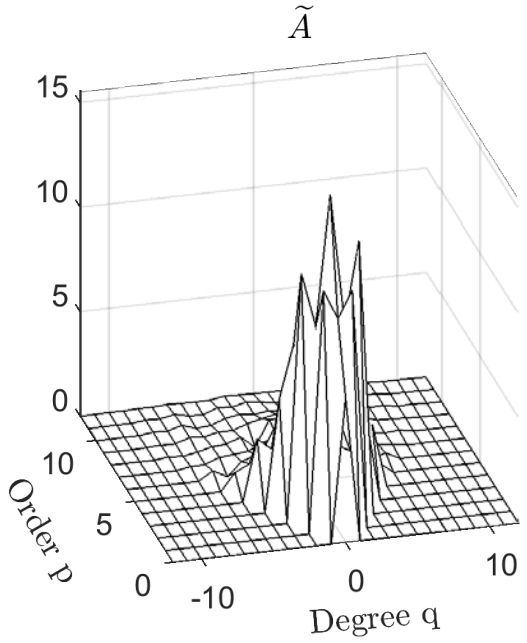

Abbildung 5.3.1: Sphärisches Amplitudenspektrum der originalen Abstrahlcharakteristik bei 1825 Hz als Betrag der komplexen Werte

Die Korrelation zwischen den sphärischen Amplitudenspektren $\widetilde{\mathbf{A}}$ und $\widetilde{\mathbf{G}}(\mathbf{x}_i)$ wird durch die Funktion `correlation_isMF` ermittelt.

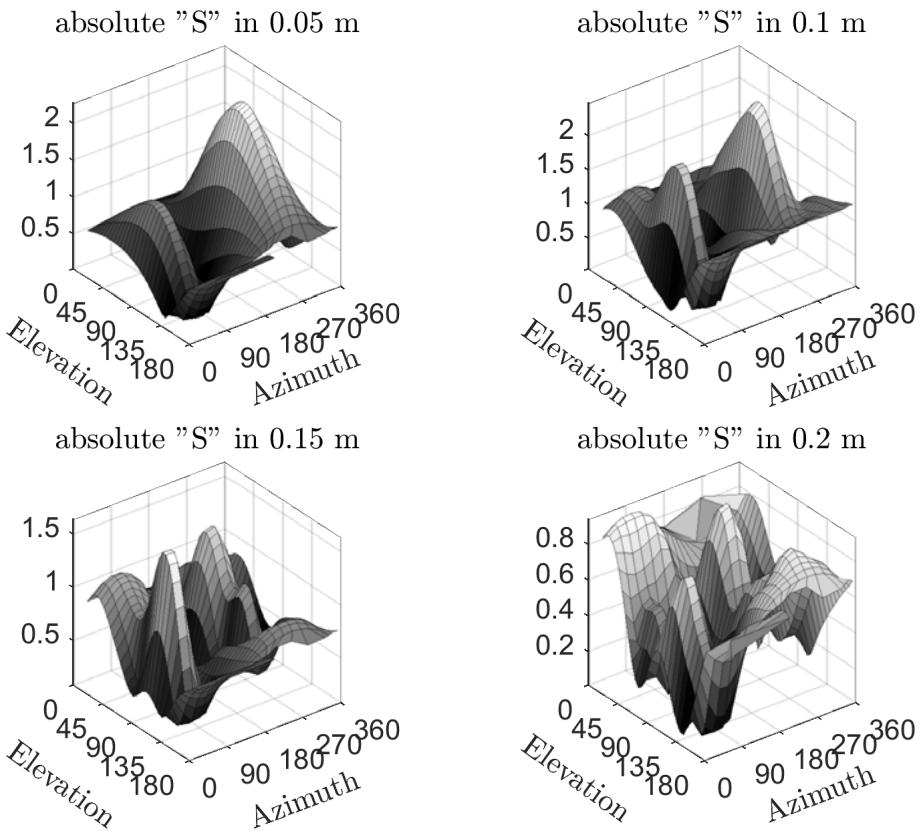

Abbildung 5.3.2: Korrelationsmatrix bei 1825 Hz als Betrag der komplexen Werte. Die Angabe in Metern gibt den Radius der Sphären an, auf denen die Monopole angeordnet sind.

Durch die Funktion `poslocmaxsph` werden die lokalen Maxima der Korrelationsmatrizen bestimmt (siehe Anhang B). Die so ermittelten Positionen werden nach Wert der Korrelation sortiert. Die Positionen mit der größten Korrelation sind am besten geeignet, um die Abstrahlcharakteristik dieser einen Frequenz zu rekonstruieren.

Tabelle 5.3.1: Positionen der Monopole bei 1825 Hz mit größter Korrelation

Azimut in °	Elevation in °	Radius in m	Korrelation S
40	95	0.080	2.64
120	100	0.112	2.03
130	30	0.114	1.27
70.56	175	0.103	1.23
70	170	0.101	1.23
20	175	0.101	1.22
5	175	0.103	1.22
10	175	0.102	1.22
40	150	0.178	1.01

Die Funktion `PosFind_FX` wird für alle rekonstruierbaren Frequenzen aufgerufen und alle gefundenen Monopole werden in der variable `giron_all` gespeichert. Mit der Funktion `FindDuplMono` werden alle doppelten und sehr nahe beieinanderliegenden Monopole zusammengeführt und nach ihrer kumulierten Korrelation sortiert (siehe Anhang C). Die Sortierung nach der kumulierten Korrelation verfolgt den Gedanken, dass Positionen, die für viele Frequenzen auftreten, wichtig sind für die Rekonstruktion der Abstrahlcharakteristik in einem breiten Frequenzbereich. Nun werden die Positionen mit der größten kumulierten Korrelation ausgewählt. Diese Positionen sollten für die Rekonstruktion am besten geeignet sein.

Tabelle 5.3.2: Positionen der Monopole für alle Frequenzen sortiert nach kumulierter Korrelation, hier beispielhaft die ersten 7 Positionen

Azimut in °	Elevation in °	Radius in m	Korrelation $\sum(S)$
130	30	0.205	56.16
120	70	0.099	54.66
130	30	0.207	53.87
40	110	0.093	51.59
40	110	0.090	51.45
40	110	0.095	50.86
130	30	0.210	49.08

5.4. Regularisierungstoolbox von Hansen (2007)

Zur Berechnung der Ansteuerung der Monopole werden Regularisierungsfunktionen aus der MatLab-Toolbox von Prof. Per Christian Hansen benutzt. Die Handhabung dieser Funktionen ist in Hansen (2007) beschrieben und ermöglicht die Anwendung der ausgewählten Regularisierungsverfahren.

Für die Regularisierungsverfahren muss die Singulärwertzerlegung berechnet werden:

`[U,s,V] = csvd(G)` dabei ist `s` $= \mathbf{\Sigma}$.

Die Tikhonov-Regularisierung wird mit der Funktion `tikhonov` durchgeführt.

```
D = tikhonov(U,s,V,direc,reg_corner)
```

`reg_corner` ist der Regularisierungsparameter λ und wird durch die L-Curve bestimmt.

```
reg_corner = l_curve(U,s,direc,'Tikh')
```

Die L-Curve zu dem Beispiel aus Abschnitt 5.3 sieht wie folgt aus.

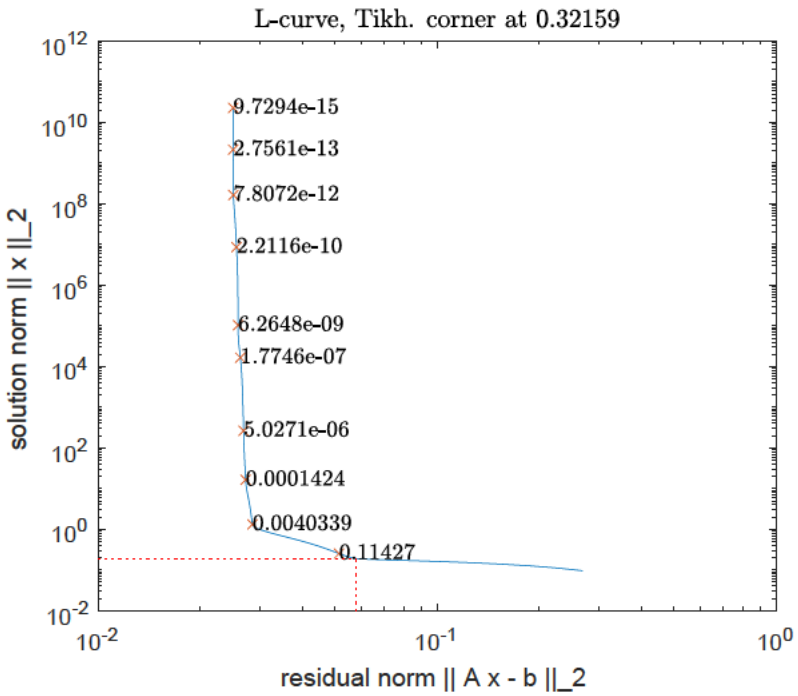

Abbildung 5.4.1: L-Kurve aus den berechneten Monopolpositionen und Abtastpunkten auf einer Sphäre mit $r = 4.3$ m

Damit ist $\lambda = 0.322$ und das Ergebnis der Tikhonov-Regularisierung sind folgende Ansteuerungsparameter.

Tabelle 5.4.1: Ansteuerungsfunktionen der Monopole nach Tikhonov-Regularisierung bei 1825 Hz, hier beispielhaft für die ersten 7 Monopole

Position	Ansteuerungsfunktionen
1	-270.58 -1094.52i
2	-6.81+ 0.35i
3	146.81+ 319.81i
4	-65.10+ 664.93i
5	102.76 - 470.18i
6	-204.86+ 289.22i
7	250.91+ 945.95i

Mit den Monopolpositionen und den dazugehörigen Ansteuerungsparametern lässt sich die Abstrahlcharakteristik durch Monopolsynthese rekonstruieren.

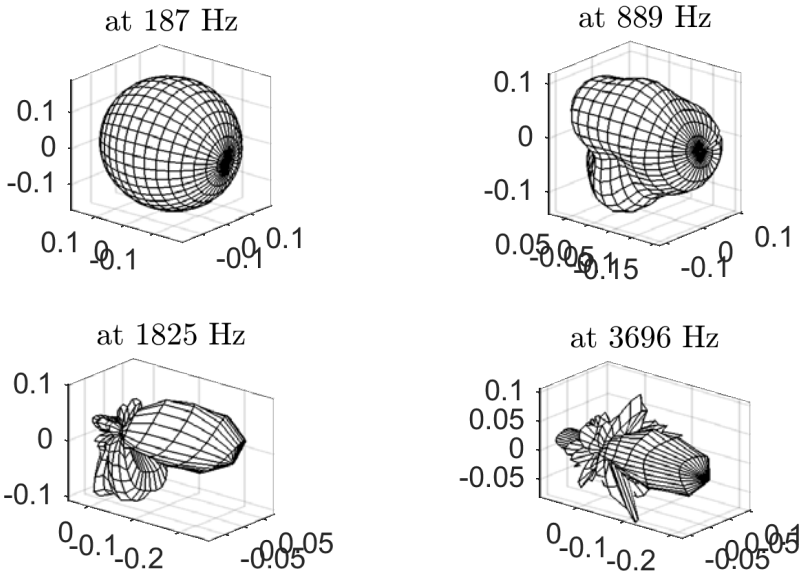

Abbildung 5.4.2: Rekonstruierte Abstrahlcharakteristik als Betrag der komplexen Werte mit 32 Monopolen nach Tikhonov-Regularisierung.

Aus berechneten Ansteuerungsfunktionen werden für die reale Wiedergabe FIR-Filter berechnet. Das ist die Aufgabe der Funktion `filterMatrixOptimalDesign`. Diese Funktion gibt die optimierten FIR-Filter als Impulsantwort aus. Dadurch erhält man folgende Ansteuerungsfunktionen.

Tabelle 5.4.2: Optimierte Ansteuerungsfunktionen der Monopole bei 1825 Hz mit Tikhonov-Regularisierung, hier beispielhaft für die ersten 7 Monopole

Position	Ansteuerungsfunktionen
1	271.83+1109.96i
2	6.91 - 0.34i
3	-150.11 - 325.75i
4	86.65 - 670.75i
5	-117.51+ 474.55i
6	214.05 - 292.03i
7	-252.42 - 958.83i

Mit den optimierten Ansteuerungsfunktionen ergeben sich folgende Rekonstruktionen der Abstrahlcharakteristik.

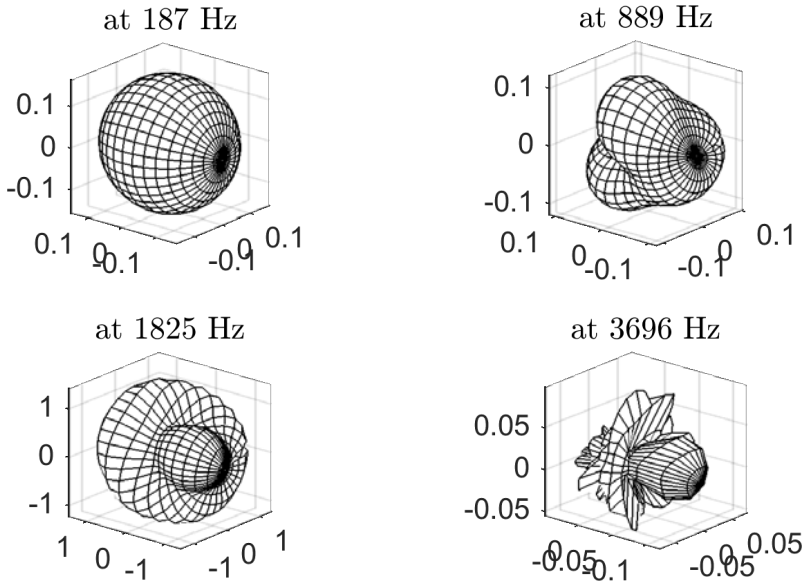

Abbildung 5.4.3: Rekonstruierte Abstrahlcharakteristik mit optimierten Ansteuerungsfunktionen als Betrag der komplexen Werte mit 32 Monopolen mit Tikhonov-Regularisierung.

Das zweite verwendete Regularisierungsverfahren ist die LSQR-Regularisierung. Diese wird durch die Funktion `lsqr_b` berechnet.

```
D = lsqr_b(G,direc,iter)
```

Tabelle 5.4.3: Ansteuerungsfunktionen der Monopole bei 1825 Hz, ermittelt mit dem LSQR-Algorithmus, hier beispielhaft für die ersten 7 Monopole und 15 Iterationen des LSQR-Algorithmus

Position	Ansteuerungsfunktionen
1	0.60 - 1.50i
2	-4.68 - 0.91i
3	0.71 - 0.81i
4	-2.46 - 1.64i
5	-5.61 - 3.25i
6	-0.24 - 0.53i
7	0.72 + 0.02i

Mit den Ansteuerungsfunktionen, welche mit dem LSQR-Algorithmus berechnet

wurden, ergeben sich folgende rekonstruierte Abstrahlcharakteristiken.

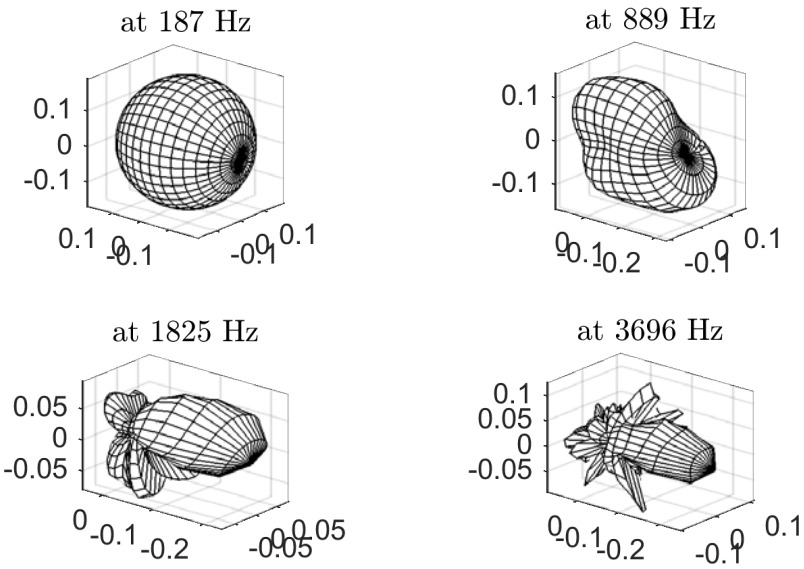

Abbildung 5.4.4: Rekonstruierte Abstrahlcharakteristik, ermittelt mit LSQR-Algorithmus, als Betrag der komplexen Werte mit 32 Monopolen.

Durch die Berechnung der FIR-Filter erhält man die folgenden Ansteuerungsfunktionen.

Tabelle 5.4.4: Optimierte Ansteuerungsfunktionen der Monopole bei 1825 Hz, ermittelt mit dem LSQR-Algorithmus, hier beispielhaft für die ersten 7 Monopole und 15 Iterationen des LSQR-Algorithmus

Position	Ansteuerungsfunktionen
1	-0.64 + 1.45i
2	4.70 + 0.94i
3	-0.73 + 0.78i
4	2.43 + 1.59i
5	5.56 + 3.15i
6	0.23 + 0.52i
7	-0.73 - 0.03i

Die rekonstruierten Abstrahlcharakteristiken sehen dann wie folgt aus.

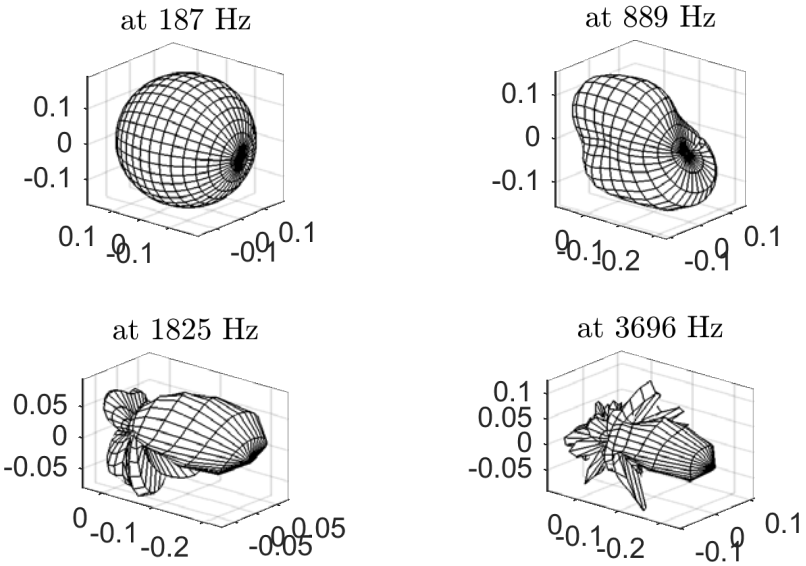

Abbildung 5.4.5: Rekonstruierte Abstrahlcharakteristik mit optimierten Ansteuerungsfunktionen als Betrag der komplexen Werte mit 32 Monopolen, ermittelt mit dem LSQR-Algorithmus.

Die in diesem Abschnitt berechneten Ansteuerungsfunktionen werden als Impulsantworten der jeweiligen Monopolpositionen gespeichert. Dazu können direkt die von der Funktion `filterMatrixOptimalDesign` berechneten Impulsantworten verwendet und als wav-Dateien gespeichert werden. Um die rekonstruierte Abstrahlcharakteristik direkt mit der originalen Abstrahlcharakteristik zu vergleichen, werden auch die Impulsantworten der rekonstruierte Abstrahlcharakteristik berechnet und als wav-Dateien gespeichert. Dazu wird für jede Abtastposition $\mathbf{x}_\mathrm{j}$ die inverse $\mathcal{FFT}$ aller rekonstruierten Abtastwerte $\mathbf{p} = p(\mathbf{x}_\mathrm{j}, f)$ berechnet.

$$p(\mathbf{x}_\mathrm{j}, f) \xrightarrow{\mathcal{IFFT}} IR(\mathbf{x}_\mathrm{j}, t) \tag{5.4.1}$$

Durch den Vergleich der rekonstruierten Abstrahlcharakteristiken (Abbildungen 5.4.2 bis 5.4.5) mit der originalen Abstrahlcharakteristik (Abbildung 5.1.1) lässt sich erkennen, dass rein mathematisch eine gute Rekonstruktion möglich ist. Die deutlichen Abweichungen der rekonstruierten Abstrahlcharakteristik in Abbildung 5.4.3 zur originalen Abstrahlcharakteristik, lassen drauf schließen, dass Ansteuerungsfunktionen der Monopole, mit großen Unterschieden zwischen den Koeffizienten, mit real umsetzbaren Filtern schlecht zu rekonstruieren sind.

Um einen bessern Überblick über die Auswirkungen der verwendeten Verfahren und deren Parameter zu bekommen, wurden die Abweichungen der rekonstruierten Abstrahlcharakteristiken $\mathbf{\Gamma}(\varphi, \vartheta, f)$ gegenüber der originalen Abstrahlcharakteristik $\mathbf{\Gamma}_{orig}(\varphi, \vartheta, f)$ berechnet. Zur Berechnung wurden die rekonstruierten Abstrahlcharakteristiken nach der Berechnung der FIR-Filter verwendet. Die Abweichung *dev* (Deviation) an einer Abtastposition $\mathbf{x} = (\varphi, \vartheta)$ und für eine Frequenzen f wurde dafür mit der Gleichung (5.4.2) ermittelt. Im zweiten Schritt wurden mit Gleichung (5.4.3) die Abweichungen über alle Positionen und Frequenzen gemittelt. Um festzustellen welche Auswirkungen vereinzelt auftretende größere Abweichung haben und ob diese vernachlässigt werden können, wurde zusätzlich auch der Median der Abweichungen mit Gleichung (5.4.4) berechnet.

$$dev = \frac{|\mathbf{\Gamma}(\varphi, \vartheta, f)|}{|\mathbf{\Gamma}_{orig}(\varphi, \vartheta, f)|} \tag{5.4.2}$$

$$dev_{mean} = \frac{1}{O} \sum dev \tag{5.4.3}$$

$$O = Anzahl\ Stützstellen + Anzahl\ Frquenzen$$

$$dev_{med} = median(dev) \tag{5.4.4}$$

Durch die Division von $\mathbf{\Gamma}(\varphi, \vartheta, f)$ und $\mathbf{\Gamma}_{orig}(\varphi, \vartheta, f)$ erhält man die geringste Abweichung, wenn $dev = 1$ ist.

Folgende Grafiken veranschaulichen die Ergebnisse aus der Berechnung der Abweichungen.

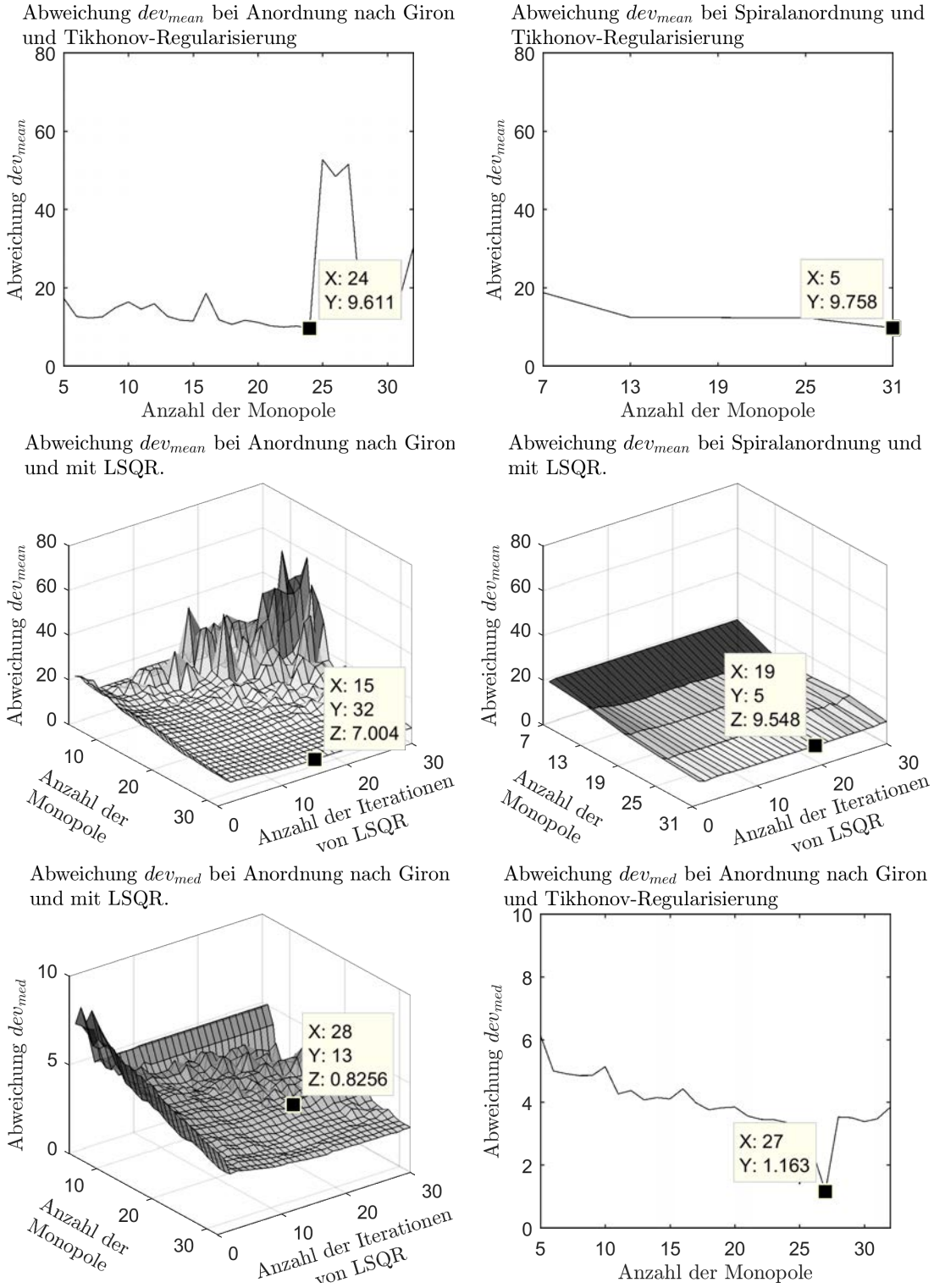

Abbildung 5.4.6: Abweichungen dev_{mean} und dev_{med} der Kombinationen aller verwendeten Verfahren und Parameter. Die angezeigten Datenpunkte geben zeigen die Kombination mit der geringsten Abweichung an.

In Abbildung 5.4.6 wird die Auswirkung der Anzahl der Monopole und die Wahl des Regularisierungsparameters auf die Abweichung dargestellt. Für die Tikhonov-Regularisierung wird der Regularisierungsparameter automatisch durch das L-Kurven Verfahren ausgewählt. Diese automatische Auswahl des Regularisierungsparameters ermöglicht nicht immer die optimale Regularisierung, wie durch einen Vergleich mit den Ergebnissen der LSQR-Regularisierung zu erkennen ist. Durch die LSQR-Regularisierung konnten zum Teil geringere Abweichungen bei gleicher Monopolanordnung erreicht werden. Die Anordnung der Monopole auf Spiralen scheint auch brauchbare Lösungen mit geringer Abweichung zu bringen.

Im Allgemeinen lässt sich feststellen, dass durch eine höhere Anzahl an Monopolen in der Regel eine geringere mittlere Abweichung dev_{mean} erzielt werden kann. Vor allem in der Abbildung der mittleren Abweichung dev_{mean} bei Verwendung des Verfahrens nach Giron und der LSQR-Regularisierung, lässt sich erkennen, dass eine zu geringe oder zu starke Regularisierung zu größeren Abweichungen führt. Gerade eine starke Regularisierung bei wenigen Monopolen führt demzufolge zu großen Abweichungen.

Da die Berechnung des Medians unempfindlich gegenüber vereinzelt auftretenden Abweichungen ist, ist der Median der Abweichungen auch dann gering, wenn vereinzelt auftretende Abweichungen sehr groß sind. Daher haben selben rekonstruierten Abstrahlcharakteristiken zum Teil einen geringen Median der Abweichungen, obwohl der Mittelwert der Abweichungen groß ist. Ob diese vereinzelt auftretenden großen Abweichungen wahrgenommen werden wird mit einem Hörversuch überprüft.

6. Hörversuch

Durch einen qualitativen psychoakustischen Hörversuch soll herausgefunden werden, ob eine Korrelation zwischen der berechneten Abweichung dev und der wahrgenommenen Hörqualität besteht. Dafür werden exemplarisch einige Kombination aus Anordnung der Monopole und Rekonstruktion mit unterschiedlichen Regularisierungsmethoden mit geringer berechneter Abweichung ausgewählt. Zur Anwendung kommt ein Testverfahren, das an das MUSHRA-Verfahren (MUlti Stimulus test with

Hidden Reference and Anchor) nach ITU-R BS.1534-1 angelehnt ist.

6.1. Design

Um die beste Methode der Anordnung und Regularisierung herauszufinden, sollen die Probanden die rekonstruierte Abstrahlcharakteristik mit der originalen Abstrahlcharakteristik vergleichen. Dabei soll auf die Richtwirkung, Unterschiede im Frequenzverhalten und Artefakte geachtet werden. In dem Versuch soll vom Probanden ein Gesamturteil im Sinne von „identisch“ oder „unterschiedlich“ gebildet werden, um generelle Aussagen über die angewendeten Verfahren und ihre Parameter treffen zu können. Für den Hörtest wurden die Kombinationen mit der geringsten Abweichung (siehe Abbildung 5.4.6) gewählt.

Tabelle 6.1.1: Auswahl der Kombinationen für den Hörversuch

Nummer	Anordnung	Regularisierung	Auswahl mit
1	Giron 13 Monopole	LSQR mit 28 Iterationen	$dev_{med} = 0.826$
2	Giron 24 Monopole	Tikh mit L-Curve	$dev_{mean} = 9.611$
3	Giron 27 Monopole	Tikh mit L-Curve	$dev_{med} = 1.163$
4	Giron 32 Monopole	LSQR mit 15 Iterationen	$dev_{mean} = 7.004$
5	Spirale 31 Monopole	LSQR mit 19 Iterationen	$dev_{mean} = 9.548$
6	Spirale 31 Monopole	Tikh mit L-Curve	$dev_{mean} = 9.758$

Durch das Wiedergabesystem werden ein paar Beschränkungen vorgegeben, die sich auf die Art und Weise der Wiedergabe der Testobjekte auswirken. Diese Beschränkungen sind in Abschnitt 6.3 näher erläutert. Aufgrund dieser Beschränkungen wurden für die Wiedergabe im Hörversuch nur Abtastpositionen auf der Azimut-Ebene verwendet ($\vartheta = 90°$) und auf 24 Positionen beschränkt ($\varphi = [0, 10, 30, 50, 70, 80, 90, 100, 110, 130, 150, 170, 180, 190, 210, 230, 250, 260, 270, 280, 290, 310, 330, 350]°$).

Für das Testverfahren wurde als Anker eine Monopolquelle genutzt, da diese mit einer kugelförmigen Abstrahlcharakteristik eindeutig definierbar ist und als größtmöglicher Unterschied zu einer gerichteten Quelle angesehen wird. Als Referenz wurde

die originale Abstrahlcharakteristik verwendet, da diese das Ziel der Rekonstruktion ist.

Als Stimuli wurden Audiodateien gewählt, welche für die Beurteilung des Unterschiedes geeignet sind. Dafür sollen sie einen möglichst großen Frequenzbereich abdecken, um Klangverfärbungen zu erkennen. Des Weiteren sind impulshaltige Stimuli gut, um Verzögerungen (z.B. verschmierte Transienten) zu erkennen. Signale die sich über die Zeit nicht oder kaum ändern, sind gut um die Richtwirkung der Quelle zu beurteilen, da eine Änderung des wahrgenommenen Schalls dann nicht durch das Signal bedingt sind. Daher wurden für den Test die vier folgenden Audiodateien ausgewählt:

Tabelle 6.1.2: Auswahl der Audiodateien für den Hörversuch

Name	Länge in Sekunden	Verwendung
PinkNoise	10	Item 1
Gesang Sopran	11	Training
Klassisches Stück: Smetana - Die Moldau	12	Item 2
Rock: Spin Doctors - Two Princes	11	Item 3

Die Dateien sind Monosignale und wurden mit 48 kHz und 24 Bit aufgenommen. Stereoaufnahmen wurden durch Addition der Kanäle auf ein Kanal reduziert.

Das rosa Rauschen (Pink Noise) wurde gewählt, da es einen großen Frequenzbereich abdeckt und bei allen Frequenzen gleich laut empfunden wird. Der Sopran-Gesang wurde wegen seiner langgezogenen Töne und wenigen Pausen gewählt, da diese das drehen der Quelle bei nahezu konstantem Signal ermöglichen. Auch das klassische Stück ist ohne Unterbrechungen und deckt einen großen Frequenzreich ab, jedoch ändert sich das Signal über die Zeit. Das Rock-Stück enthält impulshaltige Klänge, welche die Beurteilung von zeitlichen Artefakten ermöglicht.

Die Impulsantworten der Abstrahlcharakteristiken der gewählten Kombinationen und der Referenz wurden nun mit den vier Audiodateien gefaltet und stehen somit für die Wiedergabe bereit. Als Anker wurde jeweils direkt die einkanalige Audiodatei verwendet.

Die 6 Kombinationen, die Referenz und der Anker werden als Testobjekte bezeichnet und in dem Wiedergabesystem den 8 Modellen zufällig zugeordnet. Die 4 Audiodateien werden als Training und Item 1 bis 3 bezeichnet.

Tabelle 6.1.3: Zufallstabelle für die Zuordnung der Testobjekte zu den Modellen für die Wiedergabe im Hörversuch

Item	Punktq.	Orig.	Kombi. 1	2	3	4	5	6
Training	8	1	5	3	6	7	4	2
Item 1	8	7	5	1	4	2	6	3
Item 2	3	6	8	5	7	2	4	1
Item 3	2	1	4	7	8	5	3	6

6.2. Aufbau

Der Versuch fand in einem separierten Raum statt, in dem der Proband durch äußere Einflüsse nicht gestört wurde.

Die Quellen wurden mit einem institutseigenen objektbasierten Wiedergabesystem binaural gerendert und über entzerrte Kopfhörer ausgegeben. Dafür werden nicht-personalisierte HRTFs (kopfbezogene Übertagungsfunktion) eines Kunstkopfes benutzt. HRTFs dienen der Lokalisierung von Schallquellen in der binauralen Wiedergabe. Die Wiedergabe basiert auf dem Spaitial Sound Wave (SSW) Algorithmus des Fraunhofer IDMT. Zur Wiedergabe werden komplexe Quellen bestehend aus einfachen Quellen benötigt. Die einfachen Quellen sind für die aktuelle Anwendung Monopole und eine komplexe Quelle besteht aus mehreren Monopolen.

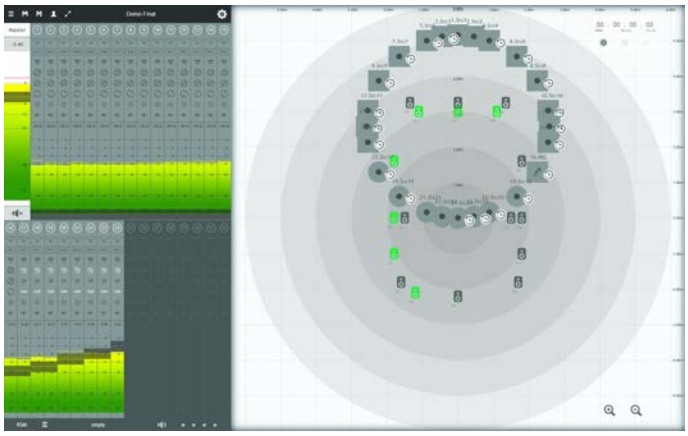

Abbildung 6.2.1: SSW-Renderer

Da wie im Abschnitt 6.3 beschrieben, die berechneten Monopole nicht direkt wiedergegeben werden können, werden die in Abschnitt 6.1 gewählten Positionen der rekonstruierten Abstrahlcharakteristik benutzt. Damit der Hörer die Schallquelle nur aus der ihm zugewandten Richtung hört, werden die Monopole in einem Öffnungswinkel von 120° eingeblendet. Die Öffnung zeigt dabei zum Hörer. Rein theoretisch dürfte der Hörer nur den Monopol hören, welcher den Schall in der Richtung beschreibt, in welcher sich der Hörer aus Sicht der komplexen Schallquelle befindet. Da das aber zu hartem Umschalten zwischen den Monopolen und somit unerwünschten Effekten (Knacken, abrupte Änderung des Signals) führen würde, werden die Monopole mit einem Amplituden-Fenster im beschriebenen Öffnungswinkel ein- und ausgeblendet.

Damit die Abstrahlcharakteristiken aus verschiedenen Richtungen gehört und beurteilt werden können, müssen die Positionen der einfachen Quellen im SSW-Renderer so geändert werden, das die komplexe Quelle virtuell gedreht werden kann. Die Positionen für die einfachen Quellen bekommt der SSW-Renderer von dem 3D-CAD-Tool Visonary Render, wodurch auch die Drehung der komplexen Quelle ermöglicht wird.

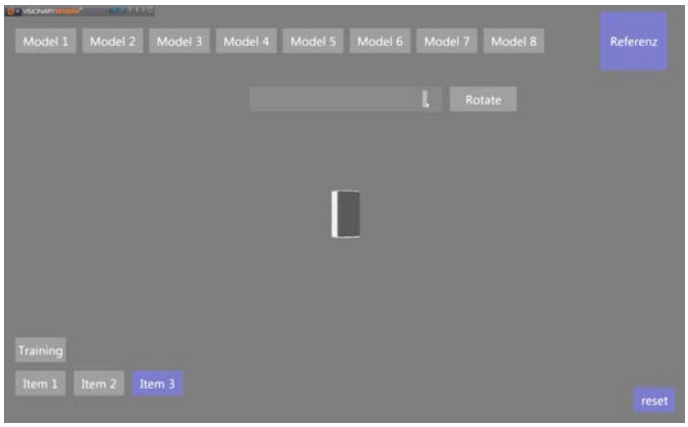

Abbildung 6.2.2: Visonary Render

Außerdem ermöglicht Visonary Render durch weitere Steuerungsoptionen die Umschaltung der Quellen im SSW-Renderer. Dadurch können mehrere komplexe Quellen (Modelle) miteinander verglichen werden. Für diese Arbeit wurden 8 Modelle untereinander und gegenüber einer Referenz verglichen. Die 8 Modelle enthalten die 6 Kombinationen aus Tabelle 6.1.1, sowie einen Monopol als versteckten Anker und die originale Abstrahlcharakteristik als versteckte Referenz. Des Weiteren ist auch die Umschaltung zwischen den vier Audiodateien möglich, welche mit Training und Item 1 bis Item 3 bezeichnet sind.

Da eine Bewertungsoberfläche entsprechend dem Testverfahren im Visonary Render nicht zur Verfügung steht, wurde vom Fraunhofer IDMT eine separate Bewertungsoberfläche (Test-GUI) entwickelt. Durch die gleiche Bezeichnung der Objekte im Visonary Render als auch in der Test-GUI können die Höreindrücke zugeordnet und bewertet werden. Die Bewertungen werden von der Test-GUI gespeichert und stehen somit zur Auswertung zur Verfügung.

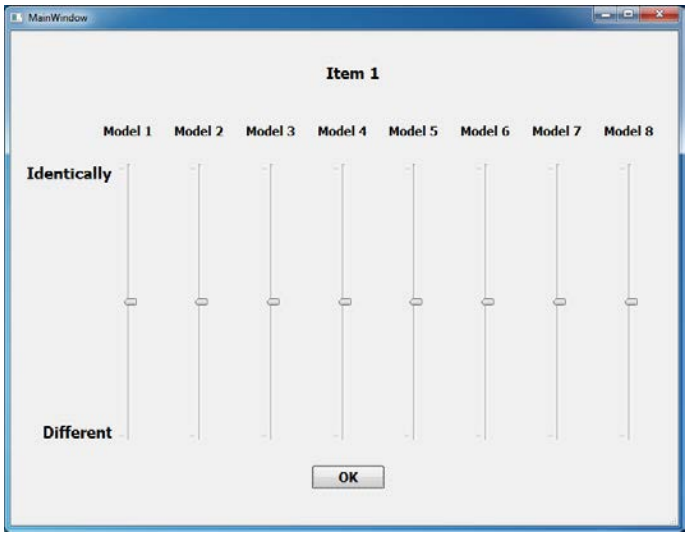

Abbildung 6.2.3: Test-GUI

6.3. Restriktionen des Wiedergabesystems

Für den Hörversuch stand das in Abschnitt 6.2 beschriebene Wiedergabesystem zur Verfügung. Mit diesem System konnte das Schallfeld, welches durch die Überlagerung der ermittelten Monopole entsteht, jedoch nicht korrekt mit der direkten Anordnung der Monopole wiedergegeben werden. Das kann daran liegen, dass numerische oder analytische Berechnungen der Schallfelder physikalische und technische Grenzen der realen Wiedergabe nicht beachten oder anders gewichten. Solche Grenzen ergeben sich zum Beispiel durch die Diskretisierung der Lautsprecher im SSW-Renderer, nicht-personalisierte HRTFs der binauralen Wiedergabe, Limitierungen des elektroakustischen Gesamtsystems oder computernumerische Rundungsfehler. Auch geringe geometrische Verschiebungen zwischen den Monopolen oder eine andere Schallgeschwindigkeit als bei der Berechnung der Monopole, aufgrund einer anderen Lufttemperatur, können zu Abweichungen in der Wiedergabe führen. Somit wurden die, aus den Monopolen rekonstruierten, Richtcharakteristiken wiedergegeben. In dem Wiedergabesystem ist zudem die Anzahl der Monopole auf 32 beschränkt, wodurch im Hörversuch nur eine Auswahl von Monopolen an Abtastpositionen auf der Azimut-Ebene verwendet wird. Da bei der Umschaltung zwischen den Monopolanordnungen alle verwendeten Monopole neu geladen werden, kommt es bei 32

Monopolen zu langen Pausen. Um die Pausen beim Umschalten auf eine akzeptable Zeit zu verkürzen, wurde die Anzahl der Stützstellen auf 24 verringert.

6.4. Versuchsdurchführung

Den Probanden wurde der Versuchsablauf mündlich erklärt und auch schriftlich vorgelegt (siehe Anhang D). Fragen zur Verwendung der Software und zur Beurteilung der Abstrahlcharakteristiken konnten direkt und während der Trainingsphase mit dem Trainingsitem geklärt werden. Während der Trainingsphase konnte sich mit der verwendeten Software vertraut gemacht werden und die Lautstärke auf einen persönlich angenehmen Pegel eingestellt werden. Nach der Trainingsphase wurde der Proband nicht gestört. Er konnte aber jeder Zeit den Versuch abbrechen oder um Hilfestellung bitten, wenn in der Trainingsphase nicht alles verstanden wurde. Der Proband sollte nun für die drei Items die acht Modelle bewerten. Die Lautstärke durfte für jedes Item angepasst werde, musste aber für die acht Modelle je Item gleich bleiben. Nach dem der Versuchsteilnehmer alle Modelle für die drei Items bewertet hatte, wurden die Daten gespeichert. Nach dem Versuch konnten auch weitere Fragen zum Zweck des Versuchs und zum Inhalt der Arbeit gestellt werden.

6.5. Ergebnisse

An dem Versuch haben 18 Probanden vorwiegend vom Fraunhofer IDMT teilgenommen. Davon waren 15 Hörer erfahren und mit psychoakustischen Hörversuchen vertraut und 3 Hörer unerfahren. Keiner der Probanden hat den Versuch abgebrochen oder nach der Trainingsphase noch Fragen zur Versuchsdurchführung gehabt. Die ermittelten Bewertungen wurden in Boxplots zusammen gefasst. Dabei war kein gravierender Unterschied zwischen erfahrenen und unerfahrenen Hörern zu erkennen, weshalb die Gruppen zusammengefasst ausgewertet werden.

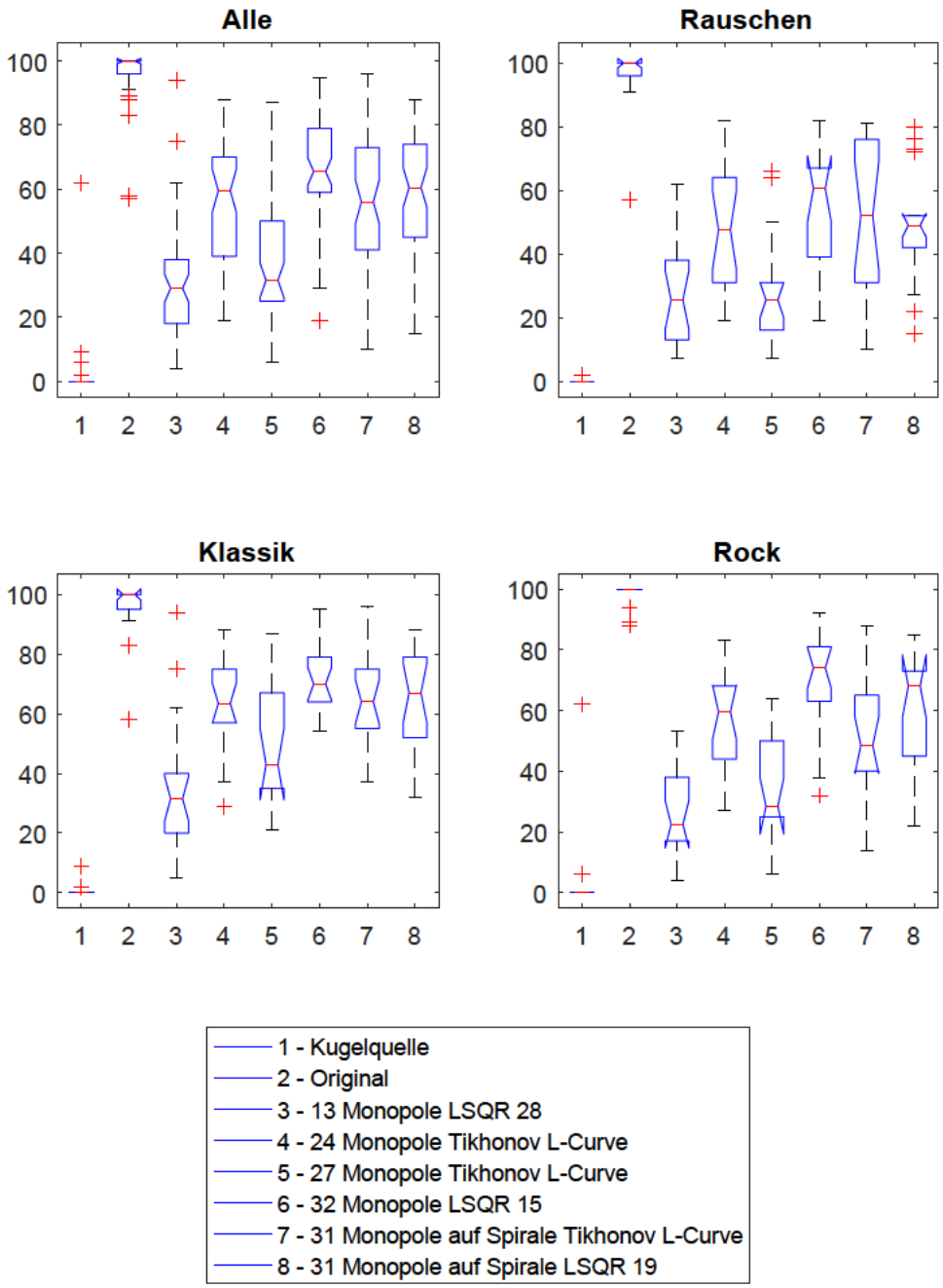

Abbildung 6.5.1: Ergebnisse des Hörversuchs, 100 bedeutet "Die Kombination hört sich an wie die Referenz und hat somit die selbe Abstrahlcharakteristik", 0 bedeutet Ës ist keine Richtwirkung zu erkennen", Alle Werte zwischen 100 und 0 geben an wie stark sich die gehörte Abstrahlcharakteristik von der Referenz unterscheidet

6.6. Auswertung und Verbesserungsvorschläge

In Abbildung 6.5.1 ist eine Tendenz zu erkennen, dass eine größere Anzahl an Monopolen auch eine bessere Rekonstruktion der originalen Abstrahlcharakteristik ermöglicht. Beim Vergleich von Model 4 und 5 fällt jedoch auf, dass mehr Monopole auch Nachteile haben können. Durch eine Vielzahl an Monopolen kann es auch schwierig werden, die Abstrahlcharakteristik korrekt zu rekonstruieren, ohne dass die Pegel der Monopole zu groß werden. Anhand von Model 7 und 8 kann beobachtet werden, dass die Regularisierungsverfahren sich in ihrem Ergebnis nur geringfügig unterscheiden. Jedoch kann bei der iterativen LSQR-Regularisierung die Variante mit einer guten Rekonstruktion und trotzdem noch moderaten Pegeln der Monopole gewählt werden. Bei der Tikhonov-Regularisierung müsste dazu der Regularisierungsparameter auf andere Weise als mit der L-Curve-Methode ermittelt werden.

Beim Vergleich der Ergebnisse des Hörtests zu den berechneten Abweichungen in Abbildung 5.4.6 kann festgestellt werden, dass bei einer geringeren mittleren Abweichung dev_{mean} die rekonstruierte Abstrahlcharakteristik ähnlich der originalen Abstrahlcharakteristik empfunden wird. Somit wird die Kombination (Modell 6) mit der geringsten gemittelten Abweichung ($dev_{mean} = 7$) als am ähnlichsten empfunden.

Die Kombinationen, welche mit dem Median der Abweichungen dev_{med} ausgewählt wurden, wurden am schlechtesten bewertet. Das lässt darauf schließen, dass vereinzelte große Abweichungen gut wahrgenommen werden und der Median der Abweichungen nicht zur Beurteilung der Abweichungen geeignet ist.

Die Streuung der Ergebnisse ist recht hoch, wodurch nur Tendenzen zu erkennen sind. Für genauere Ergebnisse müsste der Hörversuch in einzelne Parameter aufgeschlüsselt werden. Denkbar wären Parameter wie Vor-Rückverhältnis, Frequenzspektrum, spatiale Artefakte. Auch die Skala wird von jedem Probanden anders eingeschätzt. Hier müssten aussagekräftige Ankerpositionen gewählt werden, woran sich der Proband orientieren kann. Außerdem wäre eine separate Untersuchung von Verfahren zur Anordnung der Monopole und zur Regularisierung sinnvoll, um feinere Abstufungen der einzelnen Verfahren zu untersuchen.

7. Einschätzung der verwendeten Verfahren und von bekannten Fehler der Implementierung

Mit den gewählten Verfahren ist es möglich Anordnungen von Monopolen zu bestimmen, mit denen durch die verwendeten Regularisierungsverfahren die Rekonstruktion von Abstrahlcharakteristiken in einem breiten Frequenzbereich möglich ist. Im mathematischen Sinne wurden bei der Berechnung dafür einige Randbedingungen vernachlässigt. So wurden die originalen Daten nicht in der Frequenz beschränkt und die Ordnung der $\mathcal{DSHT}$ wurde auf 12 festgesetzt und nicht mit der Gleichung (2.3.7) berechnet. Das führt dazu, dass im Hörversuch breitbandige Audiodateien verwendet werden können. Jedoch können dadurch auch räumliche Aliasing-Effekte auftreten. Die Auswirkung solcher Artefakte konnten mit dem durchgeführten Hörversuch nicht geklärt werden und sind auch nicht erkannt worden. Daher sind weitere, detailliertere Hörversuche zur Verbesserung des Verfahrens notwendig.

7.1. Spiralanordnung der Monopole

Die Anordnung der Monopole auf Spiralen bietet eine gute Möglichkeit um Abstrahlcharakteristiken gerade in einem kleinen Frequenzbereich zu Rekonstruieren. Für einen größeren Frequenzbereich sind mehr Monopole notwendig, beziehungsweise werden bei gleicher Anzahl an Monopolen die Lücken in der Anordnung größer und eine gute Rekonstruktion ist nicht mehr möglich. Da die Anzahl der Monopole durch das Wiedergabesystem begrenzt ist, kann diese auch nicht ohne Ende erhöht werden und wiederum sind auch zu viele Monopole ungünstig, da das Finden von realistischen Ansteuerungsfunktionen durch die gegenseitige Beeinflussung der Monopole erschwert wird. Somit muss das Verfahren und die Einstellung seiner Parameter noch weiter untersucht werden. Der große Vorteil dieses Verfahrens ist die geringe Berechnungszeit der Monopolpositionen gegenüber der Monopoldekomposition nach Giron.

7.2. Monopoldekomposition nach Giron

Das Verfahren nach Giron bietet momentan, die wohl beste Möglichkeit Monopolpositionen für beliebig geformte Abstrahlcharakteristiken zu finden. Die Implementierung wurde direkt nach der Dissertation von Franck Giron erarbeitet und mit der Beispielberechnung aus der Dissertation getestet. Bei der Verwendung mit realen mit Monkey Forest gemessenen Daten kam es jedoch zu Schwierigkeiten bei der Rekonstruktion. Die Rekonstruktion von Abstrahlcharakteristiken klappte gut, war jedoch um einen Azimut-Winkel gedreht. Daher wird nun der Ansatz verfolgt, die Winkel der Abtastpunkte zu verwenden, um die sphärischen Harmonischen und somit die Monopolpositionen zu bestimmen. Eine Überarbeitung der Winkelabhängigkeiten ist jedoch weiterhin ratsam, um eine korrekte Lösung unabhängig von der Quellkonvention zu erhalten.

Wie oben erwähnt wurden einige Randbedingungen vernachlässigt, welche aber je nach Ordnung der $\mathcal{DSHT}$ erst bei hohen Frequenzen zu Aliasing führen. Aliasing Artefakte wurden im durchgeführten Hörversuch nicht separat untersucht. Daher sollte ein einem weiteren Hörversuch überprüft werden, wie stark dieses Aliasing wahrgenommen wird.

7.3. Tikhonov-Regularisierung mit Regularisierungsparameter nach L-Kurven Verfahren

Bei der Tikhonov-Regularisierung scheint die Wahl des Regularisierungsparameters durch das L-Kurven Verfahren nicht die beste Wahl zu sein. Häufig ist die Rekonstruktion der Abstrahlcharakteristik damit zwar sehr gut, die Norm der Ansteuerungsfunktion ist aber dementsprechend recht groß, weshalb die so gefundenen Ansteuerungsfunktion schlecht realisiert werden können.

7.4. LSQR-Regularisierung

Die LSQR-Regularisierung bietet ein gutes Verfahren, um einen gute Balance zwischen Fehlernorm und Norm des Lösungsvektors zu finden.

8. Fazit und Ausblick

Die in dieser Arbeit untersuchten und weiterentwickelten Verfahren ermöglichen die Zerlegung eines realen Datensatzes von Abstrahlcharakteristiken in ein Set von Monopolen. Diese Anordnung der Monopole ist für den verwendeten Frequenzbereich konstant und die Ansteuerungsfunktionen können mit geringer Abweichungen zwischen den Koeffizienten berechnet werden. Daher kann das Schallfeld für einen großen Frequenzbereich durch die Monopolsynthese berechnet werden und mit einem existierenden 3D-Audio-System wiedergegeben werden. Die mit den verwendeten Verfahren erzeugten Monopole, konnten auf Grund von Artefakten jedoch nicht direkt in einem 3D-Audiosystem wiedergegeben werden. Die Rekonstruktion der Abstrahlcharakteristik aus den Monopolen zeigt aber, dass sich die Monopole theoretisch Korrekt überlagern. Daher kann die Berechnung der Monopole weiter verbessert werden, um auch eine korrekte Wiedergabe zu erreichen.

Mit einem psychoakustischen Hörversuch wurden die gewählten Parameter der verwendeten Verfahren, im Bezug auf die Abweichung der rekonstruierten Abstrahlcharakteristik zur ursprünglichen Abstrahlcharakteristik, untersucht. Für den Hörversuch wurde die gesamte Signalverarbeitungskette vom realen Datensatz über die Monopolzerlegung und Monopolsynthese bis hin zur Wiedergabe aufgestellt. Dadurch wurde die Wiedergabe in einem realen Umfeld getestet. Die während des Hörversuch getesteten Abstrahlcharakteristiken haben rechnerisch zum Teil eine große Übereinstimmung mit der ursprünglichen Abstrahlcharakteristik. Die Bewertung der Abstrahlcharakteristiken durch die Probanden lässt die Tendenz erkennen, dass diese Übereinstimmung so wahrgenommen wurde. Abstrahlcharakteristiken, die rechnerisch stärker von der ursprünglichen Abstrahlcharakteristik abweichen, wurden von den Probanden tendenziell auch mit einer größeren Abweichung bewertet. Somit können Tendenzen in der Korrelation zwischen den Parametern und den Ergebnissen des Hörversuchs beobachtet werden. Bei allen untersuchten Verfahren lässt sich erkennen, dass durch die Verwendung von mehr Monopolen sich die Abweichungen verringert. Bei einer zu starken oder zu schwachen Regularisierung werden die Abweichungen größer. Es muss also bei den Regularisierungparametern ein guter

Kompromiss zwischen Abweichung der Abstrahlcharakteristik und Abweichungen in den Ansteuerungsfunktionen gefunden werden.

Durch die Spiralanordnung können die Richtcharakteristiken gut rekonstruiert werden. Die minimalen Abweichungen sind jedoch nicht kleiner als die der Anordnung der Monopole durch das Verfahrens nach Giron (1996). Ein Vorteil der Spiralanordnung ist die geringe Berechnungszeit der Monopolpositionen, gegenüber des Verfahrens nach Giron. Hier muss der Anwender zwischen einer kurzen Berechnungszeit oder einer genauen Rekonstruktion wählen.

Die erarbeiteten MatLab-Skripte stehen für weitere Forschungen zur Verfügung und können natürlich optimiert werden. Sie ermöglichen grundlegende Berechnungen der Monopoldekomposition und Monopolsynthese von realen Abstrahlcharakteristiken.

Mit der vorliegenden Arbeit wurde eine Lücke in der Signalverarbeitungskette, von einem realen Datensatz von Abstrahlcharakteristiken hin zur Wiedergabe von gerichteten Schallquellen mit 3D-Audio-Systemen, geschlossen.

Abbildungsverzeichnis

1.0.1.Mögliches Szenario der virtuellen Begutachtung 3
2.1.1.Koordinatensystem 6
2.1.2.Fast-Fourier-Transformation 7
2.2.1.Prinzipielle Anordnung zur Monopolsynthese 8
3.1.1.Spirale mit exponentiell wachsendem Radius 16
3.1.2.Spiralen in drei Dimensionen 17
3.2.1.Zusammenfassung der Schritte zum Finden der Monopole 18
4.3.1.L-Kurve 23
5.1.1.Originale Abstrahlcharakteristik 25
5.3.1.Sphärisches Amplitudenspektrum 27
5.3.2.Korrelationsmatrix 28
5.4.1.L-Kurve des Beispiels 31
5.4.2.Rekonstruierte Abstrahlcharakteristik Tikhonov 32
5.4.3.Rekonstruierte Abstrahlcharakteristik mit optimierten Ansteuerungsfunktionen Tikhonov 33
5.4.4.Rekonstruierte Abstrahlcharakteristik mit LSQR 34
5.4.5.Rekonstruierte Abstrahlcharakteristik mit optimierten Ansteuerungsfunktionen mit LSQR 35
5.4.6.Abweichung *dev* 37
6.2.1.SSW-Renderer 42
6.2.2.Visonary Render 43
6.2.3.Test-GUI 44
6.5.1.Ergebnisse des Hörversuchs 46

Tabellenverzeichnis

5.3.1.Positionen der Monopole bei 1825 Hz 29
5.3.2.Positionen der Monopole für alle Frequenzen 30
5.4.1.Ansteuerungsfunktionen Tikhonov 31

5.4.2.Optimierte Ansteuerungsfunktionen Tikhonov 32
5.4.3.Ansteuerungsfunktionen mit LSQR 33
5.4.4.Optimierte Ansteuerungsfunktionen mit LSQR 34
6.1.1.Auswahl der Kombinationen für den Hörversuch 39
6.1.2.Auswahl der Audiodateien für den Hörversuch 40
6.1.3.Zufallstabelle 41
A.0.1.Verwendete MatLab Funktionen vom Fraunhofer IDMT 58

Abkürzungsverzeichnis

MUSHRA	**MU**lti **S**timulus test with **H**idden **R**eference and **A**nchor
ITU	**I**nternational **T**elecommunication **U**nion
GUI	**G**rafical **U**ser **I**nterface (Grafische Benutzeroberfläche)
RAR	**R**eflexions**A**rmer **R**aum
CAD	**C**omputer **A**ided **D**esign
GCV	**G**eneralized **C**ross-**V**alidation
FIR	**F**inite **I**mpulse **R**esponse
LTI	**L**inear **T**ime **I**nvariant
HRTF	**H**ead-**R**elated **T**ransfer **F**unction

Glossar

Abstrahlcharakteristik $\Gamma(\varphi, \vartheta, f)$, auch Richtcharakteristik einer Schallquelle:	Frequenz- und winkelabhängige Darstellung einer Schallfeldgröße (z.B. Schalldruck) im Fernfeld eines Schallstrahlers normiert auf die Bezugsrichtung (Azimut φ und Elevation $\vartheta = 0°$) (DIN 1320:2009)
Frequenzbin	diskreter Abtastwert bei einer Frequenz, ergibt sich durch die $\mathcal{FFT}$ eines digitalisierten Zeitsignals

Normen und Standards

DIN 1320:2009	Akustik – Begriffe: Definitionen von Begriffen in der Akustik
ITU-R BS.1534-1	Method for the subjective assessment of intermediate quality level of coding systems (including MUSHRA)

Literatur

Bergner, Jakob; Tobias Clauss; Zhykhar Albert; Sladeczek Christoph und Brix Sandra (2016): „Application of Wave Field Synthesis in Virtual Acoustic Engineering.“ In: *Proceedings of INTER-NOISE 2016*, vol. 45. Hamburg: Deutsche Gesellschaft für Akustik e.V, S. 3509–3516.

Bergner, Jakob; Christoph Sladeczek und Johannes Redlich (2017): „Perception-based Investigations on the Monopole Synthesis for Reproduction of Directional Sound Sources.“ In: *Fortschritte der Akustik - DAGA 2017*. Kiel: Deutsche Gesellschaft für Akustik e.V, S. 1426–1429.

Deuflhard, Peter und Andreas Hohmann (2008): *Numerische Mathematik. 1: Eine algorithmisch orientierte Einführung*, Band 1 von *De-Gruyter-Lehrbuch*. 4., überarb. und erw. Aufl.. Berlin: de Gruyter.

Driscoll, James R. und Dennis M. Healy, Jr. (1994): „Computing Fourier Transforms and Convolutions on the 2-Sphere.“ In: *Advances in Applied Mathematics*, 15(2) S. 202–250.

Escolano, José; José J. López und Basilio Pueo (2007): „Directive sources in acoustic discrete-time domain simulations based on directivity diagrams.“ In: *The Journal of the Acoustical Society of America*, 121(6) S. EL256–EL262.

Giron, Franck (1996): *Investigations about the directivity of sound sources*. Ph.D. thesis, Ruhr-Universität Bochum, Bochum.

Hansen, Per Christian (1994): „Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems.“ In: *Numerical Algorithms*, 6(1) S. 1–35.

Hansen, Per Christian (2007): „Regularization Tools version 4.0 for Matlab 7.3.“ In: *Numerical Algorithms*, 46(2) S. 189–194.

Kaschenz, Julia (2006): *Regularisierung unter Berücksichtigung von Residuentoleranzen.* Ph.D. thesis, Technische Universität Berlin, Berlin.

Müller-Trapet, M. und M. Vorländer (2014): „Berücksichtigung der Quellenrichtcharakteristik bei der in-situ Messung von Absorptionsgraden.“ In: *Fortschritte der Akustik - DAGA 2014.* Oldenburg: Deutsche Gesellschaft für Akustik e.V, S. 232–233.

Münch, Hannes; Stefan-Georg Backhaus und Stefan Becker (2017): „Modellentwicklung für die Auralisation von Industrieprodukten für 3D-VR-Anwendungen.“ In: *Fortschritte der Akustik - DAGA 2017.* Kiel: Deutsche Gesellschaft für Akustik e.V, S. 1073–1076.

Ochmann, Martin (1990): „Die Multipolstrahlersynthese – ein effektives Verfahren zur Berechnung der Schallabstrahlung von schwingenden Strukturen beliebiger Oberflächengestalt.“ In: *Acta Acustica united with Acustica*, 72(4) S. 233–246.

Ochmann, Martin (1995): „The source simulation technique for acoustic radiation problems.“ In: *Acta Acustica united with Acustica*, 81(6) S. 512–527.

Opdam, Rob; Kira Latzke; Gottfried Behler und Michael Vorländer (2016): „Directivity as Monopole Decomposition for Active Noise Reduction.“ In: *Fortschritte der Akustik - DAGA 2016.* Aachen: Deutsche Gesellschaft für Akustik e.V, S. 461–463.

Paige, Christopher C. und Michael A. Saunders (1982): „LSQR: An algorithm for sparse linear equations and sparse least squares.“ In: *ACM transactions on mathematical software*, 8(1) S. 43–71.

Papula, Lothar (2009a): *Mathematik für Ingenieure und Naturwissenschaftler*, Band. 1. 12 Aufl. Wiesbaden: Vieweg+Teubner.

Papula, Lothar (2009b): *Mathematik für Ingenieure und Naturwissenschaftler*, Band. 2. 12 Aufl. Wiesbaden: Vieweg+Teubner.

Slavik, Karl M. und Stefan Weinzierl (2008): „Wiedergabeverfahren, 3D Audio.“ In: Stefan Weinzierl (Hrg.) *Handbuch der Audiotechnik*, 1 Aufl. Berlin Heidelberg: Springer, S. 657–685.

Williams, Earl G. (1999): *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography.* San Diego [u.a.]: Academic Press.

Anhang A Verwendete MatLab Funktionen vom Fraunhofer IDMT

Die vom Fraunhofer IDMT zur Verfügung gestellten und in dieser Arbeit verwendeten Funktionen sind in folgender Tabelle aufgelistet. Einige der Funktionen sind von Jakob Bergner schon vor seiner Arbeit beim Fraunhofer IDMT implementiert worden.

Tabelle A.0.1: Verwendete MatLab Funktionen vom Fraunhofer IDMT

Funktion	Verwendungszweck
`besselh_sph`	Berechnung der sphärischen Hankelfunktionen erster und zweiter Art.
`besselj_sph`	Berechnung der sphärischen Besselfunktion der ersten Art.
`filterMatrixOptimalDesign`	Berechnung von FIR-Filtern; benötigt die CVX Toolbox (http://cvxr.com/cvx/).
`getLogSpiral`	Erzeugung der kartesischen Koordinaten von logarithmischen Doppelspiralen im zwei- und dreidimensionalen Raum.
`getMonopoleDistribution`	Erzeugung von Monopolpositionen auf Spiralen im zwei- und dreidimensionalen Raum.
`getTransferDistances`	Berechnung der Abstände zwischen den Monopolen und den Abtastpunkten der Abstrahlcharakteristik.
`mySph2Cart`	Umrechnung von sphärischen in kartesische Koordinaten; mit Azimut $[0, 2\pi)$ und Elevation $[0, pi]$.
`sphHarm_N`	Berechnung der sphärischen Harmonischen bis zu einer Ordnung `N`.

Anhang B Ermittelung der lokalen Maxima in einer dreidimensionalen Matrix

Die lokalen Maxima werden durch den Gradienten aller Werte der umliegenden Punkte zu dem beobachteten Punkt bestimmt. Nur wenn kein Wert der benachbarten Punkte größer ist als der beobachtete, ist der Wert des beobachteten Punktes ein lokales Maximum. Als benachbarte Punkte zählen sowohl die auf der aktuellen Sphäre mit dem Azimutabstand `dph` und dem Elevationsabstand `dth` und die direkt darunter und darüber liegenden Punkte mit dem Radiusabstand `dr`. Die Abstände entsprechen in der Regel der Verschiebung um einen Punkt. Größere Abstände führen bei dieser Funktion zu Punkten, die nicht dem lokalen Maximum entsprechen, sondern darum liegen und nur an einer Stelle auch dem tatsächlichen lokalen Maximum entsprechen. Die Funktion wurde nach der Grundlage von Appendix B aus Giron (1996) geschrieben.

```
%POSLOCMAXSPH research of positive local extreme values in a 3D-Matrix
%   with spherical coordinates
%   extreme = POSLOCMAXSPH(S,dr,dph,dth)

%   (C) 2017 JOHANNES REDLICH - Audiocommunication Group, TU Berlin
%   johannes@redlichs.de
%   developed due to master's thesis
%   Refernce: Giron, Franck (1996): Investigations about the
%   directivity of sound sources. Ph.D. thesis, Ruhr-Universitaet
%   Bochum, Bochum, Appendix B.

function extreme = poslocmaxsph(S,dr,dph,dth,extreme)
if nargin<2
    dr = [1 0 0];
    dph = [0 1 0];
    dth = [0 0 1];
end
if nargin<3
    dph = [0 1 0];
    dth = [0 0 1];
end
if nargin<4
    dth = [0 0 1];
end

ng = ones(10,3);
i = 0;
```

```
Sextreme = zeros(10,1);

for rrr = 1:size(S,3)
    for ph = 1:size(S,1)
        for th = 1:size(S,2)
            r = [rrr,ph,th];
            ng(1,:) = r-dr;
            ng(2,:) = r+dr;
            ng(3,:) = r-dph;
            ng(4,:) = r+dph;
            ng(5,:) = r-dth;
            ng(6,:) = r+dth;
            ng(7,:) = r-dth-dph;
            ng(8,:) = r-dth+dph;
            ng(9,:) = r+dth-dph;
            ng(10,:) = r+dth+dph;
            % Randprobleme beachten, -> wieder zum anderen Ende springen
            ng(ng(:,1)<1,1) = 1;
            ng(ng(:,2)<1,2) = size(S,1)-dph(2)+ph;
            ng(ng(:,3)<1,3) = size(S,2)-dth(3)+th;
            ng(ng(:,1)>size(S,3),1) = size(S,3);
            ng(ng(:,2)>size(S,1),2) = 1+dph(2)-(size(S,1)-ph);
            ng(ng(:,3)>size(S,2),3) = 1+dth(3)-(size(S,2)-th);
            for z = 1:10
                % Vorzeichenprobleme bei um die 90 Grad Elev beachten
                Sextreme(z) = abs(S(r(2),r(3),r(1)))-...
                    abs(S(ng(z,2),ng(z,3),ng(z,1)));
            end
            if min((Sextreme))>0
               i = i+1;
               extreme(i,:) = [r, S(r(2),r(3),r(1))];
            end
        end
        if ph == 1 && all(abs(S(1,:,r(1)))==abs(S(1,1,r(1)))) && ...
                abs(S(1,1,r(1)))>abs(S(1,1,ng(1,1))) && ...
                abs(S(1,1,r(1)))>abs(S(1,1,ng(2,1))) && ...
                all(abs(S(1,1,r(1)))>= abs(S(2,:,r(1))))
            i = i+1;
            extreme(i,:) = [r, S(r(2),r(3),r(1))];
        end
    end
end
```

Anhang C Monopole zusammenfassen

```
%% FindDuplMono searches for duplicate monopoles and gives back the
% monopoles with the highest correlation, limited by maxmono2

function giron_all = FindDuplMono(giron_all_1,maxmono2)

% Sorted by correlation
giron_edit = flipud(sortrows(reshape(permute(giron_all_1,[2,1,3]),...
    4,[])',4));
giron_same = zeros(size(giron_edit,1),size(giron_all_1,1));
corr = zeros(size(giron_edit,1),1);
for i = 1:size(giron_edit,1)
    % Tolerance of Radius r is +-10%
    gedi = find(all((giron_edit(:,1:2)==giron_edit(i,1:2))').*...
        (giron_edit(:,3)<=(giron_edit(i,3)+giron_edit(i,3)*0.1)&...
        giron_edit(:,3)>=(giron_edit(i,3)-giron_edit(i,3)*0.1))');
    giron_same(i,1:length(gedi)) = gedi;
    corr(i) = sum(giron_edit(gedi,4));
end
giron_uni = unique(giron_same(giron_same(:,1)>0),'first','rows');
giron_edit1 = giron_edit(giron_uni(:,1),:);
% sorting by sum of correlation of duplicate positions
giron_edit1(:,4) = corr(giron_uni);
giron_edit1 = flipud(sortrows(giron_edit1,4));
%% Choosing number of monopoles for reconstruction
giron_all = ones(maxmono2,4,size(giron_all_1,3))*(-1);
gg = size(giron_edit1,1);
if gg>maxmono2
   gg = maxmono2;
end
for fn = 1:size(giron_all_1,3)
    giron_all(1:gg,:,fn) = giron_edit1(1:gg,:);
end

end
```

Anhang D Versuchsanleitung

Die Probanden haben zu dem Hörversuch folgende kurze Versuchsanleitung bekommen. Fragen zum Versuchsablauf konnten gestellt werden und der Versuch wurde, auch mündlich erklärt.

Versuchsanleitung:

Willkommen zum Hörversuch:

Der Versuch ist an das MUSHRA-Verfahren angelehnt.

Es soll die originale Richcharakteristik (Referenz) mit den verschiedenen Testobjekten (Model 1...8) verglichen werden.

Dabei kommt es nur auf die wahrgenommene Übereinstimmung zwischen der Referenz und dem Testobjekt an. Womöglich auftretende Artefakte oder Unterschiede im Frequenzgang können in die Beurteilung des Unterschiedes mit einfließen.

In dem Testfenster sehen Sie 8 Slider entsprechend den 8 Modellen, die verglichen werden sollen. Die Slider reichen von different (unterschiedlich) bis identically (identisch). Diese Regler sollen Sie nach Ihrer persönlichen Einschätzung des Unterschiedes anpassen.

Begonnen wird mit einem Trainingsobjekt, daraufhin folgen 3 Testobjekte.

Hintergrund:

Eine Richtcharakteristik ist die Richtwirkung einer Schallquelle, das bedeutet welcher Schallpegel wird von der Quelle in eine bestimmte Richtung abgestrahlt. Die Richtcharakteristik ist frequenzabhängig und kann daher für verschiedene Frequenzen unterschiedlich sein. Im Umkehrschluss bedeutet das, dass eine Schallquelle aus unterschiedlichen Richtungen unterschiedlich klingt, nicht nur leiser oder lauter, es kann sich auch das Frequenzspektrum oder die Phase ändern. Die zuletzt genannten Kriterien gehen dann in die Bewertung des Unterschiedes mit ein.