

Schalleistungsberechnung von HATS-Daten der DirPat-Datenbank und Klangfarbenanalyse von Gesangsaufnahmen im Doppelkreis-Mikrofon-Array

Paul Armin Bereuter

Betreuung: Univ. Prof. DI Dr. Alois Sontacchi

Graz, Juni 2018



institut für elektronische musik und akustik



Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am 24. Juni 2018

.....
(Unterschrift)

Zusammenfassung

Diese Bachelorarbeit gliedert sich in zwei größere Kapitel. Im ersten Kapitel wird eine mit der Softwareapplikation Matlab durchgeführte Schalleistungsberechnung dokumentiert. Bevor der Berechnungsprozess in Matlab genauer erläutert wird, werden theoretische Hintergründe und Vergleiche mit relevanten Normen durchgeführt. Die Schalleistungsberechnung und Diskussion der Ergebnisse beziehen sich ausschließlich auf einen vermessenen Kunstkopf (engl. Head and Torso Simulator). Der Kunstkopf ist in einem speziellen Doppelkreis Mikrofon Array vermessen worden. Die dazu notwendige Datenaufbereitung wird im dazugehörigen Unterkapitel nach den theoretischen Hintergründen erklärt. Im gleichen Mikrofonaufbau sind auch Gesangsaufnahmen gemacht worden. Die Aufnahmen dienen dazu, die richtungsabhängige Klangfarbenänderung genauer untersuchen zu können. Dieses Themengebiet bildet den Inhalt des zweiten Kapitels dieser Arbeit. Die Klangfarbenanalyse wird anhand von Audiodeskriptoren durchgeführt. Diese Audiodeskriptoren werden mittels einer speziell auf die Klangfarbenanalyse zugeschnittenen Toolbox in Matlab berechnet. Es werden auch im zweiten Kapitel wieder theoretische Hintergründe angeführt. Danach werden durchgeführte Nachbearbeitungen der Aufnahmen im Unterpunkt der Datenaufbereitung angeführt. Abschließend wird die Struktur des für die Berechnung der Deskriptoren geschriebenen Matlab-Codes angeführt, ehe dann die eigentliche Klangfarbenanalyse im Unterpunkt Ergebnisse durchgeführt wird. Ziel dieser Arbeit ist es, einen einfachen Berechnungsweg zur Schalleistungsberechnung im Doppelkreis Mikrofon Array zu erstellen und dessen Genauigkeit mithilfe einer Fehlerabschätzung zu dokumentieren. Weiters soll eine systematische Methode zur Klangfarbenanalyse der Gesangsaufnahmen im Doppelkreis Mikrofon Array mit Matlab erarbeitet werden. Abgeschlossen wird die Arbeit durch das Anführen der positiven und negativen Aspekte, die bei Verwendung der auf Klangfarbenanalyse spezialisierten Matlab-Toolbox streten.

Abstract

The present bachelor thesis is subdivided into two major chapters. The first one deals with the documentation of sound power calculations which were conducted by means of the software application Matlab. Before the calculation process in Matlab is elaborated, theoretical backgrounds and comparisons with relevant norms are discussed. The sound power calculations as well as the discussion of the findings exclusively relate to a head and torso simulator which was measured in a special double circle microphone array. An explanation of the necessary data elaboration can be found in the subchapter following the theoretical background. The same microphone array was used to produce vocal recordings: These were used to examine changes of timbre dependent on different directions. The second chapter of this thesis deals with this topic. The analysis of timbres is carried out with the help of audio descriptors which were calculated by means of a toolbox in Matlab which was especially designed for this purpose. The second chapter, similar to the first one, comprises a survey of theoretical backgrounds which is followed by a subchapter on data processing where the editing/postprocessing of the recordings are summed up. Eventually, the structure of the Matlab Code used for the calculation of the audio descriptors is cited. Additionally, the actual analysis of timbres and tones is carried out in chapter "findings". Firstly, the purpose of the present thesis is to find a simple method of the calculation of sound power in the double circle microphone array and document its precision by estimating errors. Moreover, Matlab shall be used to find a systematic method to analyse the timbres and tones of the vocal recordings in the double circle microphone array. The thesis concludes with a list of positive and negative aspects accompanying the use of the Matlab-toolbox designed for the analysis of timbres.

Inhaltsverzeichnis

1	Einleitung	7
2	Schalleistungsberechnung	8
2.1	Theorie	8
2.1.1	DirPat-Database and Viewer	8
2.1.2	Das SOFA-Format	10
2.1.3	Die Schalleistung	11
2.1.4	Messung der Schalleistung	13
2.2	Datenaufbereitung	17
2.2.1	Vom SOFA-Format zu Impulsantworten	17
2.2.2	Aufteilen der Impulsantworten in Terzbänder	18
2.2.3	Von Impulsantwort zu Schalldruckwert	22
2.3	Durchführung der Berechnung	22
2.3.1	Struktur des Matlabcodes	23
2.3.2	Die Flächeneinteilungen der Kugel	25
2.4	Ergebnisse	30
3	Klangfarbenanalyse	33
3.1	Theorie	33
3.1.1	Klangfarbendefinition	34
3.1.2	Die Timbre Toolbox	36
3.1.3	Verwendete Werkzeuge und Deskriptoren	37
3.2	Datenaufbereitung und Durchführung	39
3.2.1	Aufteilung der Gesangsaufnahme	39
3.2.2	Einsatz der Timbre Toolbox	40
3.2.3	Darstellung der Ergebnisse	41
3.2.4	Struktur des Matlabcodes	42
3.3	Ergebnisse	43
4	Diskussion und Ausblick	48
4.1	Schalleistungsberechnung	48
4.2	Klangfarbeanalyse	49

5	Schlusswort	51
6	Anhang	52

1 Einleitung

Im Bereich der Musik spielt die Erfahrung eine große Rolle. So basierte der historische Instrumentenbau vorerst nur auf Erfahrungswerten. Erst mit dem Fortschritt der Naturwissenschaften fanden die Mathematik und die Physik auch Einzug in die Musik.

Ein namhaftes Beispiel zur Verknüpfung zwischen Musik und Technik lässt sich bei Theobald Böhm finden. Er entwarf in der ersten Hälfte des 19. Jahrhunderts ein neues Griff- und Lochsystem für Flöten, welches zur Gänze auf akustischen Grundlagen und mathematischen Berechnungen basiert. Das von ihm entworfene Konzept hob sich durch bessere Spielbarkeit und einen volleren Klang von bisherigen Flötenbausystemen ab [Mel16]. Böhms System findet bis zum heutigen Tag Anwendung. Er bewies, dass die wissenschaftliche Herangehensweise an akustische Begebenheiten und Probleme eine systematische und nachvollziehbare Möglichkeit liefert, akustische Eigenschaften von Instrumenten zu verbessern. So ist auch einer dieser Aspekte die Abstrahlcharakteristik von Instrumenten.

Die Abstrahlcharakteristik ist für einen Tonmeister von großer Bedeutung und hat maßgeblichen Einfluss auf die Positionierung von Mikrofonen. Sie ist jedoch nicht nur bei Instrumenten zu beachten, sondern auch bei der Gesangsstimme. Die Bestimmung bzw. Berechnung von Abstrahlcharakteristika speziell von Gesangsstimmen und deren Beeinflussungsmöglichkeiten durch den Sänger mit z.B. der Mundform, ist Gegenstand aktueller Forschungsgebiete am Institut für Elektronische Musik und Akustik der Universität für Musik und darstellende Kunst in Graz. So hat beispielsweise Fabian Hohl ein Kugelmikrofonarray entworfen [Hoh09], um Abstrahlmuster zu vermessen. Da der Aufbau dieses Kugelmikrofonarrays ziemlich komplex und zeitintensiv ist, wurden weitere einfachere Mikrofonarrays entworfen.

Eines dieser simpleren Arrays wird in [BFR18a] verwendet und beschrieben. Es handelt sich dabei um ein Doppelkreis-Mikrofon-Array, welches im Unterkapitel 2.1.1 genauer beschrieben wird. Mit diesem Array wurden sowohl Messungen als auch Aufnahmen durchgeführt, welche die Grundlage dieser Arbeit bilden. Das erste Kapitel der folgenden Arbeit stützt sich auf die Messungen, die in [BFR18a], mit Konzentration auf die Messdaten des künstlichen Sängers, durchgeführt wurden. Der künstliche Sänger wird im Zuge der Arbeit auch oft mit der englischen Bezeichnung HATS (Head and Torso Simulator) benannt. Von den in [BFR18a] berechneten Abstrahlmustern wird auf die Schallleistung zurück gerechnet. Im zweiten Kapitel wird auf Basis von Aufnahmen, die im beschriebenen Mikrofonarray [BFR18a], erstellt worden sind, eine richtungsabhängige Klangfarbenanalyse durchgeführt. Die Klangfarbe wird dabei mithilfe der Timbre Toolbox (siehe Kapitel 3.1.2) analysiert. Die Toolbox wurde eigens für die Klangfarbenanalyse anhand von Audiodeskriptoren entworfen. Für die Analyse in dieser Arbeit werden einzelne Kennwerte, die mit der Toolbox berechnet wurden, ausgewählt und grafisch aufgearbeitet, um eine systematische und übersichtliche Analyse zu ermöglichen. Weiters werden sowohl der Umgang mit der Toolbox dokumentiert als auch die Analyse von Gesangsaufnahmen mit der Toolbox kritisch hinterfragt (siehe Kapitel 4).

2 Schalleistungsberechnung

Im ersten Kapitel dieser Bachelorarbeit wird die durchgeführte Schalleistungsberechnung dokumentiert. Zu Beginn werden theoretische Hintergründe aufgearbeitet. Es wird die Erstellung der Messdaten, auf welchen die Berechnungen basieren, rekonstruiert. Danach wird in einem separaten Unterpunkt das Format, in welchem die Messdaten gespeichert sind, näher erläutert. Die mathematischen Hintergründe zur Bestimmung der Schalleistung und eine kurze Einführung in relevante Normen bilden den Abschluss des Theorieteils. Bevor die Durchführung der Berechnungen in Matlab erläutert werden, wird darauf eingegangen, wie die Messdaten aufbereitet wurden, sodass die Schalleistungsberechnung durchgeführt werden konnte. Am Ende werden die berechneten Ergebnisse grafisch präsentiert.

2.1 Theorie

2.1.1 DirPat-Database and Viewer

Im Zuge der 144. Audio Engineering Society Convention in Mailand im Mai 2018 wurde in [BFR18a] die DirPat-Datenbank und die dazugehörige Darstellungsapplikation (engl. Viewer) veröffentlicht. Die Datenbank dient dazu, vermessene 3D und 2D Richtcharakteristika zu sammeln. Es wurden sowohl für Schallquellen als auch für Schallempfänger Richtmustermessungen durchgeführt und gesammelt. Eine in Matlab programmierte Visualisierungsapplikation ermöglicht die Veranschaulichung der Abstrahlmuster in 3D und 2D. Da für die Schalleistungsberechnung lediglich Schallquellen in Frage kommen, wird nur der Messvorgang für Schallquellen rekonstruiert. Die Datenbank und Visualisierungsapplikation, die online offen zu Verfügung gestellt wird, kann unter [BFR18b] gefunden werden.

Messaufbau

Für diese Arbeit ist lediglich die Vermessung des HATS relevant. Für die Vermessung von Schallquellen bzw. des HATS wurde ein Doppelkreis-Mikrofon-Array verwendet. Das Mikrofon besteht aus zwei 21 mm breiten hölzernen Ringen mit gleichem Durchmesser. Der Durchmesser von Außenkante zu Außenkante beträgt 2,46 Meter. Auf jedem Ring können bis zu 32 Messmikrofone angebracht werden. Ein Ring ist vertikal in der Medianebene angeordnet und der zweite horizontal in der Azimutalebene. Die maximale Anzahl an Mikrofonen ist 62, da sich die Ringe an zwei Stellen überschneiden. Die Seite des 21 mm dicken Ringes, die zum Messobjekt gerichtet ist, ist mit Dämmmaterial bestückt, um Reflexionen am Ring zu vermeiden. Die Mikrofone sind am Ring in $11,25^\circ$ Winkelabstand zueinander angeordnet. Durch diese Anordnung ergibt sich vom Mittelpunkt der Ringe aus ein Abstand von 1 m zu den Messmikrofonen. Als Messmikrofone wurden Mikrofone des Typs NTI MA 2230 verwendet, diese wurden wiederum mit einem Andiamo.MC Directout Technologies Vorverstärker betrieben (vgl. [BFR18a]). In Abbildung 1 ist eine schematische Darstellung des Messaufbaus ersichtlich.

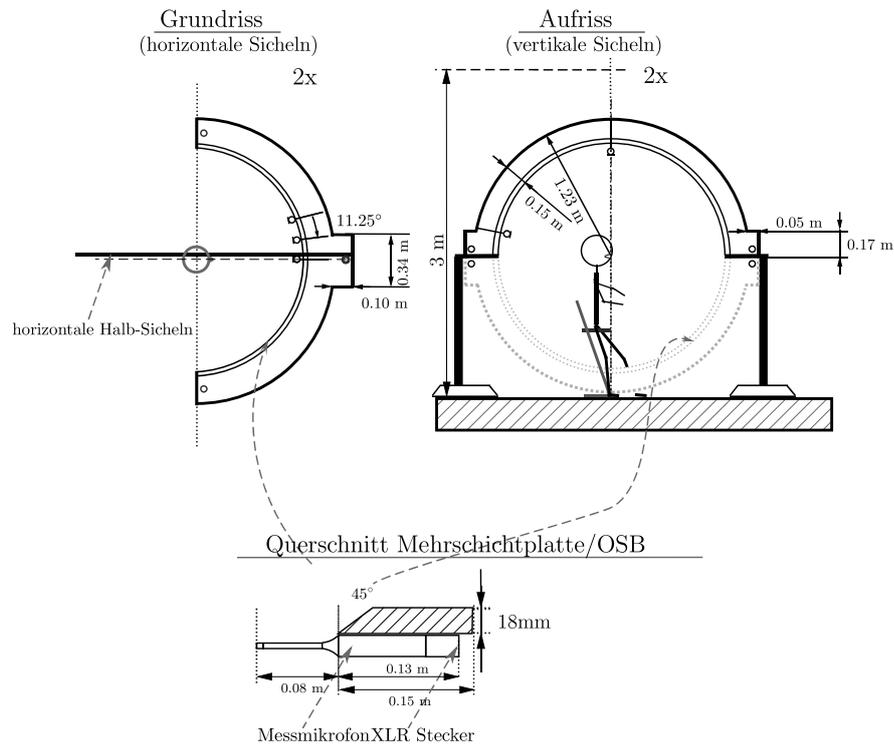


Abbildung 1 – Darstellung des Messaufbaus (Quelle: [BFR18a])

Kanalbelegung

Die Kanalbelegung des Messaufbaus wird folgend durchgeführt. Kanal 1 bis 32 sind die Mikrofone des Horizontalrings im Uhrzeigersinn. Für spätere Referenzierung auf die Mikrofone, die speziell in Kapitel 3 notwendig wird, wird an dieser Stelle folgende Referenzierungskonvention eingeführt. Das erste Mikrofon des Horizontalrings, das waagrecht normal auf die Schallquelle gerichtet ist (0° Azimuth und 90° Elevation), wird mit Mikrofon an der Stelle 0° horizontal benannt, das letzte Mikrofon des Horizontalrings (Kanal 32) würde also mit $346,75^\circ/-11,25^\circ$ horizontal referenziert werden. Selbiges gilt für die vertikal angeordneten Mikrofone. Die Kanäle 33 - 62 sind den Mikrofonen im Vertikalring gegen den Uhrzeigersinn zugeordnet. Das erste Mikrofon des Vertikalrings kann also mit $11,25^\circ$ vertikal referenziert werden, dieses Mikrofon befindet sich $11,25^\circ$ über dem vorherig angeführten 1. Mikrofon (0° horizontal). In Abbildung 2 ist die Kanalbelegung und Referenzierungskonvention für die in der Klangfarbenanalyse verwendeten Mikrofone in zwei beispielhaften Kreisen dargestellt.

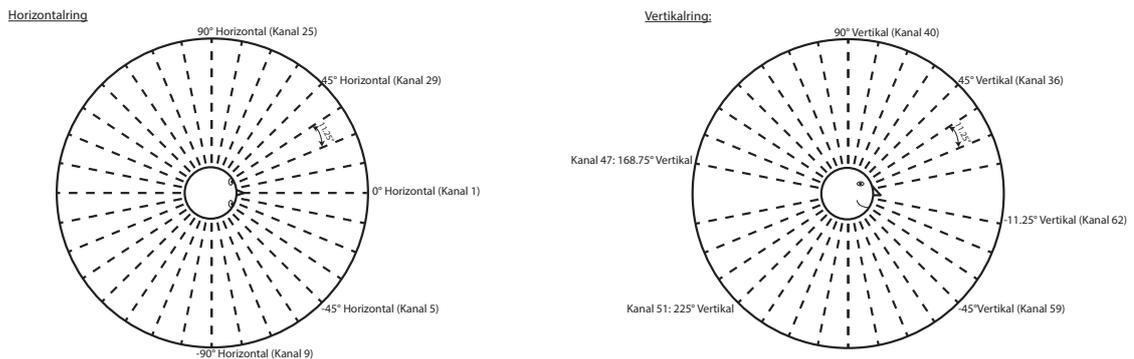


Abbildung 2 – Schematische Darstellung der Kanalbelegung

Vermessung des HATS

Das verwendete Kunstkopfmodell trägt die Bezeichnung 4128C und wird vom Unternehmen Brüel und Kjaer hergestellt. Zur Vermessung wurde der HATS auf eine Drehscheibe gestellt und in 10° Schritten (von 0° Azimuth bis 170°) wird 18 mal jeweils ein einzelner logarithmischer Sweep abgespielt und gemessen. Für die Vermessung der Lautsprecher bzw. des HATS werden nur die Mikrofone des vertikalen Rings verwendet [BFR18a]. Zenith und Nadir werden aus praktischen Gründen nicht mit gemessen bzw. berechnet. Es werden also 30 Mikrofone des Vertikalrings pro 10° Schritt verwendet. So entstehen 30 neue Messwerte pro 10° Schritt. Am Ende ergeben sich also

$$18 [\text{Messungen}] \cdot 30 \left[\frac{\text{neue Messwerte}}{10^\circ \text{ Schritt}} \right] = 540 [\text{Messwerte}] \quad (1)$$

Dadurch wird das vollständige 3D Abstrahlverhalten abgetastet und die Berechnung der Abstrahlcharakteristik in drei Dimensionen ermöglicht. Durch Entfaltung im Frequenzbereich wird aus den gemessenen Sweepantworten und den Messsignalen die Impulsantwort berechnet [BFR18a]. Lediglich Zenith und Nadir werden aus praktischen Gründen nicht vermessen und deren Impulsantwort nicht berechnet. Die gemessenen Impulsantworten werden in einem Datensatz im von der AES standardisierten SOFA Format (siehe Kapitel 2.1.2) abgespeichert und in der DirPat-Datenbank veröffentlicht.

2.1.2 Das SOFA-Format

SOFA bedeutet Spatially Oriented Format for Acoustics und ist ein Dateiformat, welches es ermöglicht, akustische Daten mit räumlichem Inhalt wie z.B. kopfbezogene Übertragungsfunktionen (engl. Head-Related Transfer Function kurz HRTF) oder wie im Falle der in Kapitel 2.1.1 beschriebenen Messung auch räumliche Impulsantworten zu speichern. Da das SOFA Format von der AES (Audio Engineering Society) standardisiert wurde, ermöglicht es den einfachen Datenaustausch zwischen Forschungsinstituten und Benutzern. Die Daten werden im Format strukturiert nach der Messung verpackt. Die Messung besteht dabei aus den Daten (z.B. Impulsantworten) und wird beschrieben durch die Metadaten und den dazugehörigen Dimensionen (siehe [MIC⁺13] für Informationen

zur Formatarchitektur). Es gibt in SOFA verschiedene Konventionen, die spezifizieren, wie die Daten und Metadaten der Messung beschaffen und zu behandeln sind, da z.B. HRTFs anders zu behandeln und auszulesen sind als räumliche Impulsantworten. Bisher wurden drei Konventionen standardisiert. Es gibt die Konventionen GeneralFIR, GeneralTF und SimpleFreeFieldHRIR. GeneralFIR ist die Allgemeine Konvention, bei welcher der Datentyp FIR (Finite Impulse Response) verwendet wird. GeneralTF ist eine allgemeine Konvention, bei welcher der Datentyp auf TF (Transfer Function) festgelegt ist. Die letzte standardisierte Konvention SimpleFreeFieldHRIR wird verwendet, wenn Frei-Feld HRTFs als Impulsantworten gespeichert werden und diese mit einer omnidirektionalen Quelle für einen einzelnen Zuhörer vermessen werden (vgl. [MNW⁺17]). Die angesprochenen Konventionen liefern Informationen zu den verschiedenen Messsetups mit Informationen über die einzelnen Objekte und Positionen im Messaufbau (Quelle, Empfänger ...). Weiters ist es wichtig zu erwähnen, dass in einem SOFA-Datensatz nicht nur Messdaten für eine Quelle oder einen Empfänger gespeichert werden. Die in [BFR18b] veröffentlichten Datensätze setzen sich aus Messdaten mehrerer Quellen bzw. Empfänger zusammen. Der Datensatz, in dem die Messdaten enthalten sind, kann über [BFR18b] geladen werden. Er trägt den Namen *LSPs_HATS_GuitarCabinets_Akustikmessplatz.sofa*. Informationen, zum Datensatzaufbau und Datensatzauslesen, folgen in Kapitel 2.2.1. Explizite Information über das SOFA Format und die vorhandenen Konventionen lassen sich in [MNW⁺17] und [MIC⁺13] finden.

2.1.3 Die Schalleistung

Das Hauptthema des ersten Kapitels bildet die Schalleistung und die damit zusammenhängende Berechnung und Messung.

„Während die mathematische Beschreibung des Schallfelds am einfachsten mit Hilfe des Schalldrucks oder Schalldruckpegels erfolgt, verwendet man zur Beschreibung der Schallquelle in erster Linie die Energiegrößen Schalleistung und - wie auch zur Beschreibung des Energiegehalts eines Schallfelds - die Schallintensität.“ ([DDHW14], S.16). Die Schallintensität berechnet sich aus Multiplikation der Effektivwerte von Schalldruck und Schallschnelle.

$$I = p_{eff} \cdot v_{eff}$$

$$\begin{aligned} p_{eff} & \dots \text{Effektivwert des Schalldruckes in } [Pa] \\ v_{eff} & \dots \text{Effektivwert der Schallschnelle in } \left[\frac{m}{s}\right] \end{aligned} \quad (2)$$

Die Schallintensität ist definiert als „[...]die Schalleistung, die durch eine Flächeneinheit, die senkrecht zur Ausbreitungsrichtung der Schallwelle steht, hindurchströmt.“ ([DDHW14], S.16). Es wird dabei angenommen, dass die Intensität I auf der Fläche konstant ist. Die Gleichung 2 kann also erweitert und auf die Schalleistung umgeformt werden.

$$\begin{aligned} I & = p_{eff} \cdot v_{eff} = \frac{P}{A} \quad \text{in } \left[\frac{W}{m^2}\right] \\ P & = I \cdot A = p_{eff} \cdot v_{eff} \cdot A \quad \text{in } [W] \end{aligned} \quad (3)$$

Wird die Annahme getroffen, dass man sich im Fernfeld der Schallquelle befindet und die Fernfeldbedingungen ([Mös10], S.346) in Gleichung 4 erfüllt sind, kann unter Einführung einer zusätzlichen Größe die Gleichung 3 weiter vereinfacht werden.

$$\begin{aligned} i) \quad & r \gg l \\ ii) \quad & r \gg \lambda \\ iii) \quad & \frac{r}{\lambda} \gg \frac{l}{\lambda} \end{aligned} \tag{4}$$

Die Fernfeldbedingungen sind frequenzabhängig, wie an der Wellenlänge $\lambda = \frac{c}{f}$ in Gleichung 4 ersichtlich ist. Das Längenmaß r steht für den Abstand zwischen der Schallquelle und dem Messpunkt, l ist die Abmessung der Schallquelle [Mös10]. Laut Kugelwellentheorie sind im Nahfeld Schalldruck und Schallschnelle nicht in Phase und steigen unterschiedlich schnell an, während sich im Fernfeld Schalldruck und Schallschnelle gleich verhalten. Je weiter man sich von der Schallquelle wegbewegt, desto mehr nähert sich die Kugelwellenfront einer ebenen Schallwelle und die im Nahfeld noch komplexe Schallimpedanz wird reellwertig und konstant. Sie wird nun auch als Schallkennimpedanz bezeichnet und berechnet sich aus der Multiplikation von der Dichte und Ausbreitungsgeschwindigkeit des Mediums. Allgemein wird die Schallimpedanz aus Quotientenbildung von Schalldruck und Schallschnelle errechnet. Für die ebene Welle werden dabei die Effektivwerte von Schalldruck und Schallschnelle verwendet. Für die Schallkennimpedanz einer ebenen Welle kann also folgende Gleichung aufgestellt werden:

$$Z_0 = \frac{p_{eff}}{v_{eff}} = \rho_0 \cdot c_0 \quad \text{in} \quad \left[\frac{Ns}{m^3} \right] \tag{5}$$

Für die Berechnung der Impedanz, die in Kapitel 2.3 durchgeführt wird, werden für eine Raumtemperatur von 20°C typische Werte von $\rho_0 = 1.2 \frac{kg}{m^3}$ und $c = 343 \frac{m}{s}$ gewählt. Wird nun Gleichung 5 auf die Schallschnelle umgeformt und in Gleichung 3 eingesetzt, ergibt sich für die Schalleistung im Fernfeld:

$$P = p_{eff} \cdot v_{eff} \cdot A = p_{eff}^2 \cdot \frac{1}{Z_0} \cdot A = \frac{p_{eff}^2}{\rho_0 \cdot c_0} \cdot A \tag{6}$$

Daraus folgt für die Schallintensität im Fernfeld:

$$I = p_{eff} \cdot v_{eff} = p_{eff}^2 \cdot \frac{1}{Z_0} = \frac{p_{eff}^2}{\rho_0 \cdot c_0} \tag{7}$$

In der unten stehenden Gleichung ist der Zusammenhang zwischen Effektivwert und Spitzenwert des Schalldruckes angeführt.

$$p_{eff} = \frac{\hat{p}}{\sqrt{2}} \quad \Rightarrow \quad p_{eff}^2 = \frac{\hat{p}^2}{2} \tag{8}$$

Um große Unterschiede in den Größen begreifbarer und anschaulicher zu gestalten, werden Schallgrößen üblicherweise in Pegelgrößen umgerechnet. Dabei sind als Bezugsgrößen der Schalldruck der Hörschwelle $p_0 = 20 \mu Pa$ und die dabei auftretende Schalleistung von $P_0 = 1 pW = 10^{-12} W$ gebräuchlich [DDHW14]. Der Schalleistungspegel errechnet sich folgend:

$$L_P = 10 \cdot \log \left(\frac{P}{P_0} \right) = 10 \cdot \log \left(\frac{P}{10^{-12} W} \right) \quad \text{in} \quad [dB] \tag{9}$$

2.1.4 Messung der Schalleistung

In der akustischen Messtechnik gibt es verschiedene Verfahren, um die Schalleistung messtechnisch zu erfassen. Das für die Berechnungen dieser Arbeit zugrunde liegende Verfahren ist das Hüllflächenverfahren, welches im nächsten Punkt genauer erläutert wird. Eine Standardisierung des Hüllflächenverfahrens ist in der Norm ISO 3745 erfasst, welche aus Vergleichsgründen im folgenden Kapitel angeführt wird.

Das Hüllflächenverfahren

Betrachtet man Gleichung 2 allgemeiner, also nicht nur für das Fernfeld, und verwendet man infinitesimal kleine Flächenelemente, ergibt sich für die Intensität folgender Zusammenhang:

$$\vec{I} = p \cdot \vec{v} \quad (10)$$

$$dP = \vec{I} \cdot d\vec{A}$$

Die gesamte, durch eine Fläche durchtretende Leistung kann also folgend berechnet werden [Mös10]:

$$P = \int dP = \int \vec{I} \cdot d\vec{A} \quad (11)$$

Handelt es sich bei der Fläche um eine geschlossene Fläche, wird das einfache Integral durch ein geschlossenes Wegintegral ersetzt:

$$P = \oint_S \vec{I} \cdot d\vec{A} \quad (12)$$

Werden die Fernfeldbedingungen vorausgesetzt, die in Gleichung 4 beschrieben werden, kann Gleichung 7 in 12 eingesetzt werden und als analytische Lösung für die Schalleistung im Freifeld kann folgende Gleichung aufgestellt werden [PZS⁺16]:

$$P = \oint_S \vec{I} \cdot d\vec{A} = \oint_S \frac{p_{eff}^2}{\rho_0 \cdot c_0} \cdot d\vec{A} = \frac{1}{\rho_0 \cdot c_0} \cdot \oint_S p_{eff}^2 \cdot d\vec{A} \quad (13)$$

Mit der Spitzenwert/Effektivwertbeziehung aus Gleichung 8 kommt man auf die Form der Gleichung, die in [PZS⁺16] angeführt wird:

$$P = \frac{1}{\rho_0 \cdot c_0} \cdot \oint_S p_{eff}^2 \cdot d\vec{A} = \frac{1}{2} \cdot \frac{1}{\rho_0 \cdot c_0} \cdot \oint_S \hat{p}^2 \cdot d\vec{A} \quad (14)$$

Für das Hüllflächenverfahren wird die analytische Gleichung diskretisiert. Es wird eine Fläche gewählt, die die Schallquelle vollständig umhüllt. Diese Hüllfläche wird in N kleinere Flächen aufgeteilt und diese werden aufsummiert, sodass sich die Leistung aus:

$$P = \frac{1}{\rho_0 \cdot c_0} \cdot \sum_{i=1}^N p_{eff,i}^2 \cdot A_i \quad \text{in [W]} \quad (15)$$

ergibt [Mös10]. Die Hüllfläche um die Schallquelle wird also abgetastet und für jeden Abtastpunkt wird der quadrierte Effektivwert des Schalldruckes auf die dazugehörige Teilfläche A_i aufmultipliziert und aufsummiert. Dieses Verfahren setzt einen reflexionsarmen Raum voraus, in welchem Freifeldbedingungen simuliert werden können. Die Messungen in [BFR18a], sind am Institut für Elektronische Musik und Akustik in einem reflexionsarmen Raum durchgeführt worden. Die Messdaten sind also qualifiziert, um eine Schalleistungsberechnung nach dem Hüllflächenverfahren durchzuführen.

ISO 3745

Die Norm ISO 3745 dient dazu, die Bestimmung der Schalleistung durch Messung von Schalldruckwerten, die von Geräuschquellen emittiert werden, zu vereinheitlichen. In der Norm werden neben räumlichen Voraussetzungen, Anforderungen an die Quellen und deren Position uvm. auch räumliche Anordnung der Abtastpunkte auf der Hüllfläche vorgeschlagen [IE03]. Als Hüllfläche wird in der Norm eine Kugel mit dem Radius r verwendet. In der folgenden Tabelle und Abbildung sind die Positionen der 20 Messmikrofone auf der Hüllfläche, wie sie in der Norm für eine Freifeldmessung vorgegeben sind, verdeutlicht.

Messmikrofon	x/r	y/r	z/r
1	-1.00	0	0.05
2	0.49	-0.86	0.15
3	0.48	0.84	0.25
4	-0.47	0.81	0.35
5	-0.45	-0.77	0.45
6	0.84	0	0.55
7	0.38	0.66	0.65
8	-0.66	0	0.75
9	0.26	-0.46	0.85
10	0.31	0	0.95
11	1.00	0	-0.05
12	-0.49	0.86	-0.15
13	-0.48	-0.84	-0.25
14	0.47	-0.81	-0.35
15	0.45	0.77	-0.45
16	-0.84	0	-0.55
17	-0.38	-0.66	-0.65
18	0.66	0	-0.75
19	-0.26	0.46	-0.85
20	-0.31	0	-0.95

Tabelle 1 – Messmikrofonpositionen laut ISO 3745 (Quelle: [IE03])

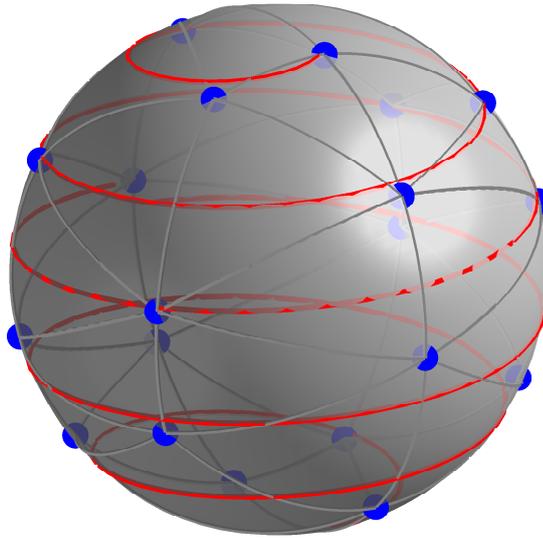


Abbildung 3 – Abtastung der Hüllfläche laut ISO 3745 (Quelle: [PZS⁺16])

Ansatz zur Flächenaufteilung

Wie in Abbildung 3 ersichtlich ist, sind die Mikrofone bei der Normpositionierung im Raum spiralförmig auf einer Kugel verteilt, dementsprechend wird, um die Schalleistung mit Gleichung 15 berechnen zu können, die Kugeloberfläche in 20 gleich große Teile aufgeteilt. Eine Teilfläche würde sich dann aus:

$$A_i = \frac{A_i}{20} = \frac{4\pi r^2}{20} \quad (16)$$

errechnen. Auf diese werden dann die einzelnen, an den Messpositionen erfassten und quadrierten Schalldruckwerte aufmultipliziert und laut Gleichung 15 aufsummiert. Die während der Vermessung des HATS erhaltenen Messwerte (siehe. Kapitel 2.1.1) sind gleichmäßig auf der Kugel, die durch das Doppelkreis Mikrofonarray aufgespannt wird (siehe Kapitel 2.1.1), verteilt. Die Messwerte können als Abtastpunkte einer kugelförmigen Hüllfläche verstanden werden, deren Verteilung in Abbildung 4 dargestellt ist.

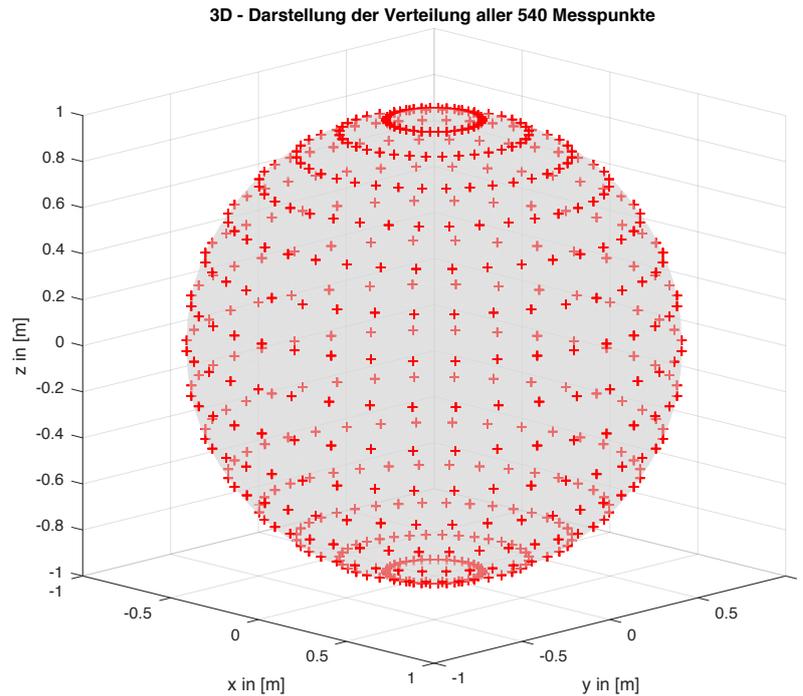


Abbildung 4 – Verteilung aller 540 Messpunkte auf der Kugel

Im Zuge dieser Arbeit werden zwei Flächeneinteilungen, auf welcher verschiedene Messpunkte der HATS Daten hochgerechnet werden, entworfen und verglichen. Wie die Flächen eingeteilt werden und wie viele Messpunkte pro Fläche aufgerechnet werden, wird im Kapitel 2.3 besprochen. Für die Berechnung der Flächen der angesprochenen Einteilungen ist jeweils die Berechnung von Kreisscheibenmantelflächen vonnöten. Diese Mantelflächen können über die Bogenlänge eines Kreises hergeleitet werden. Es wird die Bogenlänge einer Strecke auf der Kugel gedacht, um eine Achse rotiert und die Mantelfläche der dadurch entstehenden Scheibe wird berechnet. Die Bogenlänge b einer Kurve $h(x)$ ist definiert als:

$$b = \int_a^b \sqrt{1 + h'(x)^2} \cdot dx \quad (17)$$

wobei :

$$h'(x) = \frac{dh(x)}{dx} \quad (18)$$

Die Mantelfläche eines Körpers, der durch Rotation einer Kurve $h(x)$ um die x-Achse entsteht, wird mit:

$$M = 2\pi \cdot \int_a^b h(x) \cdot \sqrt{1 + h'(x)^2} \cdot dx \quad (19)$$

berechnet. Der Querschnitt entlang einer Ebene, die durch den Mittelpunkt einer Kugel verläuft, ist immer ein gleich großer Kreis. Die Kreise unterscheiden sich nur durch die Koordinaten in ihrer funktioniellen Beschreibung. Der entstehende Kreis kann nun mit der Kreisfunktion, die die dazugehörigen Raumkoordinaten enthält, beschrieben werden.

$$\begin{aligned} i) \quad & r^2x^2 + r^2y^2 = r^2 \\ ii) \quad & r^2x^2 + r^2z^2 = r^2 \\ iii) \quad & r^2y^2 + r^2z^2 = r^2 \end{aligned} \tag{20}$$

i) beschreibt einen Kreis in der x-y-Ebene. *ii)* beschreibt einen Kreis in der x-z-Ebene und *iii)* einen in der y-z-Ebene. Je nachdem, wie die Flächeneinteilung gewählt wird, muss die richtige Kreisgleichung in Gleichung 19 eingesetzt werden. Die Integrationsvariable beschreibt die Achse, um die rotiert werden soll. Die Integrationsgrenzen a und b sind auf der Rotationsachse zu wählen und bestimmen, welcher Abschnitt des Kreisbogens um eine Achse rotiert werden soll. Die in dieser Arbeit berechneten Flächen werden einmal durch Rotation um die x-Achse und einmal durch Rotation um die z-Achse berechnet. Es werden keine Flächen durch Rotation um schräge Achsen berechnet, deshalb sind die in Gleichung 20 angeführten Kreisfunktionen als Kurvenfunktionen ausreichend.

2.2 Datenaufbereitung

In diesem Unterpunkt werden die zur Berechnung mit dem Hüllflächenverfahren notwendigen Schritte der Datenaufbereitung beschrieben. Zuerst wird darauf eingegangen, wie aus dem Datensatz, in welchem die HATS Messdaten enthalten sind und welcher im SOFA Format (siehe Kaptiel 2.1.2) gespeichert ist, die räumlichen Impulsantworten extrahiert werden können. Daraufhin wird noch erwähnt, wie die Impulsantworten weiter verarbeitet werden, um auch eine Frequenzabhängigkeit in die Schalleistung zu bringen. Dadurch wird es am Ende auch ermöglicht, einen Vergleich der Leistungsspektren beider Varianten der Schalleistungsberechnung mit den unterschiedlichen Flächeneinteilungen durchzuführen.

2.2.1 Vom SOFA-Format zu Impulsantworten

Bei der im Kapitel 2.3 erläuterten und in Matlab durchgeführten Schalleistungsberechnung werden Funktionen verwendet, die mit der DirPat Visualisierungsapplikation (siehe Kapitel 2.1.1) veröffentlicht wurden. Im Speziellen geht es um eine Funktion, die den SOFA Datensatz ausliest und beim Öffnen des Datensatzes eine Abfrage startet, ob es sich bei den Messdaten im Datensatz um Messdaten einer Quelle oder eines Empfängers handelt. Die Funktion ist in den Matlabcodes, die auf [BFR18b] veröffentlicht sind, enthalten und trägt den Namen `fun_loadfromSOFA2IEM3DViewer.m`. Diese Funktion wird für die Schalleistungsberechnung angepasst. Sowohl für die angepasste als auch die ursprüngliche Form der Funktion ist das in [MIC⁺13] vorgestellte SOFA-API Voraussetzung. API bedeutet Application Programming Interface, was auf deutsch soviel wie Anwendungsprogrammierschnittstelle bedeutet. Diese Schnittstelle wurde programmiert, um das SOFA Format in Matlab zu integrieren. Die API kann über [MNW⁺17] geladen werden. Der

Ordner der geladenen API wird dann an den gewünschten Ort kopiert und zum Matlab-pfad mit dem Befehl „Add Folder“ (ohne Unterordner) im Pfadmenü von Matlab hinzugefügt. Danach kann mit dem Befehl *SOFAstart* die API gestartet werden. Die angepasste Funktion wird *fun_loadfromSOFA2MAT.m* benannt. In ihr wird die SOFA-API gestartet und mit dem Befehl *SOFAload* wird der SOFA-Datensatz von der API ausgelesen. Die Daten werden in eine von Matlab interpretierbare **.mat*-Datei gespeichert. Die einzige Abänderung, die an der Funktion gegenüber dem Original vorgenommen wird, ist, dass die Abfrage, ob es sich beim Datensatz um eine Quelle oder ein Empfänger handelt, überbrückt wird. Da die Schallleistung nur von Schallquellen berechnet werden kann, wird in der Funktion der Datensatztyp *fix* auf den Typ Schallquelle gesetzt. Dies wird erreicht, indem die Variable *type* im Code *fix* auf den Buchstaben *'s'* gesetzt wird. Da das Hauptgebiet des ersten Kapitels dieser Arbeit die Schallleistung und deren Berechnung ist, wird nicht genauer auf die Struktur der SOFA-API und -Format eingegangen. Im Kapitel 2.1.2 wird auf weiterführende Literatur verwiesen. Nachdem der SOFA-Datensatz ausgelesen und in eine **.mat*-Datei gespeichert wird, wird die **.mat*-Datei geladen. Es kann nun aus der **.mat*-Datei eine 4-D-Matrix ausgelesen werden. Sie trägt den Namen *h_alle*. Die Dimensionen von *h_alle* sind 2048 samples x 15 Elevation x 36 Azimut x 12 Schallquellen. Es sind also, wie in Kapitel 2.1.1 beschrieben, insgesamt 15 Elevation x 36 Azimut = 540 Impulsantworten für 12 vermessene Schallquellen mit jeweils 2048 samples im geladenen Datensatz enthalten.

Im Datensatz *LSPs_HATS_GuitarCabinets_Akustikmessplatz.sofa* liegen die Messdaten für den HATS an 3. Stelle und können also mit *h_alle(:, :, :, 3)* in Matlab referenziert werden. Man erhält dann eine 3D-Matrix mit 540 Impulsantworten, die den gleichmäßig auf der Kugel verteilten Messpunkten zugeordnet werden können (siehe Abbildung 4). Die Verteilung der Messpunkte ist in Abbildung 4 ersichtlich.

Es können jetzt in azimuthaler Richtung pro 10°-Schritt (beginnend bei 0° in Spalte 1) jeweils 15 verschiedene Impulsantworten in Elevationsrichtung ausgelesen werden. Die Zeilen der Matrix entsprechen also den Messpunkten in der vertikalen Richtung der Kugel, beginnend in der ersten Zeile mit einem Elevationswinkel von 11,25° (Elevation vom Zenith nach unten gemessen). Dies bedeutet, dass in der 8. Zeile (Elevation 90°) der Matrix die Messpunkte der Horizontalebene, welche die Kugel horizontal halbiert, liegen. Dies wird im Kapitel 2.3 noch genauer besprochen und mit Abbildungen erläutert.

2.2.2 Aufteilen der Impulsantworten in Terzbänder

Um die Frequenzabhängigkeit bei der Schallleistungserfassung zu berücksichtigen werden die Impulsantworten in Terzfrequenzbänder aufgeteilt. Es wird dazu die Matlabfunktion *oct3dsgn.m* verwendet. Diese Funktion retourniert Filterkoeffizienten für Terzbandfilter, die der Norm ANSI S1.1-1986 unterliegen. Der Funktion muss neben den Mittenfrequenzen der jeweiligen Filter auch die passende Samplingfrequenz übergeben werden. Die Filterordnung wird standardmäßig mit $N = 3$ festgelegt. Die Mittenfrequenzen der Terz-

bänder kann mit

$$f_m = f_r \cdot 10^{\frac{n-30}{10}}, \quad \text{wobei } n \in \mathbb{N} \quad \& \quad 10 \leq n \leq 43 \quad (21)$$

f_m ...Mittenfrequenz des Terzbandfilters in [Hz]
 f_r ...Referenzfrequenz in [Hz]

berechnet werden [Wes18]. Die Mittenfrequenzen werden für $n = 23$ bis $n = 43$ und einer Referenzfrequenz von $f_r = 1\text{kHz}$ berechnet. Dies liefert die Mittenfrequenzen der Terzbänder von 200 Hz bis 15849 Hz. In diesem Frequenzbereich befinden sich 20 Mittenfrequenzen. Aufgrund der geringen Frequenzauflösung der Impulsantworten im Frequenzbereich unter 200 Hz wird alles, was darunter liegt, nicht berücksichtigt. Es wurden mit der Matlabfunktion *oct3dsgn.m* die Filterkoeffizienten (Zähler- und Nennerkoeffizienten) für die Terzbandfilter mit den berechneten Mittenfrequenzen berechnet. Der Matlabfunktion *filter* werden diese Koeffizienten übergeben und alle 540 Impulsantworten werden terzbandgefiltert. Die Impulsantworten werden so in 20 Terzbandstücke geteilt. Die entworfenen Filter arbeiten am genauesten im Bereich:

$$\frac{f_s}{200} < f_m < \frac{f_s}{5}. \quad (22)$$

Bei der Vermessung des HATS wurde eine Samplingfrequenz von $f_s = 44100\text{Hz}$ verwendet. Der Bereich, in welchem Filterung also am besten funktioniert, ist:

$$\frac{44100\text{Hz}}{200} = 220.5\text{Hz} < f_m < 8820\text{Hz} = \frac{44100\text{Hz}}{5} \quad (23)$$

Da die Koeffizienten der Filter der Norm genügen, werden die Impulsantworten für den gesamten Frequenzbereich, der durch die berechneten Terzbandmittenfrequenzen abgedeckt wird, gefiltert. In Abbildung 5 ist der Unterschied zwischen einer Impulsantwort die aus den gefilterten 20 Bändern zusammengerechnet wird, und der ursprünglichen Impulsantwort im Frequenzgang ersichtlich. Es wird deutlich, dass im höherfrequenten Bereich (ab ca. 15 kHz) die Abweichungen zwischen der zerteilten und wieder zusammengestückelten Impulsantwort zum Original doch feststellbar sind, aber davon nur das höchste Terzband betroffen ist. Die Impulsantwort für Darstellung 5 ist jene des Messpunktes bei 0° Azimut und 90° Elevation (HATS ist am Messbeginn gerade darauf gerichtet).

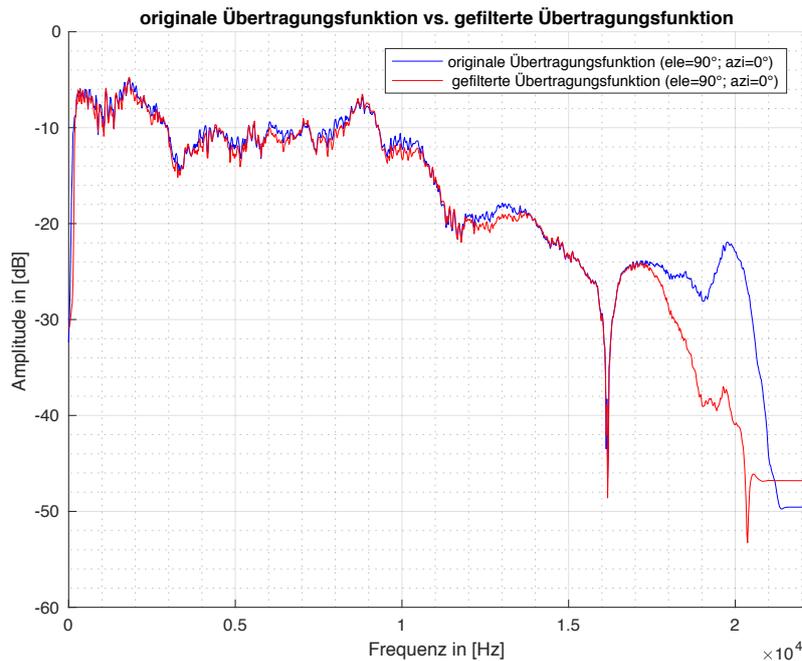


Abbildung 5 – Vergleich gefilterte Impulsantwort vs. Original (im Frequenzbereich)

Um die Ergebnisse der Schalleistungsberechnung und etwaig auftretende Fehler besser abschätzen zu können, ist es notwendig die berechneten Mittenfrequenzen auf die Fernfeldbedingungen aus Gleichung 4 zu überprüfen. Wie in Abbildung 1 ersichtlich beträgt der Abstand zwischen Schallquelle und Messmikrofonen im Doppelkreis-Mikrofon-Array 1 Meter (r). Die Lautsprecherabmessung der Mundöffnung des HATS kann aus dem dazugehörigen Datenblatt abgelesen werden. Die Mundöffnung (l) beträgt in der Breite 30 mm. Die letzte Größe die zur Überprüfung der Fernfeldbedingungen benötigt wird ist die Wellenlänge λ . Sie wird mit dem Zusammenhang $\lambda = \frac{c}{f_m}$ berechnet, wobei für die Ausbreitungsgeschwindigkeit $c = 343 \frac{m}{s}$ angenommen wird. In Tabelle 2 ist die Überprüfung der Fernfeldbedingungen für die Mittenfrequenzen angeführt.

Größen zur Überprüfung der Fernfeldbedingungen:

$$\begin{aligned} l &= 0,03 \text{ m} \\ r &= 1 \text{ m} \\ c &= 343 \frac{\text{m}}{\text{s}} \end{aligned} \quad (24)$$

Parameter				Fernfeldbedingungen		
f_m in [Hz]	$\lambda = c/f_m$ in [m]	r/λ	l/λ	$r \gg l$	$r \gg \lambda$	$r/\lambda \gg l/\lambda$
200	1,715	0,583	0,017	✓	✗	✓
251	1,367	0,732	0,022	✓	✗	✓
316	1,086	0,922	0,028	✓	✗	✓
398	0,862	1,16	0,035	✓	✓	✓
501	0,685	1,46	0,044	✓	✓	✓
631	0,544	1,838	0,055	✓	✓	✓
794	0,432	2,315	0,069	✓	✓	✓
1000	0,343	2,915	0,087	✓	✓	✓
1259	0,272	3,676	0,11	✓	✓	✓
1585	0,216	4,63	0,139	✓	✓	✓
1995	0,172	5,814	0,174	✓	✓	✓
2512	0,137	7,299	0,219	✓	✓	✓
3162	0,108	9,259	0,278	✓	✓	✓
3981	0,086	11,628	0,349	✓	✓	✓
5012	0,068	14,706	0,441	✓	✓	✓
6310	0,054	18,519	0,556	✓	✓	✓
7943	0,043	23,256	0,698	✓	✓	✓
10000	0,034	29,412	0,882	✓	✓	✓
12589	0,027	37,037	1,111	✓	✓	✓
15849	0,022	45,455	1,364	✓	✓	✓

Tabelle 2 – Überprüfung der Fernfeldbedingungen für die Terzbandmittenfrequenzen

In der Tabelle ist ersichtlich, dass die zweite Fernfeldbedingung (Gleichung 4) für die drei tiefsten Terzbandmittenfrequenzen verletzt wird. Die Auswirkungen der nicht eingehaltenen Fernfeldbedingung auf die Schallleistungsberechnung wird im Kapitel 4.1 erneut aufgefasst.

2.2.3 Von Impulsantwort zu Schalldruckwert

Um die Schalleistung zu berechnen, wird der Schalldruck benötigt. Da in den veröffentlichten Messdaten nur Impulsantworten für die Messpunkte gegeben sind [BFR18b], kann ohne Kenntnis des Kalibrierfaktors der Messmikrofone nicht auf die physikalisch, bei der Messung an den Messpunkten, vorhandenen Schalldrücke zurückgerechnet werden. Was jedoch korrekt in den Impulsantworten wiedergespiegelt wird, sind die Verhältnisse, wie sich der Schalldruck auf der abgetasteten Kugel verteilt. Wird also der Effektivwert durch Quadrieren, Mitteln und Wurzelziehen aus den Impulsantworten berechnet, erhält man auch die richtigen Schalldruckeffektivwertverhältnisse. So wie die Effektivwerte der Impulsantworten verhalten sich auch die Schalldruckeffektivwerte der einzelnen Messpunkte. Die aus der Impulsantwort errechneten Effektivwerte können für die Berechnung mit den Schalldruckeffektivwerten gleichgesetzt werden, da die Messung im reflexionsarmen Raum durchgeführt worden ist und die errechneten Werte über dem Grundgeräuschpegel liegen. Um die tatsächlichen physikalisch an den Messpunkten vorhandenen Schalldruckwerte zu erhalten, müsste nun lediglich der berechnete Effektivwert mit dem dazugehörigen Kalibrierfaktor multipliziert werden. Der Faktor ist jedoch nicht bekannt, deshalb wird die Schalleistungsberechnung ohne die kalibrierten Daten durchgeführt. Bei der Schalleistungsberechnung mit beiden Flächeneinteilungen wird mit den ident aus den Impulsantworten berechneten Effektivwerten gerechnet. Deshalb können beide Berechnungsergebnisse miteinander verglichen werden, um eine Aussage über die Genauigkeit der Flächeneinteilung zu treffen. Die tatsächliche vom HATS emittierte Schalleistung bei der Messung kann jedoch nur mit dem Kalibrierfaktor errechnet werden. Für die Berechnung der Schalleistung mit der Gleichung 15 kann p_{eff} folgendermaßen bestimmt werden.

$$p_{eff} = \sqrt{\frac{\sum_{n=1}^M (h[n])^2}{M}} \quad (25)$$

M ...Länge der Impulsantwort in [*samples*]

Für die HATS Impulsantworten ist $M = 2048$ [*samples*].

2.3 Durchführung der Berechnung

Der theoretische Hintergrund zur Berechnung der Schalleistung wurde in den vorherigen Kapiteln geklärt. In diesem Kapitel wird nun, da die Schalleistung zur Vermessung des HATS vollständig in Matlab durchgeführt wird, genauer auf die Struktur des Matlabcodes, inklusive den verwendeten Unterfunktionen sowie die gefundenen Flächeneinteilungen eingegangen.

2.3.1 Struktur des Matlabcodes

In Abbildung 6 ist die Struktur des Matlabcodes ersichtlich. Über das Schalleistungsberechnungsskript wird die Berechnung durchgeführt. Die in Kapitel 2.2.1 beschriebene Funktion *fun_loadfromSOFA2MAT* wird gestartet.

Über eine Benutzereingabe kann der SOFA-Datensatz ausgewählt werden. In der Abbildung wird dieser mit dem Dateinamen *Sources.sofa* abgekürzt. Die SOFA-API wird gestartet, der Datensatz ausgelesen und in eine **.mat*-Datei gespeichert. Diese Datei wird geladen und alle Impulsantworten aller Quellen, die im Datensatz enthalten sind, in die Variable *h_alle* geladen. Über eine erneute Benutzerabfrage wird dann abgefragt, auf welcher Position die gewünschten Messdaten liegen. Für die HATS-Messdaten muss, wie in Kapitel 2.2.1 erläutert, die 3. Position gewählt werden. Danach werden die Terzbandfilterkoeffizienten mit den berechneten Mittenfrequenzen erstellt und die extrahierten Impulsantworten werden in Terzbänder unterteilt. Der nächste Schritt ist die Berechnung der Flächeneinteilung, diese wird im nächsten Kapitel noch genauer beschrieben. Für beide Flächeneinteilungen wird dann die jeweilige Schalleistung berechnet. Abgeschlossen wird die Berechnung durch grafische Aufarbeitung der Ergebnisse durch extra erstellte Plot-Funktionen. Es werden sowohl Erklärungsabbildungen zu den Flächeneinteilungen als auch die grafische Darstellung der Schalleistung in Terzbändern erstellt. Mit einer abschließenden Benutzereingabe können die Erklärungsabbildungen aber übersprungen werden. Die letzte Benutzereingabe ist nicht in der unten angeführten Abbildung enthalten.

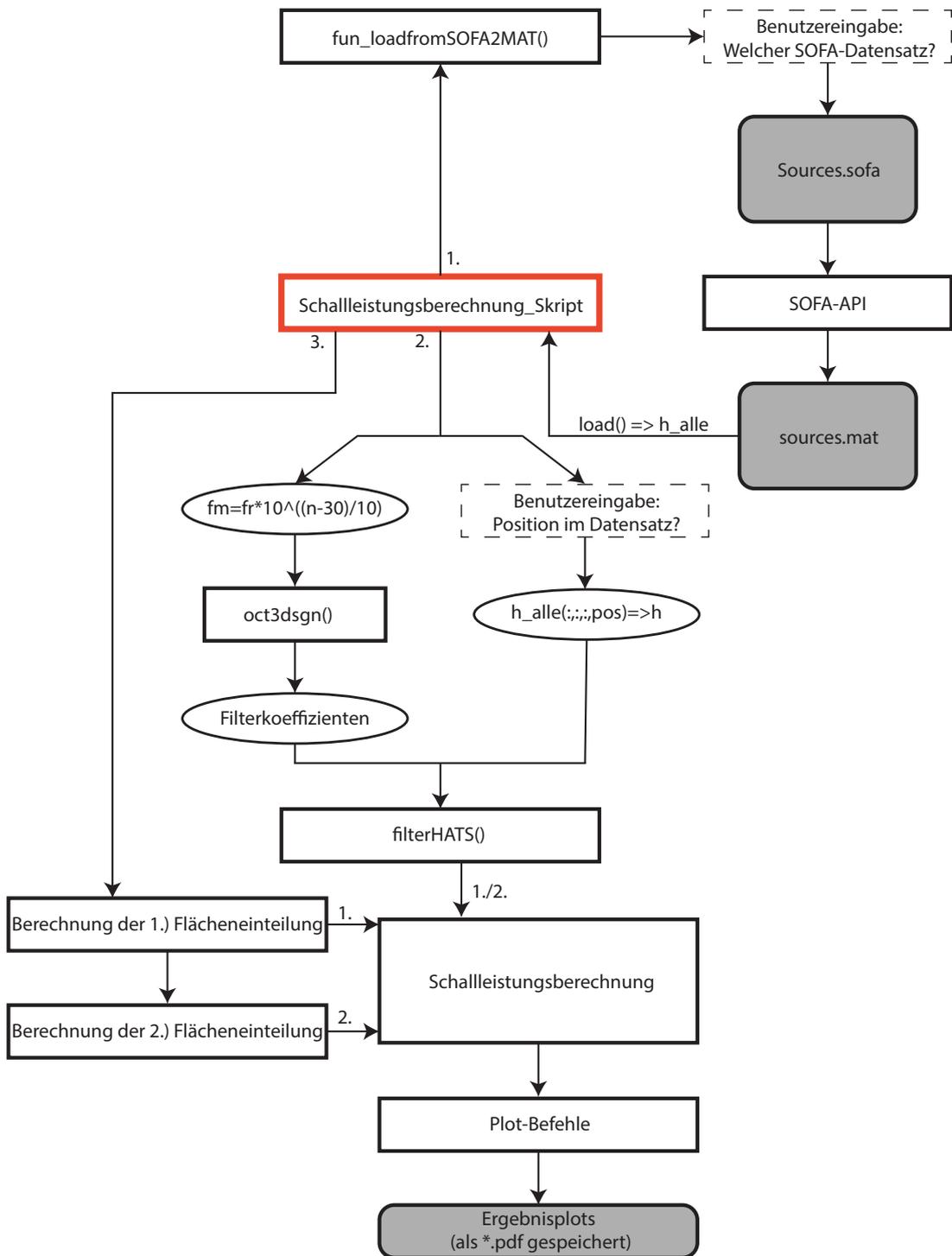


Abbildung 6 – Struktur des Matlabcodes

2.3.2 Die Flächeneinteilungen der Kugel

Die erste Flächeneinteilung

Für die erste Flächeneinteilung war das Ziel, die Schalleistung anhand der Messpunkte auf der Horizontal- und Vertikalachse zu berechnen. Die 540 vorhandenen Messpunkte wurden also auf 64 Messpunkte reduziert. In Abbildung 7 ist die reduzierte Messpunktverteilung dargestellt.

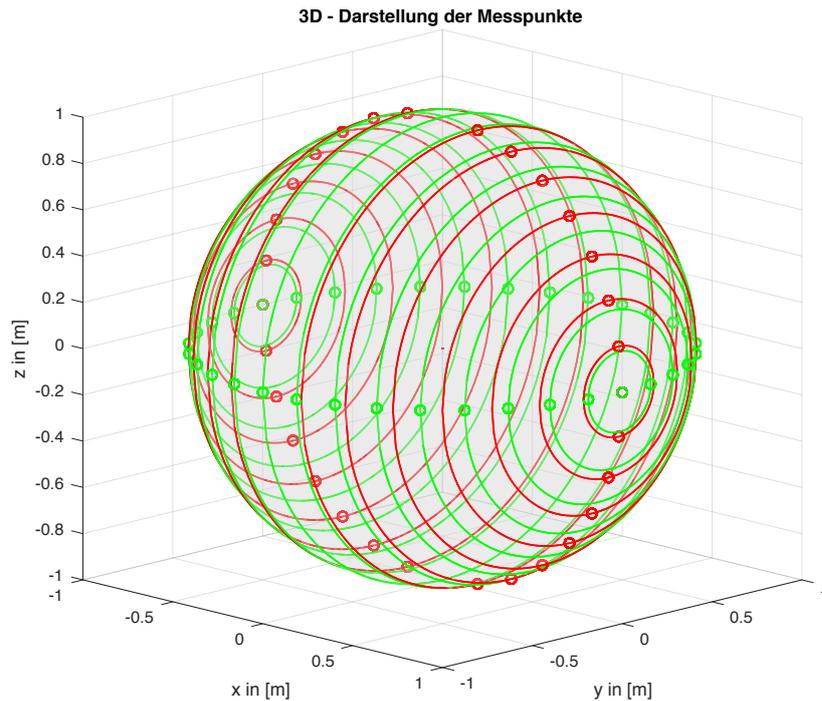


Abbildung 7 – Messpunkte auf Hauptachse

Die rot eingezeichneten Messpunkte sind die der Vertikalachse mit einem Winkelabstand von $11,25^\circ$ zueinander. In grün sind die Messpunkte der Horizontalen mit 10° Winkelabstand zueinander abgebildet. Auf den eingezeichneten Kreisen liegen jeweils zwei Messpunkte. Es wird nun eine Flächeneinteilung gesucht, die möglichst gleich viele Kreise bzw. Messpunkte pro Fläche einschließt und deren Teilflächen möglichst gleich groß sind. Dazu werden alle Messpunkte auf eine Achse (die x-Achse) projiziert. Man kann sich dazu vorstellen, dass die Horizontalen Messpunkte (grün) entlang ihrer Kreisbahnen auf der Kugel in die x-z-Ebene gedreht werden (alle Messpunkte liegen nun in einer Ebene), nun können mithilfe des Radius der Kugel und den zu den Messpunkten gehörigen Winkeln die Projektionen auf die x- und z-Achse berechnet werden. Mit der Cosinus-Funktion wird die Projektion auf die x-Achse und mit der Sinus-Funktion die Projektion auf die z-Achse berechnet. Genauso wird es in der Matlabfunktion in einer eigens dafür erstellten Sektion durchgeführt. In Darstellung 8 sind die Projektionen der Messpunkte auf die x-Achse abgebildet.

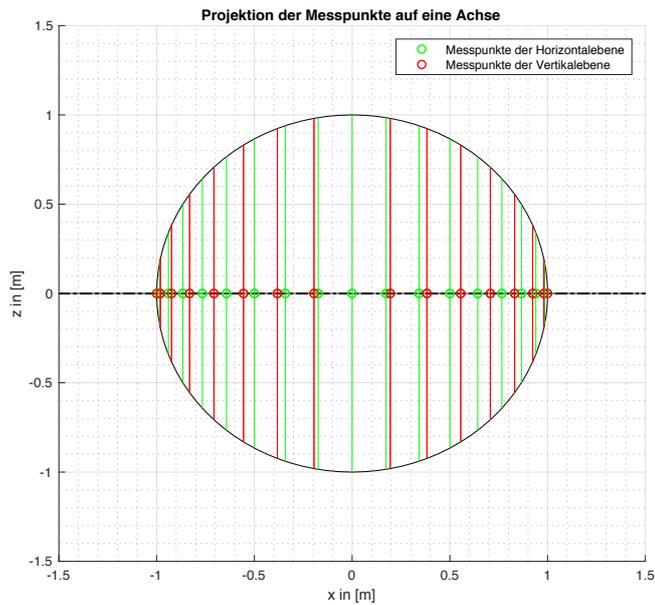


Abbildung 8 – Messpunkte auf Hauptachse

Es wird nun zwischen jedem Projektionspunkt auf der x-Achse der Abstand zum nächsten berechnet und die Position der Mittelpunkte auf der x-Achse berechnet (siehe Abbildung 9).

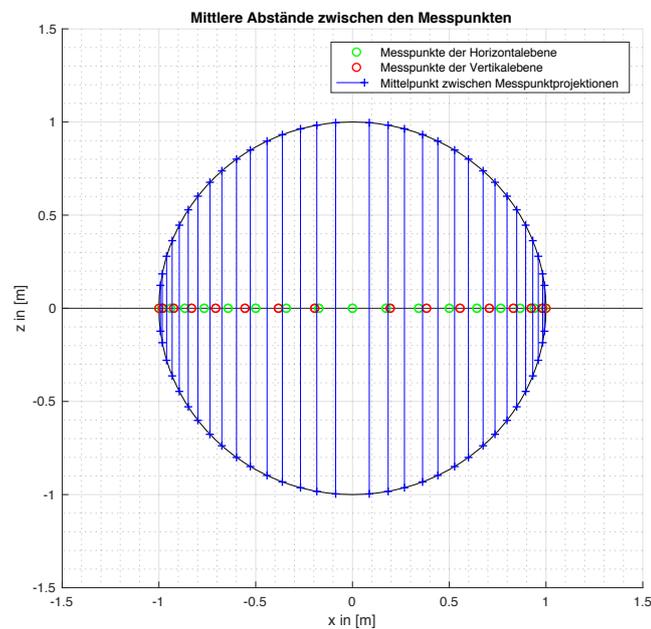


Abbildung 9 – Messpunkte auf Hauptachse

Die Abstände zwischen den Projektionen der Messpunkte werden nach ihrer Größe gereiht. Diese Reihung wird für die rechte Halbkugel aufgestellt. Es werden die Mittelpunkte der fünf größten Abstände als Projektionen der Trennpunkte für die Flächeneinteilung gewählt. Danach werden zwei weitere Mittelpunkte als Trennpunkte für Flächen ausgewählt. Bei deren Auswahl sollen möglichst gleich viele Messpunkte auf den Flächen liegen. Die Reihung und die Mittelpunkte, die als Trennpunkte für die Flächen festgelegt werden, sind in der unten angeführten Abbildung ersichtlich.

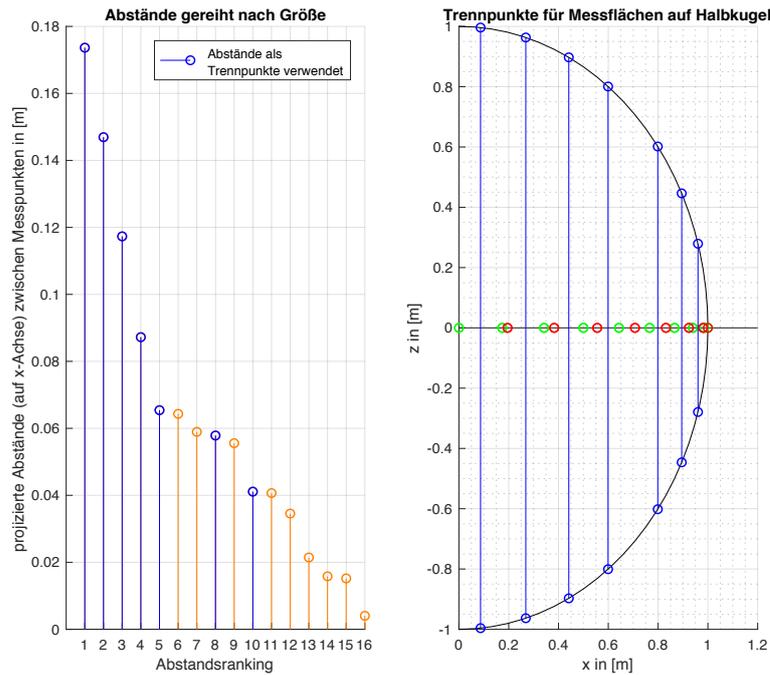


Abbildung 10 – Abstandsreihung und gewählte Trennpunkte (blau)

Die Auswahl der Trennpunkte wird nur für die rechte Halbkugel durchgeführt. Durch die Symmetrie der Kugel können die Trennpunkte auf die linke Halbkugel gespiegelt werden. In Matlab werden die Trennpunkte in der Matrix MP festgehalten. In der ersten Zeile sind die Abstände angeführt, die verwendet werden, um die Mittelpunkte festzulegen. Diese wurden nur aus Identifikationszwecken eingefügt, für die weitere Berechnung sind diese nicht weiter notwendig. Es werden ausschließlich die zweite Zeile der Matrix, in welcher die x-Positionen der Trennpunkte eingetragen wurden, und die dritte Zeile der Matrix, die die z-Werte der jeweiligen Trennflächen enthalten, für die Flächenberechnung benötigt. In den folgenden zwei Abbildungen (11 und 13) ist die gefundene Flächeneinteilung in 2D und 3D angeführt.

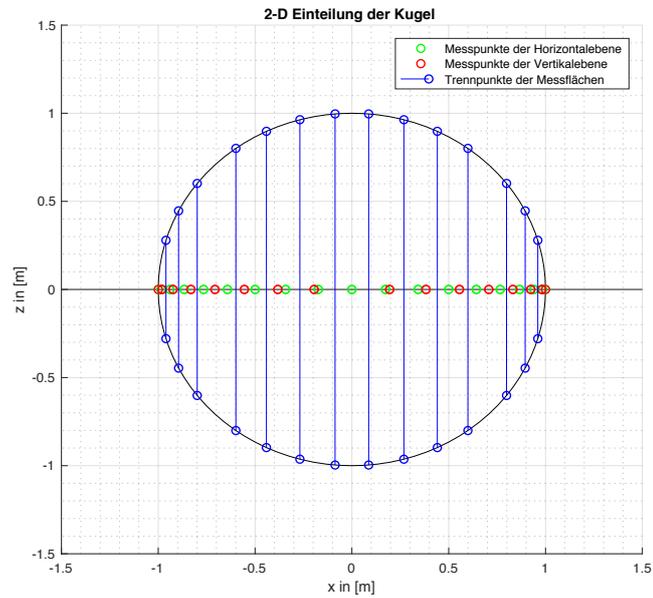


Abbildung 11 – 1.) Flächeneinteilung (reduziert) in 2D-Darstellung

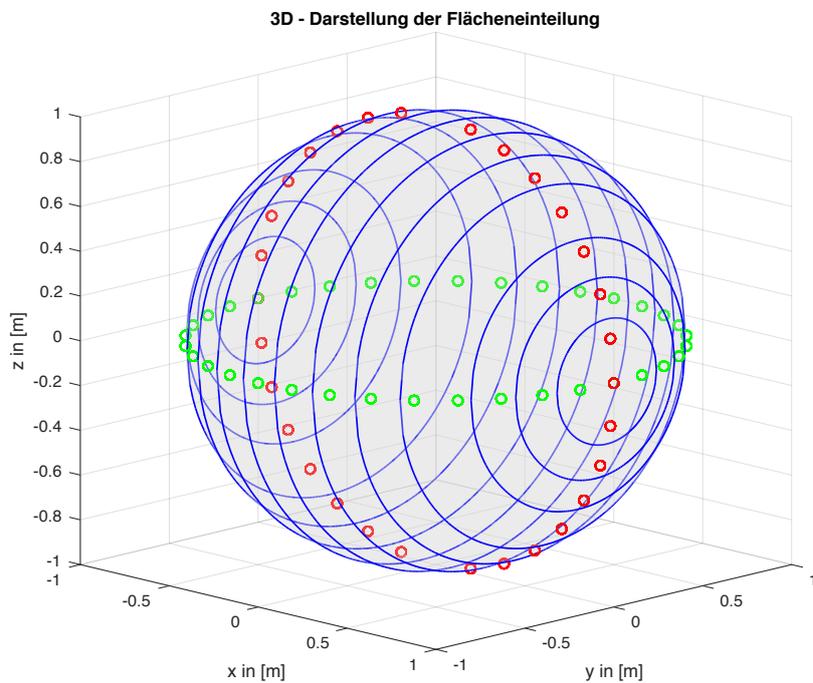


Abbildung 12 – 1.) Flächeneinteilung (reduziert) in 3D-Darstellung

Mit der gefundenen Flächeneinteilung in 15 Flächen befinden sich im Schnitt 4,2667 Messpunkte auf einer Fläche.

$$\frac{64 \text{ Messpunkte}}{15 \text{ Flächen}} = 4.2667 \left[\frac{\text{Messpunkte}}{\text{Fläche}} \right] \quad (26)$$

Durch die ungleichmäßige Aufteilung der Messpunkte und die unterschiedlichen Größen der Flächen entsteht eine teilflächenabhängige Gewichtung der dazugehörigen Messpunkte. Im Kapitel 2.4 werden die Auswirkungen der Gewichtung sichtbar. Im nächsten Abschnitt des Matlabcodes werden die Flächen berechnet. Wie in Kapitel 2.1.4 beschrieben wird der Flächeninhalt der 15 Einzelflächen durch Rotation des Kreisabschnittes um die x-Achse berechnet. Die Integrationsgrenzen bei Rotation um die x-Achse sind auch auf der x-Achse zu suchen. Die Integrationsgrenzen stellen gleichzeitig die Trennpunkte der Flächen projiziert auf die x-Achse dar. Sie befinden sich im 2. Zeilenvektor von *MP* und werden in die in Gleichung 19 eingesetzt. Zusätzlich müssen für die äußersten Teilflächen die Grenzen -1 und 1 hinzugefügt werden. Das Integral zur Mantelflächenberechnung aus Gleichung 19 wird dann in Matlab mit Hilfe von symbolischer Integration gelöst. Das daraus berechnete *sym* Ergebnis wird mittels Casting zum *double* konvertiert. Diese Berechnungsvariante hat zwar eine längere Berechnungsdauer, ist allerdings genauer als die numerischen Approximationsalternativen von Matlab (z.B. trapz) und bei insgesamt zwei Integralen, die im ganzen Code gelöst werden, ist die Verlängerung der Rechendauer durch die symbolische Integration nur minimal. Sind die Flächen berechnet, werden zuerst die Schalldruckeffektivwerte mit der Gleichung 25 berechnet. Für die Mittelung wird die Matlabfunktion *mean* verwendet. Nachdem die Schalldruckwerte berechnet sind, werden mit der Unterfunktion (*matchpanda()*) die Schalldruckwerte den jeweiligen Flächen zugeordnet. Dies wird mit den Projektionen der Messpunkte auf der x-Achse durchgeführt. Es wird überprüft, welche Messpunkte zwischen den Trennpunkten, die aus den Abständen zwischen den Projektionen der Messpunkte auf der x-Achse berechnet werden, liegen. Die Schalldruckwerte werden in der 3-D Matrix *pmatch* gespeichert. Jede Zeile der Matrix entspricht einer Fläche, beginnend in Zeile 1. Mit der Fläche die bei $x = -1$ beginnt, wandert man mit ansteigender Zeilennummer zu den Flächen in Richtung $x = 1$. Die dritte Dimension der Matrix sind die unterschiedlichen Frequenzen (20 Bänder). Die Schalldruckwerte, die einer Fläche zugeordnet sind, werden gemittelt und quadriert und dann auf die zugehörige Fläche multipliziert. Wie in Gleichung 15 beschrieben, werden alle aufmultiplizierten quadrierten Schalldruckeffektivwerte aufsummiert und mit dem konstanten Faktor $\left(\frac{1}{\rho_0 \cdot c}\right)$ multipliziert um die Leistung zu berechnen. Es ergibt sich so für jedes Frequenzband ein Leistungswert.

Die Referenzflächeneinteilung

Für die zweite Flächeneinteilung sollen alle 540 Messwerte verwendet werden und auf jeder Fläche sollen sich gleich viele Messpunkte befinden, sodass eine geringere Gewichtung entsteht. Die Flächen sollen nun allerdings durch Rotation um die z-Achse berechnet werden. Es wird die Mitte des Abstandes zwischen jedem Kreis aus Messwerten, wie sie sich in Abbildung 4 ergeben, als Trennkreis verwendet. Die Kreise sind in einem Winkelabstand von $11,25^\circ$ zueinander entfernt. Der Winkel zum Trennkreis ergibt sich aus

$11,25 + \frac{11,25^\circ}{2} = 16,875^\circ$ Mit den Winkelfunktionen und dem Kugelradius können nun die Projektionen der Trennpunkte auf die z-Achse berechnet werden. Es ergibt sich also eine Referenzflächenverteilung mit folgendem Aussehen:

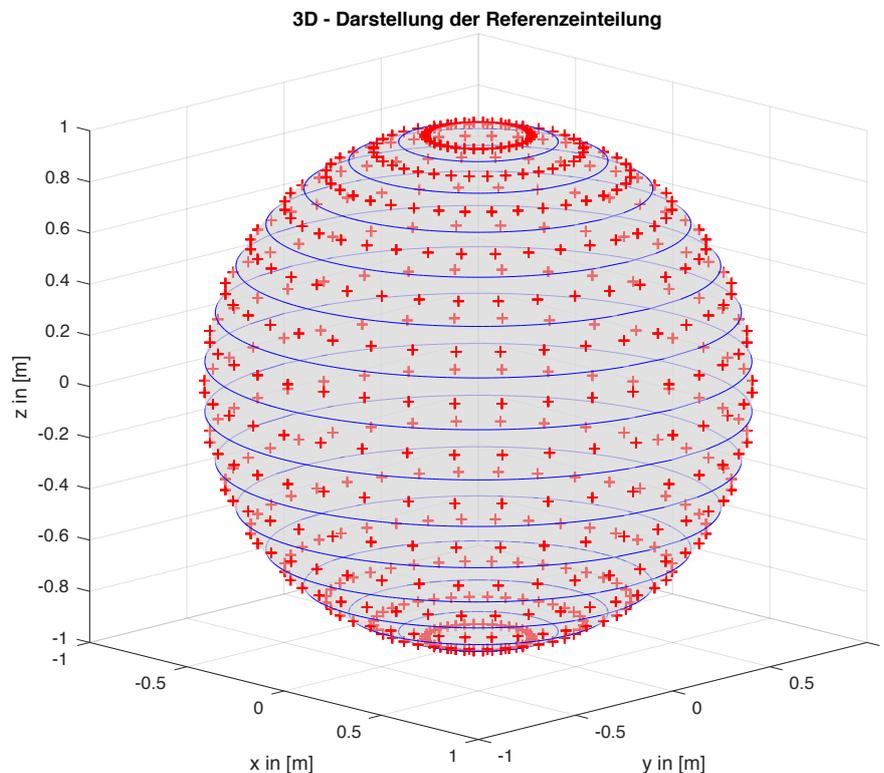


Abbildung 13 – Referenzflächeneinteilung als 3D-Darstellung

Die Schalleistung wird auf dem selben Weg wie bei der ersten Flächenverteilung errechnet, nur diesmal mit allen 540 zu Verfügung stehenden Werten.

2.4 Ergebnisse

Die frequenzabhängigen Berechnungsergebnisse der Schalleistung beider Flächeneinteilungen werden im letzten Schritt in Matlab in einem Balkendiagramm gegenübergestellt. Mit Gleichung 9 wird der Leistungspegel errechnet. Dazu wurde die in der Praxis gebräuchliche Bezugsgröße 10^{-12} Watt verwendet. Der Leistungspegel ist in einem Subplot unter der Leistungsgegenüberstellung angeführt (siehe Abbildung 14). Abgeschlossen wird die Darstellung der Ergebnisse mit einer Differenzenabbildung, welche die Unterschiede der Schalleistung in den Terzbändern zeigt. Auch hier wird zusätzlich der Pegelunterschied dargestellt (siehe Abbildung 15). Auf den folgenden zwei Seiten sind beide erwähnten Darstellungen ersichtlich. Die Diskussion der Ergebnisse und Ursache für die Differenzen in beiden Schalleistungsverläufen/-spektren finden in Kapitel 4.1 statt.

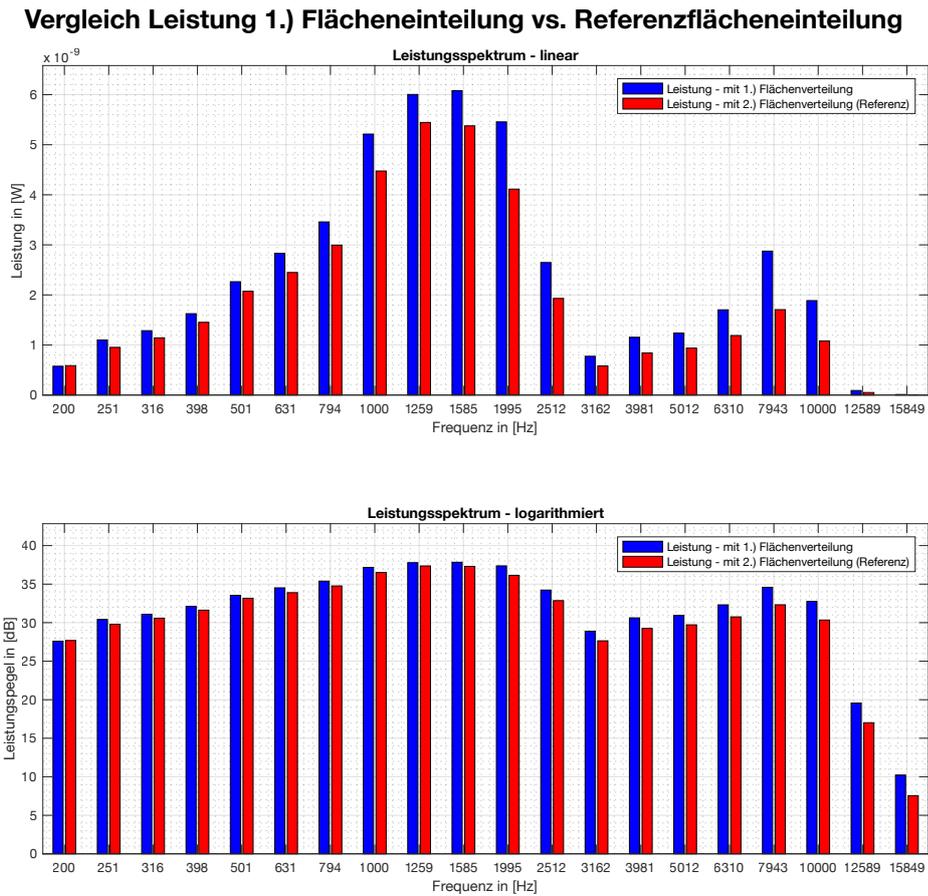


Abbildung 14 – Vergleich der mit beiden Flächeneinteilung berechneten Schalleistung: Berechnung mit reduziertem Messgrid (blau) vs. Berechnung mit 540 Messpunkten (rot)

Differenz der Leistung zwischen 1.) Flächeneinteilung und Referenzflächeneinteilung

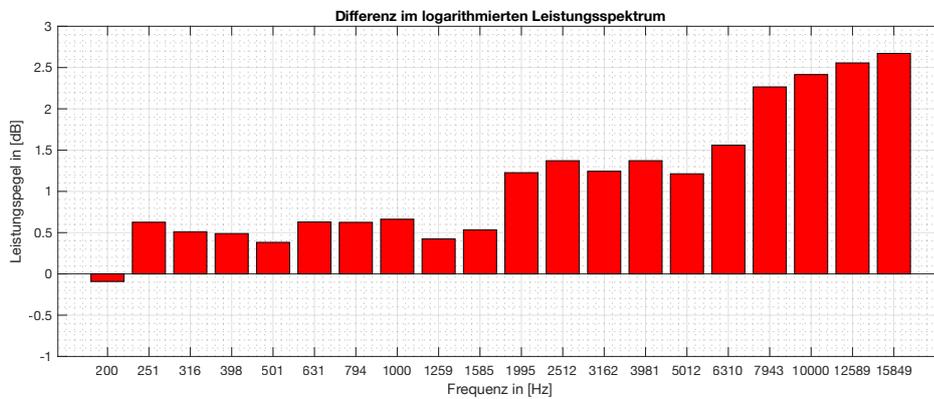
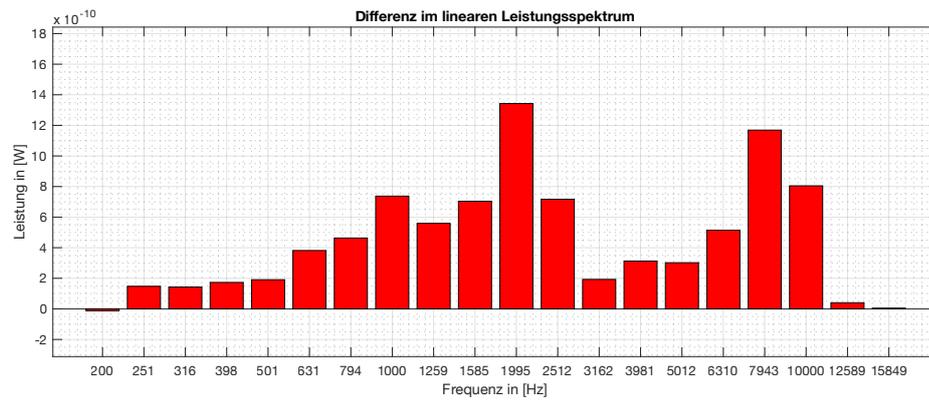


Abbildung 15 – Differenzabbildung der Schallleistung: (blau - rot) Schallleistung, rezipierter Messgrid - Schallleistung, 540 Messpunkte

3 Klangfarbenanalyse

Der zweite Teil dieser Bachelorarbeit befasst sich mit einer richtungsabhängigen Klangfarbenanalyse, die anhand von einer Gesangsaufnahme, die mit der in Kapitel 2.1.1 beschriebenen Mikrofonanordnung ausgeführt wurde, erörtert wird. Die richtungsabhängigen Klangfarbenänderungen sollen genauer untersucht werden. Dazu wird die in Abbildung 2 eingeführte Referenzierungskonvention der Mikrofone um Farben erweitert, um die Linien bzw. Punkte in den Analyseabbildungen den Richtungen genau zuordnen zu können. In Abbildung 16 ist die Farbzuzuweisung zu den Kanälen ersichtlich.

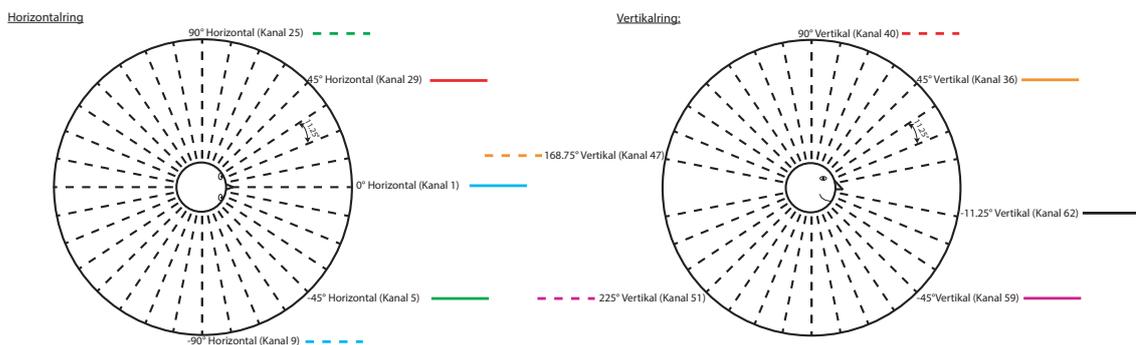


Abbildung 16 – Farbzuzuweisung der Kanäle

Bevor jedoch auf die Analyse eingegangen wird, werden zuerst die relevanten theoretischen Hintergründe aufgearbeitet.

3.1 Theorie

Der Begriff der Klangfarbe bzw. des Timbre wurde das erste Mal Mitte des 18. Jahrhunderts schriftlich beschrieben. So stellte beispielsweise Jean-Jacques Rousseau 1767 im „Dictionnaire de Musique“, das erste Mal die Theorie auf, dass das Timbre eines Instrumentes nicht von der Zahl, der Größe oder der Stärke der Schwingungen abhängt. Vielmehr müsse eine andere Ursache gefunden werden. Ungefähr 100 Jahre später beschreibt Hermann von Helmholtz in seinem Werk „Die Lehre von den Tonempfindungen“, dass sich musikalische Klänge vorwiegend durch die Eigenschaften Lautstärke, Tonhöhe und Klangfarbe beschreiben lassen [UCH⁺13]. Im folgenden werden, klassische und modernere Ansätze der Klangangfarbendefinition angeführt und einander gegenübergestellt. Im Anschluss wird auf die für die Klangfarbenanalyse notwendigen Modelle und Werkzeuge eingegangen.

3.1.1 Klangfarbendefinition

Klassische Definitionen

Laut Helmholtz ist die Klangfarbe: „[...] diejenige Eigenthümlichkeit, wodurch sich der Klang einer Violine von dem einer Flöte, oder Clarinette, oder menschlichen Stimme unterscheidet, wenn alle dieselbe Note in derselben Tonhöhe hervorbringen. [...] Wenn wir nun fragen, welcher äusseren physikalischen Verschiedenheit der Schallwellen die verschiedenen Klangfarben entsprechen, so haben wir gesehen, dass die Weite der Schwingung der Stärke, die Dauer der Schwingung der Tonhöhe entspricht. Von beiden kann die Klangfarbe nicht abhängig sein. Dann bleibt keine andere Möglichkeit übrig, als dass die Klangfarbe abhängt von der Art und Weise, wie die Bewegung innerhalb jeder einzelnen Schwingungsperiode vor sich geht.“ ([Hel77], S.20,31f). Weiters führte Helmholtz eine erste Charakterisierung und Einteilung verschiedener Klangfarben von Instrumenten in seinem Werk an. Carl Stumpf fasste diese Charakterisierung auf und erweitert sie in seiner Veröffentlichung „Tonpsychologie“. In der unten stehenden Tabelle sind die Klassifizierungen angeführt.

Nr.	Klangfarbenklasse	Instrumente	Eigenschaften	Anmerkung
1.	einfache Töne	Stimmgabeln Resonanzröhren weite gedackte Orgelpfeifen	weich, angenehm unkräftig,dumpf (tief)	gilt nicht für Tonhöhen in höchsten Oktave (des Klaviers)
2.	Klänge mit gleich- mäßig starkem 1. - 6. Teilton	Klavier, mensch. Stimme, Orgelpfeifen Horn, Flöten	klangvoll, reicher musikalisch, prächtiger	tonhöhenabhängig; Klangeindruck ansteigend mit tiefer werdender Grundtonhöhe
3a.	Klänge mit ungeradzahligen Teiltönen	gedackte Orgelpfeifen, Klaviersaiten (Mitte), Klarinette	hohl	Kombinationstöne werden auf den Positionen der nicht vorhandenen geradzahligen Teiltöne hörbar
3b.	Klänge mit vielen Obertönen		näselnd	
3c.	Klänge mit starkem Grundton		voll	
4.	Klänge mit Obertönen oberhalb des 6. Obertons	Zungenpfeifen, Oboe, Fagott mensch. Stimme Blechblasinst.	scharf, rau	tonhöhenabhängig; in der Schärfe ansteigend mit tiefer werdender Grundtöne

Tabelle 3 – Klangfarbenklassen nach Helmholtz und Stumpf (Quelle: [UCH⁺13], S. 99)

Die am meisten verwendete Definition für die Klangfarbe stammt von der American Standards Association und wurde 1960 veröffentlicht. Sie beinhaltet, dass die Klangfarbe „(...) als das Merkmal, welches es dem Hörer ermöglicht, zwei Schallereignisse, die unter gleichen Umständen mit gleicher Lautheit und Tonhöhe erklingen, zu unterscheiden“ [ASA60], verstanden werden kann.

Die Formanten

Die menschliche Stimme muss differenziert von anderen Instrumenten betrachtet werden. Maßgeblich für die Klangfarbenprägung der Stimme sind die Formantbereiche die speziell bei Vokalen beobachtet werden können. Formantbereiche sind Frequenzbereiche in welchen die Vokale, unabhängig vom Grundton, durch Resonanzen im Mund-, Nasen- und Rachenraum zusätzlich verstärkt werden [DDHW14]. Für die Langvokale der deutschen Sprache, den Gesangsformant und den Nasalformanten wurden in [VHH98] folgende Formantbereiche festgestellt:

Vokal:	Formantbereich:
U	200 bis 400 Hz und 650 bis 1000 Hz
I	150 bis 300 Hz und 1900 bis 2600 Hz
O	350 bis 550 Hz und 650 bis 1000 Hz
E	300 bis 500 Hz und 2000 bis 2600 Hz
A	600 bis 1000 Hz und 950 bis 1400 Hz
Nasalformant	1500 bis 2000 Hz
Gesangsformant	2400 bis 3200 Hz

Tabelle 4 – Formantbereiche (Quelle: [VHH98])

Bei weiblichem Gesang wurde festgestellt, dass sich die Formantbereiche der Grundtonhöhe bis zu einem gewissen Grad hin anpassen. Dies führt zu einer Verdunklung des Vokalcharakters. Der Gesangsformant kann vermehrt bei männlichem Gesang festgestellt werden [DDHW14]. Für die Analyse der weiblichen Gesangsaufnahme, die in dieser Arbeit durchgeführt wird, bedeutet dies, dass etwaige Formantverstärkungen in den Spektren in gewissem Maße immer als grundtonabhängig betrachtet werden müssen.

Klangfarbe als multidimensionale Größe

Neben Tonhöhe, Lautstärke, räumlicher Positionierung und Dauer gibt es noch viele anderen Größen psychoakustischer und musikalischer Herkunft, die zur Wahrnehmung bzw. Differenzierung von Klangfarben beitragen. Die Klangfarbe ist also multidimensional [McA99]. Vom Schallereignis bis zur binauralen Verarbeitung spielen eine Kette von Faktoren und Prozessen, die die Wahrnehmung einer Klangfarbe beeinflussen, wichtige Rollen. Es ist Gegenstand aktueller Forschungsgebiete, diese multidimensionale Wahrnehmung zu modellieren und strukturieren. In [McA99] wird beschrieben, wie die Dimensionierung von Klangfarbenräumen (sogenannte Timbre Spaces) durchgeführt werden kann. Es wird versucht, sowohl akustische als auch psychoakustische Faktoren in

Beziehung mit quantitativen Modellen, die die wahrnehmungsabhängigen Beziehungen beschreiben, in Verbindung zu setzen. Dies kann beispielsweise mit dem Verfahren der multidimensionalen Skalierung oder dem semantischen Differential geschehen [Gün12].

3.1.2 Die Timbre Toolbox

In dieser Arbeit wird auf eine bestehende Toolbox [PGS⁺11] zugegriffen. Die Toolbox wurde aufbauend auf der Multidimensionalität des Klangfarbenbegriffes von einem Forschungsteam aus Frankreich und Kanada für Matlab zusammengestellt. Diese Toolbox soll es ermöglichen, die Klangfarbe mit sogenannten Audiodeskriptoren, die mit verschiedenen Berechnungsmodellen errechnet werden können, zu beschreiben. Die Audiodeskriptoren sind so gewählt, dass die Multidimensionalität der Klangfarbe weitgehend abgedeckt wird [PGS⁺11].

Es wird unterschieden zwischen globalen und zeitabhängigen Deskriptoren. Die Berechnungsmodelle, die für die Berechnung der Deskriptoren in dieser Arbeit verwendet wurden, sind das gehörangepasste ERB-FFT Modell und das harmonische Modell. Sie werden in den folgenden Absätzen näher erläutert.

Das ERB-FFT Modell

ERB bedeutet Equivalent Rectangular Bandwidth (dt. Äquivalente Rechteckige Bandbreite) und bildet ein Konzept, mit welchem das Analysieren von Schallereignissen im auditiven System simuliert werden kann. Es wird davon ausgegangen, dass Schallereignisse bei der Wahrnehmung eine Bank von Bandpassfiltern, dessen Bandbreiten von der Mittenfrequenz abhängig sind, durchlaufen. Ähnlich wird es auch beim Konzept der kritischen Bandbreiten vorgeschlagen [PGS⁺11]. Schallereignisse, deren Frequenzen innerhalb dieser Bandbreiten liegen, werden vom Gehör gleich verarbeitet bzw. ausgewertet. Die Bandbreiten bei bekannter Mittenfrequenz können mittels folgender Näherung berechnet werden [DPJB13].

$$ERB = 0.108f + 24.7 \quad (27)$$

Die ERB, die dadurch berechnet wird, entspricht der eines idealen Rechteckfilters mit gleicher Bandbreite und Höhe. Bei der Berechnung der Deskriptoren wird also davon ausgegangen, dass sich das Frequenzspektrum von Schallereignissen in N hintereinander gereichte, auf dem menschlichen Gehör basierende ERB Filter aufteilt. In der Toolbox werden für eine Samplingfrequenz von $f_s = 44.1kHz$, $N = 77$ ERB-Filter zur Berechnung der Deskriptoren verwendet. Die Transformation in den Frequenzbereich wird dabei mit der Fast-Fourier-Transformation (FFT) durchgeführt [PGS⁺11]. Das ERB-FFT Modell wird für die Klangfarbenanalyse ausgewählt, um die Größen Spectral-Centroid und Spectral-Spread (siehe Kapitel 3.1.3) zu berechnen.

Das harmonische Modell

Beim harmonischen Modell wird davon ausgegangen, dass ein Audio-Signal aus Summation einzelner Sinuskomponenten mit leicht variierender Frequenz und Amplitude zu-

sammengesetzt werden kann [PGS⁺11].

$$s(t_n) \simeq \sum_{h=1}^H a_h(t_m) \cos(2\pi f_h(t_m) + \Phi_{h,0}(t_m)) \quad (28)$$

Die leicht variierende Frequenz $f_h(t_m)$ und die Amplitude $a_h(t_m)$ werden dabei in der Toolbox mithilfe einer Fensteranalyse mit einem 100 ms langem Blackmann-Fenster bestimmt [PGS⁺11]. Um die aus dem Sinusmodell abgeleiteten harmonischen Deskriptoren zu berechnen, wird auch eine Reihung der partiellen Teiltöne benötigt. Für die Berechnung der Tristimuluswerte (siehe Kapitel 3.1.3) muss zum Beispiel bekannt sein, wie die harmonischen Teiltöne gereiht sind. Dazu wird davon ausgegangen, dass sich die partiellen Frequenzen $f_h(t_m)$ aus Vielfachen einer Fundamentalfrequenz $f_0(t_m)$ ergeben. Um die leichte Inharmonizität, die in der Realität immer auftritt, auch abzubilden, wird die Abhängigkeit zwischen Teilton und Fundamentalfrequenz als:

$$f_h(t_m) = f_0(t_m)h\sqrt{1 + \alpha h^2} \quad (29)$$

dargestellt. Dabei ist α als Inharmonizitätsfaktor zu verstehen [PGS⁺11]. Das harmonische Modell wird für die Klangfarbenanalyse verwendet, um die Tristimuluswerte zu berechnen (siehe Kapitel 3.1.3)

3.1.3 Verwendete Werkzeuge und Deskriptoren

Mit Hilfe der Timbre-Toolbox werden die Größen Spectral-Centroid, Spectral-Spread und die Tristimulus-Werte berechnet. Da die Toolbox keine Möglichkeiten zur einfachen Visualisierung der Audiosignale bietet, wurde zusätzlich mit dem Matlabbefehl *spectrogram* ein Spektrogramm erstellt. Wie dies und die einzelnen Werte der Toolbox berechnet werden, wird in den folgenden drei Absätzen geklärt.

Spektrogramm

Das mit dem Matlabbefehl *spectrogram* erstellte Spektrogramm wird in Matlab mithilfe der Kurzzeit-Fourier-Transformation erstellt. Diese liefert einen Verlauf des Frequenzspektrums über die Zeit. Auf der x-Achse ist die Zeit aufgetragen, auf der y-Achse die Frequenz. Durch die Farbe des Frequenzverlaufs in der Abbildung wird die Leistung, die in einer Frequenz enthalten ist, angegeben. Welchen Zahlenwerten der Farbe zugeordnet sind, können aus der Balkenangabe auf der rechten Seite der Abbildung abgelesen werden [Mat18]. Es wird eine Fenstergröße von $N = 1024$ [samples] mit einer halben Fensterüberlappung zur Erstellung des Spektrogrammes gewählt.

Spectral-Centroid

Der Spectral-Centroid gibt Auskunft über den spektralen Schwerpunkt des Audiosignals. Dies bedeutet, der Wert wird größer, je mehr energiereiche Obertöne im Signal vorhanden sind. Für den Klang heißt das, dass obertonreichere Klänge heller und klangvoller klingen, teilweise sogar schärfer (siehe Tabelle 3). Signale mit niedrigerem Spectral-Centroid

klingen eher dumpfer und sind weniger obertonreich [Gün12]. Der Spectral Centroid errechnet sich aus:

$$\mu_1(t_m) = \sum_{k=1}^K f_k \cdot p_k(t_m) \quad \text{in [Hz]} \quad (30)$$

wobei $p_k(t_m)$ die normalisierte Amplitude im k ten (durch f_k beschrieben) ERB-Filter zur Zeit t_m ist. Sie wird mit:

$$p_k(t_m) = \frac{[a_k(t_m)]}{\sum_{k=1}^K a_k(t_m)} \quad (31)$$

berechnet [PGS⁺11]. Da sich bei Klängen mit stark wechselnden Formanten bzw. Frequenzspektren die Beschreibung der Klangfarbe mittels Spectral-Centroid problematisch gestalten kann, da der Schwerpunkt des Spektrums dann oft in Lücken bei Formantwechseln liegt [UCH⁺13], wird die Gesangsaufnahme auf einzelne Formanten bzw. Samples zusammengeschnitten. Die Bearbeitung der Aufnahme wird in Kapitel 3.2.1 genauer beschrieben.

Spectral-Spread

Die Kenngröße Spectral-Spread (dt. spektrale Spreizung) muss in Zusammenhang mit dem Spectral-Centroid gesehen werden. Sie ist definiert als:

$$\mu_2(t_m) = \left(\sum_{k=1}^K (f_k - \mu_1(t_m))^2 \cdot p_k(t_m) \right)^{\frac{1}{2}} \quad (32)$$

$\mu_1(t_m)$ ist dabei der Spectral-Centroid zur Zeit t_m . f_k bestimmt wiederum den k ten ERB-Filter, für welchen der aktuelle Teilwert berechnet wird. $p_k(t_m)$ steht für die normalisierte Amplitude aus Gleichung 31 [PGS⁺11]. Der Wert beschreibt, wie weit sich das Spektrum um den Mittelwert (Spectral-Centroid) spreizt. Ein höherer Wert deutet auf ein obertonreicheres und breiteres Spektrum hin.

Tristimulus

Die drei Tristimuluswerte sind abgeleitet aus der Optik, wo sie das Farbverhältnis zwischen Rot, Grün und Blau festhalten. Sie liefern eine genaue Beschreibung der harmonischen Teiltöne eines Spektrums und errechnen sich folgendermaßen:

$$\begin{aligned} T1(t_m) &= \frac{a_1(t_m)}{\sum_{k=1}^K a_k(t_m)} \\ T2(t_m) &= \frac{a_2(t_m) + a_3(t_m) + a_4(t_m)}{\sum_{k=1}^K a_k(t_m)} \\ T3(t_m) &= \frac{\sum_{k=5}^K a_k(t_m)}{\sum_{k=1}^K a_k(t_m)} \end{aligned} \quad (33)$$

Nr.	Tonhöhe	Formant
1. Sample	As'	A
2. Sample	C''	I
3. Sample	B'	U
4. Sample	Ab'	I
5. Sample	C''	E

Tabelle 5 – Tonhöhe und Formant der Samples

Die Gesangsaufnahme ist als 64 kanalige *.wav-Datei zusammengefasst. Die ersten 62 Kanäle sind, wie in Abbildung 16 angeführt, den Mikrofonen zugeordnet. Die zwei letzten Kanäle sind Trackingdaten, die die Kopfbewegung widerspiegeln und die Aufnahmen eines direkt vor dem Mund platzierten Referenzmikrofons. Die 64 kanalige Datei wurde mittels einer *for-Schleife* in Matlab in einem separaten Skript und den Funktionen *audio-read* und *audiowrite* auf 64 einzelne Dateien aufgeteilt. Die 62 relevanten Audiodateien wurden dann im Audioverarbeitungsprogramm Reaper in die 5 Samples aufgeteilt und zugeschnitten. Die Samples wurden so geschnitten, dass der gesungene Ton schon voll ausgebildet ist. Damit soll erreicht werden, dass Ein- bzw. Ausschwingungsvorgänge keinen Einfluss auf das Frequenzspektrum haben. Zur Analyse werden die Aufnahmen der in Abbildung 16 angeführten Kanäle bzw. Mikrofone verwendet, analysiert und gegenübergestellt.

3.2.2 Einsatz der Timbre Toolbox

Die Timbre Toolbox kann in [PGS⁺16] geladen werden. Anleitungen zur korrekten Installation der Toolbox sind in den geladenen Dateien enthalten. Im Ordner „doc“ der Toolbox befinden sich Beispielprogramme, die die Anwendung der Toolbox beschreiben. Das Matlabfile *get_descriptors_multifile_example.m* wird für die Berechnung der Deskriptoren herangezogen. Dieser Matlabcode wird verwendet, um die Audiodeskriptoren aller *.wav-Dateien, die sich in einem Quellverzeichnis befinden, zu berechnen. Dieser Code wird zu der Funktion *get_descriptors_multifile.m* umgeschrieben. Sie wird im Hauptfile *Timbre_Analysis.m* aufgerufen. Der Funktion wird das Quellverzeichnis übergeben, in welchem die zu analysierenden Dateien abgelegt sind, das Verzeichnis, in welches die Resultate gespeichert werden sollen und ein Vektor, welcher festlegt, welche Berechnungsmodelle verwendet werden sollen. Im Vektor wird eine 1 gesetzt, wenn das Modell verwendet werden soll. Die vierte und fünfte Position im Vektor repräsentieren das Harmonische Modell und das ERB-FFT Modell. Die Funktion ruft die Toolbox auf und berechnet alle Kennwerte, die mit den gewünschten Modellen berechnet werden können. Für jede Aufnahme werden drei *.mat-Dateien erstellt und in das übergebene Ergebnisverzeichnis geschrieben. In der einen Datei (Dateiname_desc.mat) sind die Deskriptoren im Zeitverlauf gespeichert. In der zweiten Datei (Dateiname_stat.mat) sind die statistischen Auswertungen der Kenngrößen enthalten. Die dritte Datei (Dateiname_rep.mat) ist für die Auswertung der Kenngrößen irrelevant.

3.2.3 Darstellung der Ergebnisse

Nach der Berechnung der Kennwerte aus den Audiodateien und der Speicherung in *.mat-Dateien werden die relevanten *.mat-Dateien im Matlabskript geöffnet. Die Deskriptoren werden in Matrizen gespeichert. Um die Architektur bzw. die Organisation der Matrizen zu verstehen, bietet es sich an, den beigelegten Matlabcode mit der Ordnerstruktur auszuführen. Zuvor muss jedoch der Matlabpfad in den Ordner Klangfarbenanalyse inklusive aller Unterordner gelegt werden. Nachdem alle Deskriptoren geladen worden sind, werden die Plot Befehle durchgeführt. In Tabelle 6 sind alle Abbildungen angeführt, die in Matlab erstellt werden, sowie die Richtungen, die in den Abbildungen miteinander verglichen werden.

Kenngröße	Abbildung	verglichene Richtungen
SpecCent 1. Sample Zeitverlauf	Abbildung-18 Abbildung-26 Abbildung-26 Abbildung-27	-11.25° vertikal, 0°, -45°, 45° horizontal -11.25°; 45°; -45° vertikal -11.25°; 90° vertikal, -90°; 90° horizontal -11.25°; 168.75°; 225° vertikal
SpecCent alle Samples Medianwert	Abbildung-19 Abbildung-28 Abbildung-29 Abbildung-29	-11.25° vertikal; 0°, -45°, 45° horizontal -11.25°, 45°, -45° vertikal -11.25°, 90° vertikal; -90°, 90° horizontal -11.25°, 168.75°, 225° vertikal
SpecSpread 1. Sample Zeitverlauf	Abbildung-20 Abbildung-30 Abbildung-30 Abbildung-31	-11.25° vertikal; 0°, -45°, 45° horizontal -11.25°, 45°, -45° vertikal -11.25°, 90° vertikal; -90°, 90° horizontal -11.25°, 168.75°, 225° vertikal
SpecSpread alle Samples Medianwert	Abbildung-21 Abbildung-32 Abbildung-33 Abbildung-33	-11.25° vertikal; 0°, -45°, 45° horizontal -11.25°, 45°, -45° vertikal -11.25°, 90° vertikal; -90°, 90° horizontal -11.25°, 168.75°, 225° vertikal
T1/T2 alle Samples Medianwert	Abbildung-22 Abbildung-34 Abbildung-34 Abbildung-35	-11.25° vertikal; 0°, -45°, 45° horizontal -11.25°, 45°, -45° vertikal -11.25°, 90° vertikal; -90°, 90° horizontal -11.25°, 168.75°, 225° vertikal
T3 alle Samples Medianwert	Abbildung-23 Abbildung-36 Abbildung-37 Abbildung-37	-11.25° vertikal; 0°, -45°, 45° horizontal -11.25°, 45°, -45° vertikal -11.25°, 90° vertikal; -90°, 90° horizontal -11.25°, 168.75°, 225° vertikal
Spektrogramm 1. Sample Spektrogramm 4. Sample	Abbildung-24 Abbildung-38	-11.25°, -45° vertikal; 0°, -45° horizontal -11.25° vertikal; 0° horizontal

Tabelle 6 – Organisation der Abbildungen

Da für die erste Note des Stückes das längste Sample mit einem stehenden Ton extrahiert werden kann, wird dieses Sample genauer untersucht. Es werden für dieses erste Sample zusätzlich die Zeitverläufe des Spectral-Centroid und des Spectral-Spread untersucht, abgebildet und verglichen. In allen Abbildungen werden immer die gleichen Richtungen miteinander verglichen, damit auch die verschiedenen Kenngrößen im gleichen Zusammenhang mit denselben Richtungen gesehen und analysiert werden. Für die Zeitverläufe des Spectral-Centroid/-Spread des ersten Samples wird eine normale Zeitabbildung mit der Zeit auf der x-Achse und der Kenngröße auf der y-Achse gewählt. Zusätzlich wird mit der Funktion *smoothdata()* der abgebildete Verlauf geglättet, um ihn übersichtlicher zu gestalten. Für die Medianwerte der gleichen Kennwerte wird auf der x-Achse das jeweilige Sample aufgetragen und auf der y-Achse die Kennwerte. Für die Darstellung der Tristimulus Werte 1 und 2 wird ein 2-D-Scatter-Plot gewählt, in welchem T1 auf der x-Achse und T2 auf der y-Achse aufgetragen wird. Für jedes Sample wird eine unterschiedliche Markerform gewählt, wie in den exemplarischen Tristimulusabbildungen im Kapitel 2.4 ersichtlich ist. Die Farben für die Marker werden wie in Abbildung 16 gewählt und auch mit diesen Farben gefüllt, lediglich für die Richtungen, die mit strichlierten Linien gekennzeichnet sind, werden die Marker als ungefüllt eingezeichnet. Für den dritten Tristimulus-Wert wird dieselbe Darstellungsvariante wie für die Medianwerte des Spectral-Centroid/-Spread gewählt.

3.2.4 Struktur des Matlabcodes

Zu Beginn des Matlabcodes wird also das Quell- und Zielverzeichnis definiert. Im Ordner „sounds“ befinden sich die zu analysierenden Audiodateien. In den Ordner „results“ sollen die Ergebnisdateien abgelegt werden. Im nächsten Schritt des Programmes wird überprüft, ob der Resultatsordner (Zielverzeichnis) leer ist. Falls dies der Fall ist, wird der Unterfunktion *get_descriptors_multifile.m* der Befehl zum Erörtern der Deskriptoren gegeben. Sind die Deskriptoren berechnet und in **.mat*-Dateien im Zielverzeichnis abgelegt, werden sie, in einer *for-Schleife* geöffnet, speziell organisierten Matrizen zugewiesen. Danach werden in folgender Reihenfolge die Analyseabbildungen erstellt:

1. Spektrogramme - 1. und 4. Sample
2. Spectral-Centroid Zeitverlauf - 1. Sample
3. Spectral-Spread Zeitverlauf - 1. Sample
4. Spectral-Centroid Medianwert - alle Samples
5. Spectral-Spread Medianwert - alle Samples
6. Tristimulus1/2 2-D-Scatterplots - alle Samples
7. Tristimulus 3 - alle Samples

3.3 Ergebnisse

In diesem Unterpunkt werden exemplarisch Abbildungen für die Zeitverläufe des Spectral-Centroid und Spectral-Spread, die Scatterplots der Tristimuluswerte angeführt und kurz besprochen. In den exemplarisch angeführten Abbildungen werden die Darstellungen gewählt, in welchen die Kennwerte aus den Richtungen 0° , 45° , -45° horizontal mit der Richtung -11.25° vertikal verglichen werden. Abschließend wird das erste Spektrogramm angeführt. Anhand der dargestellten Ergebnisse wird die Diskussion in Kapitel 4.2 aufgebaut. Alle anderen erstellten Abbildungen werden im Anhang (Kapitel 6) angeführt. Begonnen wird mit den zeitlichen Verläufen des Spectral-Centroids und Spectral-Spreads des ersten Samples.

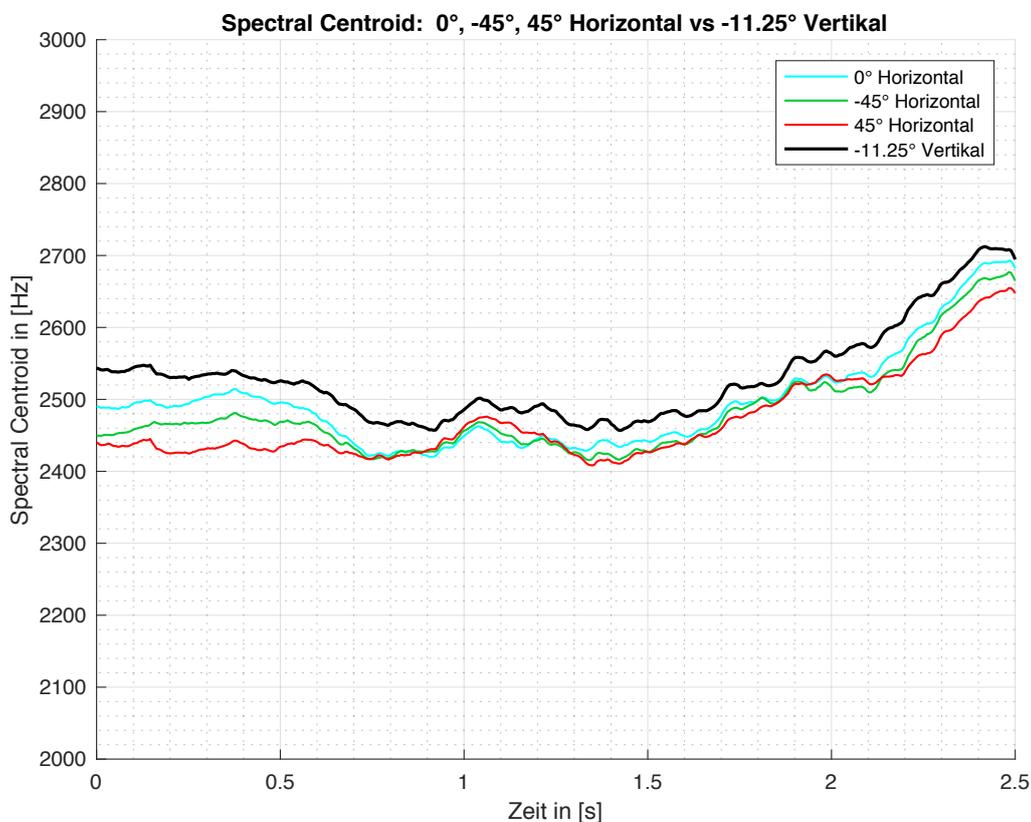


Abbildung 18 – Vergleich des Zeitverlaufs des Spectral-Centroids des ersten Samples

Es ist klar ersichtlich, dass der Verlauf der Kurve des Audiosignals, welches aus einem vertikalen Winkel von -11.25° aufgenommen wird, über den ganzen Tonverlauf die höchsten Werte aufweist. Weiters auffallend ist der Anstieg des Verlaufs der Kenngröße in den letzten 0.5 Sekunden des Samples. Auch wenn die Medianwerte aller Samples betrachtet werden, fällt auf, dass aus derselben Richtung der Spectral Centroid deutlich höher ist (siehe Abbildung 19).

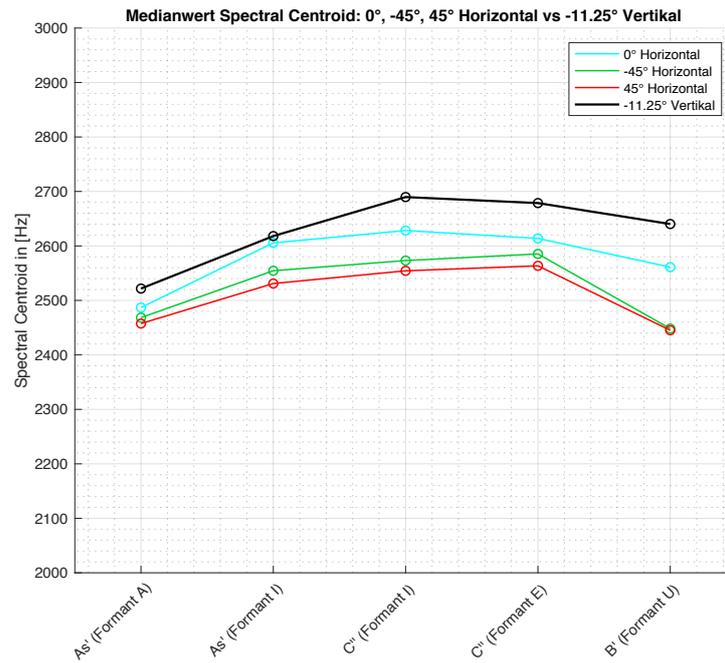


Abbildung 19 – Vergleich der Medianwerte des Spectral-Centroids des ersten Samples

Wird der Spectral Spread betrachtet, fällt auf, dass die Kurven schon enger beinander liegen. Der Anstieg der Werte zum Schluss ist wieder ersichtlich.

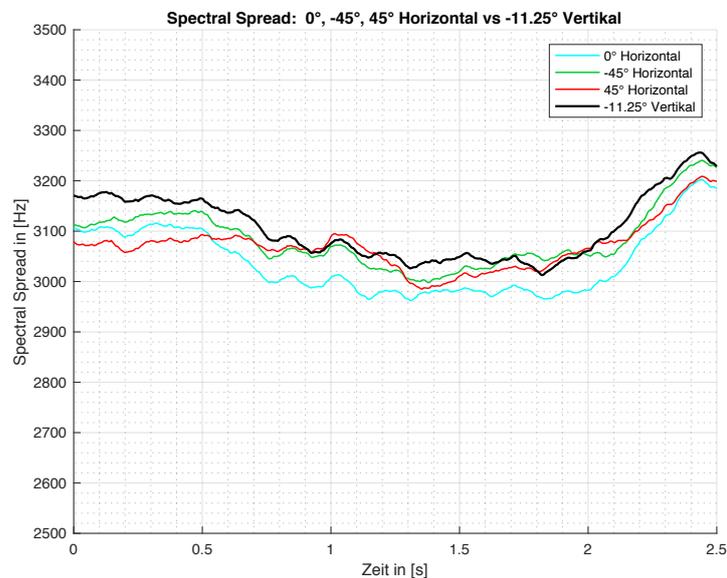


Abbildung 20 – Vergleich des Zeitverlaufs des Spectral-Centroids des ersten Samples

Betrachtet man allerdings die Medianwerte der Richtung -11.25° vertikal kann wieder festgestellt werden, dass diese deutlich höher als die der anderen Richtungen sind.

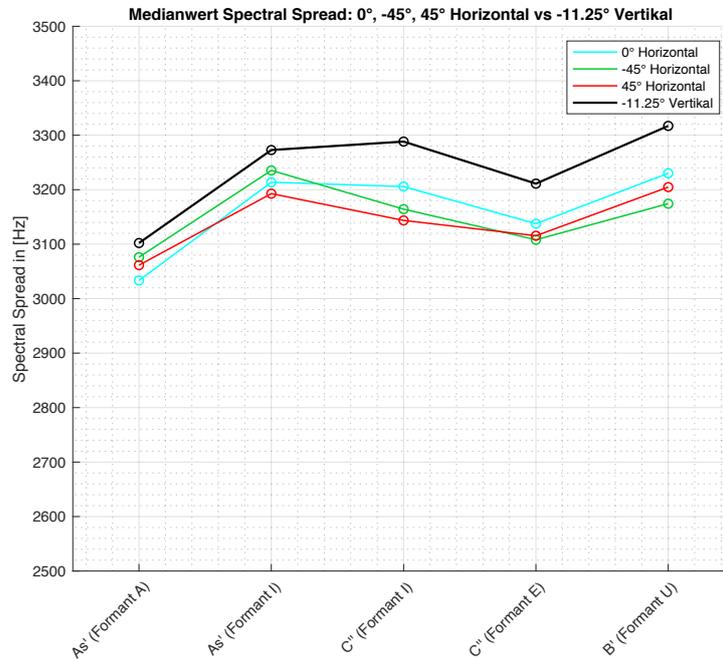


Abbildung 21 – Vergleich des Zeitverlaufs des Spectral-Spread des ersten Samples

In der Gegenüberstellung des ersten und des zweiten Tristimuluswert fällt auf, dass der T2-Wert des ersten Samples aus der Richtung -11.25° vertikal niedriger als die T2-Werte der anderen Richtungen ist. Jedoch ist der T1 Wert aus Richtung -11.25° vertikal größer. Für Sample 5 (Marker mit Diamantenform), wiederum ist der T2-Wert sogar höher als die der anderen Richtungen. Der Tristimulus 3-Wert ist für alle analysierten Samples sehr niedrig. Jedoch für die Richtung aus -11.25° kann festgestellt werden, dass er minimal höher ist, speziell bei den Samples 3, 4 und 5.

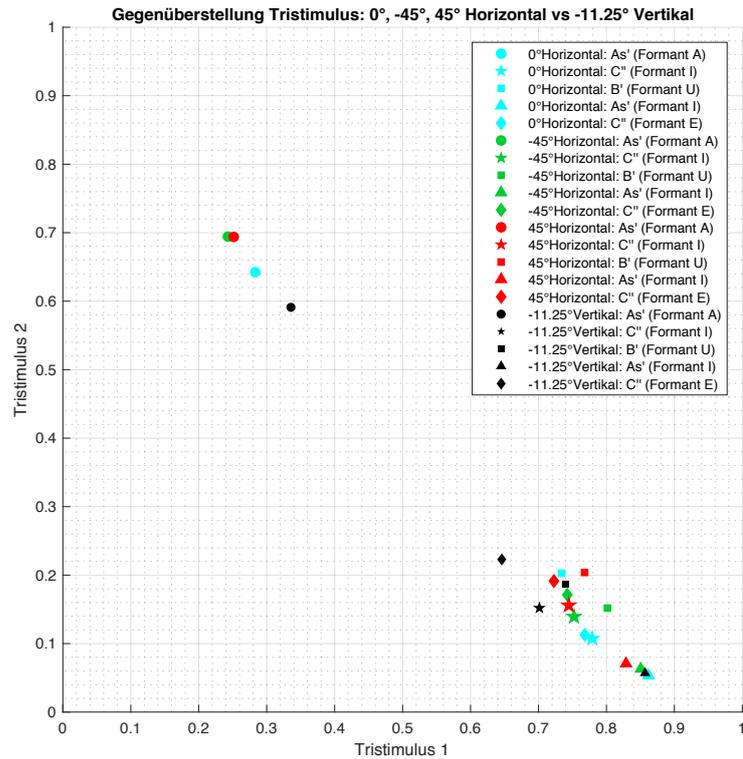


Abbildung 22 – Gegenüberstellung der Tristimuluswerte

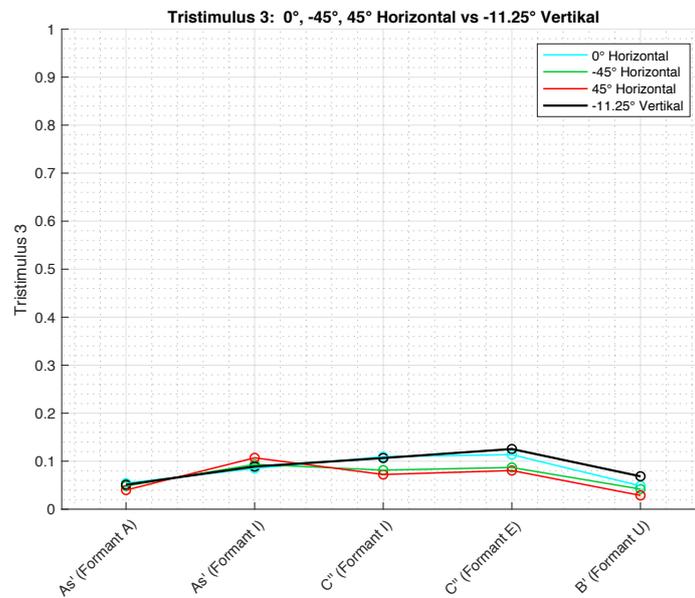


Abbildung 23 – Tristimulus 3 für alle Samples

In der unten angeführten Vergleichsabbildung der Spektrogramme der Aufnahmen aus den unterschiedlichen Richtungen ist ersichtlich, dass gegen Ende der Aufnahme bestimmte Frequenzbereiche verstärkt hervortreten. Gegen Ende der Aufnahme ist auch das in der Aufnahme hörbare Vibrato, welches durch Frequenzmodulation entsteht, sichtbar.

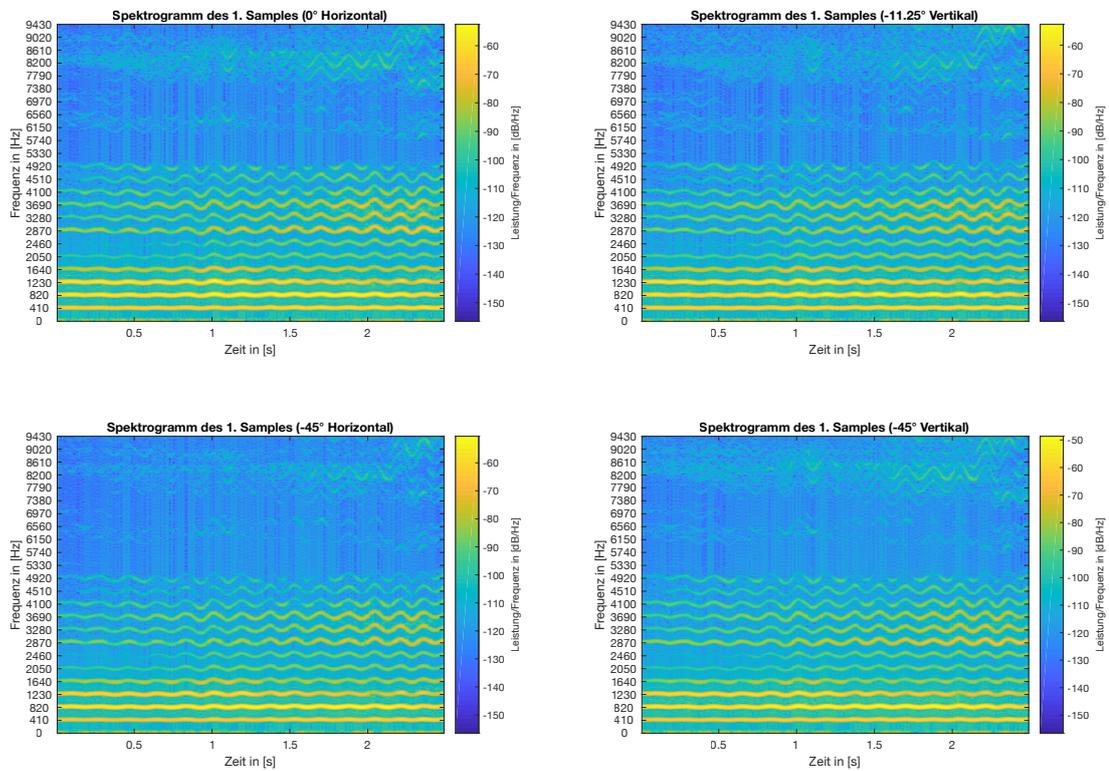


Abbildung 24 – Spektrogramme für 1. Sample

4 Diskussion und Ausblick

4.1 Schalleistungsberechnung

Vergleicht man Abmessungen und die Positionen der Mikrofone des Vertikalrings vom Doppel-Kreis-Mikrofon-Array mit den in Kapitel 2.1.4 angeführten dem Normvorgaben aus Tabelle 1 können normähnliche Mikrofonpositionen gefunden werden. Der Radius des Doppelkreis-Mikrofon-Arrays liegt in der Größenordnung der Norm. Die Messung mit dem Doppelkreis Mikrofon-Array ist jedoch nicht normzulässig, da nicht alle 540 vorhandenen Messpunkte gleichzeitig aufgenommen worden sind. Allerdings kann durch die gute Reproduzierbarkeit der Messung die Leistungsberechnung mittels Hüllflächenverfahren durchaus durchgeführt werden. Es muss allerdings beachtet werden, dass für die drei untersten Terzbänder die zweite Fernfeldbedingung verletzt wird (siehe Tabelle 2). Dies führt zu einem Fehler in der Berechnung der Schalleistung in diesen Terzbändern. Da die Fernfeldbedingungen nicht eingehalten wird gilt die diskretisierte Fernfeldgleichung (Gleichung 15) nicht mehr. Mit einer Nahfeld Abschätzung, wie sie in [PZS⁺16] durchgeführt wird, könnte der durch die Verletzung der Fernfeldbedingung entstehende Fehler verringert werden. Um die beiden Flächeneinteilungen miteinander zu vergleichen, wird die zweite Einteilung, in welcher alle 540 Messpunkte verwertet worden sind, als Referenzeinteilung angesehen. Aufgrund der geringeren Gewichtung (siehe Kapitel 2.3) liefert die Schalleistungsberechnung mit der zweiten Flächeneinteilung ein genaueres Ergebnis. Vergleicht man die beiden Flächeneinteilungen miteinander (Abbildung 14), fällt auf, dass mit der reduzierteren Variante die Schalleistung eigentlich überschätzt wird. Bis auf das erste Frequenzband wird, bei Berechnung der Schalleistung mit der reduzierten Einteilung, durchgehend ein höherer Wert ermittelt. Betrachtet man die Differenz der Schalleistungspegel, die in Abbildung 15 dargestellt sind, ist ersichtlich, dass mit zunehmender Frequenz die Differenz der Pegel nahezu linear steigt. In den höchsten Frequenzbändern beträgt die Differenz zwischen 2 und 2,7 dB. Werden allerdings die Frequenzbereiche, in denen sich die menschlichen Stimme bewegt (Tabelle 3), in die Analyse miteinbezogen, sind die höherfrequenten Frequenzbänder zur Analyse eher irrelevant. Würde also in Zukunft auch die Schalleistung von menschlichen Sängerinnen und Sängern mit Hilfe des Doppelkreis-Mikrofon-Arrays durchgeführt werden, könnten in den, für die Stimme relevanten Frequenzbändern, mittels der ersten Flächeneinteilung durchaus schon brauchbare Berechnungsergebnisse, bzw. Abschätzungen der Schalleistung erreicht werden. Weiters könnte mit einer Erweiterung des horizontalen Mikrofonrings um 4 Mikrofone, sodass die Mikrofone nun im 10° Winkel zueinander stehen, die Messpunktanordnung, wie sie in der reduzierten Flächeneinteilung angeführt worden ist, auch in der Praxis realisiert werden. Es wäre nun möglich, die in der ersten Berechnungsvariante der Schalleistung durchgeführte Abtastung der Kugelhüllfläche direkt mit einer Messung zu erreichen. Die Schallquelle müsste dazu lediglich direkt auf den Schnittpunkt beider Mikrofonringe gerichtet werden.

4.2 Klangfarbeanalyse

Zur Diskussion der Ergebnisse der Klangfarbenanalyse wurden in Kapitel 3 exemplarische Abbildungen angeführt, es wurden dabei die Richtungen ausgewählt, die klanglich den besten Eindruck vermitteln. Im angeführten Spektrogramm (Abbildung 24) ist erkennbar, dass der Grundton mit der meisten Energie beim ersten Sample bei ca. 410 Hz liegt. Die Abstrahlcharakteristik des in [BFR18a] vermessenen HATS kann mit der Darstellungsapplikation aus [BFR18a] für eine ähnliche Frequenz wie die des 1. Samples für 0° vertikal und 50° vertikal erstellt werden. Die Aufgabe des HATS ist es den Oberkörper

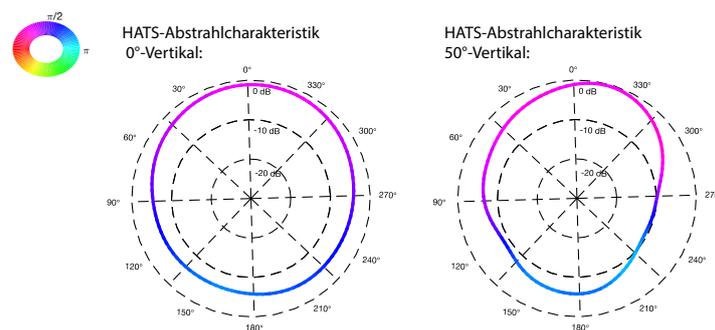


Abbildung 25 – HATS-Abstrahlcharakteristik 0° und 50° vertikal

eines Menschen zu simulieren und es wird angenommen, dass sich die Abstrahlmuster für einen menschlichen Oberkörper ähnlich verhalten. Im ersten Abstrahlmuster ist ersichtlich, dass bei 0° vertikal auch in der Horizontalen, im Bereich von $\pm 45^\circ$, keine maßgeblichen Unterschiede erkennbar sind. Deshalb werden für die exemplarischen Abbildungen in Kapitel 3.3 die Gegenüberstellungen von $-11,25^\circ$ vertikal mit 0° , -45° , 45° horizontal gewählt. Für diese Richtungen sind ähnliche Ergebnisse zu erwarten. Auch im vertikalen Bereich (45° vertikal) ist mit ähnlichen Ergebnissen zu rechnen, betrachtet man die zweite Abstrahlcharakteristik in Abbildung 25. Die Auswertung von der Richtung 45° vertikal ist im Anhang angeführt. Die Vergleiche der anderen Richtungen, die zur Analyse ausgewählt worden sind (siehe Abbildung 16), sind im Anhang hinterlegt. Anhand der Abbildungen des Spectral Centroids und des Spectral Spreads (Abbildung 18 und 20) kann klar gesehen werden, dass die Richtung $-11,25^\circ$ vertikal die höchsten Verläufe aufweist. Auch bei den Medianwerten über alle Samples betrachtet, schneidet dieselbe Richtung am besten ab. Auch die im Anhang angeführten Vergleiche liefern dasselbe Ergebnis. Betrachtet man nur die Kennwerte Spectral-Centroid und Spectral-Spread, schneidet die Richtung $-11,25^\circ$ am positivsten ab. Wird jedoch die Tristimulusgegenüberstellung aus Abbildung 22 miteinbezogen, zeigt sich, dass speziell beim ersten Sample (kreisförmiger Marker) ein geringerer T2-Wert, jedoch ein höherer T1-Wert im Vergleich zu den anderen Samples abgelesen werden kann. Es scheint, als ob die Richtung $-11,25^\circ$ vertikal ein weniger energiereiches Obertonspektrum besitzt. Die Tristimulus-Werte müssen jedoch immer in Relation gesehen werden. Der T2-Wert kann auch steigen für ein Signal, dessen Grundton weniger energiereich vorhanden ist. Es ist also vielmehr ein ausgewogenes Mittelmaß zwischen T1 und T2, was einen vollen und runden Klang ausmacht. Genauere Aussagen

über das Obertransspektrum, mittels Tristimulusgegenüberstellung wären möglich, wenn man für alle Samples den Grundton auf das gleiche Lautstärkeniveau bringen würde. Dies wurde im Zuge der in dieser Arbeit durchgeführten Analyse vernachlässigt. Die Tristimuluswerte dienen eher dazu einen Vergleich der unterschiedlichen Tonhöhen und Formanten zu ermöglichen. Aus den Abbildungen der T3-Werte, die zusätzlich im Anhang angeführt sind, kann keine aussagekräftige Schlussfolgerung getroffen werden, da die Kurven für alle Richtung in den gleichen, sehr niedrigen Bereichen liegen. Es kann also daraus geschlossen werden, dass ab dem fünften Oberton keine maßgeblichen Beiträge der Teiltöne zur Klangfarbe mehr vollzogen werden. Das abschließende Spektrogramm liefert eine mögliche Erklärung für den Anstieg des Spectral Centroids bzw. des Spectral Spreads in den letzten 0,5 s des ersten Samples. Gegen Ende des gehaltenen Tons kann in der Aufnahme gehört werden, dass die Sängerin eine Formantänderung im Gesang vollzieht. Diese Änderung führt wohl dazu, dass die Frequenzbereiche von 2870 Hz bis 4100 Hz gegen Ende des Samples stärker hervortreten. Der energiereichere und höherfrequente Bereich, der sich gegen Ende des Samples abbildet, trägt also zum Anstieg des Spectral-Centroids und des Spectral-Spreads bei. Würde man die Samples weiter kürzen d.h. würde der Teil des Samples in welchem die Formantveränderung auftritt weggeschnitten werden, könnte der Anstieg im Zeitverlauf des Spectral Centroid und des Spectral Spread verhindert werden. Im Anhang ist für das 4. Sample ein zusätzliches Spektrogramm angeführt, bei welchem die gleiche Tonhöhe wie beim 1. Sample, welches im Spektrogramm in Kapitel 3.3 analysiert wird, gehalten wird. Im 4. Sample wird zwar dieselbe Tonhöhe (As') untersucht, jedoch ein unterschiedlicher Formant. Im ersten Sample wird ein „A“ gesungen und im 4. Sample ein „I“ (siehe Tabelle 5). Die unterschiedlich stark hervortretenden Obertonbereiche sind in den Abbildung ersichtlich, speziell wenn die erste Hälfte des Spektrogrammes aus Kapitel 3.3 mit dem Spektrogramm aus dem Anhang verglichen wird. Die Obertöne bei 2000 - 2500 Hz sind im 4. Sample (Formant I) präsenter als im 1. Sample (Formant A).

Betrachtet man nur die Ergebnisse der durchgeführten Klangfarbenanalyse, decken sich diese mit den subjektiven Empfindungen die beim Anhören der einzelnen Samples aus verschiedenen Richtungen erlangt wurden. Die Richtung $-11,25^\circ$ vertikal schneidet bei der Auswertung der Kenngrößen am positivsten ab. Dieselbe Richtung wurde auch bei der Hörprobe, vom Verfasser dieser Arbeit als die Richtung mit dem vollsten Klang favorisiert. Für die Praxis würde dies bedeuten, dass für eine volle Klangfarbe bei weiblichen Gesangsaufnahmen die Mikrofone nicht zwingend normal (0° horizontal) auf die Sängerin gerichtet sein müssen. Es kann mit der Positionierung eines Mikrofones welches vertikal leicht nach unten gerichtet ist, ein gutes bzw. sogar ein besseres Ergebnis in Bezug auf die Klangfarbe erzielt werden. Es soll jedoch an dieser Stelle erneut angemerkt werden, dass die Auswertung der Trackingdaten, der Kopfbewegungen während der Aufnahme, nicht durchgeführt worden ist. Die angeführte Schlussfolgerung kann also nicht für die Praxis verallgemeinert werden, jedoch ist sie für die Aufnahmen die im Doppelkreis-Mikrofon-Array durchgeführt und analysiert wurden zutreffend. Weitere Untersuchungen der Abstrahlcharakteristika von menschlichem Gesang, sowie deren Beeinflussungsmöglichkeiten durch die Sängerin/den Sänger können weitere Erkenntnisse für die Praxis liefern.

Bezüglich der Anwendung der Timbre-Toolbox würde es sich anbieten, eine größere Datenbank an Aufnahmen von einzelnen gesungenen Tönen, mit unterschiedlichen gesungenen Formanten und Mundstellungen, über eine Oktave verteilt, in verschiedenen Dynamikstufen zu erstellen, wie es in [Gün12] für verschiedenste Instrumente durchgeführt worden ist. Für die Analyse von Aufnahmen, bei welchen ein Ton konstant gehalten wird, bietet die Timbre-Toolbox durchaus einen strukturierten Ansatz zur Klangfarbenanalyse. Die Kennwerte, die mit der Toolbox berechnet werden können, sind aussagekräftiger für weniger abwechslungsreichere Aufnahmen. Deshalb wurde auch versucht, durch Zuschneidung der Aufnahme die Gesangsaufnahme weitestgehend zu vereinfachen. Dadurch, dass die Aufnahmen so beschnitten werden, dass lediglich voll ausgebildete Töne in ihnen enthalten sind, sind allerdings auch keine Ein- bzw. Ausschwingvorgänge in den Aufnahmen erfasst. In Bezug auf die Klangfarbe können so lediglich Aussagen in Hinblick auf das Frequenz- bzw. Obertonspektrum getroffen werden. Ein statistischer Vergleich der Genauigkeit von den verschiedenen Berechnungsmodellen der Timbre-Toolbox kann in [KEDS17] eingesehen werden.

5 Schlusswort

Durch das Erstellen dieser Arbeit wurde dem Verfasser Einblick in neueste Forschungsgebiete gewährt. Auf Basis der durchgeführten Schalleistungsberechnung und Klangfarbenanalyse wurden theoretische Gebiete aufgearbeitet, die sowohl die Forschung als auch die Berufspraxis im Bereich der Akustik und des Toningenieurwesens maßgeblich prägen. Ein weiterer wichtiger Punkt, welcher durchaus als erfolgreich und prägend für den Verfasser dieser Arbeit angesehen werden kann, ist die Herangehensweise an wissenschaftliche Themen und Forschungsgebiete. Von der Literaturrecherche bis zur Darstellung der Ergebnisse wurde stets darauf geachtet, dass alles strukturiert und wissenschaftlich akkurat durchgeführt wurde. Die Fertigstellung dieser Arbeit wäre ohne die Unterstützung des Instituts für Elektronische Musik und Akustik an der Universität für Musik und darstellende Kunst in Graz nicht möglich gewesen. Im speziellen gilt ein großer Dank Herrn Univ. Prof. DI Dr. Alois Sontacchi, der fachlich hochwertige und wichtige Hilfestellungen auf alle Anfragen, sowohl fachtheoretischer als auch organisatorischer Natur, gegeben hat. Ein weiterer Dank gilt Herrn DI Bsc Manuel Brandner für die ausführliche Erläuterung des Messprozesses und Messaufbaus.

6 Anhang

Die Abbildungen, die mit Zeitverlauf deklariert sind, sind Auswertungen, die sich auf das erste Sample (siehe Kapitel 3) beziehen.

A1: Spectral-Centroid (Zeitverlauf und Medianwerte):

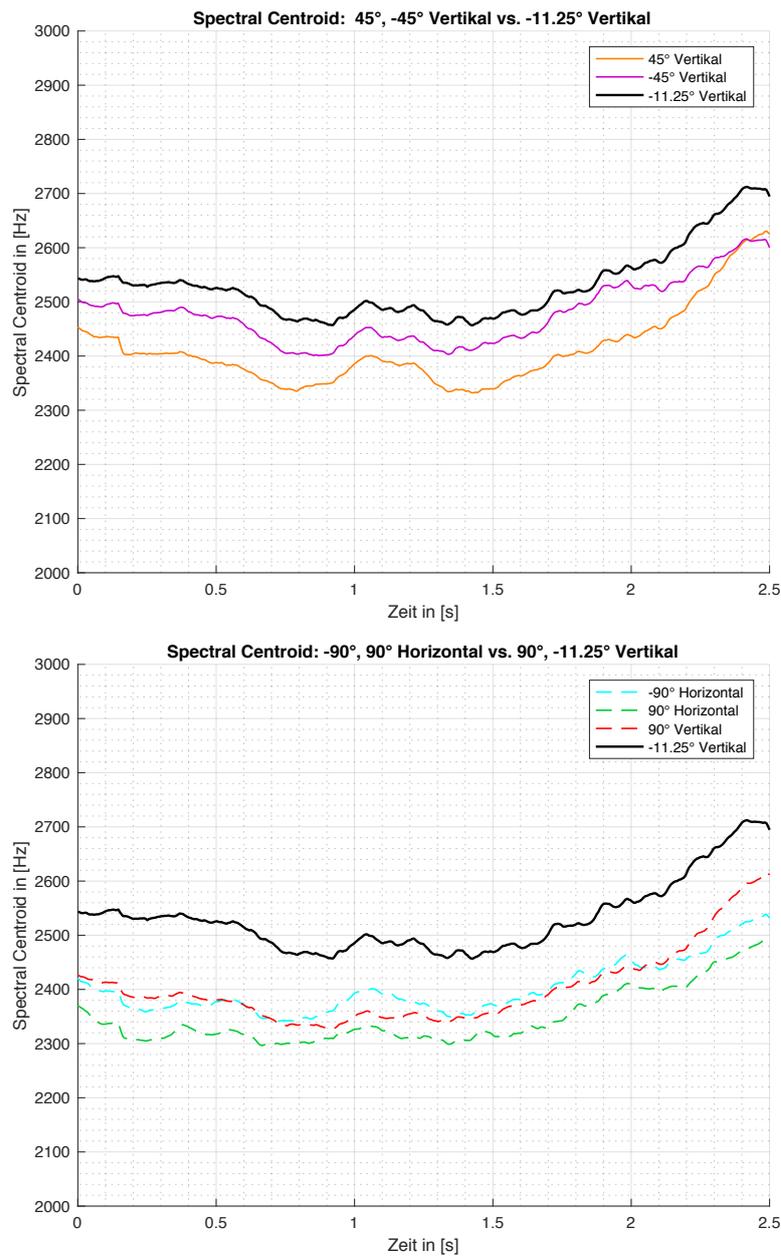


Abbildung 26 – Spectral Centroid Zeitverlauf zusätzliche Auswertungen 1

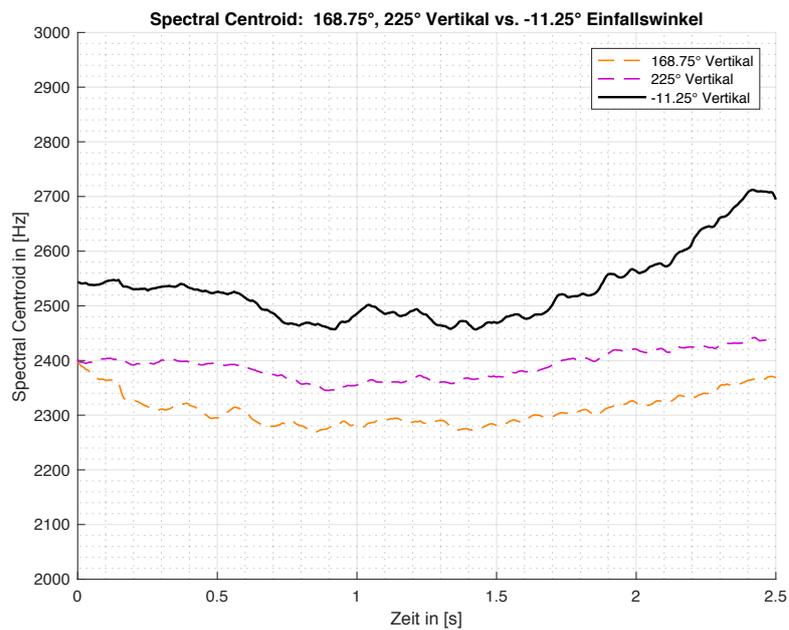


Abbildung 27 – Spectral Centroid Zeitverlauf zusätzliche Auswertung 2

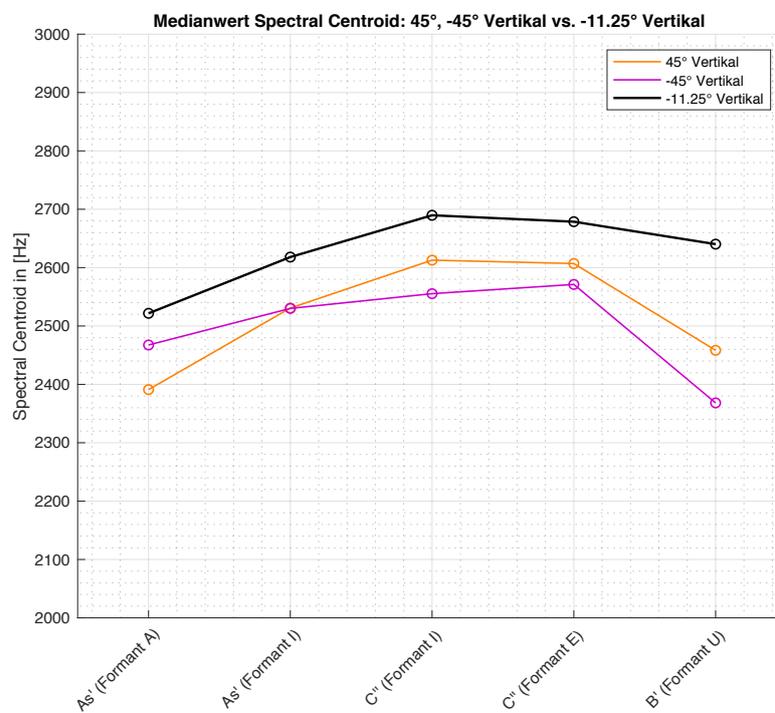


Abbildung 28 – Spectral Centroid Medianwert zusätzliche Auswertung 1

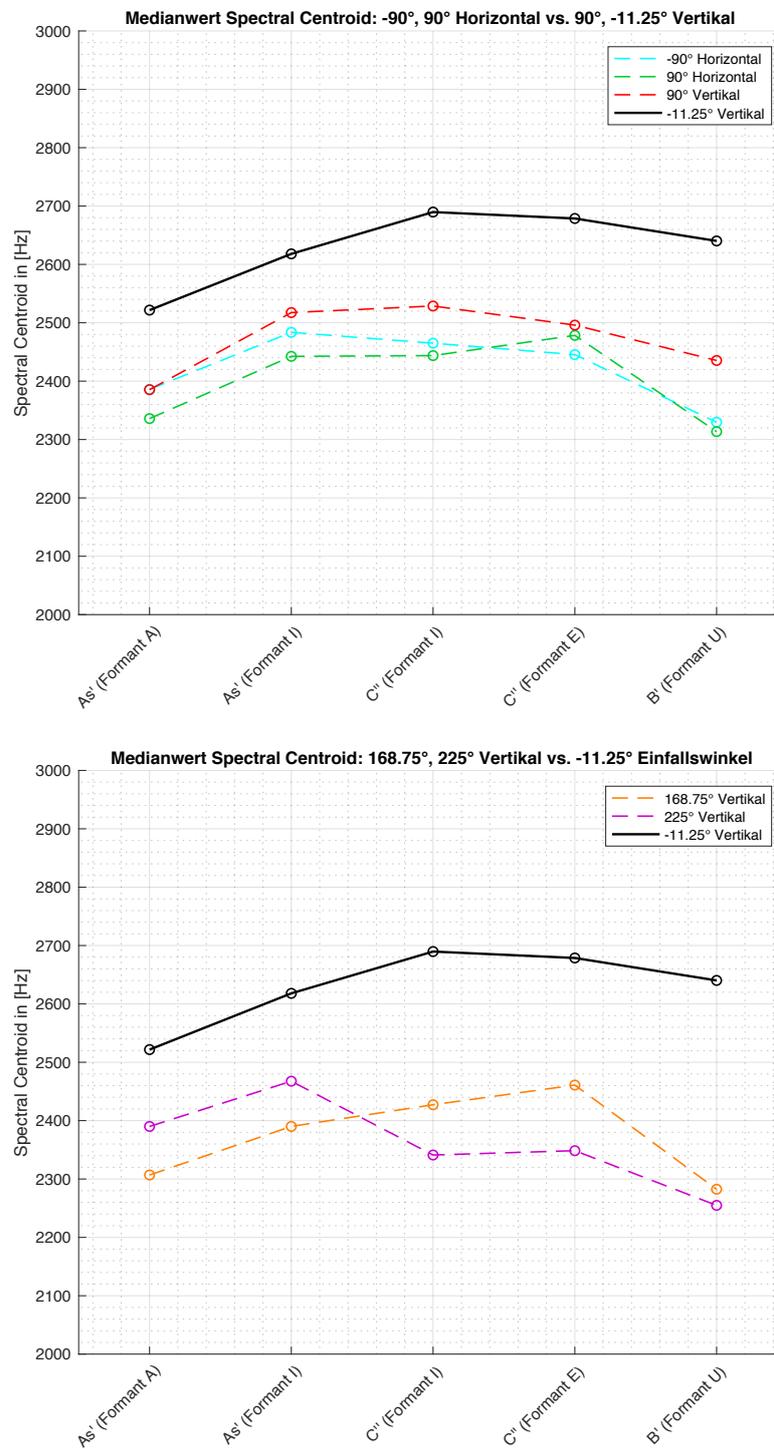


Abbildung 29 – Spectral Centroid Medianwert zusätzliche Auswertung 2

A2: Spectral-Spread (Zeitverlauf und Medianwerte):

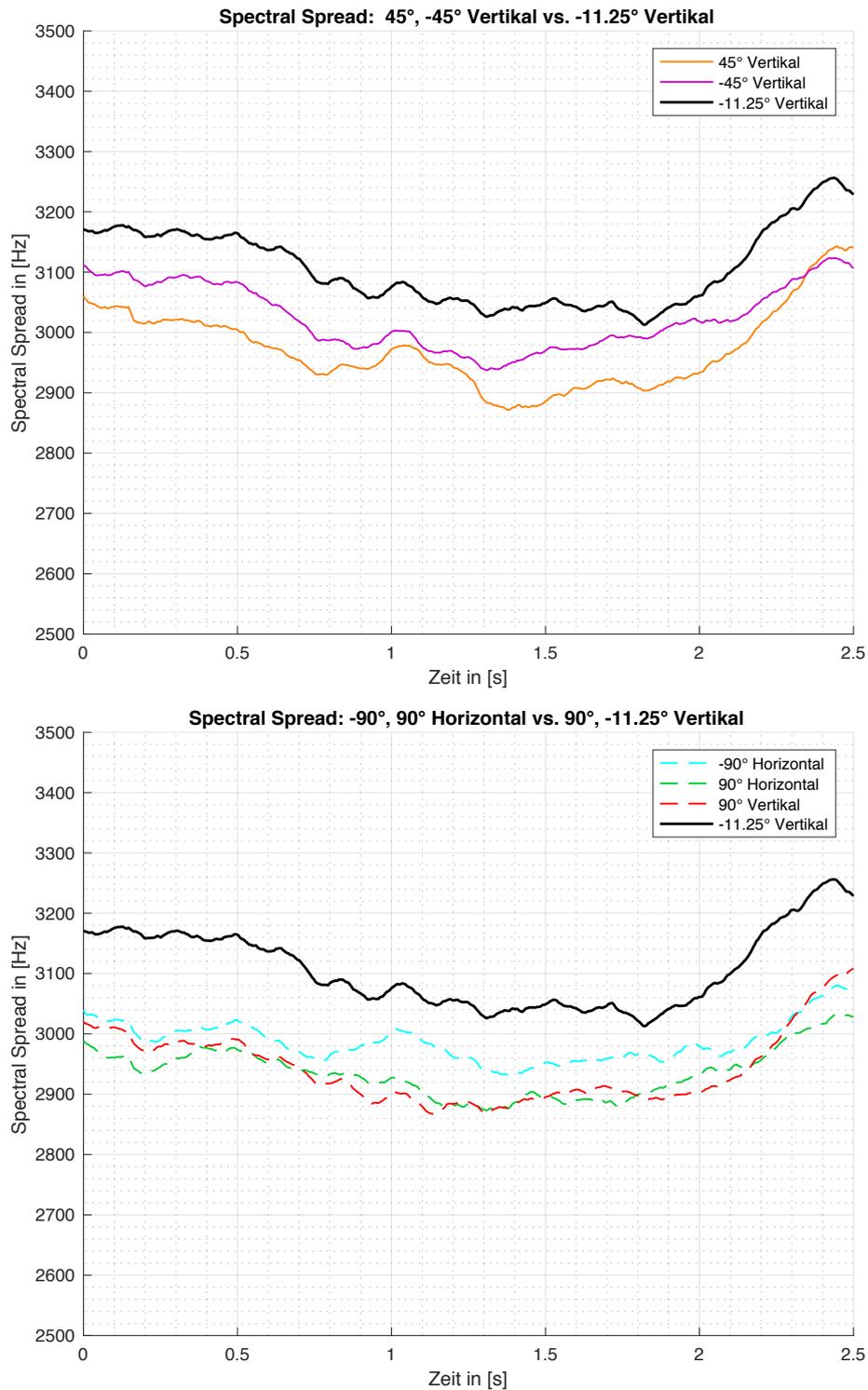


Abbildung 30 – Spectral Spread Zeitverlauf zusätzliche Auswertung 1

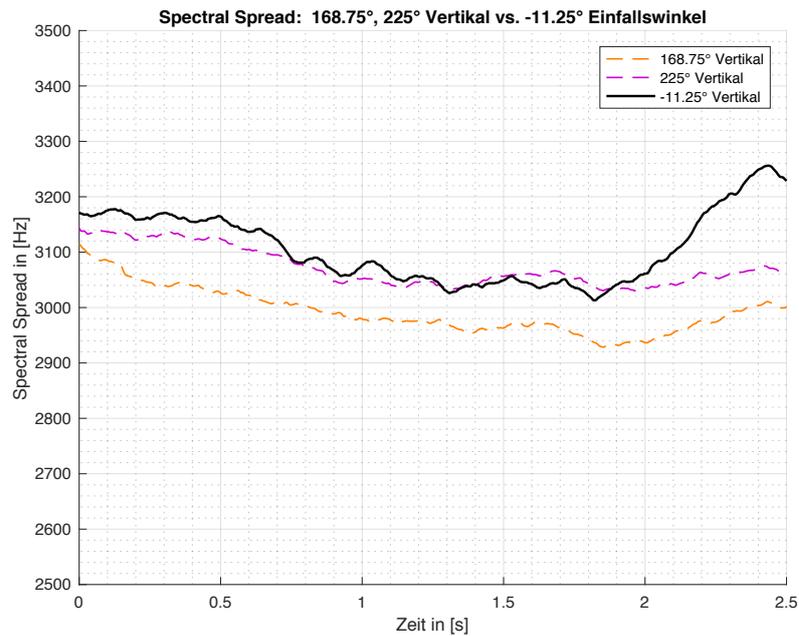


Abbildung 31 – Spectral Spread Zeitverlauf zusätzliche Auswertungen 2

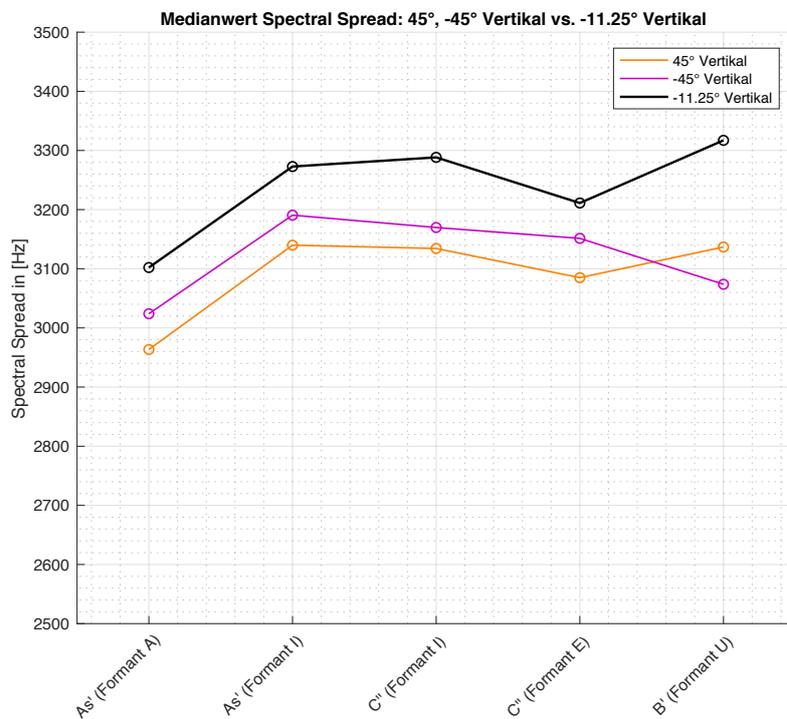


Abbildung 32 – Spectral Spread Medianwert zusätzliche Auswertungen 1

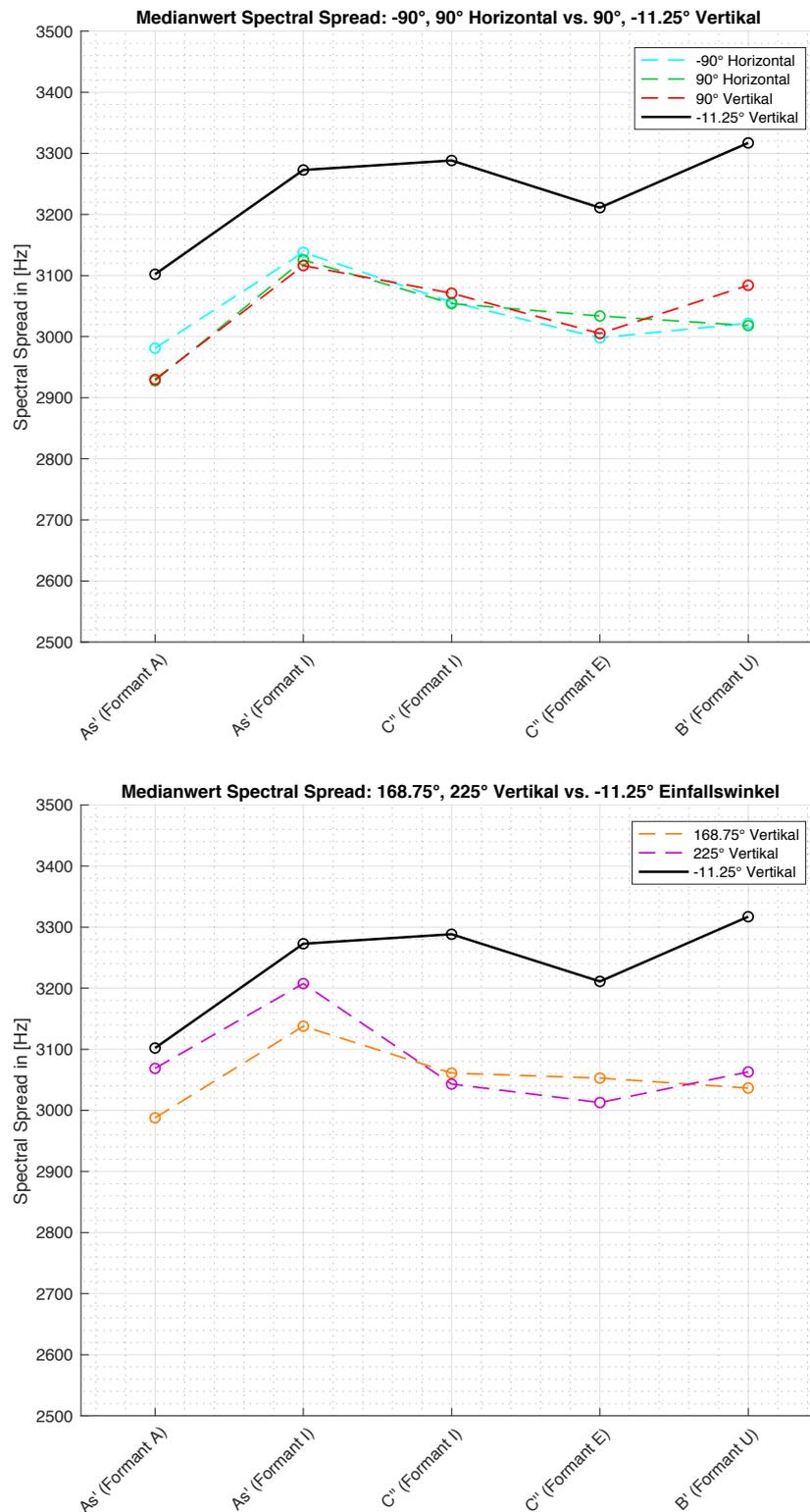


Abbildung 33 – Spectral Spread Medianwert zusätzliche Auswertungen 2

A3: Tristimulusgegenübersellung T1/T2:

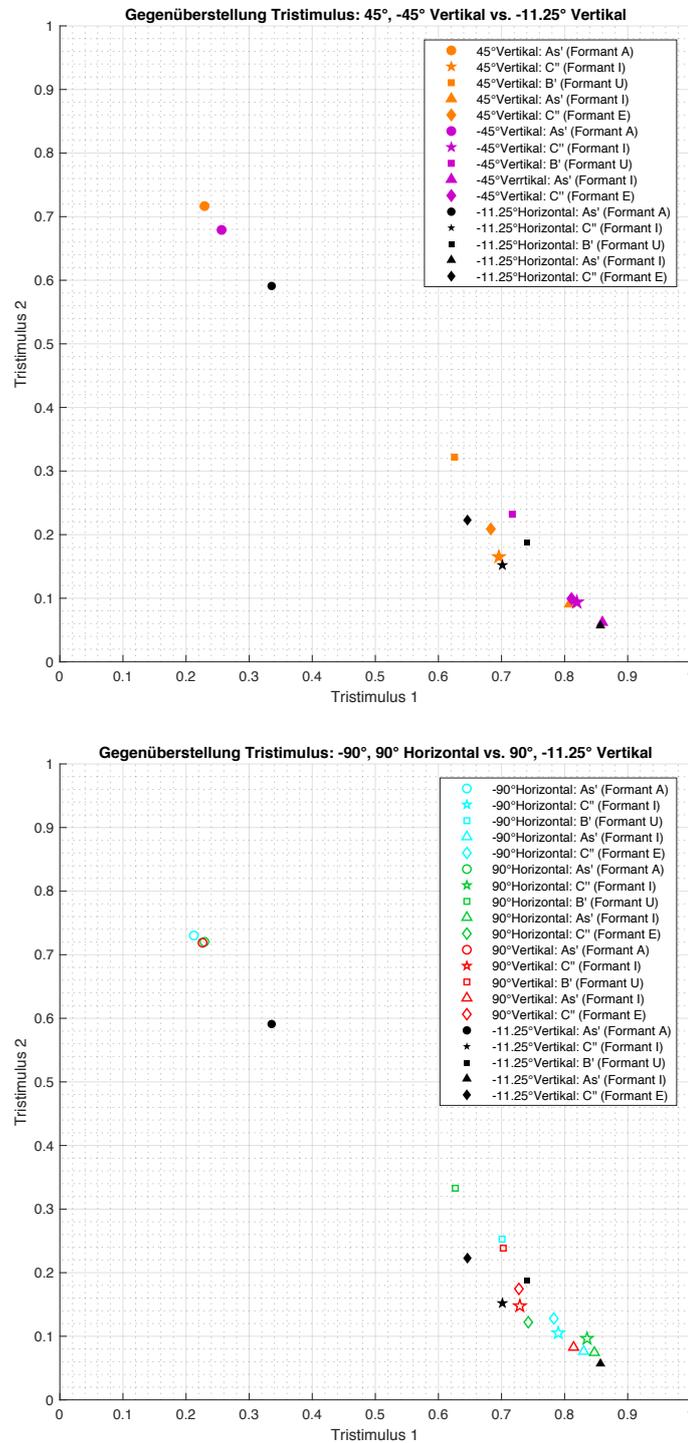


Abbildung 34 – Tristimulus Gegenüberstellung zusätzliche Auswertungen 1

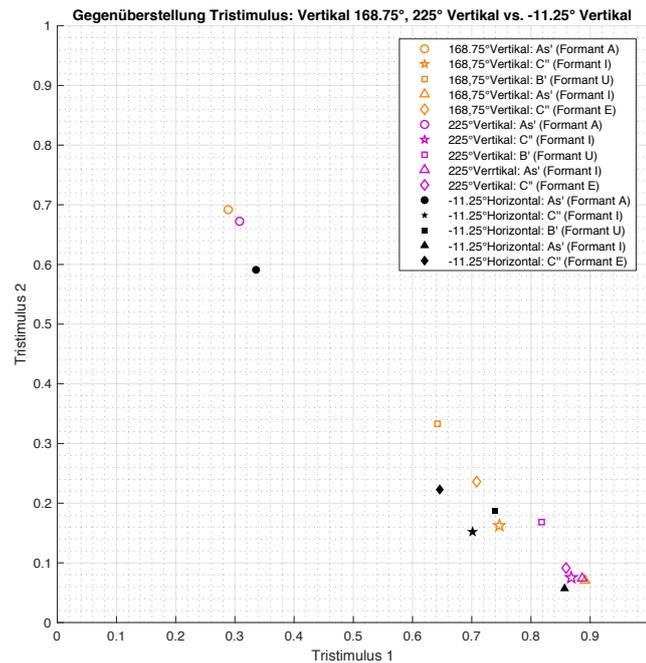


Abbildung 35 – Tristimulus Gegenüberstellung zusätzliche Auswertungen 2

A4: Tristimulus 3:

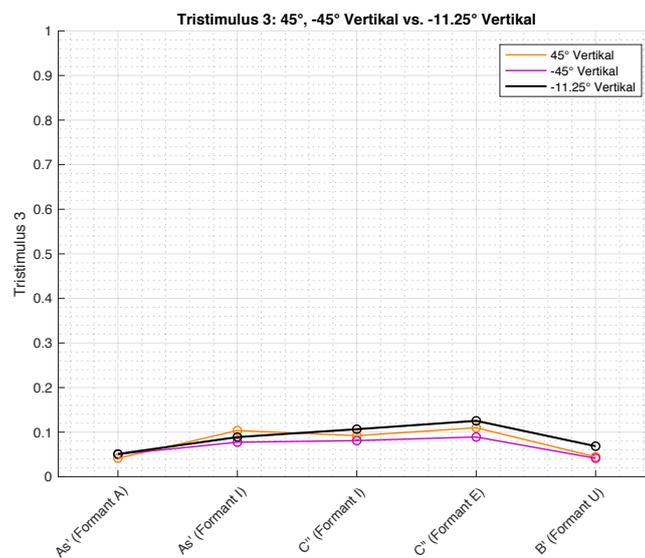


Abbildung 36 – Tristimulus 3 zusätzliche Auswertungen 1

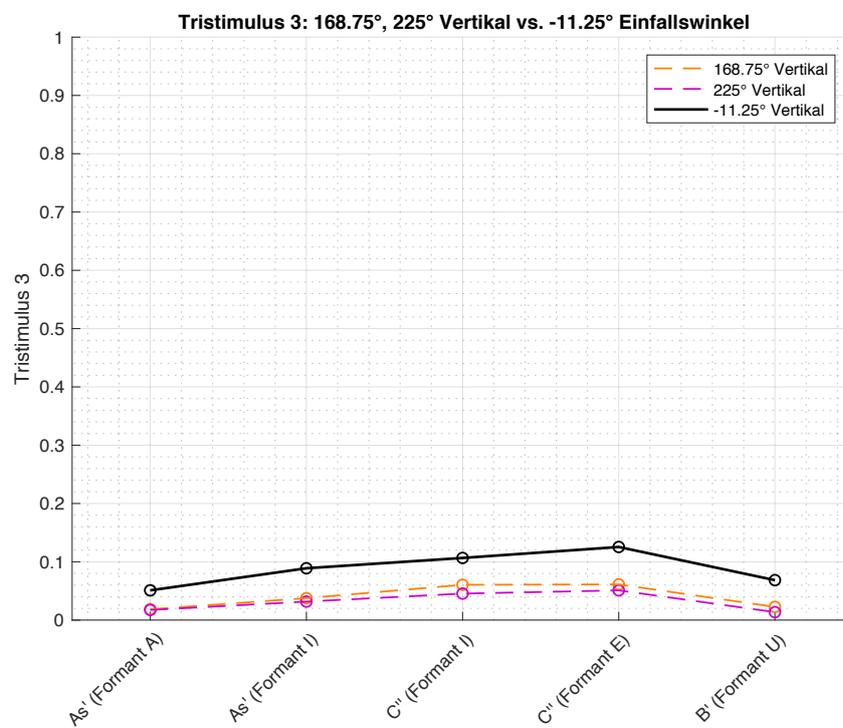
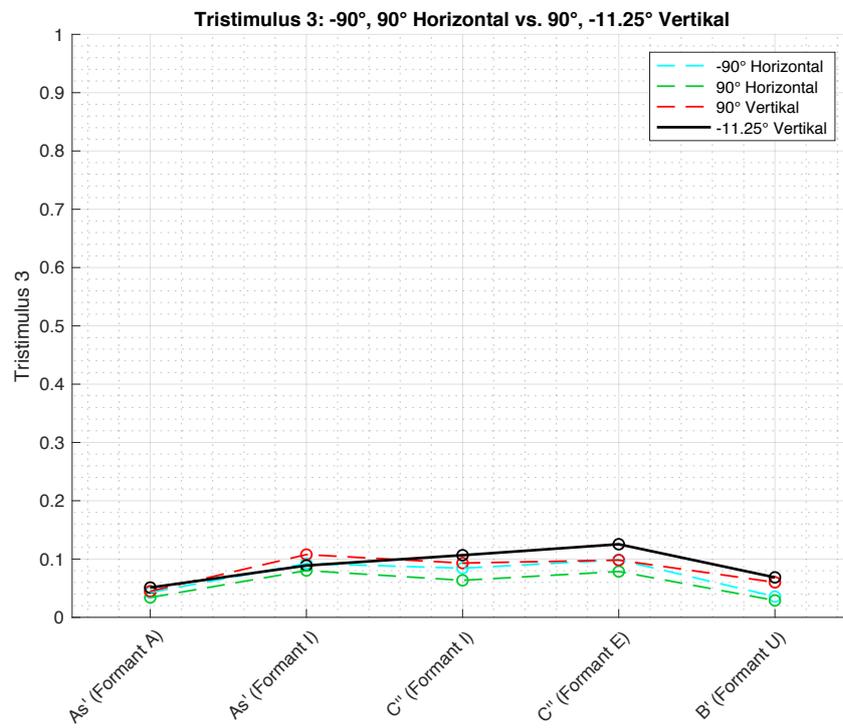


Abbildung 37 – Tristimulus 3 zusätzliche Auswertungen 2

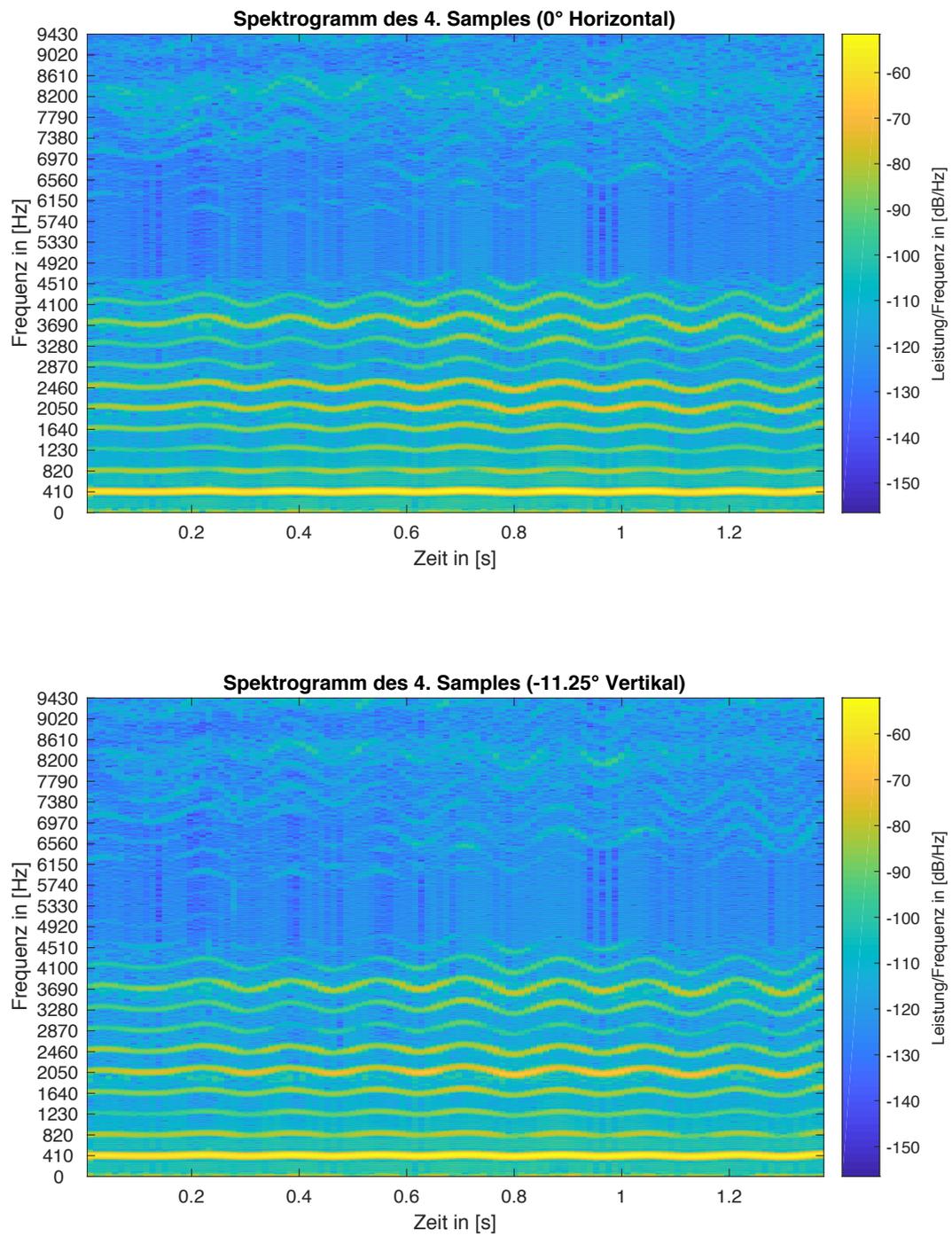
A5: Spektrogramm des 4. Samples:

Abbildung 38 – Spectrogram Sample 4 (Ab' mit Formant I)

Literatur

- [ASA60] “American Standard - Acoustical Terminology,” American Standards Association, USA, Standard, 1960.
- [BFR18a] M. Brandner, M. Frank, and D. Rudrich, “Dirpat - database and viewer of 2d/3d directivity patterns of sound sources and receivers,” in *Audio Engineering Society Convention 144*, May 2018. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=19538>
- [BFR18b] —, “Dirpat - database and visualization and analysis tools,” Mai 2018, accessed: 06.06.2018. [Online]. Available: <https://opendata.iem.at/projects/dirpat/>
- [DDHW14] M. Dickreiter, V. Dittel, W. Hoeg, and M. Wöhr, *Handbuch der Tonstudio-technik, Band 1, 8. Auflage*. De Gruyter Saur, 2014.
- [DPJB13] D. Davis, E. Patronis Jr., and P. Brown, *Sound System Engineering, 4th Edition*. Focal Press, 2013.
- [Gün12] S. Günther, “Zur Klangfarbe historischer und moderner Orchesterinstrumente,” Thesis, Institut für Sprache und Kommunikation, Fachgebiet Audiokommunikation, Technische Universität Berlin, Fakultät I, Berlin, GER, 2012.
- [Hel77] H. Helmholtz, *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. Friedrich Vieweg und Sohn, Braunschweig, 1877.
- [Hoh09] F. Hohl, “Kugelmikrofonarray zur Abstrahlungsvermessung von Musikinstrumenten,” Thesis, IEM Institut für Elektronische Musik, Kunstuniversität Graz, Austria, Technische Universität Graz, Austria, Graz, A, 2009.
- [IE03] ISO-3745:2003(E), “Acoustics - Determination of sound power levels of noise sources using sound pressure - Precision methods for anechoic and hemi-anechoic rooms,” International Organization for Standardization, Geneva, CH, Standard, Dec. 2003.
- [KEDS17] S. Kazazis, N. Esterer, P. Depalle, and M. Stephen, “A performance evaluation of the timbre toolbox and the mirtoolbox on calibrated test sounds,” in *2017 Symposium on Musical Acoustics*, Montreal, Canada, 2017.
- [Mat18] MathWorks, “Spectrogram using short-time fourier transform,” 2018, accessed: 10.06.18. [Online]. Available: <https://www.mathworks.com/help/signal/ref/spectrogram.html>
- [McA99] S. McAdams, “Perspectives on the contribution of timbre to musical structure,” *Computer Music Journal*, vol. 23, no. 3, 1999.
- [Mel16] P. Melanija, “Theobald Boehm - Flötist, Flötenbauer und Komponist,” Thesis, Institut Oberschützen, Universität für Musik und darstellende Kunst Graz, Graz, A, 2016.
- [MIC⁺13] P. Majdak, Y. Iwaya, T. Carpentier, R. Nicol, M. Parmentier, A. Roginska, Y. Suzuki, K. Watanabe, H. Wierstorf, H. Ziegelwanger, and M. Noisternig,

- “Spatially oriented format for acoustics: A data exchange format representing head-related transfer functions,” in *Audio Engineering Society Convention 134*, May 2013. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=16781>
- [MNW⁺17] P. Majdak, M. Noisternig, H. Wierstorf, M. Mihocic, H. Ziegelwanger, and T. Carpentier, “Sofa (spatially oriented format for acoustics),” November 2017, accessed: 06.06.2018. [Online]. Available: [https://www.sofaconventions.org/mediawiki/index.php/SOFA_\(Spatially_Oriented_Format_for_Acoustics\)](https://www.sofaconventions.org/mediawiki/index.php/SOFA_(Spatially_Oriented_Format_for_Acoustics))
- [Mös10] M. Möser, *Messtechnik der Akustik*. Springer-Verlag, 2010.
- [PGS⁺11] G. Peeters, B. L. Giordano, P. Susini, N. Misdariis, and S. McAdams, “The timbre toolbox: Extracting audio descriptors from musical signals,” *Journal of the Acoustical Society of America*, vol. 130, no. 5, 2011.
- [PGS⁺16] —, “Github mondaugen/timbretoolbox,” July 2016, accessed: 20.03.18. [Online]. Available: <https://github.com/mondaugen/timbretoolbox>
- [PZS⁺16] H. Pomberger, F. Zotter, A. Sontacchi, M. Brandner, M. Resch, S. Brandl, and R. Höldrich, “Free-field pressure-based sound power measurement procedure with low spatial-sampling- and near-field- induced uncertainty,” in *Inter-Noise 2016*, Hamburg, Germany, 2016.
- [UCH⁺13] C. Utz, T. Christensen, D. A. Hanninen, A. Hamilton, C. Reuter, N. Urbanek, L. Haselböck, E. Vlitakis, E. Ungeheuer, L. Holtmeier, T. Janz, and M. Neuwirth, “Organized sound,” 2013.
- [VHH98] P. Vary, U. Heute, and W. Hess, *Digitale Signalverarbeitung*. B.B. Teubner Stuttgart, 1998.
- [Wes18] W. Weselak, “Akustische Messtechnik, VO, Version 12.0,” Ausgabe WS 2017/18.