

Hochschule der Medien Stuttgart  
Fakultät **Electronic Media**

---

# MASTERARBEIT

---

## **Konzeption und Erprobung eines Stützmikrofonverfahrens zur ausgedehnten Abbildung akustischer Instrumente in mehrdimensionalen Audiomischungen**

Autor:  
**Leon Hofmann**

Studiengang:  
Audiovisuelle Medien (Master of Engineering)

Matrikelnummer:  
37033

Erstprüfer:  
Prof. Dr. Frank Melchior, Hochschule der Medien

Zweitprüfer:  
Prof. Oliver Curdt, Hochschule der Medien

Praxisbetreuer:  
M.Eng. Benjamin Müller, Fraunhofer-Institut für Bauphysik IBP

vorgelegt am 14. November 2020

---

## **Ehrenwörtliche Erklärung**

Hiermit versichere ich, Leon Hofmann, ehrenwörtlich, dass ich die vorliegende Masterarbeit mit dem Titel: „Konzeption und Erprobung eines Stützmikrofonverfahrens zur ausgedehnten Abbildung akustischer Instrumente in mehrdimensionalen Audiomischungen“ selbstständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden.

Ich habe die Bedeutung der ehrenwörtlichen Versicherung und die prüfungsrechtlichen Folgen (§ 23 Abs. 2 Master-SPO der HdM) einer unrichtigen oder unvollständigen ehrenwörtlichen Versicherung zur Kenntnis genommen.

Stuttgart, 14. November 2020 - Leon Hofmann

---

## Kurzfassung

Die vorliegende Arbeit beschäftigt sich mit der Abbildung auditiver Ausdehnung einzelner akustischer Instrumente auf einem mehrdimensionalen Wiedergabesystem, mithilfe eines geeigneten Stützmikrofonverfahrens. In Vorbereitung zur Konzeption dieser Mikrofonanordnung werden die zugrunde liegenden psychoakustischen Aspekte der Richtungslokalisierung und Ausdehnungswahrnehmung herausgestellt. Im Anschluss werden die relevanten Eigenschaften akustischer Instrumente untersucht. Hierbei liegt der Fokus auf charakteristischen spektralen Merkmalen sowie winkel- und frequenzabhängiger Schallabstrahlung. Weiterführend wird auf konventionelle Mikrofonierungs- und Wiedergabeverfahren zur mehrdimensionalen Darstellung virtueller Schallquellen eingegangen, die auf dem Prinzip der Stereophonie beruhen. Daran anknüpfend werden die Phänomene beschrieben, die bei der Wahrnehmung ausgedehnter virtueller Schallquellen in einem mehrdimensionalen Lautsprechersystem zu berücksichtigen sind. Nach der Definition weiterer künstlerischer und audiatechnischer Anforderungen an das Mikrofonverfahren wird die konzipierte koinzidente Anordnung vorgestellt. Zur Erprobung des Verfahrens wird eine Auswahl an Instrumenten im reflexionsarmen Raum an jeweils drei Positionen mikrofoniert und aufgezeichnet. Dieser Prozess wird im Rahmen der Arbeit dokumentiert. Die generierten Audiodaten und begleitenden Videoaufzeichnungen werden aufbereitet und zur Evaluation des Verfahrens, sowie zu weiteren Forschungs- und Demonstrationszwecken, zur Verfügung gestellt. Abschließend werden erste Eindrücke und Ansätze für zukünftige Untersuchungen durch eine Testmischung gewonnen.

---

## Abstract

The present thesis deals with the topic of reproducing the auditory spread of individual acoustical instruments on multi-dimensional playback systems with a suitable close mixing technique. In preparation for the conception of this microphone arrangement, the underlying psychoacoustic aspects of directional localization and perception of source extent are discussed. Afterwards, the relevant properties of acoustical instruments are examined, essentially regarding characteristic features in sound spectra as well as angle- and frequency-dependent sound radiation. In addition, conventional methods for recording and reproducing multi-dimensional virtual sound sources based on the principle of stereophony are discussed. Subsequently, various phenomena are described that have to be taken into account when perceiving extended virtual sources in a multi-dimensional loudspeaker system. After defining further requirements concerning aesthetics and signal properties, the designed coincident microphone arrangement is presented. In order to test the developed method, a recording of selected instruments is carried out in an anechoic chamber. This process is documented as part of the present work. The generated audio files as well as accompanying video recordings are processed and made available for evaluation, demonstration and further research. Finally, first impressions and approaches for future investigations are gained by mixing a test session.

---

## Danksagung

Die Entstehung dieser Arbeit und insbesondere die Durchführung der praktischen Untersuchungen konnten nur durch die Unterstützung einiger Personen und Institutionen gelingen, welchen ich hiermit meinen herzlichen Dank aussprechen möchte.

Zu aller erst bedanke ich mich bei Herrn Prof. Dr. Frank Melchior für die sehr engagierte, hilfsbereite und motivierende Betreuung meiner Arbeit. Das Thema der Thesis wurde nicht zuletzt durch die inspirierenden Gespräche mit ihm geformt und die Schritte im Verlauf des Arbeitens wurden stets mit Interesse und einem offenen Ohr begleitet. Ich danke weiterhin Herrn Prof. Oliver Curdt für seine Hilfsbereitschaft bei verschiedensten audiatechnischen und tonmeisterlichen Fragen, sowie Jörg Bauer für die Unterstützung bei meinen Anliegen rund um die Tonstudios der Hochschule der Medien.

Ein ganz herzlicher Dank gilt der Gruppe SoundLab des Fraunhofer-Instituts für integrierte Schaltungen IIS in Erlangen. Stellvertretend möchte ich Daniela Rieger, Dr. Ulli Scuda, Christian Simon und Yannik Grewe nennen. Die Gruppe hat mich, neben zahlreichen Anregungen und Tipps, durch großzügige Equipment-Leihgaben unterstützt. Dem Fraunhofer-Institut für Bauphysik IBP in Stuttgart gebührt als Kooperationspartner, unter anderem für die Bereitstellung von Equipment und die Benutzung des Freifeldraumes, ebenfalls mein großer Dank. An dieser Stelle auch ein Dankeschön an meine Kolleginnen und Kollegen aus der Gruppe kognitive Ergonomie und Psychoakustik, stellvertretend Noemi Martin. Ganz herzlich bedanken möchte ich mich bei Benjamin Müller, der mich als Praxisbetreuer vielseitig durch Ratschläge, Korrekturlesen und bei der Durchführung des praktischen Teils unterstützt hat.

Ich habe mich des Weiteren sehr über die Bereitschaft, das Engagement und die Zeit gefreut, die mir Seitens der Musiker entgegengebracht wurde. Sie haben damit den wesentlichen Beitrag zur erfolgreichen Durchführung der Untersuchungen geleistet. Daher möchte ich mich ganz herzlich bedanken bei: Martina Lechner, Cornelia Schenker, Thorsten Hohensee, Jan Kappes, Jonas Martin, Johannes Pflaum, Johannes Reinhuber und Johannes Weik.

Für eine wunderbare und spannende Studienzeit möchte ich mich vielmals bei meinen aufgeschlossenen, humorvollen und hilfsbereiten Mitstudierenden bedanken. Namentlich möchte ich an dieser Stelle Tobias Kurzweg nennen, dem mein großer Dank für die Konstruktion und Fertigung der Mikrofonhalterungen gilt. Außerdem möchte ich Daniela Rieger und Marcel Malchau für das offene Ohr und das Korrekturlesen, sowie Marcel Remy für die fachliche Unterstützung und Inspiration danken. Zuletzt gilt ein von Herzen kommender Dank meiner Partnerin, sowie meinem Bruder und meinen Eltern, die mir immer den nötigen Rückhalt gegeben haben und auf deren Unterstützung ich stets zählen konnte.

---

## **Bibliografische Angaben**

Hofmann, Leon: Konzeption und Erprobung eines Stütz mikrofonverfahrens zur ausgedehnten Abbildung akustischer Instrumente in mehrdimensionalen Audiomischungen, 79 Seiten, 49 Abbildungen, 4 Tabellen, Hochschule der Medien, Fakultät Electronic Media

Masterarbeit, 2020

Aus Gründen der besseren Lesbarkeit wird in dieser Masterarbeit an einigen Stellen die Sprachform des generischen Maskulinums verwendet. Es wird ausdrücklich darauf hingewiesen, dass sämtliche Personenbezeichnungen gleichermaßen für alle Geschlechter gelten.

# I. Inhaltsverzeichnis

<b>Inhaltsverzeichnis</b>	<b>I</b>
<b>Abbildungsverzeichnis</b>	<b>II</b>
<b>Tabellenverzeichnis</b>	<b>III</b>
<b>Abkürzungsverzeichnis</b>	<b>IV</b>
<b>Glossar</b>	<b>V</b>
<b>1 Einleitung und Zielsetzung</b>	<b>1</b>
<b>2 Psychoakustische Grundlagen</b>	<b>3</b>
2.1 Lokalisation einer Schallquelle . . . . .	4
2.1.1 Head-Related Transfer Function . . . . .	4
2.1.2 Horizontale Merkmale . . . . .	6
2.1.3 Vertikale Merkmale . . . . .	8
2.1.4 Weitere Prozesse . . . . .	11
2.2 Lokalisation mehrerer Schallquellen . . . . .	12
2.3 Wahrnehmung einer ausgedehnten Schallquelle . . . . .	13
2.4 Zusammenfassung . . . . .	14
<b>3 Schallabstrahlung akustischer Instrumente</b>	<b>16</b>
3.1 Gesangsstimme . . . . .	18
3.2 Holzblasinstrumente . . . . .	19
3.2.1 Querflöte . . . . .	20
3.2.2 Oboe . . . . .	21
3.2.3 Klarinette . . . . .	22
3.2.4 Tenorsaxofon . . . . .	23
3.3 Streichinstrumente . . . . .	25
3.3.1 Violine . . . . .	26
3.3.2 Kontrabass . . . . .	27
3.4 Akustische Westerngitarre . . . . .	29
3.5 Schlaginstrumente und Percussion . . . . .	30
3.5.1 Shaker . . . . .	31
3.5.2 Congas . . . . .	31
3.5.3 Bassdrum . . . . .	32
3.5.4 Snaredrum . . . . .	32
3.5.5 Becken . . . . .	33

---

3.6	Bewegungen der Musiker . . . . .	33
3.7	Zusammenfassung . . . . .	34
<b>4</b>	<b>Aufnahme und Wiedergabe ausgedehnter Schallquellen</b>	<b>36</b>
4.1	Stereofonie . . . . .	37
4.2	Lautsprecheranordnung 4+7+0 . . . . .	38
4.3	Mikrofonierung . . . . .	40
4.4	Lokalisation und Ausdehnung in mehrdimensionalen Lautsprechersystemen . . . . .	43
4.5	Konventionelles Mikrofonierungs- und Mischkonzept . . . . .	46
4.6	Kompatibilität und Flexibilität . . . . .	48
4.7	Konzipierte Mikrofonanordnung MHV . . . . .	48
<b>5</b>	<b>Testaufnahmen zur Untersuchung der MHV</b>	<b>51</b>
5.1	Vorversuche . . . . .	51
5.2	Arrangiertes Musikstück . . . . .	53
5.3	Aufnahmeraum . . . . .	53
5.4	Aufnahmeverfahren . . . . .	54
5.5	Instrumente, verwendetes Equipment, Abhörraum . . . . .	56
5.6	Mikrofonpositionen . . . . .	59
5.6.1	Gesangsstimme . . . . .	64
5.6.2	Holzblasinstrumente . . . . .	64
5.6.3	Streichinstrumente . . . . .	68
5.6.4	Gitarre . . . . .	70
5.6.5	Percussion und Schlagzeug . . . . .	70
5.7	Bereitstellung der Daten zur Evaluation . . . . .	73
<b>6</b>	<b>Testmischung und Eindrücke</b>	<b>74</b>
<b>7</b>	<b>Fazit und Ausblick</b>	<b>77</b>
	<b>Literaturverzeichnis</b>	<b>79</b>
<b>A</b>	<b>Bilder zur Dokumentation</b>	<b>84</b>
A.1	Weiblicher Gesang . . . . .	84
A.2	Männlicher Gesang . . . . .	86
A.3	Querflöte . . . . .	88
A.4	Oboe . . . . .	90
A.5	Klarinette . . . . .	92
A.6	Tenorsaxofon . . . . .	94
A.7	Violine . . . . .	96
A.8	Kontrabass . . . . .	98
A.9	Gitarre . . . . .	100
A.10	Shaker . . . . .	102

---

A.11 Congas . . . . .	103
A.12 Schlagzeug . . . . .	104
<b>B Datenblätter</b>	<b>105</b>
B.1 Freifeldraum . . . . .	105
B.2 Mikrofone . . . . .	107
<b>C Partitur Komposition</b>	<b>110</b>
<b>D Datenträger</b>	<b>122</b>

## II. Abbildungsverzeichnis

2.1	Dreidimensionales, kopfbezogenes, Polarkoordinatensystem . . . . .	4
2.2	Frequenzgänge der linken und rechten HRTF für Schalleinfall von links . . . . .	5
2.3	ITD je nach Schalleinfallswinkel bei durchschnittlicher Kopfgröße . . . . .	6
2.4	Darstellung des Cone of Confusion . . . . .	8
2.5	Richtungsbestimmende Frequenzbänder in der Medianebene . . . . .	9
2.6	Peilbewegungen des Kopfes zur Detektion der Schallquellenposition . . . . .	11
2.7	Summenlokalisierung, Präzedenz-Effekt und Echo . . . . .	13
3.1	Gemessenes Abstrahlverhalten der Gesangsstimme . . . . .	18
3.2	Berechnetes Abstrahlverhalten der Harmonischen der Querflöte . . . . .	21
3.3	Abstrahlverhalten von Klarinette und Oboe . . . . .	22
3.4	Quellorte für Oboe und Klarinette im 1 kHz-Oktavband bei unterschiedlichen Tönen	24
3.5	Quellorte für Tenorsaxofon im 1 kHz-Oktavband bei unterschiedlichen Tönen . .	25
3.6	Quellorte für Kontrabass zwischen 4,2 und 8,4 kHz . . . . .	28
3.7	Quellorte für Gitarre im 1 kHz-Oktavband bei unterschiedlichen Tönen . . . . .	30
4.1	Lautsprecheranordnung 4+7+0 . . . . .	39
4.2	Illustration MHV-Anordnung mit Richtcharakteristika . . . . .	50
5.1	Mikrofonanordnung MHV auf Stereoschiene für Vorversuche . . . . .	52
5.2	Konstruktionszeichnung und Ansichten der MHV-Halterung . . . . .	52
5.3	Freifeldraum P22 des Fraunhofer-Instituts für Bauphysik IBP . . . . .	54
5.4	Grundriss des Aufnahmerraums . . . . .	55
5.5	Abhörplatz im Halbfreifeldraum P21 . . . . .	59
5.6	Aufgezeichnetes Hauptmikrofon OCT + Lh und Rh . . . . .	60
5.7	Virtuelle Aufstellung des Ensembles . . . . .	60
5.8	Positionen der MHV für weiblichen Gesang . . . . .	64
5.9	Positionen der MHV für männlichen Gesang . . . . .	65
5.10	Positionen der MHV für Querflöte . . . . .	65
5.11	Positionen der MHV für Oboe . . . . .	66

---

5.12	Positionen der MHV für Klarinette . . . . .	67
5.13	Positionen der MHV für Tenorsaxofon . . . . .	67
5.14	Positionen der MHV für Violine . . . . .	68
5.15	Positionen der MHV für Kontrabass . . . . .	69
5.16	Positionen der MHV für Gitarre . . . . .	69
5.17	Positionen der MHV für Shaker . . . . .	70
5.18	Positionen der MHV für Congas . . . . .	71
5.19	Positionen der MHV für Schlagzeug . . . . .	72
5.20	Panning der dekodierten MHV Signale . . . . .	73
A.1	Bilder weiblicher Gesang . . . . .	85
A.2	Bilder männlicher Gesang . . . . .	87
A.3	Bilder Querflöte . . . . .	89
A.4	Bilder Oboe . . . . .	91
A.5	Bilder Klarinette . . . . .	93
A.6	Bilder Tenorsaxofon . . . . .	95
A.7	Bilder Violine . . . . .	97
A.8	Bilder Kontrabass . . . . .	99
A.9	Bilder Gitarre . . . . .	101
A.10	Bilder Shaker . . . . .	102
A.11	Bilder Congas . . . . .	103
A.12	Bilder Schlagzeug . . . . .	104
B.1	Technische Daten Mikrofone . . . . .	109

---

## III. Tabellenverzeichnis

4.1 Lautsprecherpositionen und Kanalbezeichnungen im 4+7+0 System . . . . .	39
5.1 Angaben zu den aufgezeichneten Instrumenten . . . . .	57
5.2 Auflistung der verwendeten technischen Komponenten . . . . .	58
5.3 Positionen der MHV-Anordnungen vor den Instrumenten . . . . .	62

## IV. Abkürzungsverzeichnis

<b>BT</b>	Bottom/Top Bus (dekodiertes MHV Signal)
<b>HRTF</b>	Head-Related Transfer Function
<b>ITD</b>	Interaural Time Difference
<b>ILD</b>	Interaural Level Difference
<b>IAC</b>	Interaural Cross-Correlation
<b>ICC</b>	Interchannel Cross-Correlation
<b>LR</b>	Left/Right Bus (dekodiertes MHV Signal)
<b>MHV</b>	Mitte-Horizontal-Vertikal (konzipierte Mikrofonanordnung)
<b>VBAP</b>	Vector Base Amplitude Panning

---

## V. Glossar

**Cone of Confusion** Darstellung der möglichen Schallquellenpositionen im Raum, bei denen gleiche Werte für interaurale Zeit- und Pegeldifferenz auftreten.

**Directional Tone Color** Beschreibung des Effekts, dass durch grundtonabhängige Vorzugsrichtungen in der Abstrahlung bei unterschiedlichen Tönen das Klangbild eines Instrumentes bei Monomikrofonierung schwankt, bzw. bei Stereomikrofonierung der lokalisierte Quellort variiert.

**Höhenmikrofone** Mikrofone innerhalb einer mehrdimensionalen Anordnung, welche Signale von oben aufzeichnen.

**Interaurale Ohrsignalmerkmale** Differenzen zwischen den akustischen Signalen an beiden Ohren, die auf eine bestimmte Schalleinfallrichtung schließen lassen.

**Lokalisation** Prozess der Richtungsbestimmung eines Schallereignisses.

**Lokalisationsunschärfe** Größe des Winkelfehlers bei der Richtungsbestimmung von Schallereignissen.

**Mehrdimensional** Bezeichnung für Mikrofon- und Lautsprecheranordnungen, welche die Aufzeichnung- bzw. Wiedergabe von Quellpositionen unterschiedlicher Elevation ermöglichen.

**MHV** Konzipiertes Stützmikrofonverfahren zur ausgedehnten Abbildung akustischer Instrumente mit den drei Komponenten Mitte, Horizontal, Vertikal.

**Monaurale Ohrsignalmerkmale** Eigenschaften des akustischen Signals an einem oder beiden Ohren gleichzeitig, die auf eine bestimmte Schalleinfallrichtung schließen lassen.

**Pitch-Height-Effekt** Beschreibung des Phänomens, dass die wahrgenommene Elevation verschiedener Frequenzbänder unabhängig vom tatsächlichen Quellort ist, mit der Grundtendenz höherer Elevation bei höherer Frequenz.

**Surround-Formate** Mehrkanalige Wiedergabe über Lautsprecher, die in der Horizontalebene um den Hörer platziert sind.

**Virtuelle Schallquelle** Von Lautsprechern aus einer bestimmten Position wiedergegebene Aufzeichnung einer realen Schallquelle.

# 1 Einleitung und Zielsetzung

Mehrdimensionale Audioformate bieten dem Hörer im Vergleich zu konventioneller Stereowiedergabe einen Mehrwert unter anderem durch einen umhüllenderen Klangeindruck (Power, Davies & Hirst, 2014). Soll die künstlerische Darbietung eines klassischen Ensembles akustischer Instrumente auf einem solchen Wiedergabesystem abgebildet werden, kommt bei konventioneller Herangehensweise das Konzept von Haupt- und Stützmikrofonierung zum Einsatz. Das mehrkanalige Hauptmikrofon dient dabei der räumlichen Abbildung der gesamten auditiven Szene im Konzertsaal, bestehend aus den auf der Bühne verteilten Instrumentengruppen, den ersten Reflexionen von Seitenwänden und Decke sowie dem Raumhall. Stützmikrofone sind meist als Monomikrofon ausgeführt und werden vor einzelnen Instrumenten platziert. Die von diesen Anordnungen aufgezeichneten Signale geben dem Tonmeister die Möglichkeit, Präsenz und dargestellte Position einzelner Instrumente hervorzuheben bzw. anzupassen (Weinzierl, 2008, S.603). Aufgrund der großen Abmessungen eines Orchesters und der daraus resultierenden großen Entfernung zwischen Klangkörper und Hauptmikrofon, ist die Abbildung eines einzelnen Instrumentes in der mehrdimensionalen Mischung auf einen Punkt reduzierbar. Daher ist die Stützmikrofonierung mit einem Monomikrofon, trotz des mehrdimensionalen Wiedergabesystems, ausreichend.

Bei der Aufzeichnung kleinerer Ensembles oder der Erstellung unkonventioneller Audio-mischungen, stellt sich dem Tonmeister die Frage, mit welchen Signalen der vorhandene Raum in der mehrdimensionalen Mischung gefüllt werden kann. Aus diesem Grund beschäftigt sich eine Vielzahl an Studien mit der Synthetisierung und Abbildung einer virtuellen Schallquelle mit Ausdehnung. Beispielsweise beschreibt Melchior (2012) ein Verfahren zur Erstellung einer Quelle mit wahrgenommener Ausdehnung durch die räumliche Verteilung einzelner Frequenzbänder eines Monosignals. Diese erzielte Ausdehnung wird weiterführend durch verschiedene Steuersignale dynamisch variiert.

Im Falle der Abbildung eines akustischen Instrumentes könnte mit dem aufgezeichneten Signal einer monophonen Stützmikrofonierung ähnlich verfahren werden. Jedoch weisen Instrumente durch verschiedene, schallabstrahlende Komponenten, bereits von Natur aus eine bestimmte akustische Ausdehnung auf, welche durch geeignete Mikrofonierung genutzt werden kann (Menzies, 2002, S. 1). Aus diesem Grund besteht die Motivation der vorliegenden Arbeit darin, ein praxistaugliches Stützmikrofonverfahren zu konzipieren, welches eine mehrdimensionale Abbildung einzelner akustischer Instrumente über ein geeignetes Wiedergabesystem ermöglicht. Dieses soll durch eine Reihe an Testaufnahmen, unter Anwendung verschiedener Positionen an unterschiedlichen Instrumenten, erprobt und beschrieben werden. Da sich die Abbildung der ausgedehnten Schallquelle

durch die Bewegungen der Instrumente beim Vortrag dynamisch ändert, sollen diese zum späteren Abgleich mit Höreindrücken durch Videoaufzeichnungen dokumentiert werden. Abschließend entsteht eine prototypische, mehrdimensionale Audiomischung, welche subjektiv bewertet wird und die Basis für weitere Untersuchungen und Tests bildet.

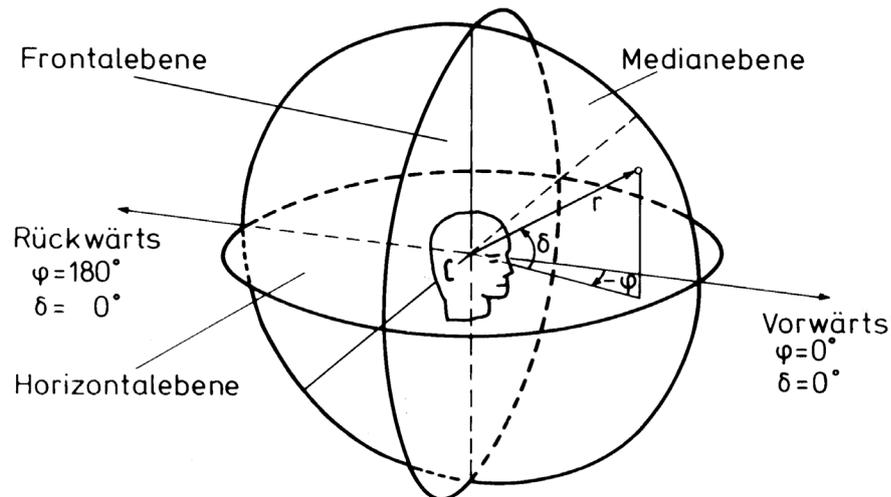
Die Konzeption des Stützmikrofonverfahrens bedarf der Aufarbeitung einiger Funktionsweisen und Phänomene der Psychoakustik, Elektroakustik, musikalischen Akustik und der Audiotechnik. Folgende Fragen sollen als Leitfaden durch die nächsten Kapitel führen:

- Wie gestaltet sich die Funktionsweise der Richtungslokalisierung von Schallquellen?
- Wie entsteht der Eindruck einer ausgedehnten Schallquelle?
- Welche charakteristischen akustischen Eigenschaften besitzen die ausgewählten Instrumente?
- Welche Effekte ergeben sich bei der Wahrnehmung virtueller Schallquellen in einem mehrdimensionalen Lautsprechersystem?
- Welche Eigenschaften sollen die aufgezeichneten Signale sowohl aus künstlerischer, als auch aus audiotechnischer Sicht besitzen?

## 2 Psychoakustische Grundlagen

Psychoakustische Grundlagen spielen eine wichtige Rolle für die Einordnung der Thematik. Im Kern geht es darum aufzuzeigen, in welcher Weise der Mensch akustische Szenen wahrnimmt. Besonderes Augenmerk liegt dabei auf der Lokalisation und Wahrnehmung einer realen, ausgedehnten Schallquelle. Ein solcher Schallsender strahlt Schallwellen ab, indem ein schwingendes Medium die angrenzenden Luftteilchen ebenfalls in Schwingung versetzt, welche wiederum ihre Nachbarpartikel in Bewegung versetzen. Auf diese Weise breitet sich eine Schallwelle durch das Medium Luft aus und trifft schließlich am Ohr ein. Im Wesentlichen beschreibt sich das am Trommelfell eintreffende Signal durch zwei Eigenschaften: Die Frequenz und die Amplitude der Schwingung. Erstere zeigt sich verantwortlich für die wahrgenommene Tonhöhe und Letztere für die Lautstärke des Tones. Diese zwei detektierten Merkmale stellen den Kern des Hörens dar. Töne in ihrer Höhe und Lautstärke zu unterscheiden bedeutet, zwischen verschiedenen Klängen differenzieren zu können. Somit ist das Hören von Musik, das Unterscheiden zwischen Stimmen verschiedener Personen und das Erkennen von Geräuschen möglich. Über die Lautstärke und spektrale Verteilung kann bei bekannten Signalen außerdem abgeschätzt werden, in welcher Entfernung sich die Quelle des wahrgenommenen Klanges befindet (Weinzierl, 2008, S. 98 f.). Die folgenden Abschnitte beschäftigen sich mit der Frage, wie die Richtung erkannt werden kann, aus welcher ein Schallereignis stammt. Das Verständnis dieses Prozesses ist grundlegend für die weiteren Untersuchungen und Überlegungen bezüglich einer mehrdimensionalen Wiedergabe von Schallquellen.

Abbildung 2.1 zeigt das kopfbezogene Koordinatensystem, auf welches sich die Richtungsangaben in den folgenden Kapiteln beziehen. Alle Richtungen des Schalleinfalls werden mit der Angabe jeweils eines Winkels auf der Horizontal- und auf der Medianebene beschrieben. Diese beiden Winkel werden als *Azimuth*  $\phi$  (horizontal) und *Elevation*  $\delta$  (median bzw. vertikal) bezeichnet. Auf der interauralen Achse in der Mitte zwischen beiden Gehörkanaleingängen befindet sich der Ursprung des gedachten Polarkoordinatensystems. Die frontale Blickrichtung ist als  $0^\circ$ -Referenz definiert. Die Entfernung der Schallquelle vom Ursprung wird mit  $r$  bezeichnet (Weinzierl, 2008, S. 88). Der Azimutwinkel  $\phi$  läuft gegen den Uhrzeigersinn positiv von  $0^\circ$  bis  $180^\circ$ , im Uhrzeigersinn negativ. Der Elevationswinkel  $\delta$  läuft in der oberen Hälfte der Medianebene positiv und entsprechend in der unteren negativ. Sofern nicht anders definiert, wird vorerst die Modellvorstellung einer einzelnen punktförmigen Schallquelle angewandt. Im Gegensatz dazu stehen die realen Schallquellen mit bestimmter Breite und Höhe. Das Ziel der theoretischen Überlegungen innerhalb dieses Kapitels bildet die Beschreibung der Wahrnehmung und Lokalisation solcher ausgedehnter Quellen.



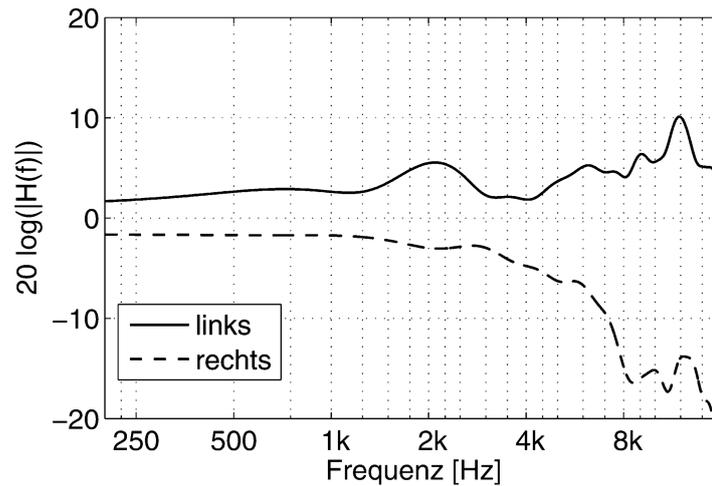
**Abbildung 2.1:** Dreidimensionales, kopfbezogenes, Polarkoordinatensystem.  
(Weinzierl, 2008, S. 88)

## 2.1 Lokalisation einer Schallquelle

Der Begriff **Lokalisation** beschreibt den Prozess des Gehirns, Richtung und Entfernung zu bestimmen, aus der die am Ohr eintreffenden Schallwellen stammen. Die vorliegende Arbeit konzentriert sich auf die ausgedehnte Abbildung von Schallquellen, weswegen sich die folgenden Ausführungen auf die Richtungsbestimmung konzentrieren, wenn der Terminus *Lokalisation* verwendet wird. Grundlage dieses Prozesses bilden die Ohrsignale, d.h. die Signale, welche auf die beiden Trommelfelle auftreffen. Um die Richtung des eintreffenden Schalles zu bestimmen, vergleicht das Gehirn die beiden Ohrsignale. Vor diesem Vergleich passieren im Ohr mehrere komplexe Umwandlungen, sodass der Luftschall über mechanische Impedanzanpassung durch die Flüssigkeit der Gehörschnecke weitergeleitet wird und schließlich als Reiz durch den Hörnerv am Gehirn ankommt (Howard & Angus, 2009, S. 74 ff.). Im Folgenden wird allerdings vereinfacht davon gesprochen, dass das Gehirn die Schallsignale am Trommelfell analysiert und vergleicht.

### 2.1.1 Head-Related Transfer Function

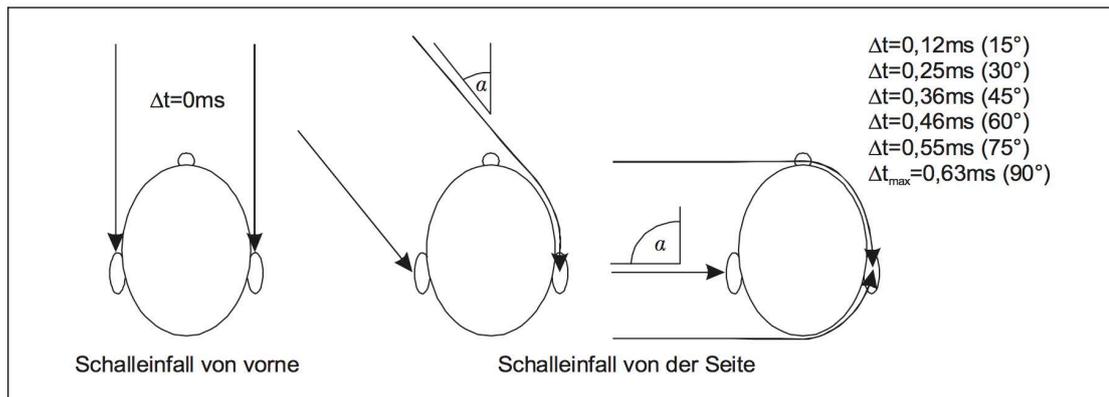
Die Schallwellen werden durch verschiedene Faktoren auf dem Weg von der Quelle bis zum Außenohr akustisch beeinflusst. Dadurch weisen die Ohrsignale Differenzen zum Ursprungssignal auf, welches von der Schallquelle abgestrahlt wurde. Diese Verzerrungen des Quellsignals werden beschrieben durch eine messbare Funktion, die sogenannte **HRTF**, Kurzform für *Head-Related Transfer Function* (deutsch: *Außenohr-Übertragungsfunktion*). Das durch diese Funktion definierte System besteht aus dem



**Abbildung 2.2:** Frequenzgänge der linken und rechten HRTF für Schalleinfall von links. (Weinzierl, 2008, S. 91)

Weg des akustischen Signales von der Schallquelle an einer bestimmten Position im Raum, durch das dämpfende Übertragungsmedium (in der Regel Luft) bis in das Außenohr. Die Verzerrungen liegen dabei in der akustisch wirksamen Anatomie des Menschen begründet. Sie äußern sich in je nach Einfallrichtung charakteristischer Veränderung von Amplituden und Phasen der Schallwellen verschiedener Frequenzen (Lerch, Sessler & Wolf, 2009, S. 425). Dieses Phänomen kommt durch die Überlagerung zeitlich versetzt am Trommelfell eintreffender Schallwellen zustande. Je nach Frequenz einer Schallwelle kommt es zu verschiedentlichen Reflexionen an Körper, Kopf oder Ohrmuscheln. Diese reflektierten Wellen gelangen mit geringer Verzögerung gegenüber der auf direktem Weg eintreffenden Wellen ebenfalls ans Trommelfell. Des Weiteren stellen Rumpf und Kopf für Schallwellen höherer Frequenzen ein Hindernis dar, sodass es bei seitlichem Schalleinfall zur Dämpfung durch Abschattung am gegenüberliegenden Ohr kommt (Blauert, 1974).

Die HRTF lässt sich mathematisch beschreiben durch  $H_l(f, r, \phi, \delta)$  und  $H_r(f, r, \phi, \delta)$  mit Frequenz  $f$ , Abstand  $r$  der Quelle zum Hörer und der Richtungsangabe durch die beiden vorgestellten Winkel Azimut  $\phi$  und Elevation  $\delta$ .  $H_l$  stellt die Funktion des linken und  $H_r$  die des rechten Ohres dar (Weinzierl, 2008, S. 90 f.). Abbildung 2.2 zeigt Beispiele für die Darstellung des Frequenzganges (Betragsspektrum) der HRTF bei verschiedenen Schalleinfallswinkeln. Dabei sind die beschriebenen Abschattungseffekte im höheren Frequenzbereich am von der Quelle abgewandten Ohr erkennbar. Diese werden in Abschnitt 2.1.2 genauer betrachtet.



**Abbildung 2.3:** ITD je nach Schalleinfallswinkel bei durchschnittlicher Kopfgröße. (Friesecke, 2014, S. 135)

### 2.1.2 Horizontale Merkmale

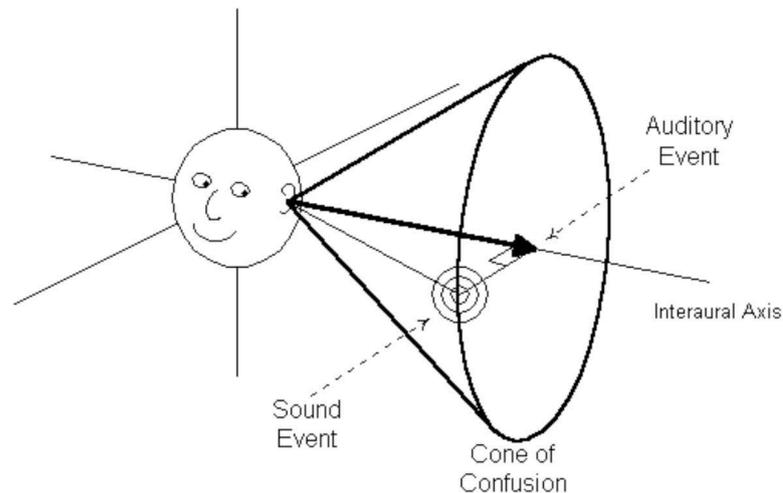
Um zu bestimmen, aus welcher Richtung der Schall eintrifft, macht sich das Gehirn die richtungsabhängigen Verzerrungen zunutze und vergleicht kontinuierlich die beiden Ohrsignale miteinander. Dabei existieren mehrere Merkmale, welche in horizontale und vertikale Merkmale unterteilt werden können. In Bezug auf die benutzten Merkmale bei der horizontalen Lokalisation kann grundlegend festgestellt werden, dass sowohl die Distanz zwischen beiden Ohren, als auch deren räumliche Trennung durch den dazwischenliegenden Kopf ausschlaggebend sind. Eine frontal auf den Kopf auftreffende Schallwelle legt, aufgrund der Symmetrie desselben, die gleiche Wegstrecke zu beiden Ohren zurück. Gleiches gilt für Schalleinfall von mittig hinter dem Kopf. Trifft der Schall jedoch seitlich ein, so ergeben sich unterschiedliche Schalllaufwege und das Signal gelangt somit später zum der Schallquelle abgewandten Ohr. Die daraus resultierende zeitliche Differenz zwischen den beiden Ohrsignalen wird als interaurale Laufzeitdifferenz bezeichnet. Geläufig ist die Verwendung der Abkürzung des englischen Begriffes *Interaural Time Delay* **ITD**. In [Abbildung 2.3](#) sind verschiedene ITD für beispielhafte Schalleinfallswinkel aufgetragen. Es ist zu erkennen, dass bei Schalleinfall von der Seite ( $\phi = 90^\circ$ ) mit 0,63 ms die maximale ITD erreicht ist. Ebenfalls wird ersichtlich, dass der Schall bei seitlichem Einfall sowohl um die Stirn herum, als auch am Hinterkopf vorbei, mit unterschiedlicher Verzögerung an das gegenüberliegende Ohr gelangt. Dadurch wird das Festlegen der ITD erschwert. Dies führt dazu, dass die sogenannte **Lokalisationsunschärfe** auf Achse der Ohren mit bis zu  $18^\circ$  relativ groß ist. Die Lokalisation in dieser Richtung besitzt also eine eher geringe Auflösung. Je frontaler der Schall eintrifft, umso genauer wird die Richtungsbestimmung. Die Lokalisationsunschärfe beträgt bei Schalleinfall von vorne nur noch ca.  $2^\circ$ . Damit hat bereits eine ITD von ca.  $30 \mu\text{s}$  eine Auslenkung der wahrgenommenen Schallrichtung aus der Frontalrichtung zur Folge. Kleinere Differen-

zen können nicht detektiert werden (Friesecke, 2014, S. 134 f.).

Die ITD zwischen den beiden Ohrsignalen wird vom Gehirn durch den Phasenversatz detektiert. Aufgrund der großen Wellenlängen ist dieser im Tiefbassbereich sehr gering. Dies erklärt die schlechte Lokalisierbarkeit sehr tieffrequenter Schallanteile (Friesecke, 2014, S. 135). Zu höheren Frequenzen hin existiert ebenfalls eine Limitation für die Funktionalität dieses Prinzips. Bei immer kleineren Wellenlängen werden die Phasenversätze zwischen beiden Ohrsignalen bei gleichem Einfallswinkel sukzessive größer. Überschreitet der Versatz bei einer bestimmten Frequenz den Wert von  $180^\circ$ , kann die Richtung für die Schallwelle dieser Frequenz nicht mehr eindeutig zugeordnet werden, da zwei Einfallswinkel denkbar wären, welche zu dieser Phasenlage führen. Damit ergibt sich die Grenzfrequenz für die Lokalisation durch ITD, je nach Schalleinfallswinkel, ab ca. 750 Hz aufwärts (Howard & Angus, 2009, S. 110 f.).

Der durchschnittliche Kopfdurchmesser eines Erwachsenen beträgt ca. 18 cm, was der Wellenlänge von knapp 2 kHz entspricht. Oberhalb eines Drittels dieser Wellenlänge, also etwa ab 640 Hz, stellt der Kopf für Schallwellen zunehmend ein Hindernis dar. Der Schall wird nicht mehr um den Kopf gebeugt, sondern reflektiert (Howard & Angus, 2009, S. 112). Bei seitlichem Schalleinfall wird aus diesem Grund die gegenüberliegende Kopfhälfte abgeschattet, wodurch das Ohrsignal des abgewandten Ohres eine deutliche Dämpfung erfährt (vgl. Abbildung 2.2). Dieser Effekt stellt das zweite Merkmal zur Lokalisation in der horizontalen Ebene dar und wird als interaurale Pegeldifferenz zwischen den beiden Ohrsignalen bezeichnet (Görne, 2011, S. 127). Im Folgenden wird die englische Abkürzung **ILD** für *Interaural Level Difference* verwendet. Der Übergang zwischen den beiden vorgestellten Merkmalen ITD und ILD gestaltet sich fließend. Der Bereich, in welchem beide Prinzipien Anwendung finden, reicht in etwa von 700 bis 2800 Hz (Howard & Angus, 2009, S. 112 f.).

Die Analyse der interauralen Differenzen kann unter bestimmten akustischen Bedingungen widersprüchliche Lokalisationsrichtungen zum Ergebnis haben. Es ist also denkbar, dass sich verschiedene Merkmale gegenseitig aufheben, oder für die selbe Schallquelle unterschiedlichen Frequenzbereichen verschiedene Hörereignisrichtungen zugeordnet werden (Weinzierl, 2008, S. 97 f.).



**Abbildung 2.4:** Darstellung des Cone of Confusion.  
(Zaar, 2010, S. 7)

### 2.1.3 Vertikale Merkmale

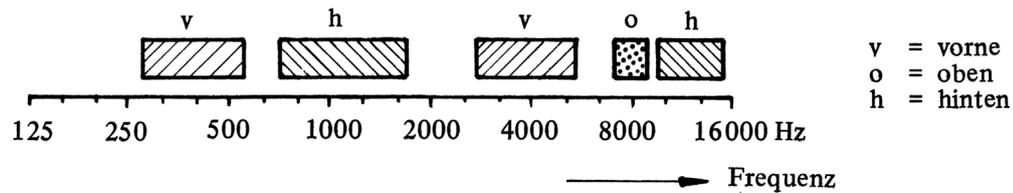
Das Funktionsprinzip der Lokalisation von Schallquellen in der horizontalen Ebene wurde in Abschnitt 2.1.2 betrachtet. Die beschriebenen Merkmale, welche dafür herangezogen werden, finden sich durch den Vergleich der beiden Ohrsignale. Es handelt sich demnach um sogenannte **interaurale Ohrsignalmerkmale** (Weinzierl, 2008, S. 95). ITD und ILD gestalten sich allerdings für verschiedenliche Einfallsrichtungen identisch. Abbildung 2.4 veranschaulicht diese Problematik. Wenn sich eine Schallquelle an einem der theoretischen Punkte im Raum befindet, die durch den aufgetragenen Zylinder beschrieben werden, so resultiert der Vergleich beider Ohrsignale jeweils in den gleichen ITD und ILD. Also muss das Gehirn noch andere Anhaltspunkte in den Signalen finden, um diesen sogenannten **Cone of Confusion**, also den „Kegel der Verwechslung“, auflösen zu können (Boren & Roginska, 2011, S. 2). Es handelt sich hierbei um Merkmale, die bereits eines der beiden Ohrsignale aufweist. Diese werden als **monaurale Ohrsignalmerkmale** bezeichnet.

Wie bereits einführend beschrieben, kommt es an Rumpf und Kopf sowie hauptsächlich an den Ohrmuscheln, zu Reflexionen der Schallwellen mit höherer Frequenz. Die reflektierten Wellen gelangen, mit leichter Zeitverzögerung im Vergleich zum Direktschall, ebenfalls ans Trommelfell. Die Überlagerung dieser Komponenten resultiert in einer für eine bestimmte Einfallsrichtung spezifischen Kammfilterstruktur im Spektrum des Ohrsignals (Friesecke, 2014, S. 135). Der Frequenzgang des Ohrsignals ändert sich also je nach Schalleinfallswinkel. Somit macht sich das Gehirn verschiedene winkelabhängige

Verfärbungen im Klangbild zu Nutze, um die Elevation einer Schallquelle zu lokalisieren. Am auffälligsten sind dabei die Pegelbrüche durch destruktive Interferenz, welche bei um  $180^\circ$  phasenverschobener Ankunft zweier Schallwellen entstehen. Diese liegen je nach Einfallswinkel an unterschiedlichen Stellen im Spektrum. Da das Ohr in Relation zu Wellenlängen des hörbaren Schalles eher kleine Abmessungen aufweist, macht sich dieser Effekt erst bei Frequenzen oberhalb von ca. 5 kHz bemerkbar (Howard & Angus, 2009, S. 113).

Die Position im Spektrum sowie Tiefe der beschriebenen Kammfilter hängen im Detail von der individuellen Anatomie des Hörers ab. Denn je nach Größe und Winkel der betreffenden Körperteile ergeben sich andere Reflexionsbedingungen. Somit besitzt jeder Mensch seine eigene charakteristische HRTF, die das Gehirn im Laufe der Zeit gelernt hat auszuwerten (Howard & Angus, 2009, S. 113 f.). Die im vorangegangenen Unterkapitel beleuchteten horizontalen Merkmale ITD und ILD hängen im Prinzip ausschließlich vom Kopfdurchmesser der betrachteten Person ab. Zum einen ist dies ein einfacher, ermittelbarer Zahlenwert und zum anderen eine Größe, die in einem überschaubaren Rahmen von Mensch zu Mensch variiert. Daher sind valide Aussagen über die horizontalen Merkmale zur Lokalisation möglich. Für die beschriebenen vertikalen Merkmale ist jedoch die Feinstruktur des Außenohres entscheidend. Die Reflexionen hoher Frequenzen entstehen an verschiedensten kleinen Flächen, deren Variationsbreite in Dimensionen und Winkelungen eher keine detaillierte Voraussagen bezüglich der exakten resultierenden Kammfilterstruktur der Ohrsignale zulässt. Nur durch aufwändigere Mess- oder Simulationsverfahren, wie beispielsweise der Ermittlung einer individuellen HRTF via Scan der Ohrmuschel, sind präzise Aussagen möglich (siehe Huttunen & Vane, 2017). Will man also für die beschriebenen vertikalen Merkmale Aussagen treffen, die für alle Hörer gelten, so muss stets beachtet werden, dass diese Approximationen darstellen und nur für einen „Durchschnittskopf“ gültig sind. Es können folgende, verallgemeinernde Aussagen getroffen werden: Ein Schallsignal wird eher unten lokalisiert, wenn Frequenzbrüche bei 6 bis 7 kHz vorliegen. Sind entsprechende Einbrüche eher bei 8 kHz vorhanden, so wird die Schallquelle auf Ohrhöhe lokalisiert. Einbrüche bei 10 bis 11 kHz deuten auf eine Schallquelle von oben hin (Friesecke, 2014, S. 136).

Blauert beschreibt durch die Ergebnisse seiner Studien eine Funktionsweise der Lokalisation, die sich *richtungsbestimmende Frequenzbänder* nennt. Demnach beruht die Lokalisation in vertikaler Richtung hauptsächlich auf einem gegenteiligen Prinzip. Anstatt die Einbrüche durch destruktive Interferenz zu betrachten, wurde je nach Dominanz bestimmter typischer Frequenzbereiche das Schallereignis in einer zugeordneten Richtung lokalisiert, ohne dass die abstrahlende Quelle sich dort befinden muss. Diese spezifischen

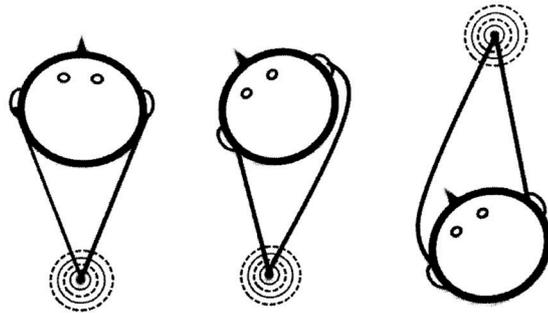


**Abbildung 2.5:** Richtungsbestimmende Frequenzbänder in der Medianebene.  
(Dickreiter, Dittel, Hoeg & Wöhr, 2014, S. 131)

Frequenzbänder sind in [Abbildung 2.5](#) dargestellt. Sie reichen bis weit unterhalb des Bereiches, ab dem der Kopf akustisch wirksam wird. Eine mögliche Erklärung hierfür wäre, dass das Gehirn auch gängige Muster durch Bodenreflexionen bei tieferen Frequenzen verwendet, da in fast allen Hörumgebungen ein solcher existiert ([Cabreria & Tilley, 2003](#), S. 6). Nach [Blauert](#) führen Überhöhungen im Bereich 300 bis 500 Hz bzw. 3 bis 5 kHz zu einer frontalen Lokalisation des Schallereignisses. Anhebungen um 8 kHz führen zur Lokalisation oben, bei 1 kHz bzw. oberhalb von 10 kHz zur Lokalisation hinter dem Kopf ([Blauert, 1974](#)).

Eine aktuellere Studie aus 2015 von [Wallis und Lee](#) konnte die richtungsbestimmenden Frequenzbänder teilweise bestätigen, deren Bandbreiten genauer eingrenzen und Rahmenbedingungen konkretisieren. Zu tieferen Frequenzen hin wurden aber auch differierende Ergebnisse erzielt. Die Versuche konnten beispielsweise die drei Drittel-Oktavbänder bei 1, 4 und 8 kHz bestätigen, wobei im Gegensatz zu [Blauert](#) in mehr als die drei Richtungen vorne, oben und hinten unterschieden wurde. Beispielsweise liegen Ergebnisse vor, die dem 8 kHz-Band eher die Richtung *hinten-oben* zuschreiben. Es konnte weiterführend der sogenannte **Pitch-Height-Effekt** beobachtet werden. Dieser beschreibt das Phänomen, dass die tieferen betrachteten Frequenzbänder 500 Hz und 1 kHz - unabhängig der tatsächlichen Elevation des abstrahlenden Lautsprechers - grundsätzlich als nicht-eleviert wahrgenommen wurden. Das Ergebnis von [Blauert](#), dass der Frequenzbereich um 500 Hz als richtungsbestimmendes Frequenzband für die Lokalisation vorne fungiert (siehe [Abbildung 2.5](#)), konnte nicht verifiziert werden.

Die Schwierigkeit bei der beschriebenen Funktionsweise der vertikalen Lokalisation besteht in den Voraussetzungen, die erfüllt sein müssen, damit das Gehirn die spektralen Merkmale überhaupt analysieren kann. Zum einen muss das Quellsignal genügend hochfrequente Anteile besitzen, damit die Kammfilterstrukturen überhaupt auftreten. Zum anderen muss das Geräusch oder der Klang dem Gehirn bekannt sein, da sonst keine spektrale Referenz vorhanden ist, mit dem der aktuell eintreffende Schall verglichen werden kann ([Görne, 2011](#), S. 126).



**Abbildung 2.6:** Peilbewegungen des Kopfes zur Detektion der Schallquellenposition. (Roginska & Geluso, 2018, S. 103)

### 2.1.4 Weitere Prozesse

Das Gehirn lernt im Laufe seines Lebens aus einer Vielzahl wiederkehrender akustischer Muster und kann diese für die Lokalisation von Schallquellen nutzen. Ein naheliegendes Beispiel stellt das Geräusch eines vorbeifliegenden Flugzeugs dar. Dieses wird automatisch eleviert verortet (Friesecke, 2014, S. 137). Ebenfalls angeführt werden kann der im vorherigen Unterkapitel 2.1.3 bereits dargelegte Prozess, aus spektralen Mustern unterschiedliche Elevationswinkel abzuleiten. Denn das Gehirn rechnet die zur Lokalisation verwendeten Klangverfärbungen aus dem Signal heraus, sodass eine Schallquelle aus verschiedenen Richtungen trotzdem etwa gleich klingt (Dickreiter et al., 2014, S. 350). Des Weiteren wurde beobachtet, dass der Kopf meist in Richtung des akustischen Signals, welchem gefolgt werden soll, gedreht wird (Howard & Angus, 2009, S. 114). Allerdings liegt an dieser Stelle die Vermutung nahe, dass die Schallquelle auch visuell zentriert wahrgenommen werden soll.

Nur kurz andauernde Schallsignale weisen eine schlechtere Lokalisierbarkeit auf als längere. Dies resultiert aus der Tatsache, dass für länger andauernde Signale ( $> 200$  ms) automatische Peilbewegungen durch Drehen des Kopfes ausgeführt werden. Durch den Abgleich der Kombinationen aus Kopfposition und detektierter ITD und ILD kann das Gehirn die unter 2.1.2 und 2.1.3 beschriebenen Unsicherheiten bezüglich vorne-hinten Lokalisation bzw. den Cone of Confusion auflösen (Weinzierl, 2008, S. 88). Abbildung 2.6 veranschaulicht das Prinzip: Wenn bei leichter Drehung des Kopfes nach links das Ohrsignal am rechten Ohr schwächer wird und zeitverzögert ankommt, so befindet sich die Quelle hinter dem Kopf.

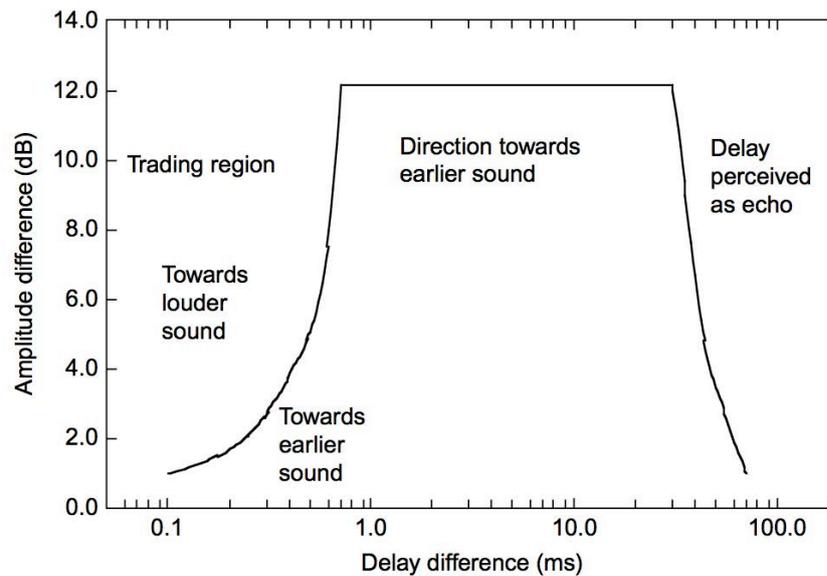
## 2.2 Lokalisation mehrerer Schallquellen

Im bisherigen Verlauf des Kapitels wurde die Funktionsweise der Lokalisation einer einzelnen Schallquelle erläutert. Der nächste Schritt ist folglich, die entstehenden psychoakustischen Effekte bei der Wahrnehmung mehrerer Schallquellen aus unterschiedlichen Richtungen zu beleuchten. Strahlen die Quellen völlig unabhängige, also vollständig dekorrelierte Signale ab, so werden sie dementsprechend als einzelne Quellen separat wahrgenommen. Haben die beiden Signale jedoch Bezug zueinander, wie beispielsweise ein Direktsignal und dessen Wandreflexion oder verschiedene schallabstrahlende Teile eines Instrumentes, so ergeben sich je nach zeitlicher Abfolge unterschiedliche Phänomene (Howard & Angus, 2009, S. 27).

Zeitlich sehr schnell aufeinanderfolgende korrelierte Signale ( $\Delta t < 0,6$  ms) werden als Einheit interpretiert und die ITD und ILD, wie in den vorangegangenen Abschnitten beschrieben, analysiert. Der wahrgenommene Quellort befindet sich auf der Strecke zwischen den beiden realen Quellorten. Besteht kein Zeitversatz und sind beide Schallquellen gleich laut, wird das Ereignis mittig dazwischen lokalisiert. An dieser Stelle muss also zwischen Richtung des Hörereignisses (Wahrnehmung) und Richtung des Schallereignisses (tatsächliche Positionen der Schallquellen) unterschieden werden. Die entstandene **virtuelle Schallquelle** wird als *Phantomschallquelle*, der Prozess als *Summenlokalisierung* bezeichnet. (Görne, 2011, S. 128; Weinzierl, 2008, S. 101) Dieses Prinzip bildet die Grundlage für die Verteilung von virtuellen Schallquellen zwischen zwei oder mehreren Lautsprechern eines Wiedergabesystems (siehe Abschnitt 4.1).

Weisen die beiden korrelierten Signale einen Zeitversatz im Bereich 0,6 bis 30 ms auf, so wird das Schallereignis in Richtung der zuerst eintreffenden Welle lokalisiert. Man spricht vom *Gesetz der ersten Wellenfront*, bzw. dem *Präzedenz-Effekt* (Dickreiter et al., 2014, S. 132). Dabei kann das später eintreffende Signal bis zu 12 dB lauter sein als das frühere, ohne dass sich die Lokalisationsrichtung auf den Quellort des zweiten Signales verschiebt (Howard & Angus, 2009, S. 114).

Kommt ein zweites Signal später als 30 ms nach dem ersten Signal am Ohr des Hörers an, so werden zwei getrennte Ereignisse wahrgenommen (Görne, 2011, S. 128). Abbildung 2.7 veranschaulicht die drei ausgeführten Bereiche. Es wird zusätzlich deutlich, dass die genannten Grenzen des Zeitversatzes je nach Pegelverhältnis zwischen beiden Signalen variieren.



**Abbildung 2.7:** Summenlokalisierung, Präzedenz-Effekt und Echo bei zeitlich versetzten, korrelierten Signalen.  
(Howard & Angus, 2009, S. 115)

## 2.3 Wahrnehmung einer ausgedehnten Schallquelle

Im bisherigen Verlauf wurden idealisierte Punktschallquellen betrachtet. Die realen Schallquellen, wie die zu untersuchenden Musikinstrumente, besitzen hingegen eine definierte Breite und Höhe. Strahlen verschiedene räumlich getrennte Flächen einer solchen Quelle Schall mit unterschiedlichen spektralen Anteilen ab (siehe Kapitel 3), resultiert daraus auch eine bestimmte wahrgenommene akustische Ausdehnung. Da dem Gehirn nur der Vergleich der beiden Ohrsignale zur Wahrnehmung des auditiven Geschehens zur Verfügung steht, stellt sich an dieser Stelle die Frage, wie diese erfasst werden kann.

Die wohl größte Ausdehnung eines akustischen Ereignisses weist der Raumhall in einem großen Gebäude oder Saal auf. Die Schallwellen von einer in einem solchen Raum befindlichen Quelle werden an verschiedenen Oberflächen reflektiert und gestreut. Diese Reflexionen werden wiederum reflektiert. Daraus ergibt sich schließlich das Diffusschallfeld, welches vom Hörer als gleichmäßiger, umhüllender Raumklang aus allen Richtungen wahrgenommen wird. Die Signale stammen dabei alle von der selben Quelle, variieren aber in Spektrum und Ankunftszeit. Es handelt sich somit um ähnliche Signale, welche stark dekorreliert sind (Santala & Pulkki, 2011, S. 1522). Demnach muss die wahrgenommene Ausdehnung einer Schallquelle auch mit der Korrelation der beiden Ohrsignale verknüpft sein. Gängigerweise wird der Koeffizient der Interaural Cross-Correlation (kurz **IAC**), welcher Werte von 0 (keine Korrelation) bis 1 (identisches Signal) annimmt,

verwendet. Ein niedriger Wert bedeutet eine große wahrgenommene Ausdehnung (Blauert, 1996). Die Bekanntheit des Signals spielt jedoch ebenfalls eine Rolle. Wird beispielsweise das Geräusch einer sehr großen Schallquelle von einem einzelnen Lautsprecher (mono) wiedergegeben, so wird dennoch aufgrund der Erfahrung eine gewisse Ausdehnung wahrgenommen (Potard & Spille, 2003, S. 2). Weiterhin ist die wahrgenommene Größe einer Schallquelle, im Gegensatz zur Richtungslokalisierung, abhängig von der Lautstärke. Ein lautes Signal wird grundsätzlich größer wahrgenommen als ein leises. Das Spektrum des Signals spielt ebenfalls eine Rolle: Ein Signal, welches sehr tiefe Frequenzen beinhaltet, vermittelt das Gefühl einer großen Schallquelle. In vertikaler Richtung macht sich außerdem der Pitch-Height-Effekt bemerkbar. Das Signal muss genügend Anteile im höherfrequenten Bereich aufweisen, damit sich eine Ausdehnung nach oben ergeben kann (Cabrera & Tilley, 2003).

Ziemer hat Messungen und Untersuchungen zur Abstrahlung verschiedener Instrumente durchgeführt, um Erkenntnisse über den Zusammenhang zwischen Ohrsignalen und wahrgenommener Ausdehnung zu gewinnen. Dabei wurden verschiedene monaurale und interaurale Parameter ausgewertet. Signifikante Ergebnisse bestanden im Zusammenhang zwischen wahrgenommener Ausdehnung und ITD unterhalb 1,2 kHz bzw. ILD bei höheren Frequenzen, sowie der Anzahl relevanter Obertöne. Diese Beobachtungen stehen im Einklang mit der Aussage, dass Signale mit sehr tieffrequenten Anteilen als größer empfunden werden. Denn je tiefer der Grundton eines Instrumentes, desto mehr hör- bzw. messbare Oberschwingungen sind potentiell im Signal enthalten (Schneider, 2017, S. 336).

## 2.4 Zusammenfassung

Zusammenfassend für das Kapitel der psychoakustischen Grundlagen ist festzuhalten, dass zur Richtungslokalisierung einer Schallquelle die Differenzen zwischen den beiden Ohrsignalen analysiert werden. Diese Signale wurden - verglichen mit dem Quellsignal - durch Reflexionen am Körper in zeitlichem Verhalten und Spektrum beeinflusst. Sie weisen daher für bestimmte Einfallrichtungen typische Amplituden- und Phasenverläufe auf. Diese Verzerrungen im Vergleich zum Ursprungssignal werden durch die HRTF dargestellt, welche aufgrund unterschiedlicher Anatomie für jeden Menschen verschieden ausfällt.

Die Lokalisation in horizontaler Richtung basiert bei tiefen Frequenzen auf aus Zeitversätzen resultierenden Phasenverschiebungen zwischen den beiden Ohrsignalen (ITD) und für hohe Frequenzen auf Pegeldifferenzen (ILD). Im Gegensatz zu diesen interauralen

Ohrsignalmerkmalen werden für die Lokalisation in vertikaler Richtung monaurale Merkmale herangezogen. Dabei handelt es sich um Muster spezifischer Einbrüche und Überhöhungen im oberen Teil des Spektrums der Ohrsignale, je nach Elevationswinkel der Schallquelle. Entsprechend sind breitbandige Signale besser lokalisierbar als schmalbandige. Des Weiteren führt der Kopf Peilbewegungen durch, hauptsächlich um zwischen Signalen von vorne und hinten unterscheiden zu können. Daraus ergibt sich die Tatsache, dass kurze Signale schlechter lokalisierbar sind. Außerdem werden bekannte Schallereignisse, aufgrund des Lernprozesses des Gehirns, präferiert in der für diese Quelle typischen Richtung lokalisiert.

Bei der Wahrnehmung mehrerer korrelierter Signale aus verschiedenen Richtungen ist der zeitliche Versatz ausschlaggebend. Bei sehr geringer Differenz kommt es zur Summenlokalisierung, bei großem Versatz zur Wahrnehmung zweier Quellorte. Im Bereich dazwischen wird in der Richtung lokalisiert, aus der die erste Wellenfront eintrifft. Eine wahrgenommene Ausdehnung einer Schallquelle entsteht, wenn die nahezu gleichzeitig eintreffenden Signale einen gewissen Grad an Dekorrelation aufweisen. Ausschlaggebend scheinen dabei für hohe Frequenzen die interauralen Pegelunterschiede (ILD) und für tiefe Frequenzen Phasenversätze zu sein. Die Größe der Ausdehnung hängt außerdem von der Lautstärke, sowie vom Obertonreichtum bzw. der unteren Grenzfrequenz des vorliegenden Spektrums ab. Es wurde des Weiteren der Pitch-Height-Effekt vorgestellt, nach welchem tiefe Frequenzen im Gegensatz zu hohen nur bedingt vertikale Ausdehnung annehmen können.

### 3 Schallabstrahlung akustischer Instrumente

Im Folgenden werden die akustischen Eigenschaften der Musikinstrumente, die im praktischen Teil der Arbeit aufgezeichnet werden, vorgestellt. Jedes Instrument besitzt einen bestimmten Klang, also eine charakteristische spektrale Verteilung bei bestimmten Tönen, sowie geräuschhafte Anteile. Vor allem das unterschiedliche Einschwingverhalten bei gleichem Grundton dient der Unterscheidung zwischen Instrumenten (Meyer, 2009, S. 23 ff.). Um jedes Instrument optimal darstellen zu können, ist es wichtig zu verstehen, wodurch sich der jeweilige Klang charakterisiert und wo er entsteht. Des Weiteren muss die Schallabstrahlung unter verschiedenen Winkeln zum Instrument betrachtet werden, da sich aus Überlagerungen des von mehreren schwingenden (Teil-)Flächen abgestrahlten Schalls gewisse Vorzugsrichtungen für bestimmte Frequenzen ergeben. So kommt es je nach Winkel zu phasenbedingten Auslöschungen und Erhöhungen. Außerdem variieren bei den meisten Instrumenten je nach gespielter Ton die Quellorte des Schalls und damit meist auch die Abstrahlcharakteristik (Jacques, 2005, S. 4). Beispielhaft sei an dieser Stelle das Abgreifen einer Gitarrensaite an verschiedenen Positionen auf dem Griffbrett genannt. Diese Effekte sorgen schließlich für die im Rahmen der vorliegenden Arbeit untersuchten Möglichkeit, die Instrumente in der Mischung ausgedehnt darzustellen und müssen daher bei den späteren Überlegungen zur Mikrofonposition (Abschnitt 5.6) bedacht werden.

Die Auswahl der Instrumente für die Testaufnahme fand unter Beachtung folgender Anhaltspunkte statt:

1. Hohe Relevanz in aktuellen Ensembles für Popular- und/oder Orchestermusik
2. Sinnvolle und praktikable Abmessungen
3. Komplexere Abstrahlcharakteristik
4. In konventionellen Mischungen mono dargestellt

Punkt 1 stellte dabei eine Grundvoraussetzung dar, da die Erkenntnisse aus der vorliegenden Arbeit Anwendung in den aktuellen Tätigkeiten von Tonschaffenden finden sollen. Die unter Punkt 2 gelisteten Abmessungen des Instrumentes sollten aus zweierlei Gründen „sinnvoll“ und „praktikabel“ sein: Einerseits muss das Instrument in den ausgewählten Aufnahmeraum transportierbar, aufbaubar und aus einer sinnvollen Distanz mikrofonierbar sein (siehe dazu Abschnitt 5.3) und andererseits sollen die Abmessungen

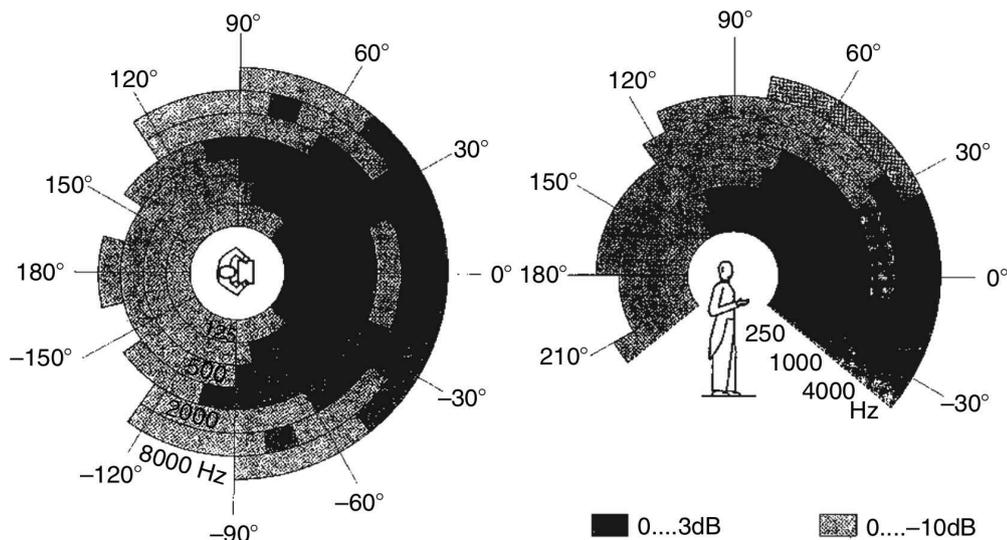
eine ausgedehnte Darstellung in der mehrdimensionalen Mischung rechtfertigen. Aus diesen Gründen werden sehr große und unhandliche Instrumente, wie Orgel oder Flügel, und sehr kleine Instrumente, wie Piccoloflöte oder Woodblocks, nicht berücksichtigt. Die Familie der Blechblasinstrumente wird, obwohl sie optisch bezüglich ihrer Dimensionen durchaus in das zu betrachtende Raster passen, ebenfalls ausgelassen. Dies erklärt sich durch Punkt 3: Die Schallabstrahlung folgt bei allen Mitgliedern dieser Instrumentenfamilie zu mittleren und hohen Frequenzen hin fast ausschließlich über den Schalltrichter (Pätynen & Lokki, 2010, S. 138). Damit haben Blechblasinstrumente nur eine schallabstrahlende Fläche mit annähernd punktförmigem Ursprung. Eine ausgedehnte Darstellung in der mehrdimensionalen Mischung verspricht daher keinen Mehrwert. Punkt 4 erklärt, dass vor allem Instrumente untersucht werden sollen, die in konventionellen Stereo-Mischungen meist an einem Punkt (mono) platziert sind und somit ohne Ausdehnung dargestellt werden. Dies gilt beispielsweise nicht für das Schlagzeug. Dennoch wird dieses Instrument betrachtet, da es eine sehr hohe Relevanz nach Punkt 1 besitzt. Selbiges gilt für die Gesangsstimme, wobei an dieser Stelle noch ein weiterer Effekt zum Tragen kommt, der ebenso für den Shaker gilt: Die Bewegungen der Musizierenden bzw. des Instrumentes an sich (siehe 3.6).

Folgende Instrumente sind schließlich ausgewählt worden, Teil des aufzuzeichnenden virtuellen Ensembles zu sein:

- Gesangsstimmen
- Querflöte, Oboe, Klarinette, Tenorsaxofon
- Violine, Kontrabass
- Akustische Westerngitarre
- Shaker, Congas, Schlagzeug

In den folgenden Abschnitten werden die verschiedenen Instrumentengruppen mit ihren akustischen Eigenschaften vorgestellt und auf die ausgesuchten Instrumente im Detail eingegangen. Es ist festzuhalten, dass die dabei zu untersuchende Abstrahlcharakteristik der Instrumente komplexer Abhängigkeiten unterliegt. Sie gestaltet sich frequenzabhängig, variiert je nach Instrument teilweise von Ton zu Ton und hängt von der Spieltechnik ab (Pätynen & Lokki, 2010, S. 139). Die getroffenen Aussagen und präsentierten Diagramme stellen daher oft die aus Mittelung gewonnenen globalen Tendenzen dar.

Im Rahmen der praktischen Untersuchungen am *Fraunhofer-Institut für Bauphysik IBP* konnten, neben der Durchführung der Testaufnahmen, für jedes Instrument Daten mit dem *Visor*-System von *HEAD acoustics* gewonnen werden. Dieses, im Folgenden als *akustische Kamera* bezeichnete, System dient der Lokalisation von Schallquellen in Echtzeit und der Visualisierung dieser.



**Abbildung 3.1:** Gemessenes Abstrahlverhalten der Gesangsstimme für alle Vokale gemittelt. (Meyer, 2009, S. 176)

Für jedes Instrument wurde eine chromatische Tonleiter über eine Oktave frontal ( $\phi \approx 0$ ) und seitlich ( $\phi \approx -60^\circ$ ) gemessen. Im Verlauf des Kapitels wird an einigen Stellen auf diese Messungen Bezug genommen und Bildschirmaufnahmen aus den entstandenen Videofiles präsentiert (siehe weiterführend Anhang D).

### 3.1 Gesangsstimme

Die menschliche Stimme zeigt sich als sehr facettenreicher, flexibler und komplexer Ton-erzeuger. Er kann verschiedenste geräuschhafte Komponenten (Konsonanten und Laute) sowie Töne verschiedener Klangfarben produzieren. Der Tonumfang der Stimme reicht maximal vom großen D bei Bassstimmen (Grundswinung 73 Hz) bis zum e''' beim hohen Sopran (1,32 kHz) (Dickreiter et al., 2014, S. 70). Die Lautstärke der gesungenen Töne nimmt im Allgemeinen innerhalb einer Stimmlage zu den höheren Registern hin zu. Höhere Stimmlagen besitzen auch einen generell nach oben verschobenen Dynamikbereich, dessen Obergrenze bis zu 125 dB erreichen kann. Der effektiv nutzbare Umfang beträgt ca. 30 dB. Die stimmhaften Konsonanten, aber vor allem die verschiedenen Vokale, werden durch einen dominierenden Formantenbereich unabhängig der gesungenen Tonhöhe charakterisiert. Bis auf sehr hohe Lagen ist dieser Frequenzbereich in der Regel 15 bis 20 dB lauter, als der eigentliche Grundton. Für den gesungenen Vokal „o“ dominieren beispielsweise die Anteile zwischen 400 und 600 Hz. Der Bereich um 2,5 kHz trägt zur Unterscheidung zwischen Gesangsstimmen und zur Durchsetzungsfähigkeit gegen andere Instrumente bei (Meyer, 2009, S. 123 ff.).

Der Mund gleicht in gewisser Weise einem Schalltrichter. Das Abstrahlverhalten der Stimme bei Vokalen gestaltet sich entsprechend ähnlich dem eines Blechblasinstrumentes. Deutlich unterhalb von 1 kHz ist die Abstrahlung kugelförmig, da hier die halben Wellenlängen größer als der Kopfdurchmesser sind. Zu höheren Frequenzen hin wird der Schall zunehmend nach vorne gebündelt, allerdings schwächer, als bei Blechbläsern. Im Allgemeinen nehmen weder die Lautstärke oder Tonhöhe, noch der gesungene Vokal maßgeblichen Einfluss auf das Verhalten. Die größten Varianzen bestehen in den Drittel-Oktavbändern 800 Hz und 1 kHz. Hier zeigt sich, dass sowohl dunklere Vokale als auch ein sehr lauter Ton zur Dämpfung in frontaler Richtung führen (Katz & d'Alessandro, 2007, S. 3 f.). In Abbildung 3.1 wird ersichtlich, dass die Vorzugsrichtung, in der alle Frequenzen ihren maximalen Pegel besitzen, nicht frontal und auf Höhe des Mundes liegt, sondern etwa 30° seitlich und nach unten versetzt.

Der Quellort des Schalles ist der Mund. Damit besitzen Gesangsstimmen, ähnlich wie Blechblasinstrumente, eher geringe Relevanz für die Motivation hinter dieser Arbeit, eine Schallquelle mit größerer Ausdehnung als solche darzustellen. Es gibt jedoch zwei maßgebliche Gründe, warum sie trotzdem untersucht werden: Zum einen stellt der Gesang meist die Hauptrolle in Populärmusikstücken dar und zum anderen wird vermutet, dass durch die Bewegungen von Kopf und Körper bei mehrdimensionaler Mikrofonabnahme interessante dynamische Effekte erzielt werden. Hierauf wird in Abschnitt 3.6 eingegangen.

## 3.2 Holzblasinstrumente

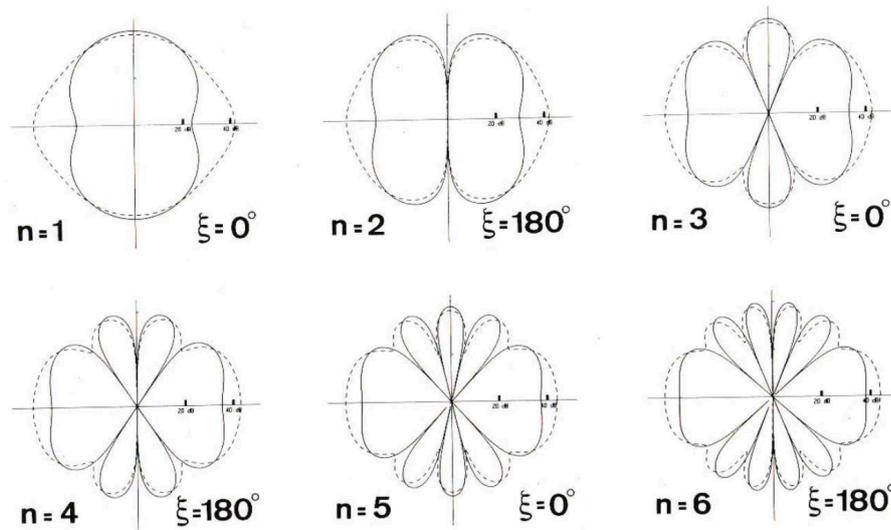
Die Klangerzeugung der zu untersuchenden Holzblasinstrumente geschieht im Falle der Flöte durch Anblasen einer Kante, während bei Klarinette, Saxofon und Oboe ein einfaches bzw. doppeltes Rohrblatt in Schwingung versetzt wird. Diese erzeugen im länglichen Körper des Instruments stehende Wellen einer Luftsäule, deren Länge je nach geöffneten Grifföffnungen variiert. Dadurch können unterschiedliche Frequenzen und somit verschiedene Tonhöhen erzielt werden (Dickreiter et al., 2014, S. 85). Die abstrahlenden Komponenten der zu untersuchenden Holzblasinstrumente, mit Ausnahme der Flöte, sind der Schalltrichter (also das offene Ende des Instrumentes) und die offenen Fingerlöcher. Bei der Flöte kommt zusätzlich das Mundstück selbst zum Tragen. Da je nach gespieltem Ton unterschiedliche Löcher geöffnet sind, ergeben sich für Holzblasinstrumente sehr komplexe akustische Wechselwirkungen, welche eine schwierige Vorhersagbarkeit der Abstrahlcharakteristik zur Folge haben (Pätynen & Lokki, 2010, S. 148).

### 3.2.1 Querflöte

Die Querflöte unterscheidet sich in vielen Punkten maßgeblich von den anderen drei ausgewählten Holzblasinstrumenten: Sie wird aufgrund ihrer früheren hölzernen Bauformen zu dieser Instrumentengruppe gezählt, besteht heute jedoch aus Metall. Das gerade Rohr der Querflöte hat eine zylindrische Form, bei sehr kleinem Durchmesser (Fletcher & Rossing, 1998, S. 538). Sie besitzt außerdem kein schnabelförmiges Mundstück, welches an das Rohr anschließt, sondern ein seitlich am Rohr angebrachtes Anblasloch. Dieses hat zur Folge, dass das Instrument seitlich bespielt wird. Ihr Tonumfang reicht vom kleinen  $h$  bis zum  $c''''$ , dies entspricht den Grundtonfrequenzen 247 Hz bis 2,1 kHz (Dickreiter et al., 2014, S. 88). Der Dynamikbereich innerhalb einer Lage beschränkt sich auf maximal 20 dB. Die Lautstärke variiert jedoch stark je nach Höhe des gespielten Tones. Ein *pianissimo* in hoher Lage entspricht mit ca. 83 dB bereits nahezu dem Pegel eines *fortissimo* in tiefer Lage (86 dB). Damit wird über alle Lagen hinweg ein Bereich von 33 dB (67 bis 100 dB) abgedeckt.

Das Spektrum der Querflöte charakterisiert sich durch einen dominanten Grundton. Die Stärke der ersten Obertöne kann durch die Spieltechnik stark variiert werden, wodurch sich unterschiedliche Klangfarben erzielen lassen (Meyer, 2009, S. 64 ff.). Die geräuschhaften Anteile, die durch den auf die Kante des Loches auftreffenden Luftstrom entstehen, sind charakteristisch für den Klang der Flöte und sorgen dafür, dass sie auch in größeren Ensembles hörbar bleibt (Bork, 1990, S. 356).

Da der Mund beim Anblasen nur anliegt, besitzt die Querflöte selbst bei geschlossenen Fingerlöchern zwei schallabstrahlende Öffnungen, weshalb ihre Abstrahlcharakteristik idealisiert als dipolförmig beschrieben werden kann (Rindel, 2020, S. 5). In der ersten Oktave entsteht durch die erste Mode des Flötenkorpuses ein gleichphasig abstrahlender Dipol (Bipol). Die erste Oberwelle entspricht der zweiten Mode, die an den beiden Enden als gegenphasiger Dipol abstrahlt. Der zweite Oberton ( $n=3$ ), wie alle weiteren ungeraden Harmonischen, strahlt wieder als phasenrichtiger Bipol ab, die geraden entsprechend als gegenphasiger Dipol. Es entstehen außerdem zusätzliche Seitenkeulen in senkrechter Richtung zur Instrumentenachse (Rindel, 2020, S. 5; Bork, 1990, S. 359 ff.). Das beschriebene Verhalten ist in den Polardiagrammen in Abbildung 3.2 dargestellt. Sie zeigen das Abstrahlverhalten der Harmonischen (Grundton  $n = 1$  bis fünfter Oberton  $n = 6$ ) in der ersten Oktave. Die horizontale Achse entspricht dabei der Ausrichtung der Flöte. Die gestrichelten Linien stellen das Verhalten im Nahfeld, die durchgezogenen das im Fernfeld dar. Es wird ersichtlich, dass die Winkelbereiche außerhalb der Vorzugsrichtungen im Nahfeld weniger starke Dämpfung aufweisen als im Fernfeld.

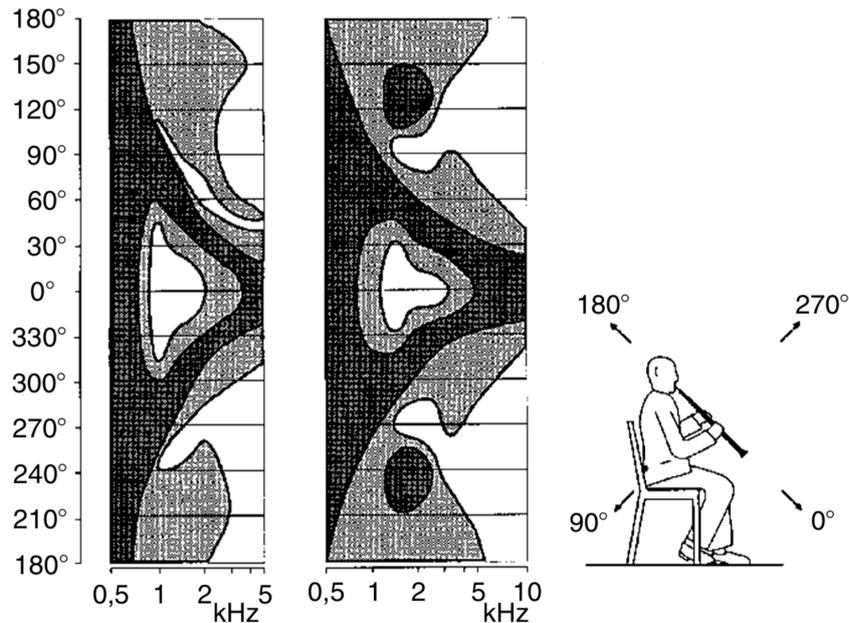


**Abbildung 3.2:** Berechnetes Abstrahlverhalten der Harmonischen der Querflöte.  
(Bork, 1990, S. 360)

Die Töne der zweiten Oktave entstehen durch Überblasen, somit wird der Grundton durch die zweite Mode erwirkt. Die Obertonreihe besteht dann nur aus den geradzahigen Moden. In der Dritten Oktave gestaltet sich die Abstrahlung für alle Töne mit vielen Seitenkeulen (in etwa wie das in Abbildung 3.2 abgebildete für  $n=5$  und  $n=6$ ) und kann damit als nahezu kugelförmig charakterisiert werden (Rindel, 2020, S. 5). Die am Mundloch entstehenden geräuschhaften Anteile können einen Frequenzbereich abdecken, der sowohl tiefer als der Grundton als auch deutlich höher als die relevanten Obertöne reicht. Diese Geräusche werden annähernd kugelförmig abgestrahlt (Meyer, 2009, S. 143).

### 3.2.2 Oboe

Die Oboe besteht im Wesentlichen aus einem konischen, nicht gebogenen Rohr, an welches sich unten der Becher (Schalltrichter) und oben das Mundstück in Form eines Doppelrohrblattes anschließen. Letzteres wird beim Anblasen des Tones mit den Lippen von beiden Seiten umschlossen und stellt damit, im Gegensatz zur Flöte, keine für die Abstrahlung relevante Schallquelle dar (Fletcher & Rossing, 1998, S. 491; Meyer, 2009, S. 143). Der Tonumfang der Oboe reicht vom kleinen  $h$  bis etwa zum  $f'''$ , entsprechend von 247 Hz bis 1,4 kHz (Dickreiter et al., 2014, S. 88). Der Dynamikumfang liegt bei etwa 33 dB für Einzeltöne (70 dB bis 103 dB). Charakteristisch für den Klang der Oboe ist die relativ starke Dämpfung höherer Oberschwingungen, sodass vor allem bei leisen Tönen in hoher Lage nur die ersten Harmonischen ausgebildet werden und der Grundton



**Abbildung 3.3:** Abstrahlverhalten von Klarinette (links) und Oboe (rechts); Schnitt durch die Medianebene.  
(Meyer, 2009, S. 147)

dominiert. Bei Tönen in mittlerer und tiefer Lage spielt der Grundton jedoch eine weniger wichtige Rolle. Der charakteristische Klang ergibt sich hier aus der starken Präsenz an Obertönen im Bereich ab 1 bis maximal 3 kHz. Bei pianissimo Passagen überwiegt der erste Oberton, während bei lauterem Tönen größere Vielfache der Grundschwingung in den Vordergrund treten (Meyer, 2009, S. 70 ff.).

Abbildung 3.3 zeigt das Abstrahlverhalten der Oboe sowie der Klarinette, da sich diese aufgrund der Ähnlichkeit in Abmessung und Aufbau in etwa entsprechen. Genauer wird in Abschnitt 3.2.3 erläutert. Auffällig sind zwei Winkelbereiche in etwa über und hinter dem Spielenden. Hier kommt es zu einer Erhöhung bei ca. 2 kHz.

### 3.2.3 Klarinette

Die Klarinette gleicht im generellen Erscheinungsbild der Oboe. Sie ist jedoch etwas breiter, besitzt einen größeren Becher und ein schnabelförmiges Mundstück, welches mit einem einfachen Rohrblatt ausgestattet ist. Außerdem ist das Rohr zylindrisch ausgeführt (Fletcher & Rossing, 1998, S. 486). Dies resultiert in einem um eine Oktave nach unten erweiterten Tonumfang vom kleinen  $d$  bis zum  $g'''$ . Dies entspricht den Grundtonfrequenzen 147 Hz bis 1,57 kHz (Dickreiter et al., 2014, S. 88). Auch der Dynamikumfang der Klarinette ist mit bis zu 50 dB für Einzeltöne in mittlerer Lage groß,

da vergleichsweise leise Töne (um ca. 60 dB) erzeugt werden können. Der charakteristische, als „dunkel“ bezeichnbare Klang der Klarinette, erklärt sich durch die Eigenschaft, dass in tiefer Lage (ca. d bis d') die ungeraden Harmonischen mindestens 25 dB leiser ausgebildet sind, als die geraden Oberschwingungen. Werden Töne in höheren Lagen gespielt, verringert sich diese Differenz, wodurch ein „wärmerer“ Klang wahrgenommen wird. Die Lautstärke des gespielten Tones nimmt ebenfalls Einfluss auf diesen Effekt (Meyer, 2009, S. 74 ff.).

Das Abstrahlverhalten der Klarinette gestaltet sich nach Meyer aufgrund ihrer Abmessungen und unter Berücksichtigung der Abschattungseffekte durch den Spielenden unterhalb von 500 Hz kugelförmig. Die Messungen von Pätynen und Lokki weisen dagegen bei 500 Hz bereits eine Dämpfung von ca. 11 dB im Vergleich zur stärksten Abstrahlrichtung auf. Zu höheren Frequenzen hin nimmt die Bündelung in Richtung des Trichters zu. Wie auch bei der Oboe ist festzustellen, dass auf dieser 0°-Achse ein relativ starker Einbruch im Bereich 1 bis 2 kHz vorhanden ist. Dieser Frequenzbereich ist bei deutlich größeren Winkeln mit höherem Pegel vertreten (Meyer, 2009, S. 146 f.). Abbildung 3.3 veranschaulicht das beschriebene Verhalten. Es ist dabei jedoch zu beachten, dass es sich bei der Darstellung um eine Mittelung der Ergebnisse aus Messungen bei verschiedenen Grundtönen handelt. Da je nach Ton unterschiedliche Fingerlöcher geöffnet sind und ebenfalls zur Abstrahlung beitragen, gestaltet sich die Abstrahlcharakteristik im Einzelnen sehr komplex. Abbildung 3.4 zeigt die entsprechende Visualisierung mit der akustischen Kamera für Klarinette und Oboe. Die verschiedenen Quellorte des 1 kHz-Oktavbands für drei verschiedene Töne, welche sich bei unterschiedlichen Griffen ergeben, sind dargestellt. Die Lautstärke, mit der ein bestimmter Ton gespielt wird, hat zwar großen Einfluss auf den Pegel verschiedener Obertöne, jedoch bleibt deren Abstrahlrichtung unabhängig davon (Pätynen & Lokki, 2010, S. 146).

### 3.2.4 Tenorsaxofon

Das Saxofon besitzt ein ähnliches Mundstück wie die Klarinette, da die Schwingungserzeugung ebenfalls über ein einfaches Rohrblatt erfolgt. Dies ist der Grund, warum das Saxofon in die Instrumentengruppe der Holzbläser eingeordnet wird - der Rest des Instrumentes besteht meist aus Messing. Das konisch ausgeführte Rohr, welches einen deutlich größeren Durchmesser als das der Klarinette besitzt, endet in einem breiten Schalltrichter. Diese Eigenschaften sorgen für einen lauten und vollen Klang mit Grundton-Dominanz, wobei gerade und ungerade Oberschwingungen gleichmäßig ausgebildet sind (Fletcher & Rossing, 1998, S. 496). Der Tonumfang eines Tenorsaxofons reicht vom großen As bis zum e''. Dies entspricht den Grundtonfrequenzen 104 Hz bis 660 Hz.



(a) Oboe, gespielter Ton: c'



(b) Klarinette, gespielter Ton: g'



(c) Oboe, gespielter Ton: g'



(d) Klarinette, gespielter Ton: c'



(e) Oboe, gespielter Ton: a'



(f) Klarinette, gespielter Ton: e'

**Abbildung 3.4:** Quellorte für Oboe und Klarinette im 1 kHz-Oktavband bei unterschiedlichen Tönen.

Das Abstrahlverhalten des Tenorsaxofons gestaltet sich ähnlich komplex wie das der Klarinette und der Oboe, da die Abhängigkeit vom gespielten Grundton besteht. Abbildung 3.5 zeigt beispielhaft drei Bildschirmaufnahmen bei verschiedenen Tönen aus den Messungen mit der akustischen Kamera.



(a) Gespielter Ton: cis'



(b) Gespielter Ton: d'



(c) Gespielter Ton: e'

**Abbildung 3.5:** Quellorte für Tenorsaxofon im 1 kHz-Oktavband bei unterschiedlichen Tönen.

### 3.3 Streichinstrumente

Die Saiten eines Streichinstrumentes werden meist mit Hilfe des Bogens, manchmal aber auch mit den Fingern in Oszillation versetzt. Diese Schwingung trägt wenig zur Schallabstrahlung bei, vielmehr wird sie über den Steg des Instrumentes auf den Korpus übertragen, welcher für die hauptsächliche Abstrahlung sorgt. Aufgrund verschiedener Dimensionen, Materialien und daraus resultierender Resonanzen, entsteht dann der für jedes Mitglied der Instrumentenfamilie typische Klang (Pätynen & Lokki, 2010, S. 158). Im Gegensatz zu den Holzblasinstrumenten, deren Klang weniger durch die verwendeten Materialien beeinflusst wird, gibt es bei den Streichinstrumenten demnach eine deutlich größere Varianz in Klang und Abstrahlverhalten zweier Instrumente, die zwar exakt gleiche Abmessungen aufweisen, aber aus unterschiedlichem Holz gefertigt wurden. Dies liegt darin begründet, dass die Schallabstrahlung eines Streichinstrumentes maßgeblich durch verschiedene Korpusresonanzen geprägt ist (Meyer, 2009, S. 85).

Durch die aufwändigen Geometrien in der Bauform und damit das Zusammenspiel verschiedener schallabstrahlender Teile, gestalten sich die Abstrahlcharakteristika komplex. Das Verhalten eines betrachteten Frequenzbereiches ist jedoch weitestgehend unabhängig von der Lautstärke und, im Gegensatz zu den Holzblasinstrumenten, ebenfalls unabhängig vom gespielten Ton (Pätynen & Lokki, 2010, S. 165).

### 3.3.1 Violine

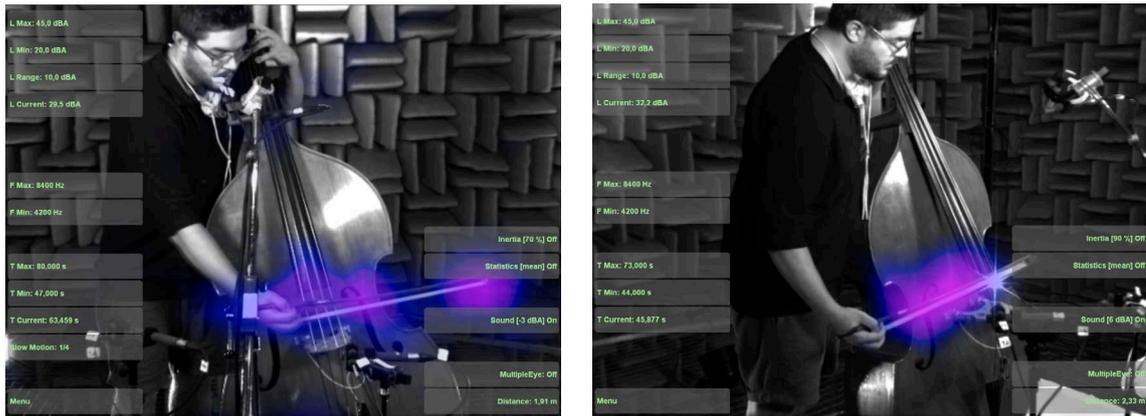
Die viersaitige Violine ist der kleinste Vertreter der Familie der Streichinstrumente (Fletcher & Rossing, 1998, S. 274). Der Tonumfang reicht vom kleinen g bis etwa zum c''', was den Grundtonfrequenzen 196 Hz bis 2,1 kHz entspricht (Dickreiter et al., 2014, S. 86). Der Dynamikumfang für Einzeltöne kann mit maximal 41 dB angegeben werden (58 bis 99 dB), verringert sich aber bei schnelleren Passagen deutlich. Für gezupfte Töne gilt ein ähnlicher Dynamikbereich, jedoch bei ca. 8 dB geringerem Pegel (Meyer, 2009, S. 90 ff.).

Die Saitenschwingung der Violine überträgt sich auf den Korpus, welcher eine Vielzahl von im Spektrum fix verankerten Resonanzfrequenzen hervorruft. In den höheren Lagen dominiert der Grundton das Spektrum der Violine. Die tiefste Resonanzfrequenz beläuft sich bei den meisten Instrumenten auf einen Wert zwischen etwa 260 und 300 Hz und entspricht damit dem Grundton von c' bzw. d'. Die Töne in der übrigen Quarte oder Quinte bis zum tiefsten Ton der Violine weisen daher einen um 20 bis 25 dB gedämpften Grundton auf (Meyer, 2009, S. 85 f.). Durch die Resonanzen höherer Ordnung entstehen bei verschiedenen Frequenzen Überhöhungen, welche die in diesem Bereich liegenden Obertöne des aktuell gespielten Tones hervorheben. Diese Bereiche werden immer dichter aufeinanderfolgend, je höher die Frequenz (Weinreich, 1997, S. 2338). Charakteristisch für den Klang der Violine zeigt sich beispielsweise der Bereich zwischen 800 und 1200 Hz, der einen Formantenbereich darstellt, welcher dem Vokal „a“ entspricht und dem Violinenklang Substanz verschafft (Meyer, 2009, S. 86 f.). Die Stärke der sich ausbildenden Obertöne hängt von vielen Faktoren ab. Beispielsweise entstehen bei offen gespielten Saiten mehr Obertöne, als bei abgegriffenen. Des Weiteren spielen Ansatzpunkt und Druck, vor allem aber die Geschwindigkeit des Bogens eine große Rolle. Mit dieser werden beispielsweise die unterschiedlichen Dynamikstufen erzielt. Im Gegensatz dazu führt erhöhter Bogendruck zu stärkeren Obertönen und folglich zu einem helleren Klang. Diese Zusammenhänge sind auf den Kontrabass übertragbar. Die am Bogen entstehenden geräuschhaften Komponenten sind ausschlaggebend für den typischen Streicher-Klang (Meyer, 2009, S. 88 f.).

Durch Abschätzungen anhand der relativ kleinen Dimensionen des Violinenkorpus wird kugelförmige Abstrahlung für Frequenzen unterhalb von 800 Hz erwartet (Weinreich, 1997, S. 2339). Die von Meyer (2009) sowie Pätynen und Lokki (2010) beschriebenen Messungen zeigen, dass die omnidirektionale Abstrahlung bis ca. 500 bzw. 630 Hz annähernd eingehalten wird. Im beschriebenen Formantenbereich um 1 kHz kommt es zu einer Verengung der Abstrahlung in die Richtung, in welche die f-Löcher zeigen. Für höhere Frequenzen kann generalisierend festgehalten werden, dass das Richtverhalten von einem zum nächsten Frequenzbereich sehr sprunghaft ist (Meyer, 2009, S. 153 f.). Dies hat einen Effekt zur Folge, der beispielsweise von Weinreich diskutiert wird und mit **Directional Tone Color**, also etwa „richtungsabhängige Klangfärbung“ der Violine bezeichnet wird. Bei fester Position von Instrument und Zuhörer werden bei einer gleichmäßig laut gespielten chromatischen Tonleiter einzelne Töne lauter und heller wahrgenommen als andere, da sich der Zuhörer im einen Moment für die momentan abgestrahlten Frequenzen der Oberschwingungen in einer Vorzugsrichtung befindet und im nächsten nicht mehr. Auch ein Vibrato (schnelles leichtes Variieren der Tonhöhe) bringt damit also eine zusätzliche dynamische Komponente in Form von Amplitudenmodulation mit sich. Dieses Phänomen führt dazu, dass bei einer monophonen Aufzeichnung der Violine, einige Töne lauter und andere leiser abgebildet werden. Bei Abnahme mit einem Stereomikrofon kommt es entsprechend der beschriebenen Vorzugsrichtungen zu Sprüngen im Panorama (Weinreich, 1997). Diese richtungsabhängigen Effekte werden daher einer der Kernpunkte der auswertenden Betrachtung in Abschnitt 6 sein.

### 3.3.2 Kontrabass

Der Kontrabass ist das größte Instrument seiner Familie und deckt einen Frequenzbereich ab, der bis an die untere Grenze des hörbaren Schalles reicht. Üblich sind Instrumente mit vier Saiten, deren Tonumfang vom Kontra-E bis ca. c' reicht (entsprechend 41 bis 260 Hz) (Dickreiter et al., 2014, S. 86). Es existieren aber auch fünfsaitige Varianten, deren Tonumfang dann nach unten teilweise bis zum Subkontra-H (31 Hz) erweitert ist. Aktuell gespielte Kontrabässe sind in ihrer Bauform deutlich weniger einheitlich, als die anderen Streichinstrumente, können aber prinzipiell anhand ihres Rückens in zwei Kategorien eingeteilt werden: Bässe mit flachem Rücken und Bässe mit gebogenem Rücken (Brown, 2004, S. 2). Diese Eigenschaft ist zusammen mit dem Aufbau der Verstrebungen im Korpus laut Brown noch ausschlaggebender für Abstrahlung und Klang des Basses, als die Wahl der Holzart (Brown, 2004, S. 79). Der Dynamikumfang bei Einzeltönen beläuft sich sowohl gezupft als auch gestrichen auf einen ähnlichen Wert (33 bzw. 34 dB), wobei die gestrichenen Töne deutlich lauter (66 bis 100 dB) als die gezupften sein können (60 bis 93 dB). Der Kontrabass ist damit generell etwas lauter als die anderen



(a) Kontrabass, Ansicht frontal

(b) Kontrabass, Ansicht seitlich

**Abbildung 3.6:** Quellorte für Kontrabass zwischen 4,2 und 8,4 kHz.

Streichinstrumente. Durch das relativ träge Ansprechen des Instrumentes und dem erforderlichen Feingefühl im Bogendruck, verringert sich der nutzbare Dynamikumfang in schnellen Passagen mit zusammenhängenden Tönen jedoch.

Der charakteristische Klang des Kontrabasses zeichnet sich durch den Grundtonbereich zwischen ca. 70 und 350 Hz (tiefer, voller Klang), einen Formantbereich je nach Lage bei ca. 500 Hz bzw. 800 Hz (dunkler „oh“- bzw. „ah“-Laut) sowie dem Bereich jenseits der Resonanzfrequenz des Stegs (ca. 1,25 kHz) aus. Letzterer wird vor allem durch die geräuschhaften Anteile geprägt, die am Bogen entstehen (Meyer, 2009, S. 100). Abbildung 3.6 zeigt zwei Bilder aus den Messungen mit der akustischen Kamera, welche dies verdeutlichen. Der Frequenzbereich 4,2 bis 8,4 kHz wurde dabei visualisiert. Diese Streich-Geräusche tragen mit ihrer direkten Ansprache dazu bei, dass auch bei schnelleren Passagen der gespielte Rhythmus hörbar bleibt, denn der Einschwingvorgang der tonalen Anteile des Basses misst teilweise weit über 100 ms (Meyer, 2009, S. 101 f.). Die tiefste Resonanzfrequenz eines Kontrabasses, welche sich aus dem Luftvolumen innerhalb des Korpus ergibt, liegt je nach Instrument zwischen 57 und 70 Hz und damit ca. eine Oktave oberhalb des tiefsten spielbaren Tones. Dies hat zur Folge, dass sich diese Lage durch die ersten Oberschwingungen charakterisiert, da der eigentliche Grundton bis zu 20 dB leiser ist. In tiefen und mittleren Lagen nimmt bei steigender Lautstärke des gespielten Tones der Anteil der Grundschwingung im Vergleich zu dem der Obertöne stärker zu. Dies liegt darin begründet, dass bei pianissimo Stellen der Druck des Bogens so schwach ist, dass die Grundschwingung der Saite nicht mehr voll ausgebildet wird. Entsprechend ist der Klang bei leisen Passagen leichter und nasaler (Meyer, 2009, S. 101 f.).

Das Abstrahlverhalten des Kontrabasses gestaltet sich durch das Zusammenspiel der Moden verschiedener Korpusteile relativ komplex. Es kann festgehalten werden, dass selbst bei tiefen Frequenzen bereits eine Richtwirkung vorhanden ist. Diese nimmt zu höheren Frequenzen hin zu und konzentriert sich auf einen Winkelbereich, der in etwa den Halbkreis vor dem Instrument beschreibt. Von einem Frequenzbereich zum nächsten treten allerdings Sprünge in der Vorzugsrichtung auf. Außerdem ist die Schallabstrahlung bei bestimmten Frequenzen auf mehrere solcher Richtungen aufgeteilt (Meyer, 2009, S. 161 ff.).

### 3.4 Akustische Westerngitarre

Die Gitarre gehört wie die Streichinstrumente zu den Saiteninstrumenten und weist im Regelfall sechs Saiten auf. Diese, im Falle der Westerngitarre, aus Stahl gefertigten Saiten werden allerdings nicht mit einem Bogen, sondern meist von Hand bzw. mit einem Plektrum in Schwingung versetzt. Der Tonumfang einer Gitarre in standardmäßiger Stimmung reicht vom großen E bis etwa zum d<sup>'''</sup>, je nach Anzahl der Bünde. Dies entspricht den Grundtonfrequenzen 82 Hz bis 1,2 kHz.

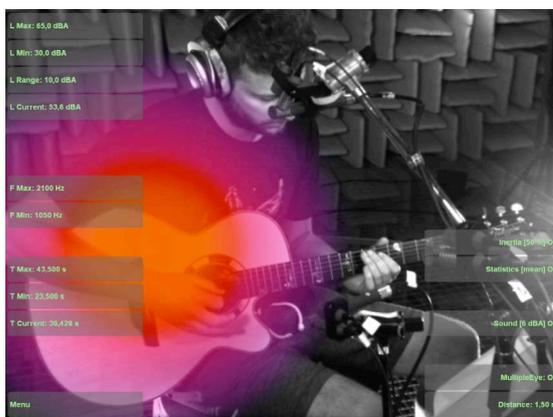
Die Schallabstrahlung erfolgt nach einem ähnlichen Prinzip und weist entsprechende Effekte auf, wie sie bei den Streichinstrumenten beschrieben worden sind. Die in Schwingung versetzte Saite trägt nur zu einem kleinen Teil zur Schallabstrahlung bei. Vielmehr überträgt sie die Schwingung über Steg und Hals auf den gesamten Korpus des Instrumentes. Die Vibrationen dieser Flächen, sowie die Oszillation der Luft durch das Schallloch, senden den Großteil der abgestrahlten Energie aus. Letzteres trägt auch noch deutlich oberhalb der Resonanzfrequenz des Hohlraums zum abgestrahlten Schallfeld bei (Bork, 1994, S. 425). Zu höheren Frequenzen hin sind nur noch der Steg und die anschließende Decke des Korpus aktiv (Fletcher & Rossing, 1998, S. 240). Neben den tonalen Anteilen ist das Klangbild der Gitarre geprägt von Geräuschen, die sowohl beim Anschlag als auch beim Rutschen der Greifhand über die Saiten entstehen. Abbildung 3.7 zeigt, wie die Schwerpunkte der genauen Quellorte für das 1 kHz-Oktavband zwischen gespielten Tönen variieren.



(a) Gespielter Ton: kleines c



(b) Gespielter Ton: kleines f



(c) Gespielter Ton: kleines gis

**Abbildung 3.7:** Quellorte für Gitarre im 1 kHz-Oktavband bei unterschiedlichen Tönen.

### 3.5 Schlaginstrumente und Percussion

Die Familie der Schlaginstrumente wird meist in zwei Kategorien gegliedert, nämlich in Instrumente mit und ohne definierter Tonhöhe. Sowohl die Einzelkomponenten des Schlagzeugs, als auch Congas und Shaker fallen in letztere Kategorie. Die einzelnen Trommeln eines Schlagzeugs folgen dem gemeinsamen Prinzip, dass sie ein Schlagfell und ein Resonanzfell besitzen, die akustisch durch das in der Trommel eingeschlossene Luftvolumen gekoppelt sind. Ein Schlagzeug kann aus sehr vielen verschiedenen Trommeln unterschiedlicher Größe bestehen. Das im Rahmen dieser Arbeit untersuchte Setup benutzt neben den Becken jedoch nur Bassdrum und Snaredrum. Diese beiden Trommeln bilden die rhythmische Basis für den Großteil aller Populärmusikstücke der letzten Jahrzehnte. Verschiedene Durchmesser des Kessels und das verwendete Material, sowie die Art der aufgezogenen Felle, sind entscheidend für den Klang. Congas besitzen im Gegensatz zu diesen beiden Trommeln nur eine Membran, das Schlagfell. Das gegenüberliegende Ende des Kessels bleibt offen (Fletcher & Rossing, 1998, S. 583).

Zu Richtwirkung und Abstrahlung von Schlag- und Percussioninstrumenten finden sich kaum Abhandlungen oder Messungen in der Literatur. Wenn, dann beziehen sie sich auf die Schlagwerke im Orchester (größere Bassdrum, Doppelbecken, siehe [Pätynen & Lokki, 2010](#)). Es wird vermutet, dass dies mitunter an der meist kurzen Dauer eines Anschlages und am geräuschhaften Klang ohne tonalem Bezug liegt. Dadurch wird eine Messung erschwert und aufgrund der fehlenden Verknüpfung zu gespielten Tönen weniger relevant. Das Abstrahlverhalten des Schlagzeugs bei verschiedenen Frequenzen zu bestimmen ist außerdem aufgrund der Einzelkomponenten, deren Spektren sich größtenteils überlappen, hinfällig. Es wird trotz seiner vielen Einzelbestandteile als ein gesamtes Instrument betrachtet und daher auch im Rahmen dieser Arbeit als eine Einheit verstanden. Im Falle des Shakers spielt die Abstrahlcharakteristik ebenfalls eine untergeordnete Rolle, da aufgrund der kleinen Abmessungen vor allem der dynamische Effekt durch die Bewegung eingefangen werden soll (siehe Abschnitt [3.6](#)). Dies gilt auch für das Anschlagen der Congas.

### 3.5.1 Shaker

Der Shaker dient meist als rasselartiges, durchlaufendes Element im Rhythmusgefüge. Er ist aufgrund seiner kleinen Abmessungen weniger dafür geeignet, als ausgedehnte Klangquelle dargestellt zu werden. Allerdings gestaltet sich die Klangerzeugung räumlich sehr dynamisch, nämlich durch Schüttelbewegungen. Es wird vermutet, dass diese Bewegungen genutzt werden können, um auch kleinen Instrumenten eine dynamische Ausdehnung und damit einen natürlicheren Platz in der mehrdimensionalen Mischung zu verschaffen. Aus diesem Grund wurde der Shaker stellvertretend für einige weitere kleine perkussive Instrumente (z.B. Rasseln und Schellenkranz) mit in die vorliegende Arbeit aufgenommen. Die Effekte, welche sich durch Bewegungen der Musiker ergeben, werden in Abschnitt [3.6](#) genauer erläutert.

### 3.5.2 Congas

Congas bestehen aus einer Fass-artigen Röhre aus verklebten Holzplatten an deren oberen Ende über einen Spannring das Fell aufgezogen ist. Obwohl die Congas zur Gruppe der Schlaginstrumente ohne bestimmter Tonhöhe zählen, kann dennoch die Grundschwingung definiert und durch den Spannring in gewissen Grenzen modifiziert werden ([Fletcher & Rossing, 1998](#), S. 618 f.). Ziel ist dabei jedoch nicht, eine bestimmte Tonhöhe zu erreichen, sondern die Resonanz des Conga-Körpers möglichst gut für einen länger anhaltenden Ton zu nutzen ([Albrecht, 2010](#), S. 102). Die Durchmesser der Schlagfelle

liegen meist zwischen 23 und 30 cm und die Instrumente produzieren einen Ton im Bereich 130 bis 260 Hz. Es entstehen allerdings auch deutlich höhere und tiefere Anteile, je nach Art des Anschlags. Am Rand des Schlagfells mit den Fingern angeschlagen, entstehen hellere, schärfere Transienten, während in der Mitte mit dem Handballen ein tiefer, dunkler und länger anhaltender Klang entsteht (Fletcher & Rossing, 1998, S. 619).

### 3.5.3 Bassdrum

Die Bassdrum wird im modernen Bandkontext, im Gegensatz zum Einsatz im Orchester, auf dem Boden stehend per am Fußpedal befestigten Schlägel gespielt. Dieser verursacht ein transientes Anschlaggeräusch. Damit der Impuls präzise bleibt und der Kessel nicht zu lange nachschwingt, wird er oft durch ein absorbierendes Material gedämpft, welches durch das Loch im Resonanzfell eingeführt werden kann. Manchmal fehlt dieses Fell jedoch komplett oder ist ohne Loch ausgeführt (Albrecht, 2010, S. 8). Im Regelfall wird das Schlagfell der Bassdrum beim Einsatz in kleineren Ensembles oder Bands etwas fester gespannt als das Resonanzfell. Es entstehen Wechselwirkungen zwischen den beiden Fellen, denn aufgrund der leicht differierenden Spannung weisen diese Moden auf, die sich in ihrer Frequenz leicht unterscheiden. Es resultiert eine Reihe harmonischer Schwingungen, die allerdings im mittleren Frequenzbereich von einer Vielzahl unharmonischer Komponenten verwaschen wird (Fletcher & Rossing, 1998, S. 599 ff.).

### 3.5.4 Snaredrum

Die Snaredrum besteht im Gegensatz zur Bassdrum immer aus einem Schlagfell (oben) und einem Resonanzfell (unten). Sie trägt ihren Namen aufgrund des unter dem Resonanzfell gespannten Gebildes aus Drahtsträngen (*Snare-Teppich*), welches die typischen schnarrenden Geräusche abstrahlt, wenn es durch die Schwingungen des Resonanzfells angeregt wird. Meist ist eine Mechanik zum Entspannen des Teppichs vorhanden, sodass dieser nicht mehr mitschwingt und dadurch ein dumpferer Klang entsteht, der anderen Trommeln ähnelt (Fletcher & Rossing, 1998, S. 602). Vor allem das Verhältnis der Spannungen der beiden Felle zueinander sowie die absolute Spannung bringen deutliche Klang- und Tonhöhenunterschiede (Albrecht, 2010, S. 17). Es existieren mehrere Spielweisen der Snaredrum. Neben dem normalen Anschlagen sind *Rimshots* und *Rimclicks* sowie Spieltechniken mit anderen Schlägeln üblich.

### 3.5.5 Becken

Ein Schlagzeug kann neben der Vielzahl an Trommeln auch aus einer großen Variationsbreite an Becken bestehen. Im Folgenden wird sich auf die Bauformen beschränkt, welche am aufzuzeichnenden Schlagzeugaufbau Verwendung finden. Die HiHat besteht aus zwei gleich großen Becken, die entgegengesetzt aufeinander liegen. Sie sind auf einem per Fußpedal steuerbaren Stativ montiert, welches unterschiedliche Abstände zwischen den beiden zulässt. Die Becken werden entweder durch diesen Mechanismus aufeinander geschlagen oder mit den Sticks in Schwingung versetzt. Durch das gegenseitige Aufeinandertreffen entsteht der typische klirrende Klang der HiHat. Sie wird in der rhythmischen Struktur meist als durchlaufendes, schnelleres Element verwendet.

Crash und Ride sind jeweils einzelne Becken, wobei das Ride meist größer ausfällt, nur mit der Spitze des Sticks angeschlagen und ähnlich der Funktion der HiHat verwendet wird. Das Crash sorgt hingegen für laute Akzente und wird mit dem dickeren Teil des Sticks angeschlagen. Generell kann festgehalten werden, dass Becken im Popmusik-kontext für den höheren Frequenzbereich des Schlagzeugs verantwortlich sind. Anteile unterhalb von ca. 400 Hz sind daher irrelevant. Der Klang eines Beckens charakterisiert sich durch den differierenden zeitlichen Verlauf verschiedener Frequenzbereiche nach dem Anschlag. Dieser entsteht durch die als zufällig einzustufenden, sich gegenseitig beeinflussenden Vibrationen des Beckens. Es kann festgehalten werden, dass nach dem Anschlag teilweise über 100 ms vergehen, bis einige Frequenzanteile hörbar sind. Vor allem der für den hellen Klang verantwortliche Präsenzbereich (ca. 3 bis 5 kHz) klingt viele Sekunden lang nach, während beispielsweise der Bereich um 400 Hz sowie zwischen 700 und 1000 Hz relativ schnell nach dem Anschlag wieder an Intensität verlieren. Erst einige Sekunden später dominiert aufgrund der schnelleren Dämpfung höherer Frequenzen wieder der Bereich um 400 Hz. Der Dynamikumfang von Becken umfasst ca. 30 dB, wobei bis über 110 dB erreicht werden (Meyer, 2009, S. 122; Fletcher & Rossing, 1998, S. 649 ff.).

## 3.6 Bewegungen der Musiker

Mit Ausnahme des Schlagzeugs und der Congas sind die untersuchten Musikinstrumente in ihrer Position nicht fixiert und werden beim Spielen bewegt. Diese Bewegungen geschehen teilweise ohne musikalischen Bezug, zum Beispiel durch eine Kopfdrehung beim Blick auf die Noten oder beim Wechsel eines bestimmten Griffes, aber oft auch als Teil des Vortrags, um beispielsweise Melodielinien oder Dynamik zu unterstreichen. Es ist bekannt, dass der bei diesen Bewegungen entstehende visuelle Eindruck einer Darbietung

nicht unerheblichen Einfluss auf die subjektive Bewertung der musikalischen Qualität desselben hat. Studien zeigen jedoch, dass die Bewegungen des Musikers auch ohne visuelle Verknüpfung gehört werden und positiver bewertet werden, als eine nicht-bewegte Darbietung (Ackermann, Böhm, Brinkmann & Weinzierl, 2019, S. 366). Die vorangegangenen Abschnitte haben gezeigt, wodurch die klangliche Dynamik bei Bewegungen des Musikers entsteht: Es sind teilweise große klangliche Unterschiede zu verzeichnen, je nach Winkel zum Instrument. Und dieser verschiebt sich durch die Bewegungen fortlaufend. In einem Konzertsaal hat dies zur Folge, dass verschiedene Frequenzkomponenten an unterschiedlichen Raumbegrenzungsflächen oder anderen Oberflächen reflektiert werden. Je nach Bewegung des Musikers verschieben sich diese Reflexionsorte auf andere Oberflächen mit unter Umständen differierendem frequenzabhängigem Absorptionsverhalten. Beim Zuhörer kommt damit ein dynamisches Klangbild an Direktschall vom Instrument an, gefolgt von Reflexionen, die in Richtung und Klangfarbe leicht variieren. Im Gesamten bleibt das Spektrum des Diffusschalles jedoch eher statisch, sofern die Raumbegrenzungsflächen keine starke Varianz im Absorptions- und Reflexionsverhalten aufweisen. Dadurch werden die deutlichen Unterschiede im Direktschall in Konzertsälen teilweise verdeckt. In Messungen wurden dennoch Pegelunterschiede von bis zu 10 dB in einzelnen Drittel-Oktavbändern festgestellt. Für Freifeldbedingungen wurden sogar bis zu 20 dB gemessen (Ackermann et al., 2019, S. 365).

Wird nun das Instrument an einem Punkt (mono) mikrofoniert, erfasst dieses meist gerichtete Mikrofon nur den Direktschall aus einem fixen Winkel zum Instrument. Je nach Bewegung des Instrumentes kommt es zu den charakteristischen Varianzen im Spektrum. Dies kann je nach Mikrofonabstand auch ungewünscht große klangliche Veränderungen mit sich bringen (vgl. beispielsweise Abnahme der Querflöte (Bork, 1990, S. 363) und (Curdt, 1998, S. 18)). Des Weiteren wird dieses Signal oft auch als Eingangssignal für einen künstlichen Halleffekt im Mischprozess benutzt. Die Frequenzbereiche, welche eher in andere Richtungen am Mikrofon vorbei abgestrahlt wurden, aber im realen Raum trotzdem vorhanden sind, fehlen damit im künstlich erzeugten Raumsignal.

### 3.7 Zusammenfassung

Dieses Kapitel hat einige wichtige akustische Eigenschaften der beispielhaft zu untersuchenden Musikinstrumente vorgestellt. Die Angaben zu Tonumfang und Dynamikbereich dienen der allgemeinen Einordnung der zu behandelnden Schallquellen. Insbesondere die untere Grenzfrequenz der Instrumente aus tieferen Registern, bzw. auf der anderen Seite die Anforderungen an Auflösung der hohen Frequenzen, sind bei der Wahl des Mikrofons zu berücksichtigen. Die für den Klang charakteristischen Frequenzbereiche der ersten

Oberschwingungen sind vor allem im Zusammenspiel mit deren Vorzugsrichtung in der Abstrahlung für die Positionierung der Mikrofone zu bedenken. Die Abstrahlcharakteristika der einzelnen vorgestellten Instrumentengruppen variieren durch verschiedenartige Tonerzeugung und Bauformen stark. Für Holzbläser kann festgehalten werden, dass tiefe Frequenzen nahezu kugelförmig abgestrahlt werden und zu hohen Frequenzen immer stärkere Bündelung auf die Instrumentenachse, also in Richtung des Trichters vorhanden ist, wobei sich für Klarinette und Oboe eine Vorzugsrichtung oberhalb und unterhalb des Instruments für Frequenzen um ca. 1,5 kHz ergibt. Die Abstrahlcharakteristik einzelner Frequenzbänder variiert jedoch je nach gespielter Tonhöhe, die Aussagen beziehen sich daher auf Mittelungen. Die Querflöte spielt aufgrund der differierenden Tonerzeugung eine besondere Rolle. Ihre Abstrahlung gestaltet sich dipolförmig und ist von starken Anblasgeräuschen geprägt. Die Saiteninstrumente weisen ein sehr komplexes Abstrahlverhalten auf, was durch die Resonanzen und Wechselwirkungen zwischen den Moden der Einzelkomponenten des Korpus hervorgerufen wird. Die Vorzugsrichtungen wechseln bei den Streichinstrumenten von einem Halbton zum nächsten sehr schnell, wodurch ein dynamisches Verhalten bei ausgedehnter Mikrofonierung erwartet werden kann. Für die Schlaginstrumente können wenige Aussagen über das Richtverhalten gemacht werden. Aufgrund der meist tiefen Grundtonfrequenzen, die tendenziell kugelförmig abgestrahlt werden, ist dies auch weniger relevant. Für Congas und Shaker ergeben sich durch die räumliche Dynamik bei der Klangerzeugung andere zu untersuchende Aspekte bei der ausgedehnten Abbildung. Die Relevanz der Gesangsstimme ist ebenfalls aufgrund dynamischer Effekte durch Bewegungen beim Vortrag gegeben.

Insgesamt ist also festzuhalten, dass die Abstrahlung verschiedener Frequenzbereiche in verschiedene Richtungen sowie der sich teilweise dynamisch ändernde Quellort bei verschiedenen Tönen Charakterisierungen der Instrumente darstellen. Dies muss bei den weiteren Überlegungen zur mehrdimensionalen Mikrofonierung in Kapitel 5, im Besonderen auch in Bezug auf die Bewegungen des Instrumentes, beachtet werden, damit ein möglichst stimmiges, ausgedehntes akustisches Bild des Instrumentes aufgenommen werden kann.

## 4 Aufnahme und Wiedergabe ausgedehnter Schallquellen

Die mehrdimensionale Darstellung von Musikinstrumenten ist über eine Vielzahl an Wiedergabesystemen, bestehend aus Lautsprechern oder über Kopfhörer realisierbar. Jedes System verfügt über gewisse Vorzüge und Nachteile bezüglich akustischer aber auch praktischer Aspekte. Im Rahmen der vorliegenden Arbeit beschränkt sich die Wiedergabe auf ein fest definiertes, geeignetes Lautsprechersystem. Aus Gründen der Vergleichbarkeit und Praktikabilität, sowie der direkten Anwendbarkeit in der Studioumgebung der Hochschule der Medien Stuttgart wurde entschieden, ein standardisiertes kanalbasiertes Wiedergabesystem zu verwenden. Auf Untersuchungen mit binauraler Kopfhörerwiedergabe wurde verzichtet, da das Ergebnis im räumlichen Eindruck von vielen Faktoren abhängt: Angefangen bei der Wahl des Kopfhöreres, über verschiedene Binauralisierungsverfahren bis hin zur individuellen HRTF des Hörers (vgl. Abschnitt 2.1.1). Ebenso wurde auf die Betrachtung aktueller objekt- oder szenenbasierter Formate verzichtet, da deren Renderingprozesse ebenfalls zusätzliche Variablen in der räumlichen Darstellung mit sich bringen. Die entstehenden Testaufnahmen können dennoch für spätere Forschung in den genannten Bereichen Verwendung finden.

Bevor Abschnitt 4.2 das Zielformat zur Wiedergabe und dessen Eigenschaften vorstellt, wird zuerst auf das zugrundeliegende Prinzip der stereofonen Positionierung virtueller Schallquellen und später auf die zugehörigen Mikrofonierungskonzepte eingegangen. Abschnitt 4.4 beschäftigt sich anknüpfend an die aus Kapitel 2 bekannten Prinzipien mit den Effekten bei der Lokalisation von virtuellen Schallquellen in einer mehrdimensionalen Lautsprecheranordnung und der Wahrnehmung von Ausdehnung. Im Anschluss werden unter 4.5 gängige Mischkonzepte und zugehörige Mikrofonierungen vorgestellt und das sich daraus ableitende Ziel des zu konzipierenden Mikrofonierungsverfahrens nochmals herausgestellt. Nachdem Abschnitt 4.6 zusätzliche technische Anforderungen an das Verfahren stellt, wird in 4.7 schließlich das entwickelte Konzept für die Mikrofonierung der zu untersuchenden Instrumente beschrieben.

## 4.1 Stereofonie

Die Lautsprecherstereofonie bildet nach der Monowiedergabe die erste Stufe und damit die Grundlage für die räumliche Darstellung virtueller Schallquellen. Das Prinzip beruht auf der Wiedergabe korrelierter Signale über zwei Lautsprecher mit der Bezeichnung L und R für links und rechts. Die Strecken zwischen Rezipient und je einem der Lautsprecher müssen dabei gleich lang sein wie der Abstand zwischen den Lautsprechern (*Lautsprecherbasis*). Die Lautsprecher werden normalerweise auf Ohrhöhe installiert und bilden somit gemeinsam mit dem Kopf des Hörers ein gleichseitiges Dreieck mit eingeschlossenen Winkeln von  $60^\circ$ . Somit ergeben sich die Positionen bei einem Stereo-Lautsprechersystem im kopfbezogenen polaren Koordinatensystem  $(\phi, \delta)$  zu: L ( $30^\circ, 0^\circ$ ) und R ( $-30^\circ, 0^\circ$ ). Die unter Abschnitt 2.2 beschriebene Summenlokalisation erklärt den entstehenden Effekt der Phantomschallquelle. Liegen zwei korrelierte Signale, die keinen zeitlichen Versatz und gleichen Pegel aufweisen, vor, so wird eine einzelne Quelle mit Hörereignisrichtung in der Mitte zwischen beiden realen Quellen wahrgenommen. Strahlen die beiden Lautsprecher mit unterschiedlichem Pegel ab, so rückt die virtuelle Quelle näher an den Lautsprecher mit höherem Pegel, bis sie schließlich ab einer Differenz von 15 dB gänzlich in Richtung der lauterer Quelle lokalisiert wird. Bei steigender Pegeldifferenz nimmt die Schärfe, mit welcher die Position der Phantomschallquelle lokalisiert werden kann, zu (Dickreiter et al., 2014, S. 221 ff.).

Alternativ gelingt die Darstellung einer Phantomschallquelle zwischen zwei Lautsprechern durch den Einsatz von Zeitdifferenzen. Die Lautsprecher geben das gleiche Signal mit gleichem Pegel wieder und die entstehende Phantomschallquelle wird durch Verzögerung des Signales eines Lautsprechers in die Richtung des früher abstrahlenden ausgeleitet. Ein Zeitversatz von ca. 1 ms führt bereits zur Lokalisation ganz außen auf einem Lautsprecher. Für größere Zeitdifferenzen kommen die unter Abschnitt 2.2 erläuterten Phänomene Präzedenz-Effekt und schließlich Echo-Wahrnehmung zum Tragen. Im Gegensatz zur Stereofonie über Pegeldifferenzen gestaltet sich die Richtungsstabilität zur Seite hin unsicher. Es ist außerdem festzuhalten, dass die Relation zwischen Pegel- oder Zeitversatz und wahrgenommener Auslenkung der Phantomschallquelle nur bei kleinen Winkeln linear verläuft. Zu größeren Winkeln hin sind zunehmend größere Unterschiede für eine gleich weite Auslenkung nötig (Dickreiter et al., 2014, S. 223 ff.).

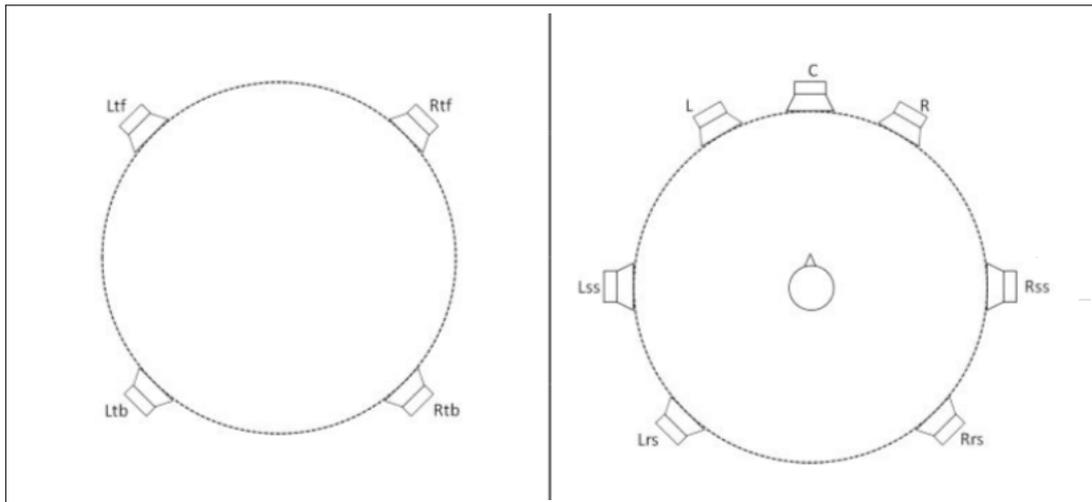
Die Problematik der Stereofonie besteht darin, dass sie nur für einen sehr kleinen Bereich, genannt Sweet Spot, das gewünschte akustische Bild liefert. Bereits kleine Drehungen und Bewegungen des Kopfes aus der Mitte zwischen beiden Lautsprechern haben starke Änderungen der Ohrsignale zur Folge. Daraus ergeben sich klangliche Verfärbungen, aber vor allem auch große Verzerrungen bezüglich der

Lokalisation (Roginska & Geluso, 2018, S. 64). Eine seitliche Verschiebung des Kopfes um ca. 30 cm lässt die eigentlich mittig befindliche Phantomquelle, entsprechend des Prinzips der Laufzeitverfahren, komplett außen auf dem nähergelegenen Lautsprecher erscheinen (Dickreiter et al., 2014, S. 223 ff.). Durch das Einfügen eines zusätzlichen Lautsprechers zwischen L und R kann der Sweet Spot in geringem Maße vergrößert werden, im Grundsatz bleibt die Problematik allerdings bestehen. Einige **Surround-Formate** fügen einen solchen zusätzlichen Center (C) Lautsprecher ein. Die Surround-Stereofonie erweitert das Prinzip der zweikanaligen Stereofonie um mehrere Lautsprecher auf Ohrhöhe, sodass virtuelle Schallquellen aus allen Richtungen in der Horizontalebene um den Hörer herum platzierbar sind. Gängig ist das durch die International Telecommunication Union (ITU) für den Rundfunk standardisierte 5.1 Format mit den Lautsprechern  $L_s$  ( $\phi = 100$  bis  $120^\circ$ ) und  $R_s$  ( $\phi = -100^\circ$  bis  $-120^\circ$ ) (ITU-R, 2012). Zwischen zwei benachbarten Lautsprechern findet das gleiche Prinzip von Phantomschallquellen durch Pegel- oder Zeitdifferenzen Anwendung (Roginska & Geluso, 2018, S. 180 ff.). Es zeigt sich jedoch, dass die Lokalisation dieser Quellen entsprechend der Erkenntnisse aus Abschnitt 2.1.2 nur in frontaler Richtung präzise funktioniert. Position und Spektrum einer Phantomquelle zwischen LS und RS, also hinter dem Kopf, und vor allem zwischen LS und L bzw. RS und R (also auf der interauralen Achse) gestaltet sich sehr instabil (Dickreiter et al., 2014, S. 227 f.).

## 4.2 Lautsprecheranordnung 4+7+0

Die gängige Bezeichnung „3D“ für verschiedentliche, den Hörer umgebende Lautsprecheranordnungen ist strenggenommen nicht ganz korrekt. Eine dreidimensionale Darstellung würde sich physikalisch gesehen aus den drei Dimensionen Breite, Höhe und Tiefe ergeben. Letztere, also die Entfernung einer Schallquelle, wird in einem solchen Lautsprechersystem über Hall- und Höhenanteil im Signal erzielt. Die Lautsprechersysteme decken physikalisch meist maximal eine zweidimensionale, gebogene Fläche in Form einer Halbkugel um den Hörer herum ab. Aus diesem Grund wird im Rahmen der vorliegenden Arbeit auch der Begriff **mehrdimensional** anstatt „3D“ verwendet. Das einzelne Instrument wird also konkret in den beiden Dimensionen Breite und Höhe ausgedehnt dargestellt und auf einer Halbkugelfläche um den Hörer herum platziert.

Nach ITU werden kanalbasierte mehrdimensionale Wiedergabesysteme in die drei Lautsprecherebenen *Upper Layer* (Lautsprecher oberhalb der Ohrhöhe), *Middle Layer* (Lautsprecherring auf Ohrhöhe) und *Bottom Layer* (Lautsprecher unterhalb der Ohrhöhe) eingeteilt. Abbildung 4.1 zeigt das für die Untersuchungen ausgewählte Wiedergabesystem J (4+7+0) nach ITU-R BS.2051-2 (ITU-R, 2018). Die drei Zahlenangaben



**Abbildung 4.1:** Lautsprecheranordnung 4+7+0 mit den Positionen im Upper Layer (links) und Middle Layer (rechts). (ITU-R, 2018)

beziehen sich auf die Anzahl an Lautsprechern pro Layer. Es handelt sich also um eine Anordnung mit vier elevierten Lautsprechern und sieben auf Ohrhöhe. Der optionale Tieftonlautsprecher für den LFE-Kanal wird nicht verwendet. Tabelle 4.1 listet die Lautsprecherpositionen mit ihren zugehörigen Kanälen nach ITU auf.

Die Positionierung der virtuellen Schallquellen erfolgt über *Vector Base Amplitude Panning*, kurz **VBAP**. Es handelt sich um eine Methode, die auf dem Grundprinzip der Stereophonie durch Pegeldifferenzen (siehe Abschnitt 4.1) aufbaut. Aufgrund der guten Lokalisation und einer geringen Anforderung an Rechenleistung stellt VBAP den verbreitetsten Panning-Algorithmus für mehrdimensionale Lautsprechersysteme dar (Franck & Fazi, 2019, S. 1) und ist beispielsweise auch im MPEG-H-Standard verankert (Herre, Hilpert, Kuntz & Plogsties, 2015). VBAP wurde grundsätzlich für eine frei wählbare Anzahl an Lautsprechern konzipiert, die sich im gleichen Abstand zum Hörer befinden. Immer drei der vorhandenen Lautsprecher bilden aneinandergrenzende virtuelle Dreiecke. Die zu positionierende Quelle befindet sich stets in einem dieser Dreiecke, was zur Folge hat, dass eben diese drei Lautsprecher das Signal wiedergeben. Je näher sich eine virtuelle Quellposition an der eines realen Lautsprechers befindet, desto höher ist der diesem Lautsprecher zugewiesene Verstärkungsfaktor, während die anderen beiden das Signal zunehmend leiser abstrahlen. Die Verstärkungsfaktoren vektorgeometrisch multipliziert mit den Richtungsvektoren, welche in Richtung der Lautsprecherposition zeigen, ergeben den resultierenden Richtungsvektor für die virtuelle Quelle. Befindet sich diese genau auf der Verbindungslinie zwischen zwei Lautsprechern, so wird der Verstärkungsfaktor für den dritten Lautsprecher zu Null. Deckt sich die virtuelle Position mit einer

**Tabelle 4.1:** Lautsprecherpositionen und Kanalbezeichnungen im 4+7+0 System. (ITU-R, 2018)

Lautsprecher	Kanal	$\phi$ [°]	$\delta$ [°]
<b>M+030</b>	Left (L)	+30 bis +45	0
<b>M-030</b>	Right (R)	-30 bis -45	0
<b>M+000</b>	Center (C)	0	0
<b>M+090</b>	Left side surround (Lss)	+85 bis +110	0
<b>M-090</b>	Right side surround (Rss)	-85 bis -110	0
<b>M+135</b>	Left rear surround (Lrs)	+120 bis +150	0
<b>M-135</b>	Right rear surround (Rrs)	-120 bis -150	0
<b>U+045</b>	Left top front (Ltf)	+30 bis +45	+30 bis +55
<b>U-045</b>	Right top front (Rtf)	-30 bis -45	+30 bis +55
<b>U+135</b>	Left top back (Ltr)	+100 bis +150	+30 bis +55
<b>U-135</b>	Right top back (Rtb)	-100 bis -150	+30 bis +55

Lautsprecherposition, so strahlt nur dieser das Signal ab. Dies hat den Vorteil, dass eine scharfe Lokalisation erreicht werden kann (Pulkki, 1997).

Nachteilig zeigt sich jedoch die Tatsache, dass aufgrund der unterschiedlichen Anzahl an aktiven Lautsprechern die wahrgenommene Schärfe der virtuellen Quelle variiert. Liegt die Position auf einem realen Lautsprecher, dann gestaltet sich die einzelne virtuelle Quelle scharf, während ihre Position in der Mitte eines Lautsprecherdreiecks verschwommener wahrgenommen wird. Je weiter die gleichzeitig abstrahlenden Lautsprecher auseinander liegen, desto größer werden diese Differenzen. Zusätzlich kommt es zu positionsabhängigen Klangverfärbungen (Pulkki, 1999). Aus diesem Grund gibt es mehrere Ansätze, wie beispielsweise von Franck und Fazi 2019 vorgestellt, zur Modifizierung von VBAP, sodass die Eigenschaften der dargestellten Quelle möglichst unabhängig von deren Positionierung konstant sind. Des Weiteren bleiben bei VBAP frequenzabhängige Effekte weitgehend unberücksichtigt, weswegen Algorithmen vorgeschlagen werden, welche unterschiedliche Frequenzbänder separat behandeln (siehe beispielsweise (Hughes, Franck, Cox, Shirley & Fazi, 2018)). Diese Entwicklungen werden im abschließenden Kapitel 7 nochmals aufgegriffen.

### 4.3 Mikrofonierung

Den Prinzipien der beiden Verfahren zur Stereophonie folgend haben sich in der audio-technischen Praxis zur räumlichen Darstellung akustischer Quellen mehrere Mikrofonierungskonzepte etabliert, welche im Folgenden vorgestellt werden. Diese bilden die Basis für das Mikrofonierungskonzept, welches als Ergebnis dieses Kapitels unter 4.7 vorgestellt wird. Die Anordnungen lassen sich in die drei Bereiche koinzidente Verfahren, Laufzeitverfahren und gemischte Verfahren gliedern.

Eine koinzidente Stereo-Mikrofonierung besteht aus zwei Mikrofonen, die je nach Schalleinfallswinkel idealerweise ausschließlich Pegelunterschiede zwischen den beiden Signalen erzeugen. Damit keine Laufzeitunterschiede entstehen, müssten sich beide Kapseln exakt am selben Ort befinden. Da bei der Stereowiedergabe meist ein Winkelbereich auf der Horizontalen räumlich abgebildet werden soll, sind die beiden Mikrofone im Regelfall so nah wie möglich übereinander angeordnet. So kommt es nur noch bei sehr hohen Frequenzen aufgrund des übrigen Kapselabstandes weniger Zentimeter zu relevanten Laufzeiten. Damit die gewünschten Pegeldifferenzen entstehen, müssen gerichtete Mikrofone verwendet werden. Klassischerweise kommt das XY-Verfahren mit zwei Mikrofonen zum Einsatz, welches direkt die Lautsprechersignale  $L = X$  und  $R = Y$  liefert. Die Mikrofonachsen sind dabei gekreuzt und liegen symmetrisch zur gedachten Mittelachse des Systems, welche auf das Zentrum des aufzuzeichnenden Klangkörpers ausgerichtet ist. Je größer der Winkel zwischen den beiden Mikrofonachsen, desto kleiner der horizontale Raumwinkel, der auf die volle Stereobreite abgebildet wird (Dickreiter et al., 2014, S. 251 f.). Eine Alternative zur XY-Mikrofonierung ist das MS-Verfahren. Hierbei kommt ein Mittenmikrofon und ein Seitenmikrofon mit Richtcharakteristik Acht zum Einsatz. Für das Mittenmikrofon wird meist ein Nieren- oder Kugelmikrofon verwendet. Aufgrund der Gegenphasigkeit der beiden empfindlichen Richtungen des Achtermikrofons kann durch Matrizierung der beiden Mikrofonensignale ein dem XY-Verfahren entsprechendes Signal für die Stereo-Lautsprecherwiedergabe produziert werden. Dabei gilt  $L = M + S$  und  $R = M + (-S)$ , wobei die phasenrichtige Seite des Achtermikrofons nach links zeigt. Dieses Verfahren hat den Vorteil, dass der zwischen den beiden Lautsprechern abzubildende Raumwinkel im Nachhinein durch die Pegelverhältnisse zwischen Mitten- und Seitensignal justiert werden kann (Rumsey, 2001, S. 161 ff.).

Die Lokalisationsunschärfe für koinzidente Verfahren gestaltet sich klein, das heißt die Richtungsbestimmung einer virtuellen Quelle im Stereobild ist relativ exakt möglich. Entsprechend der Erkenntnisse aus Abschnitt 2.1.2 ist jedoch zu erwarten, dass dies nur für Signale mit genügend Anteilen oberhalb von ca. 1 kHz gilt, da die Lokalisation deutlich tieferer Frequenzen über Zeitversätze funktioniert.

Das Stereobild von Laufzeitverfahren weist dagegen eine größere Lokalisationsunschärfe auf, sodass die virtuelle Quellrichtung auf der Lautsprecherbasis weniger scharf bestimmt werden kann. Allerdings bieten Laufzeitverfahren ein räumlicheres Bild der aufgezeichneten auditiven Szene (Dickreiter et al., 2014, S. 250). Reine Laufzeitverfahren weisen idealerweise keine Pegelunterschiede zwischen den beiden Kanälen auf und bestehen daher aus zwei ungerichteten Mikrofonen (Kugelcharakteristik) mit ausreichend Abstand zueinander, welcher für die gewünschte Laufzeitdifferenz sorgt. Je größer dieser horizontale Abstand zwischen den Kapseln, desto kleiner der Raumwinkel, welcher auf der vollen Stereobreite abgebildet wird. Man spricht von AB-Verfahren (Dickreiter et al., 2014, S. 256 f.).

Alle Anordnungen, welche sowohl Abstände zwischen den Kapseln aufweisen, als auch gerichtete Mikrofone mit Winkelversatz zueinander verwenden, werden als gemischte Verfahren bezeichnet. Beispielhaft sei an dieser Stelle die ORTF-Mikrofonierung, mit zwei Nierenkapseln im Abstand von 17 cm und Öffnungswinkel  $110^\circ$  zueinander, genannt (Rumsey, 2001, S. 168).

Für Surround-Anordnungen existiert ebenfalls eine Vielzahl an Mikrofonierungsverfahren unter Verwendung von Pegel- oder Laufzeitdifferenzen und Mischformen daraus. Beispielhaft für koinzidente Anordnungen kann das Doppel-MS Verfahren angeführt werden, welches eine Stereo-MS Anordnung mit Nieren-Mittensignal um ein zweites Nierenmikrofon in entgegengesetzter Richtung erweitert. Dieses liefert durch Matrizierung mit dem Seiten Signal die Surroundkanäle. Als Laufzeitverfahren kann das Decca-Tree mit Kugelkapseln für die drei Front-Kanäle L, C und R angeführt werden. Die meisten Anordnungen arbeiten jedoch mit einer Mischung aus beiden Prinzipien (Friesecke, 2014, S. 451). Beispielhaft sei an dieser Stelle das dreikanalige OCT-Verfahren genannt, welches aus zwei in entgegengesetzte Richtung zeigenden Supernieren für die Kanäle L und R sowie einer um 8 cm nach vorne versetzten Niere für den Center-Kanal besteht. Der Abstand zwischen den beiden Supernieren kann zwischen 40 und 80 cm, je nach Breite der aufzuzeichnenden Quelle, justiert werden. Zwei zusätzliche Nierenkapseln, welche einen größeren horizontalen Abstand zueinander haben und 40 cm nach hinten versetzt positioniert werden, liefern die Signale für die Surroundkanäle SCHOEPS (2020c).

Sollen nun für die Wiedergabe über ein mehrdimensionales System, mit Lautsprechern im Upper Layer, Signale für diesen generiert werden, so kommen wiederum zusätzliche Mikrofone zum Einsatz. Dabei kann in die zwei Kategorien von Systemen mit und ohne Abstand zwischen Middle und Upper Layer eingeteilt werden. Das beschriebene OCT System für Surround wird beispielsweise um vier senkrecht nach oben gerichtete Supernieren erweitert, welche sich koinzident an der selben Position befinden, wie die vier

Mikrofone für L und R, bzw. LS und RS. Damit beruht die Anordnung in der vertikalen Ebene rein auf Pegeldifferenzen. Alternativ können die vier **Höhenmikrofone** auch bis zu 100 cm nach oben abgesetzt werden, sodass zusätzliche Laufzeitdifferenzen auftreten (Theile & Wittek, 2011). Geluso führte 2012 die Idee eines MS-3D-Verfahrens ein, welches Positionen einer Surround-Anordnung ein vertikal ausgerichtetes Mikrofon mit Achtercharakteristik koinzident hinzufügt, sodass Signale von oben und unten aufgezeichnet werden. Ein von Lee und Gribben durchgeführter Hörtest zum Vergleich verschiedener Hauptmikrofone mit Höhenmikrofonen kam zu dem Ergebnis, dass ein Abstand zwischen den beiden Mikrofonebenen nicht zu einem verbesserten räumlichen Eindruck führt. Für transiente Signale präferierten die Probanden die in vertikaler Richtung koinzidente Anordnung (Lee & Gribben, 2015).

In diesem Abschnitt wurden verschiedene Mikrofonanordnungen zur räumlichen Darstellung einer auditiven Szene vorgestellt. Auch komplexere Systeme für mehrdimensionale Lautsprecheranordnungen basieren auf dem Grundprinzip der Stereophonie über Pegel- oder Laufzeitdifferenzen zwischen den Mikrofonen. Abschnitt 4.5 wird nochmals auf diese Anordnungen eingehen und deren Verwendung bei der Aufzeichnung realer Ensembles herausstellen, sowie die Weiterverarbeitung im Mischprozess erläutern. Zuvor wird untersucht, welche Effekte bei der Wahrnehmung virtueller Quellen in einem mehrdimensionalen Wiedergabesystem bezüglich Lokalisation und Ausdehnung auftreten und inwieweit diese das zu konzipierende Stützmikrofonverfahren beeinflussen.

## 4.4 Lokalisation und Ausdehnung in mehrdimensionalen Lautsprechersystemen

Die wichtigsten Erkenntnisse aus Kapitel 2 bezüglich der Lokalisierung von Schallquellen können zu folgenden Kernpunkten zusammengefasst werden: In horizontaler Richtung zeigen sich für tiefe Frequenzen ITD und für hohe Frequenzen ILD verantwortlich, während in vertikaler Richtung spektrale Merkmale dominieren. Dabei besteht das Phänomen des Pitch-Height-Effekts, wonach Signale im Frequenzbereich deutlich unterhalb 1 kHz nicht eleviert lokalisiert werden können, während andere höhere Frequenzbereiche, trotz tatsächlichem Ursprungsort auf Ohrhöhe, eleviert wahrgenommen werden. Weiterführend ist zu bemerken, dass eine virtuelle Quelle, die als horizontale Phantomschallquelle durch Wiedergabe über zwei Lautsprecher erzeugt wurde, ebenfalls eleviert wahrgenommen wird. Die Höhe der Elevation hängt von Lautsprecherabstand und Spektrum des Signals ab. Transiente oder breitbandige Signale werden höher wahrgenommen als tieffrequente, konstante (Lee, 2017). Somit ist die Wahrnehmung der

Elevation frequenzabhängig (Mironovs & Lee, 2016) und entspricht, vorallem bei Dominanz tieffrequenter Anteile, nicht der tatsächlichen Quellposition. Dies gilt insbesondere für virtuelle Quellen auf der Medianebene über und hinter dem Kopf, wegen der nicht vorhandenen ILD und ITD (Mironovs & Lee, 2017; Mironovs & Lee, 2018).

Abschnitt 4.1 hat dargelegt, dass die Positionierung von Phantomschallquellen in der horizontalen Ebene zwischen zwei Lautsprechern durch Pegel- oder Zeitdifferenzen zweier korrelierter Signale entsteht. Aufgrund fehlender ILD und ITD zwischen zwei Quellen mit gleichem Azimutwinkel und unterschiedlicher Elevation bilden sich demnach keine stabilen Phantomquellen in vertikaler Richtung aus (Barbour, 2003). Die Studien von Lee aus 2011 sowie Wallis und Lee aus 2016 haben zum Ergebnis, dass auch der Präzedenz-Effekt zwischen zwei Lautsprechern in vertikaler Richtung nicht anwendbar ist. Tregonning und Martin kamen 2015 zu einem gegenteiligen Ergebnis, nach welchem ein vertikaler Präzedenz-Effekt vorhanden und mit 5 ms Verzögerung des Lautsprechersignales auf Ohrhöhe die maximale Elevation der wahrgenommenen Quelle erreicht wird.

In Abschnitt 2.3 wurde gezeigt, dass die Wahrnehmung von großer Ausdehnung einer Schallquelle dann erzielt wird, wenn der IAC-Koeffizient klein ist. Das bedeutet, dass dekorrelierte, ähnliche Signale aus mehreren Richtungen am Ohr eintreffen. Zusätzliche Eigenschaften, wie hohe Obertondichte eines Instrumentes, hoher Schalldruckpegel sowie Bekanntheit der realen Ausdehnung der wahrgenommenen Quelle, spielen ebenfalls eine Rolle. Soll ein beliebiges Monosignal eine wahrgenommene Ausdehnung erhalten, so muss dieses Signal demnach dekorreliert werden. Mit Dekorrelation ist der Prozess gemeint, aus dem Ursprungssignal zwei oder mehrere ähnliche Signale zu erstellen, die sich aber in ihrer Wellenform unterscheiden. Diese werden dann von unterschiedlichen Lautsprechern wiedergegeben und sorgen somit künstlich für ein ausgedehntes akustisches Bild durch die reduzierte **ICC** (*Interchannel Cross-Correlation*), welche wiederum einen kleinen IAC-Koeffizienten zur Folge hat. Die scharfe, schmale Phantomschallquelle auf der Stereobasis wird dadurch unscharf und breit (Kendall, 1995). Das Prinzip der horizontalen Signaldekorrelation findet sich bereits seit vielen Jahrzehnten in der Aufnahme- und Mischtechnik von Stereo-Musikproduktionen der Populärmusik wieder und ist unter dem Begriff *Doubling* bekannt. Beispielsweise werden doppelt aufgenommene Gesangs- oder Gitarrenspuren links und rechts in der Stereomischung platziert, damit ein breites auditives Bild entsteht (Schneider, 2017, S. 310 f.).

Es existiert eine Vielzahl an Studien, die sich mit Algorithmen zur Synthetisierung dekorrelierter Signale zur Verbreiterung der wahrgenommenen virtuellen Quelle beschäftigen. Das Grundprinzip beruht meist auf der frequenzmäßigen Fensterung des Ursprungssignals. In den Versuchen von [Boueri und Kyriakakis \(2004\)](#) wird das Signal auf beiden Kanälen durch Rechteckfilterung in mehrere Frequenzbänder unterteilt, welche jeweils um einen zufälligen Wert zwischen  $-20$  ms und  $+20$  ms zeitlich verschoben werden. Die Einführung eines Zeitversatzes im breitbandigen Signal hätte gemäß Präzedenz-Effekt eine Verschiebung des gesamten Signals auf den früher wiedergebenden Lautsprecher zur Folge. Durch die Filterung und zufällige Verschiebung ergibt sich jedoch der gewünschte Dekorrelationseffekt und es konnte der Verlust der scharfen Phantomquelle aufgrund des deutlich reduzierten IAC-Koeffizienten festgestellt werden. Allerdings bringt diese Methode leichte Klangverfärbungen durch Interferenzeffekte, wegen der eingeführten frequenzabhängigen Phasenverschiebung, mit sich ([Boueri & Kyriakakis, 2004](#)). [Schmele und Sayin](#) kommen bei ihren Versuchen 2018 mit Wiedergabe über ein Ambisonics-System zu ähnlichen Ergebnissen. [Gribben und Lee](#) untersuchen in ihren Studien aus 2017, 2019 und 2020 die Verbreiterung virtueller Quellen speziell in vertikaler Richtung, durch Verringerung der ICC, indem komplementäre Kammfilterung bzw. Phasenrandomisierung Anwendung finden. Sie kommen zu dem Ergebnis, dass eine Dekorrelation erst ab 500 Hz aufwärts empfehlenswert ist, da die Klangverfärbungen im Vergleich zum Ursprungssignal geringer ausfallen und die wahrgenommene Ausdehnung in vertikaler Richtung trotzdem gleich groß wie bei breitbandiger Dekorrelation ausfällt. Als Voraussetzung für die Verbreiterung des Bildes wurden abermals hohe Frequenzbereiche (konkret das 8 kHz-Oktavband) identifiziert ([Gribben & Lee, 2019](#)), sodass die Wirkung des Prinzips abhängig vom Energiegehalt der darzustellenden Quelle in diesem Frequenzband ist. Es wurde angeführt, dass beispielsweise die vertikale Ausdehnung eines auf beschriebener Weise dekorrelierten Schlagzeugsignals im Probandentest gute Ergebnisse erzielte ([Gribben & Lee, 2020](#)). Es ist jedoch anzumerken, dass andere Tests von [Gribben und Lee](#) in 2014 bereits gezeigt hatten, dass die Unterschiede in Koeffizienten der vertikalen ICC deutlich geringe Varianzen in der wahrgenommenen Ausdehnung nach sich ziehen, als entsprechende Unterschiede in der horizontalen Ebene.

Die Studie von [Lee](#) aus 2016 belässt den zeitlichen Verlauf des Signals unverändert. Stattdessen soll eine Verbreiterung in vertikaler Richtung erfolgen, indem die einzelnen Oktavbänder des Signals entweder einem horizontalen Lautsprecherpaar auf Ohrhöhe, oder einem um  $30^\circ$  elevierten Paar zugeordnet wird. Damit bleibt neben dem Phasenverlauf auch der Amplitudenverlauf des Signalspektrums im Gesamten gleich und es treten keine Klangverfärbungen durch Kammfilter auf. Das Verfahren wird als *Perceptual Band Allocation* bezeichnet. Im ersten Test wurde Rauschen in den einzelnen

Oktavbändern zufällig vom Stereo-Lautsprecherpaar im Upper oder Middle Layer abge- spielt. Dabei wurde das dem bereits beschriebenen Pitch-Height-Effekt entsprechende Ergebnis erzielt, dass die wahrgenommene Elevation stärker vom Frequenzband als von der tatsächlichen Wiedergabeposition abhängt und grundsätzlich die Tendenz gilt, dass höhere Frequenzbereiche höher lokalisiert werden. Eine Studie von [Kamekawa und Marui](#) aus 2020 kommt entsprechend zum Ergebnis, dass der Upper Layer eines 22.2 Systems unterhalb von 400 Hz hochpassgefiltert werden kann, ohne dass im räumlichen Eindruck ein Unterschied zur breitbandigen Wiedergabe aller Lautsprecher hörbar war. Im zweiten Test von [Lee](#) wurden verschiedene Varianten der Verteilung der Frequenzbänder auf die beiden Lautsprecherebenen untersucht. Die größte wahrgenommene vertikale Ausdehnung erzielten dabei zwei Varianten. Bei Variante 1 wurden die Frequenzbänder unterhalb 2 kHz nur auf Ohrhöhe und die Bänder oberhalb ausschließlich auf dem Upper Layer platziert. Variante 2 bildete die Wiedergabe des kompletten breitbandigen Signals über den Upper Layer ([Lee, 2016](#)). Allerdings muss angemerkt werden, dass die Kopfposition der Probanden fixiert war, was nicht der gewöhnlichen Hörgewohnheit eines Rezipienten einer mehrdimensionalen Musikmischung entspricht. Des Weiteren wurden die virtuellen Positionen durch horizontale Phantomschallquellen produziert, welche prinzipiell eleviert erscheinen.

Das Erkennen bestimmter ausgedehnter Formen von virtuellen Schallquellen zeigt sich als eher schwierig. Wenn, dann wurden frontal abgebildete Formen richtig erkannt. Dabei musste das Signal genügend hochfrequente Anteile aufweisen ([Schneider, 2017, S. 302](#)). Es kann festgehalten werden, dass hauptsächlich die Ränder der Ausdehnung einer virtuellen Schallquelle lokalisierbar sind und diese maßgeblich von ILD bestimmt werden ([Santala & Pulkki, 2011, S. 1530](#)).

## 4.5 Konventionelles Mikrofonierungs- und Mischkonzept

Im Verlauf des Kapitels wurden verschiedentliche Varianten gängiger Mikrofonanordnungen besprochen, welche für die Abbildung räumlicher auditiver Szenen verwendet werden. Des Weiteren wurden diverse Effekte bezüglich der Wahrnehmung solcher, über ein mehrdimensionales Lautsprechersystem wiedergegebener Szenen beleuchtet. Im Folgenden werden gängige Verfahren zum Erstellen von konventionellen mehrdimensionalen Musikmischungen beschrieben und die sich daraus ableitenden Anforderungen an das zu konzipierende Stützmikrofonverfahren herausgestellt.

Die beschriebenen Mikrofonverfahren finden auf zwei verschiedene Arten bei der Aufzeichnung eines Ensembles Anwendung. Eine im Fernfeld platzierte Anordnung dient zur Abbildung der Position einzelner Instrumente oder Instrumentengruppen sowie zur Aufzeichnung des Raumhalls. Anordnungen im Nahfeld eines Instrumentes sollen dessen Einzelsignal, teilweise mit Ausdehnung aufzeichnen (Schneider, 2017, S. 309). Man spricht vom Hauptmikrofon (fern) und den einzelnen Stützmikrofonen (nah). Dieses Konzept kommt hauptsächlich bei der Aufzeichnung von Konzerten klassischer Ensembles zum Tragen. In diesem Fall soll meist das auditive Bild, welches dem Besucher der Veranstaltung dargeboten wird, eingefangen und dem Hörer der Aufzeichnung möglichst ähnlich dargeboten werden. Die Hauptmikrofone sind daher darauf ausgelegt, den Klangeindruck und das räumliche Bild des Ensembles von der Bühne, im Falle von Surround- oder mehrdimensionalen Anordnungen, aber auch den Raumhall im Speziellen einzufangen (Rumsey, 2001, S. 154 f.). Die Höhenmikrofone bei einer Orchesterproduktion sind daher meist für die Aufzeichnung von Deckenreflexionen bzw. generellem Raumhall konzipiert. Die virtuelle Bühne wird meist zwischen den Lautsprechern M+030 und M-030 platziert. Der Aufnahmewinkel der vorderen Hauptmikrofone in der unteren Ebene mehrdimensionaler Anordnungen werden entsprechend gewählt. Die Stützmikrofone vor Einzelinstrumenten bzw. Instrumentengruppen dienen hauptsächlich dafür, zum Signal des Hauptmikrofons beigemischt zu werden. Dabei können Lautstärke- und Klangbalancen, sowie Entfernungseindrücke zwischen den Instrumenten angepasst werden (Weinzierl, 2008, S. 603). Aufgrund der großen Dimensionen eines solchen Ensembles und der Entfernung des Zuhörers, ist die Ausdehnung eines einzelnen Instrumentes kaum wahrnehmbar. Deshalb kommt meist ein einzelnes Mono-Stützmikrofon zum Einsatz.

Bei der Aufzeichnung kleinerer Ensembles, insbesondere im Kontext von Jazz- oder Pop-Studioproduktionen in Stereo, wird in der Regel auf ein Hauptmikrofon verzichtet und ein einzelnes Instrument mit nahen Mikrofonanordnungen aufgezeichnet. Diese sollen verschiedene Klangbestandteile separat einfangen - beispielsweise Korpusklang und Griffgeräusche bei einer Akustikgitarre. Ebenso üblich ist die Verwendung eines Stereoverfahrens, wie in Abschnitt 4.3 beschrieben. Durch entsprechende Positionierung im Panorama entsteht schließlich der Eindruck einer ausgedehnten Quelle. Die Mischung erscheint dadurch breiter und voller. Dabei geht es im Gegensatz zur Orchesteraufzeichnung weniger um die korrekte Darstellung des realen auditiven Bildes (Schneider, 2017, S. 310). Mehrdimensionale Mischungen bieten eine deutlich vergrößerte Fläche, auf welcher die Platzierung virtueller Schallquellen denkbar ist. Insbesondere in vertikaler Richtung kann diese genutzt werden, um einem einzelnen Instrument nicht nur wahrgenommene Breite, sondern auch Höhe zu verleihen. Das zu konzipierende mehrdimensionale Stützmikrofonverfahren soll Signale generieren, welche dies möglich machen.

Dabei soll die Anordnung sowohl zur direkten und alleinigen Abnahme eines Instrumentes in Studioproduktionen oder Live-Mitschnitten, als auch in Kombination mit einem Hauptmikrofon Anwendung finden. In beiden Fällen passiert die Mikrofonierung nah am Instrument, weswegen innerhalb der vorliegenden Arbeit der geläufige Terminus *Stützmikrofon* verwendet wird.

## 4.6 Kompatibilität und Flexibilität

Neben den in Abschnitt 4.5 vorgestellten klanglichen und räumlichen Anforderungen an das zu konzipierende Stützmikrofonverfahren, sind noch weitere audioteknische Eigenschaften zu betrachten, welche mit Kompatibilität und Flexibilität überschrieben wurden. Gemeint ist damit, dass der Aufbau der Anordnung im praktischen Anwendungsfall möglichst einfach zu bewerkstelligen ist sowie keine Unannehmlichkeiten für den aufnehmenden Musiker darstellt. Das heißt konkret, dass die Anzahl an Mikrofonen und deren räumliches Ausmaß begrenzt sein sollte. Ebenso ist ein möglichst universell einsetzbares System gewünscht, dessen Funktion beispielsweise nicht auf ein spezifisches Instrument begrenzt ist.

Weiterhin soll das von der Mikrofonanordnung aufgezeichnete Signal möglichst vielseitig und systemübergreifend anwendbar sein sowie eine intuitive Modifikation der Ausdehnung der aufgezeichneten Quelle im Mischprozess ermöglichen. Ausschlaggebend ist außerdem eine gute Abwärtskompatibilität. Wenn beispielsweise aus einer mehrdimensionalen Mischung mit Lautsprechern im Upper Layer eine Surround-Mischung abgeleitet wird, passiert dies oft durch Addition der Signale für Lautsprecher mit ähnlichem Azimutwinkel, also beispielsweise  $L_{0+7+0} = L_{4+7+0} + L_{4+7+0}$ . Dadurch ergeben sich für Mikrofonierungsverfahren mit Laufzeitdifferenzen Auslöschungen in Form von Kammfilterstrukturen und daraus resultierend ein verzerrtes Klangbild. Koinzidente Verfahren betrifft dieses Problem entsprechend nicht (Dickreiter et al., 2014, S. 277 f.).

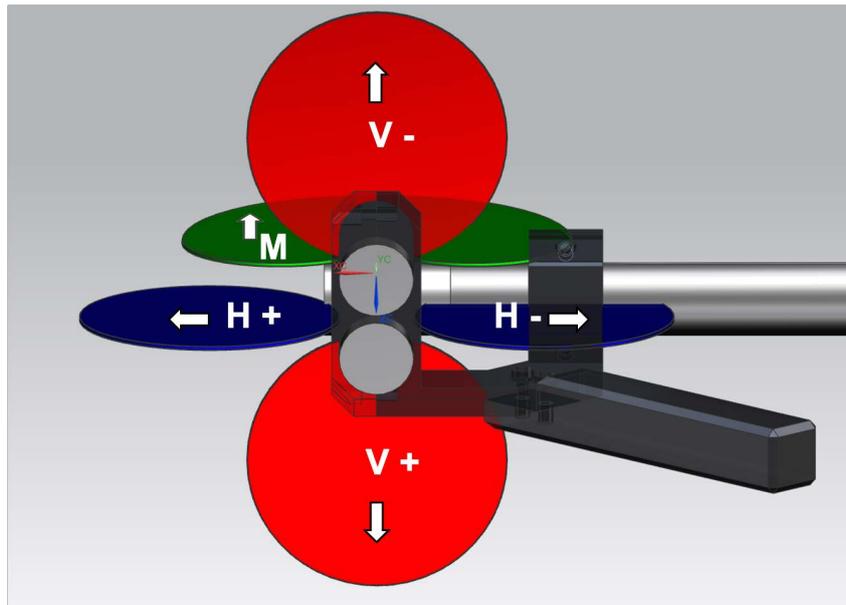
## 4.7 Konzipierte Mikrofonanordnung MHV

Insbesondere die im letzten Abschnitt 4.6 gestellten Anforderungen an das Stützmikrofonverfahren führten zu dem Entschluss, eine koinzidente Anordnung basierend auf dem MS-Prinzip zu verwenden, welche eine hohe Flexibilität im Mischprozess verspricht. Durch den Ausschluss von Laufzeitdifferenzen ist außerdem verzerrungsfreie Abwärtskompatibilität garantiert. Es wurde die in Abschnitt 4.7 vorgestellte Idee von Geluso aufgegriffen, ein vertikal ausgerichtetes Achtermikrofon einzuführen, welches die Ausdehnung des Instrumentes in dieser Richtung erfasst. Insgesamt finden drei Mikrofone

Anwendung: Ein Mittensignal in Form eines auf das Instrument ausgerichteten Nierenmikrofons, sowie ein horizontales und ein vertikales Signal von jeweils einem Achtermikrofon. Abbildung 4.2 zeigt schematisch den konzipierten Aufbau aus der Rückansicht, also mit Blickrichtung zum Instrument. Die Anordnung wird **MHV** für die drei Komponenten Mitte, Horizontal und Vertikal benannt. Die vielleicht offensichtlichere Bezeichnung *Doppel-MS* wird bereits von der unter Abschnitt 4.3 beschriebenen Anordnung für Surround-Signale verwendet, eine ähnliche Problematik besteht beim Begriff *MS-3D*. Damit die Kapseln so dicht wie möglich aneinander platziert sind, wurde eine Anordnung gewählt, bei der M- und H-Mikrofon übereinander platziert werden und das V-Mikrofon orthogonal dazu mit der Kapsel vor das H-Mikrofon, und damit auf dessen unempfindlicher Seite, platziert wird (siehe dazu auch Abschnitt 5.1). Im Mischprozess werden die drei Mikrofonsignale zu zwei Stereo-Bussen LR (left/right) und BT (bottom/top) dekodiert, wobei  $L = M + H$ ;  $R = M + (-H)$ ,  $B = M + V$  und  $T = M + (-V)$  gilt. Das negative Vorzeichen entspricht dem invertierten Mikrofonsignal. Durch Positionierung der beiden Signalpaare im beispielhaften 4+0+7 Lautsprechersystem entsteht eine mehrdimensionale Phantomschallquelle mit variabler absoluter Breite und Höhe. Durch Variieren des jeweiligen Achtermikrofon-Pegels kann im Nachhinein der Aufnahmewinkel justiert werden (siehe Abschnitt 4.3).

Abschnitt 4.4 hat gezeigt, dass für eine große wahrgenommene Ausdehnung mehrere dekorrelierte Signale vorliegen müssen. Dementsprechend wird die MHV-Anordnung relativ nah am Instrument platziert, um verschiedene schallabstrahlenden Teile desselben mit unterschiedlichem Pegel in unterschiedlichen Mikrofonen aufzuzeichnen. Zusätzlich sollen die in Kapitel 3 erläuterten dynamischen Effekte durch Bewegungen der Musiker in Kombination mit den komplexen Abstrahlcharakteristika der Instrumente eingefangen werden. Wird beispielsweise der Kontrabass gestrichen, wandert der Quellort für hohe Frequenzen des Streich-Geräusches horizontal hin- und her (siehe Abbildung 3.6). Der Korpus (unten) strahlt tiefere Frequenzen ab und am Griffbrett (oben) entstehen schnarrende Saitengeräusche. Zusätzlich bewegt sich der Musiker und damit auch das Instrument. Dadurch gelangen spektral dynamisch variierende Signale an die unterschiedlichen Kapseln der MHV-Anordnung. Aus diesem Grund wird ein Effekt erwartet, der in etwa der unter 4.4 beschriebenen Dekorrelation durch Perceptual Band Allocation entspricht.

Abschnitt 4.2 hat unter anderem die Nachteile des VBAP Panning-Verfahrens dargestellt: Die Ausdehnung von virtuellen Quellen variiert je nach Positionierung dieser, da sie entweder von nur einem, zwei oder drei aktiven Lautsprechern als Phantomschallquelle dargestellt wird. Werden die aus MHV dekodierten Signale LR und BT jeweils zwischen zwei Lautsprechern positioniert, so entsteht die wahrgenommene Mitte der virtuellen Quelle praktisch aus drei Stufen der Phantomschallquellenbildung (VBAP, horizontale



**Abbildung 4.2:** Illustration MHV-Anordnung mit Richtcharakteristika.

Phantommitte, vertikale Phantommitte). Es ist folglich davon auszugehen, dass die entstehenden Phantomschallquellen je nach Positionierung in ihrer Ausdehnung variieren und eine gewisse Lokalisationsunschärfe aufweisen werden. Nach Abschnitt 4.4 kann die Ausdehnung einer Schallquelle insbesondere in vertikaler Richtung jedoch ohnehin nicht scharf lokalisiert werden.

## 5 Testaufnahmen zur Untersuchung der MHV

Zur Evaluation des konzipierten MHV-Verfahrens wurden Testaufnahmen mit verschiedenen Instrumenten durchgeführt. Dabei sind bis zu drei mögliche Positionen der Anordnung für jedes Instrument individuell bestimmt und aufgezeichnet worden. Einen wichtigen Aspekt der Untersuchungen stellen die Bewegungen der Musiker dar. Daher wurden die Aufnahmen von zwei Kameras begleitet, welche die Musiker in frontaler und seitlicher Einstellung aufzeichneten. Da die MHV-Anordnung sowohl als eigenständiges Mikrofonierungsverfahren, als auch in Kombination mit einem Hauptmikrofon getestet werden soll, wurde ein solches zusätzlich aufgezeichnet (siehe 5.6). Allerdings liegt der Hauptaugenmerk dieser Arbeit auf der eigenständigen Anwendung der MHV-Anordnung. Die Positionen und Winkel der Mikrofonanordnungen werden dokumentiert und durch Fotoaufnahmen illustriert, sodass schließlich eine übersichtliche Datenbank zu weiterführenden Forschungs- und Demonstrationszwecken entsteht. Die Bereitstellung der Daten wird in Abschnitt 5.7 beschrieben.

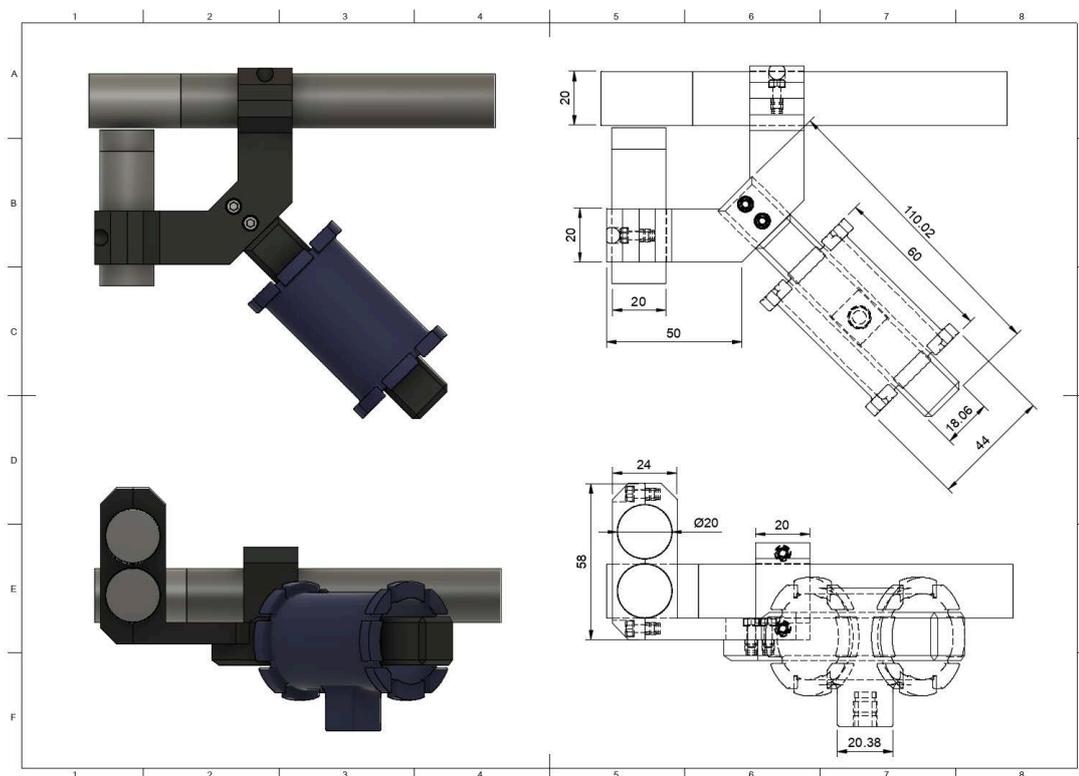
### 5.1 Vorversuche

Um Erfahrungen bezüglich Positionierung der konzipierten Mikrofonanordnung sowie erste Höreindrücke zu sammeln, wurden in Vorbereitung auf die eigentlichen Testaufnahmen Vorversuche mit akustischer Gitarre und Gesangsstimme im großen Aufnahmerraum Studio D der Tonstudios U54 der Hochschule der Medien Stuttgart durchgeführt. Die aufgezeichneten Signale wurden in Regie U48 in zwei Stereosignale dekodiert und zum Abhören mittig zwischen den vier Lautsprechern M+030, M-030, U+030 und U-030 positioniert (siehe Abbildung 5.20). Die Vorversuche hatten folgende Erkenntnisse zum Ergebnis:

Damit genügend Anteile von den beiden Mikrofonen H und V aufgezeichnet werden, sodass eine wahrgenommene Ausdehnung der Gitarre entsteht, muss die MHV näher am Instrument als zuerst angenommen platziert werden. Dabei ist in horizontaler Richtung eine exakte Positionierung sehr wichtig, da ansonsten der wahrgenommene Schwerpunkt des ausgedehnten Instrumentes immer weiter in eine Richtung wandert, wenn man die Ausdehnung durch Höherpegeln der Achten vergrößert. Eine Positionierung frontal vor dem Schallloch würde dieses Kriterium erfüllen, bringt allerdings Schwierigkeiten im Klangbild durch Überbetonung der Hohlraumresonanz mit sich. Die Gesangsstim-



**Abbildung 5.1:** Mikrofonanordnung MHV auf Stereoschiene für Vorversuche.



**Abbildung 5.2:** Konstruktionszeichnung und Ansichten der MHV-Halterung für 3D-Druck.

me liefert die unter Abschnitt 3.6 beschriebenen dynamischen Effekte. Auch hier muss sehr nahe mikrofoniert werden. Dadurch kommt es bei frontaler Mikrofonierung zu Pop-Geräuschen, welche mit einem entsprechenden Pop-Schutz eliminiert werden. Es wurde weiterhin festgestellt, dass die koinzidente Positionierung der drei Kapseln so nah wie möglich beieinander am besten mit Kleinmembranmikrofonen zu realisieren ist. Dennoch müssen Kompromisse eingegangen werden, da zwangsweise bei der Zusammensetzung aus drei einzelnen Mikrofonen die Verdeckung kleiner Bereiche nicht komplett verhindert werden kann. Die in Abbildung 5.1 gezeigte Anordnung mit vertikaler Acht vor horizontaler Acht wurde schließlich definiert. Die auf dem Bild dargestellte Installation der Mikrofone auf einem Stativ mit Hilfe einer Stereoschiene wurde als störanfällig gegenüber Verschiebung und Verdrehung eingestuft. Die exakte Rekonstruktion beispielsweise nach Wechsel des Mikrofonstatives zwischen zwei unterschiedlich hohen Instrumenten würde sich schwierig gestalten. Aus diesem Grund wurde eine individuell konstruierte Halterung, welche die Mikrofone über den gesamten Zeitraum der Testaufnahmen fixiert, in Auftrag gegeben. Abbildung 5.2 zeigt die Konstruktionszeichnung der Halterung, welche von einer Mikrofonspinne zur besseren Entkopplung von Körperschall gehalten wird.

## 5.2 Arrangiertes Musikstück

Ein wichtiges Ziel der vorliegenden Arbeit war, dass das erzeugte Audiomaterial veröffentlicht und für weitere Forschung verwendet werden kann. Damit weiterführend neben der Begutachtung der Aufnahmen einzelner Instrumente auch eine mehrdimensionale Testmischung entstehen konnte (siehe Abschnitt 6), wurde ein Arrangement eines Titels erstellt, welches möglichst viele und typische Facetten der einzelnen Instrumente abdeckt, um etwaige Auswirkungen auf die ausgedehnte Darstellung analysieren zu können. Dabei wurden Tonumfang, sowie verschiedene Dynamiken und Spieltechniken berücksichtigt. Als Stück wurde das Volkslied *Scarborough Fair* ausgewählt. Die Partitur des erstellten Arrangements kann im Anhang C eingesehen werden. Die Aufnahmen entstanden zwischen dem 05. und 17. August 2020, die Spielzeit der Spuren beläuft sich auf 04:09 min.

## 5.3 Aufnahmeraum

Damit sich die spätere Analyse der Testaufnahmen ausschließlich auf die Abbildung der Instrumente konzentrieren kann, sollten alle Störfaktoren ausgeschlossen werden. Sowohl Übersprechen von anderen Instrumenten, als auch hörbare Raumreflexionen und Nachhall hätten das wahrgenommene Bild des einzelnen Instrumentes verzerrt. Daher fand die

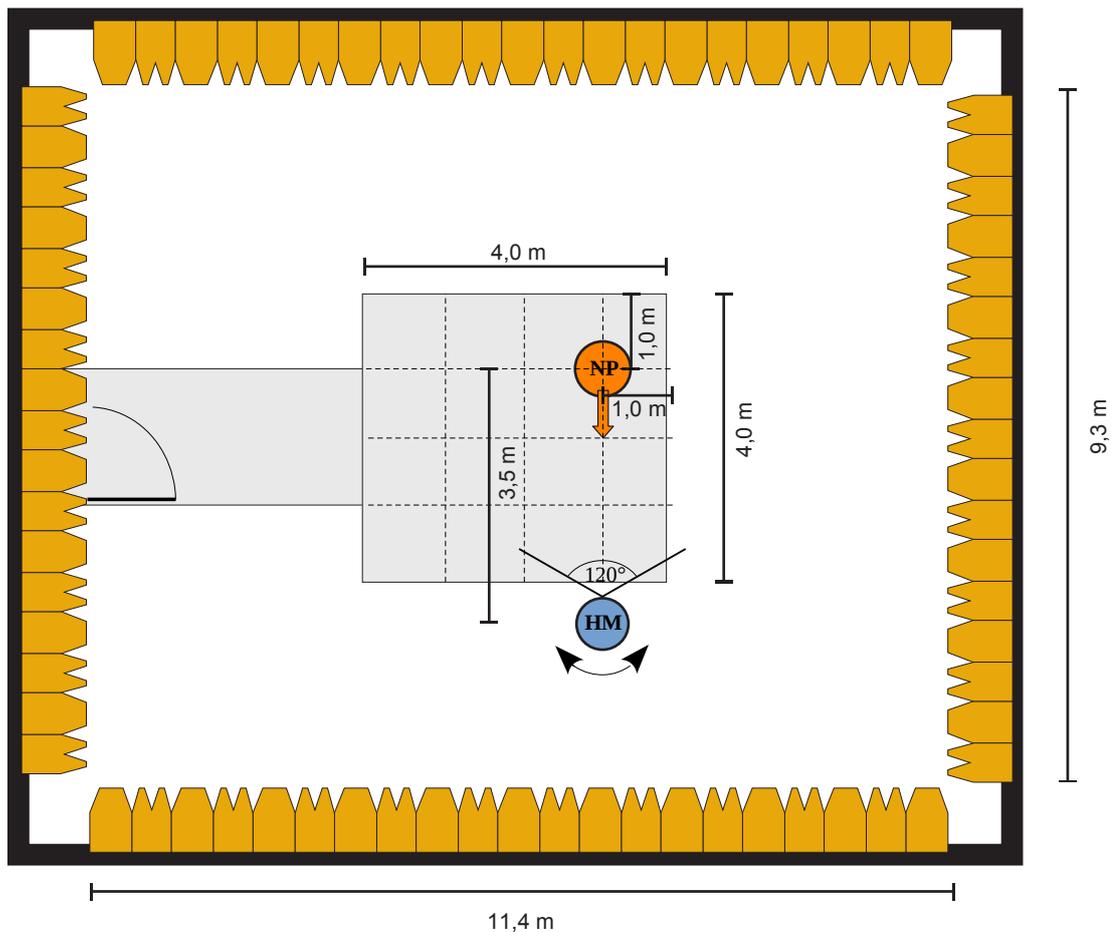


**Abbildung 5.3:** Freifeldraum P22 des Fraunhofer-Instituts für Bauphysik IBP.

Aufnahme der Instrumente nicht als Gruppe, sondern einzeln statt und wurde zudem nicht in der natürlichen Umgebung eines Vortragssaales, sondern im reflexionsarmen Freifeldraum P22 des *Fraunhofer-Instituts für Bauphysik IBP* in Stuttgart durchgeführt (siehe [Abbildung 5.3](#)). [Anhang B.1](#) enthält weitere technische Daten des Raumes. In [Abbildung 5.4](#) ist der Grundriss aufgetragen und der Punkt markiert, welcher für jedes Instrument die *Nullposition* darstellt. Die Instrumente waren wie eingezeichnet stets in Richtung Hauptmikrofon ausgerichtet. Die Positionen der MHV sind in [Abschnitt 5.6](#) dokumentiert. Zur Minimierung von Reflexionen des Gitterbodens wurde dieser auf den nicht verwendeten Flächen ausgebaut und aus dem Raum entfernt. An den benötigten Stellen nahe der Mikrofone, als Untergrund für die Stative, wurden um diese herum ca. 10 cm dicke Caruso-Absorber ( $20 \text{ kg/m}^3$ ) platziert (siehe [Abbildung 5.3](#)). Der Notenständer wurde ebenfalls mit einem solchen präpariert.

## 5.4 Aufnahmeverfahren

Neben den unter [5.3](#) erläuterten Vorteilen, die sich bei der Einzelspur-Aufnahme in der wissenschaftlichen Umgebung eines reflexionsarmen Raumes ergeben, müssen auch die Nachteile beachtet und Gegenmaßnahmen ergriffen werden. Da die aufgenommenen Spu-



**Abbildung 5.4:** Grundriss des Aufnahmerraums mit Nullposition der Instrumente (NP, orange) sowie Position des Hauptmikrofons (HM, blau).

ren im Nachgang zu einer testweisen Gesamtmischung zusammengesetzt werden (siehe Kapitel 6), ist neben technischer Sauberkeit des Signales und für das jeweilige Instrument repräsentativem Klang, auch das musikalische ineinandergreifen der Aufnahmen wichtig. In einem Studiokontext würde dies durch das überwiegend gleichzeitige Einspielen des Titels und Positionierung der Musiker mit Sichtverbindung für eine gute musikalische Kommunikation realisiert werden (Albrecht, 2010). Da die beteiligten Musiker allerdings kein bereits eingespieltes Ensemble darstellen und aus organisatorischen Gründen im Zuge der Corona-Pandemie vorab keine gemeinsame Probe realisierbar war, bestand dieses Problem ohnehin. Um dennoch möglichst gutes musikalisches Zusammenwirken über die Aufnahmen hinweg zu ermöglichen, wurde den Musikern im Vorhinein neben ihren Stimmauszügen die gesamte Partitur, sowie ein mit Softwareinstrumenten erstelltes Demo-Playback in Stereo zu Übungszwecken zur Verfügung gestellt (siehe Anhang D). Letzteres war für jeden Musiker zusätzlich in Varianten mit angehobenem Pegel bzw. mit deaktivierter Spur des eigenen Instrumentes, sowie mit und ohne Metronomspur

vorhanden. Das Metronom diente auch während der Aufnahme an bestimmten Stellen im Stück als Ersatz für die Kommunikation beim Zusammenspiel, nämlich wenn das Tempo verlangsamt oder beschleunigt wurde bzw. das Metrum sich änderte.

Des Weiteren ist festzuhalten, dass der Freifeldraum neben dem visuellen Eindruck vor allem durch seine Akustik einen sehr ungewohnten und fordernden Aufnahmeraum darstellt. Aufgrund der fehlenden akustischen Rückmeldung durch Reflexionen und tragendem Nachhall wird angenommen, dass die Musiker dazu tendieren werden, langsamere Passagen zu schnell und leisere Passagen zu laut anzusetzen (Curdt, 1998, S. 53). Die Musiker wurden aus diesem Grund im Voraus über die ungewohnte Umgebung und die beschriebenen Effekte aufgeklärt und es wurde genügend Zeit eingeplant, sodass vor der Aufnahme eine ausreichende Einspielzeit möglich war. In dieser Zeit fand auch die Messungen mit der akustischen Kamera statt, welche in Auszügen in Kapitel 3 präsentiert wurden und die auf dem beiliegenden Datenträger hinterlegt sind (siehe Anhang D). Des Weiteren wurde über das Kopfhörer-Monitoring künstlicher Raumhall zugespielt, der die Natürlichkeit gewohnter musikalischer Vortragsumgebungen imitiert. Als Grundlage für das Monitoring der Musiker diente das Demo-Playback. Die Spuren der virtuellen Software-Instrumente wurden dann im Laufe der Aufnahmen sukzessive durch die neu hinzukommenden Spuren der realen Instrumente ersetzt. Dabei wurde für das Monitoring der Musiker stets das Mittensignal einer MHV-Anordnung verwendet.

## 5.5 Instrumente, verwendetes Equipment, Abhörraum

Tabelle 5.1 beinhaltet Angaben zu den aufgezeichneten Instrumente. Tabelle 5.2 stellt eine Auflistung des relevanten verwendeten technischen Equipments dar. Dabei sind die Gerätschaften in der Reihenfolge des Audio-Signalfusses angeführt. Zum Aufzeichnen und Abhören wurde im angrenzenden Halbfreifeldraum P21 ein Regieplatz eingerichtet. Die Audiosignale wurden in der DAW Reaper mit Samplerate 48 kHz und 24 bit Samplingtiefe aufgezeichnet. Zum direkten mehrdimensionalen Abhören wurde ein temporärer Lautsprecheraufbau mit vier *Genelec 8351A* Fullrange Studiomonitoren eingerichtet (siehe Abbildung 5.5), welche mithilfe des *Genelec GLM 3*-Systems entzerrt und laufzeitangepasst wurden. Die Lautsprecher waren relativ zur Ohrhöhe im Sitzen an den Positionen M+030 ( $\delta = 0^\circ$ ), M-030 ( $\delta = 0^\circ$ ), U+030 ( $\delta = 37^\circ$ ) und U-030 ( $\delta = 37^\circ$ ) installiert. Es erfolgte eine Dekodierung der MHV Signale auf die Busse LR und BT. Diese wurden im Gegensatz zum späteren freien Panning (siehe Abschnitt 5.7) statisch den entsprechenden Lautsprecherpaaren zugeordnet.

**Tabelle 5.1:** Angaben zu den aufgezeichneten Instrumenten.

<b>Instrument</b>	<b>Hersteller, Typ</b>	<b>Baujahr</b>	<b>Details</b>
<b>Gesang (w)</b>	-	-	Alter: 33, Stimmlage: Mezzosopran
<b>Gesang (m)</b>	-	-	Alter: 21, Stimmlage: Tenor
<b>Querflöte</b>	Pearl	k.A.	Moderne Standardquerflöte in C
<b>Oboe</b>	Marigaux	1991	Transponierendes Instrument in B
<b>Klarinette</b>	Hammerschmidt	2017	Deutsches System, französisches Mundstück, transponierendes Instrument in B
<b>Tenorsax.</b>	Selmer MK6	1961	Transponierendes Instrument in B
<b>Violine</b>	Friedrich Kleverkaus (Kronstadt)	1958	Bogen: J-F Daber Lyon, Saiten: Evah Pirazzi Gold
<b>Kontrabass</b>	k.A.	k.A.	Instrument ca. 20 Jahre alt, Bogen: Robert Reichel, Saiten: Evah Pirazzi weich (E und A), Pirastro oliv. (D und G)
<b>Westerngit.</b>	Takamine Santa Fe TSF 48C	k.A.	Saiten: Ernie Ball EB2078 Paradigm Phosphor Bronze .011 bis .052
<b>Shaker</b>	Meinl SH5R	k.A.	Studio Shaker
<b>Congas</b>	Sonor	k.A.	Durchmesser Schlagfelle: Ca. 29 cm und 31 cm
<b>Schlagzeug</b>	Set: TAMA Superstar Hyperdrive, Becken: Meinl, Crash: Soundcaster Fusion Trash, HiHat: Soundcaster Fusion Medium, Ride: Byzance Sand	2010	Bassdrum: 22'' (ca. 56 cm), Snaredrum: 14'' (ca. 36 cm), Crash: 16'' (ca. 41 cm), HiHat: 14'' (ca. 36 cm), Ride: 20'' (ca. 51 cm)

**Tabelle 5.2:** Auflistung der verwendeten technischen Komponenten.

<b>Bezeichnung</b>	<b>Hersteller</b>	<b>Modell</b>	<b>Details</b>
<b>Mikrofon</b>	SCHOEPS	CCM 5	4x Kleinmembran Kondensator kompakt, umschaltbare Charakteristik Kugel/Niere
<b>Mikrofon</b>	SCHOEPS	CCM 8	3x Kleinmembran Kondensator kompakt, Charakteristik Acht
<b>Mikrofon</b>	SCHOEPS	MK 8 auf CMC6	3x Kleinmembran Kondensator, Charakteristik Acht
<b>Mikrofon</b>	SCHOEPS	CCM 41	4x Kleinmembran Kondensator kompakt, Charakteristik Superniere
<b>Mikrofon</b>	Beyerdynamic	TG D71	Grenzfläche, Kondensator für Bassdrum
<b>Vorverstärker, AD-Wandler</b>	Midas	DL16	16 Mikrofonvorverstärker und Wandlung für Übertragung via AES50
<b>Mischpult, Interface</b>	Behringer	X32 Rack	USB 2.0 Audiointerface und Mischpult für Monitormischung
<b>KH-Verstärker</b>	Presonus	HP4	-
<b>Kopfhörer</b>	Beyerdynamic	DT-990	Monitoring der Musiker
<b>Rechner</b>	Apple	MacBook Pro mid 2015	OS: Mojave 10.14.6, CPU: Intel Core i7-4980HQ CPU @ 2.80GHz, 16 GB RAM
<b>DAW</b>	Cockos	Reaper	v. 6.13
<b>Abhörsystem</b>	Genelec	8351A	4x Fullrange Studiomonitor



**Abbildung 5.5:** Abhörplatz im Halbfreifeldraum P21.

## 5.6 Mikrofonpositionen

Unter Berücksichtigung der Erkenntnisse aus Abschnitt 4.3 wurde für das Hauptmikrofon ein OCT-System mit zwei koinzidenten Höhenmikrofonen gewählt. Da sich die Untersuchungen auf die Abbildung des Direktschalls der Instrumente beschränkt und im Aufnahmeraum Freifeldbedingungen herrschen, wurde auf die rückwärtigen Mikrofone verzichtet. Die Anordnung befand sich in 3,5 m Entfernung vom Nullpunkt der Instrumente (siehe Abbildung 5.4) und wurde mittels Stativ auf 1,9 m Höhe positioniert. Abbildung 5.6 zeigt den installierten Aufbau. Da die Aufnahmen der Instrumente nacheinander und am selben Punkt im Raum stattfanden, wurde eine virtuelle Bühnenaufstellung kreiert (Abbildung 5.7). Vor jeder Aufnahme wurde das Hauptmikrofon in die für das jeweilige Instrument entsprechende Richtung gedreht. Der Abstand der beiden Supernierenkapseln war so gewählt, dass die volle Breite der virtuellen Bühne innerhalb der frontalen  $\pm 30^\circ$  horizontal abgebildet wird (siehe Abschnitt 4.5). Für den passenden Aufnahmewinkel von  $120^\circ$  ergibt sich ein Abstand von 60 cm zwischen den Supernieren L und R (SCHOEPS, 2020c).

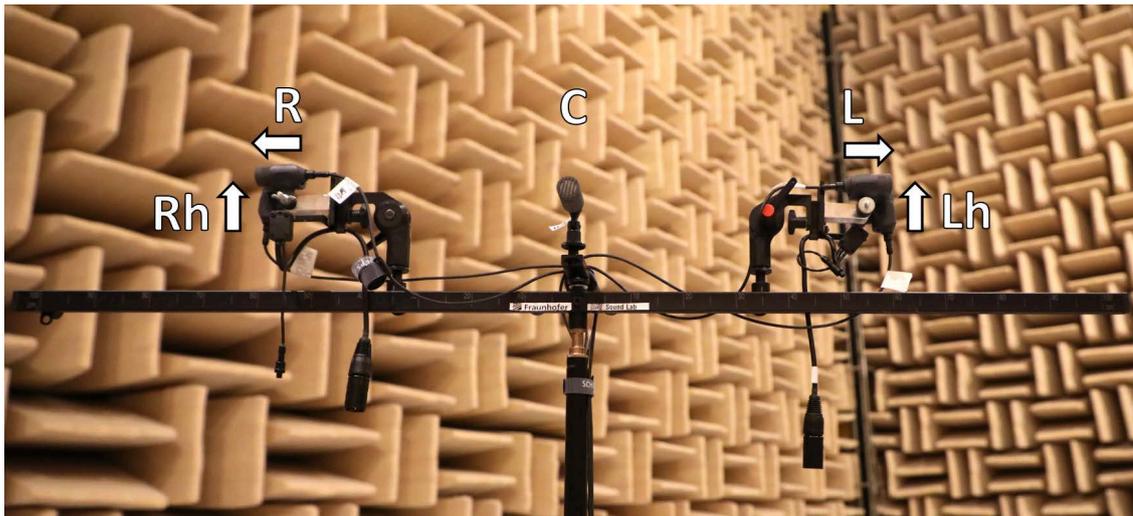


Abbildung 5.6: Aufgezeichnetes Hauptmikrofon OCT + Höhenkanäle Lh und Rh.

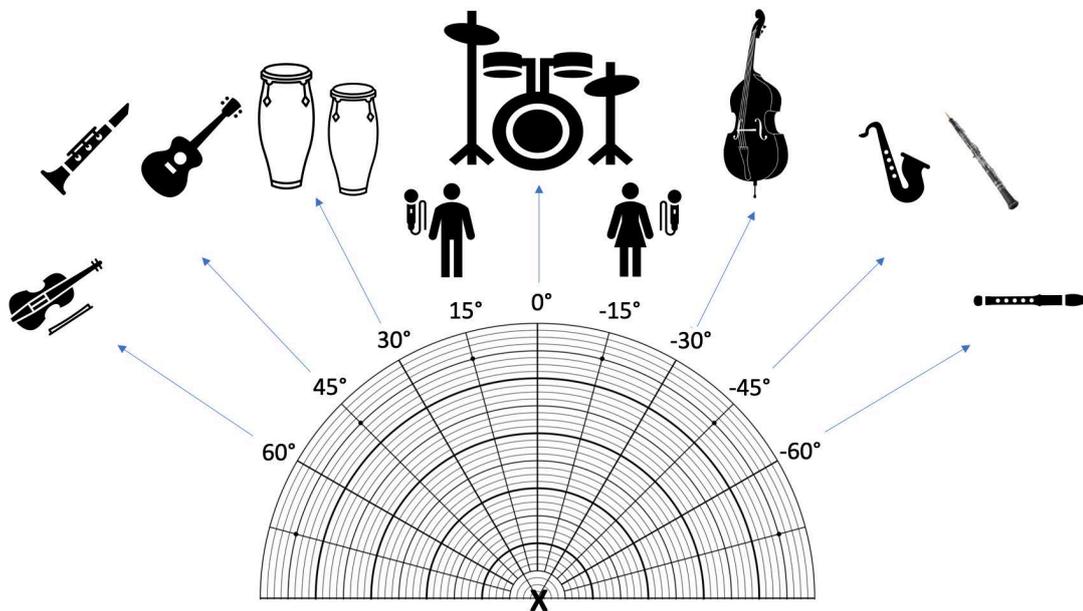


Abbildung 5.7: Virtuelle Aufstellung des Ensembles.

In der Vorbereitung wurden unter Berücksichtigung der in Kapitel 3 dargelegten akustischen Eigenschaften der aufzuzeichnenden Instrumente, geeignete Positionen für die MHV-Anordnungen bestimmt. Hauptaugenmerk lag dabei auf einer guten Balance zwischen Positionen, von denen aus das M-Mikrofon allein bereits ein ausgewogenes Klangbild aufzeichnet (weiter entfernt) und Positionen, bei denen genügend dekorrelierte Signale auf die empfindlichen Seiten der Achtermikrofone treffen, sodass die Ausdehnung in der Mischung wahrgenommen wird (näher am Instrument). Die Positionen wurden am Tag der Aufnahme am jeweiligen Instrument getestet und in kleinem Umfang nachjustiert. Dabei wurde insbesondere darauf geachtet, dass das Zentrum der virtuellen Quelle - bei realem Instrument in Ruhelage am Nullpunkt - horizontal mittig dargestellt wird. Um diese Ruhelage herum wird das auditive Bild anschließend, je nach Bewegung des Musikers, schwanken. In Tabelle 5.3 finden sich die finalen, für die Aufnahmen verwendeten Positionen MHV\_1 bis MHV\_3 an jedem Instrument. Die Höhenangabe der Instrumente bezieht sich auf das definierte akustische Zentrum des Instruments am Nullpunkt und wurde, wie auch die Höhe der Anordnungen, vom Boden aus gemessen. Die Entfernung beschreibt die direkte Distanz zwischen Vorderkante des M-Mikrofons und Nullpunkt des Instruments. Die beiden Winkelangaben sind aus der Sicht des in Richtung Hauptmikrofon blickenden Musikers angegeben, wobei der Nullpunkt des Instruments den Ursprung des Polarkoordinatensystems darstellt. Es handelt sich um errechnete Werte aus den ausgemessenen kartesischen Positionen. Die letzte Spalte der Tabelle verbalisiert den definierten Nullpunkt der Instrumente bzw. für die einzelnen MHV-Anordnungen den Teil des Instruments, auf welchen die Hauptachse des M-Mikrofons ausgerichtet ist. Die folgenden Abschnitte erläutern die MHV-Positionen am jeweiligen Instrument im Detail. Zusätzlich dokumentierende Bilder finden sich im Anhang A.

**Tabelle 5.3:** Positionen der MHV-Anordnungen vor den Instrumenten.

	Höhe	Entfernung	Azimut	Elevation	Ausrichtung /
	h [m]	r [m]	$\delta$ [°]	$\phi$ [°]	Nullpunkt
<b>Gesang (w)</b>	<b>1.50</b>				<b>Mund</b>
MHV_1	1.30	0.40	-46	-30	Mund
MHV_2	1.50	0.40	0	0	Mund
MHV_3	1.58	0.95	0	5	Mund
<b>Gesang (m)</b>	<b>1.60</b>				<b>Mund</b>
MHV_1	1.40	0.40	-46	-30	Mund
MHV_2	1.60	0.40	0	0	Mund
MHV_3	1.75	0.80	0	11	Mund
<b>Querflöte</b>	<b>1.60</b>				<b>Mitte Flöte</b>
MHV_1	1.45	0.30	-74	-30	Mitte Flöte
MHV_2	1.60	0.25	0	0	Mitte Flöte
MHV_3	1.85	0.35	0	46	oberer Teil Flöte
<b>Oboe</b>	<b>1.30</b>				<b>Mitte Oboe</b>
MHV_1	1.30	0.25	-53	0	hinter Trichter
MHV_2	1.00	0.35	0	-59	oberhalb Trichter
MHV_3	1.80	0.60	0	56	Mitte Oboe
<b>Klarinette</b>	<b>1.30</b>				<b>Mitte Klarinette</b>
MHV_1	1.25	0.25	-55	-12	Mitte Klarinette
MHV_2	0.95	0.40	0	-61	Trichter
MHV_3	1.60	0.40	0	49	Mitte Klarinette
<b>Tenorsax.</b>	<b>1.15</b>				<b>Trichter</b>
MHV_1	1.30	0.35	-39	25	hinter Trichter
MHV_2	1.35	0.23	0	60	Mitte Saxofon
MHV_3	1.80	0.65	0	90	oberer Bogen
<b>Violine</b>	<b>1.10</b>				<b>Steg</b>
MHV_1	1.25	0.30	-35	30	Steg; Achse gekippt

MHV_2	1.05	0.40	0	-7	Steg
MHV_3	1.60	0.70	0	46	Steg
<b>Kontrabass</b>	<b>0.80</b>				<b>Steg</b>
MHV_1	0.95	0.50	-71	17	Steg
MHV_2	0.85	0.30	0	10	zw. Steg und Griffbrett
MHV_3	1.50	0.80	-40	61	unterer Teil Griffbrett
<b>Gitarre</b>	<b>0.80</b>				<b>Schalloch</b>
MHV_1	1.15	0.70	0	30	Steg
MHV_2	0.90	0.35	0	17	Steg
MHV_3	1.20	0.42	-51	72	Steg
<b>Shaker</b>	<b>1.55</b>				<b>Nullposition</b>
MHV_1	1.50	0.20	90	-14	gerade auf Nullpunkt
MHV_2	1.70	0.15	0	90	gerade auf Nullpunkt
MHV_3	1.35	0.35	-44	-35	gerade auf Nullpunkt
<b>Congas</b>	<b>1.00</b>				<b>Mitte zw. Schlagfellen</b>
MHV_1	1.05	0.30	0	10	gerade auf Nullpunkt
MHV_2	1.20	0.21	0	72	gerade auf Nullpunkt
MHV_3	1.20	1.00	0	12	gerade auf Nullpunkt
<b>Schlagzeug</b>	<b>0.70</b>				<b>Snaredrum</b>
MHV_1	1.60	1.00	-27	64	Snare; Achse gekippt
MHV_2	0.75	0.85	-21	3	Snare; Achse gekippt



**Abbildung 5.8:** Positionen der MHV für weiblichen Gesang.

### 5.6.1 Gesangsstimme

Die Abbildungen 5.8 und 5.9 zeigen die drei MHV-Positionen für die Aufzeichnung der Gesangsstimmen. Bis auf Anpassung an die Mundhöhe wurden die Positionen für die Sängerin und den Sänger gleich gewählt. MHV\_2 nimmt eine mittige Position mit geringem Abstand zum Mund - ähnlich einer konventionellen Monomikrofonierung im Studio - ein. MHV\_1 ist im gleichen Abstand schräg rechts unterhalb des Mundes positioniert. Sie soll zum einen die klangliche Variante in der Vorzugsrichtung aller Frequenzanteile darstellen (vgl. Abschnitt 3.1) und zum anderen können aus diesem Winkel Bewegungen nach vorne und hinten in horizontaler Richtung der Anordnung erfasst werden. MHV\_3 stellt eine Variante von MHV\_2 in größerem Abstand dar.

### 5.6.2 Holzblasinstrumente

Die Positionen der MHV-Anordnungen für die Aufzeichnung der Querflöte werden in Abbildung 5.10 ersichtlich. MHV\_1 befindet sich seitlich rechts auf das offene Ende der Flöte zeigend in einer der beiden Vorzugsrichtungen der dipolförmigen Abstrahlung (siehe Abbildung 3.2). Es wird vermutet, dass bereits kleine Bewegungen mit der Flöte aus diesem Winkel sowohl vertikal als auch horizontal dynamische Effekte in der ausgedehnten



**Abbildung 5.9:** Positionen der MHV für männlichen Gesang.



**Abbildung 5.10:** Positionen der MHV für Querflöte.



**Abbildung 5.11:** Positionen der MHV für Oboe.

Abbildung erzielen. Die beiden Anordnungen MHV\_2 und MHV\_3 sind jeweils mittig positioniert und befinden sich außerhalb der Vorzugsrichtung der Abstrahlung. Allerdings wird auf diese Weise die horizontale Ausdehnung des Instrumentes ausgeglichen erfasst. Zudem schwächen die Positionierung im Nahfeld sowie die Bewegungen des Musikers den Auslöschungseffekt in der Mitte des Dipols ab. MHV\_2 befindet sich frontal auf Höhe der Flöte, während MHV\_3 eleviert positioniert ist und einen größeren Anteil der Klappen- und Anblasgeräusche aufzeichnen soll.

Die drei ausgewählten Positionen für Oboe, Klarinette und Saxofon gestalten sich aufgrund der in etwa übereinstimmenden Abstrahlcharakteristika ähnlich (siehe Abbildungen 5.11, 5.12 und 5.13). MHV\_2 nimmt die mittige, frontale Position ein, die sich im Falle von Oboe und Klarinette knapp unterhalb und beim Saxofon oberhalb des Schalltrichters befindet. Dadurch nimmt das M-Mikrofon jeweils den präsenten Klang der Vorzugsrichtung aus dem Trichter auf. MHV\_3 stellt eine elevierte gekippte Variante dar, welche ein von den abstrahlenden Klappen geprägtes Bild zum Ziel hat. MHV\_1 befindet sich seitlich rechts und stellt eine alternative Ausrichtung der dynamischen Effekte durch Bewegung sowie eine klangliche Variante dar.



**Abbildung 5.12:** Positionen der MHV für Klarinette.



**Abbildung 5.13:** Positionen der MHV für Tenorsaxofon.



**Abbildung 5.14:** Positionen der MHV für Violine.

### 5.6.3 Streichinstrumente

Abbildung 5.14 zeigt die MHV-Positionen an der Violine. MHV\_1 befindet sich dabei mittig vor dem Instrument, mit der Intension, die Ausdehnung symmetrisch zu erfassen. Aufgrund der schrägen Orientierung der Violine beim Spielen ist die Achse der Anordnung gekippt und die Position entspricht nicht der frontalen Blickrichtung, welche wiederum MHV\_2 auf Höhe des Instrumentes einnimmt. MHV\_3 befindet sich oberhalb, in größerer Entfernung zum Instrument.

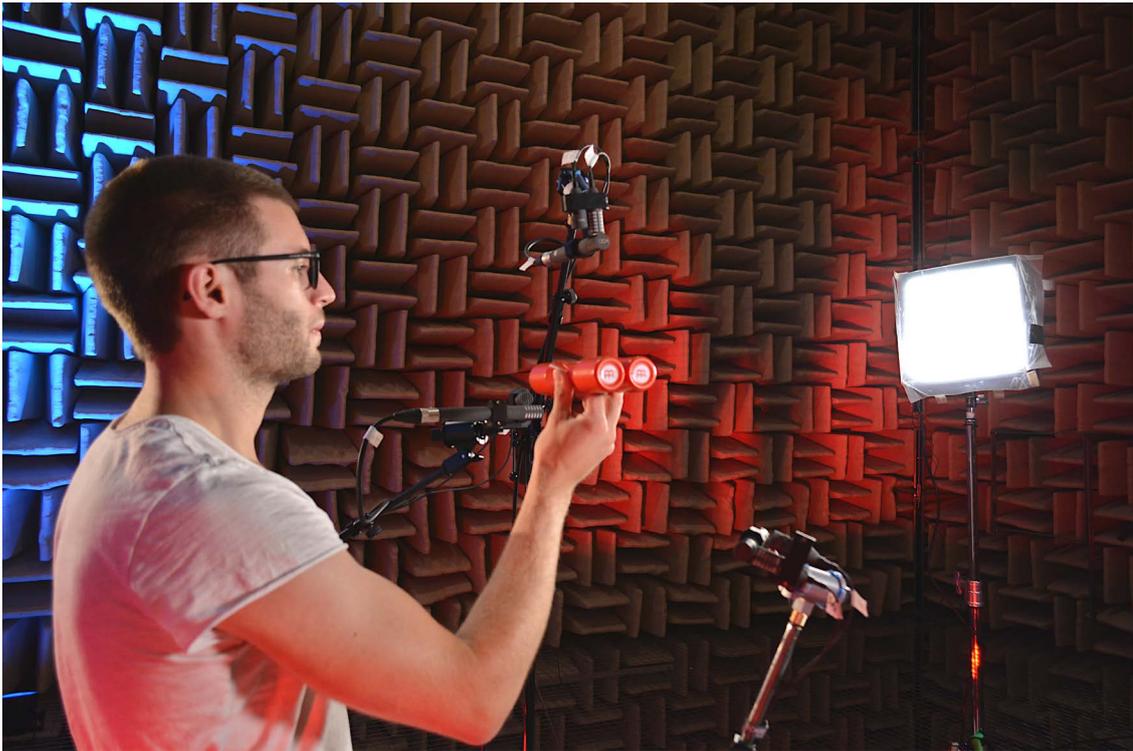
Die Anordnung MHV\_1 am Kontrabass befindet sich seitlich ca. auf Höhe des Stegs und stellt eine klangliche Variante zur frontal positionierten MHV\_2 dar (siehe Abbildung 5.15). Diese ist auf die Stelle oberhalb des Stegs ausgerichtet, wo der Bogen geführt wird und soll somit ein symmetrisches Bild des Instrumentes und insbesondere der horizontalen Dynamik des Streichens aufzeichnen. MHV\_3 ist leicht aus der Mitte versetzt und deutlich eleviert positioniert sowie auf den unteren Teil des Griffbretts ausgerichtet. Dabei wurde das Ziel verfolgt, ein Klangbild mit mehr geräuschhaften Anteilen durch Zupfen und Saiten-Schnarren zu erhalten.



**Abbildung 5.15:** Positionen der MHV für Kontrabass.



**Abbildung 5.16:** Positionen der MHV für Gitarre.



**Abbildung 5.17:** Positionen der MHV für Shaker.

## 5.6.4 Gitarre

Anordnung MHV\_1 und MHV\_2 sind in unterschiedlicher Entfernung frontal zur Gitarre angeordnet. Die in Abschnitt 5.1 beschriebenen Vorversuche haben gezeigt, dass eine Ausrichtung auf das Schallloch nicht das gewünschte klangliche Ergebnis liefert. Die Ausrichtung auf den Steg schien geeigneter und das dargestellte Bild der Gitarre wirkte so auch horizontal zentrierter. MHV\_3 stellt mit der elevierten Positionierung eine klangliche Variante mit weniger Anteil der Korpusresonanz dar. Abbildung 5.16 zeigt die beschriebenen Positionen.

## 5.6.5 Percussion und Schlagzeug

Wie in Abschnitt 3.6 erläutert, ist das Ziel der Mikrofonierung für den Shaker die Abbildung der dynamisch variierenden Quellorte durch die Schüttel-Bewegung. Aus diesem Grund wurden drei verschiedene Orientierungen der Anordnungen gewählt (siehe Abbildung 5.17). MHV\_1 befindet sich ca. auf Höhe des Shakers auf der linken Seite und nimmt die beschriebenen Bewegungen in horizontaler Richtung auf. MHV\_2 ist direkt über der Ruhelage des Instruments positioniert und nimmt die Bewegungen in vertikaler Richtung auf. MHV\_3 stellt eine Zwischenvariante von rechts unterhalb dar.



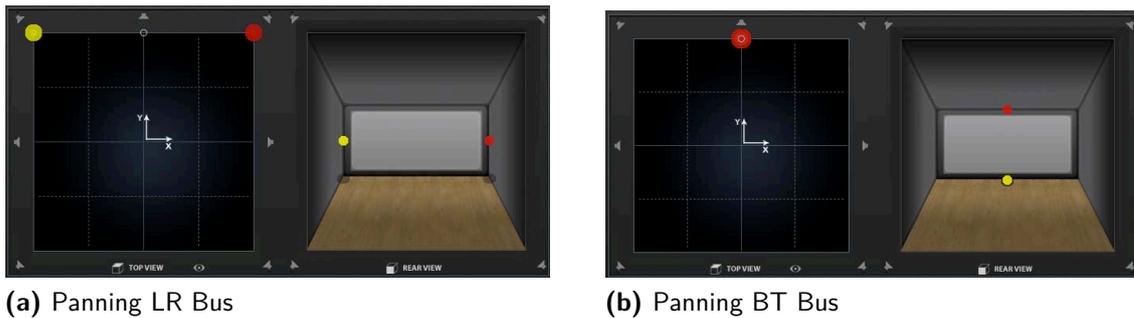
**Abbildung 5.18:** Positionen der MHV für Congas.

Die einzufangenden Signale der Congas lassen sich in die beiden Bereiche Anschlag (oben) und Ausschwingen des Kessels (unten) einteilen (siehe Abschnitt 3.5.2). Der je nach gewünschtem Klang variierende Anschlagpunkt verspricht außerdem dynamische Positionen in der Ausdehnung. Aus diesem Grund wurden zwei nahe (siehe 5.18) und eine entfernte Anordnung gewählt. MHV\_1 befindet sich leicht eleviert nahe der Ränder der Schlagfelle und soll das vertikale Bild von Anschaggeräuschen mittig bis oben und Ausschwingen des Kessels unten erfassen. MHV\_2 befindet sich eleviert oberhalb der Schlagfelle und bildet dadurch hauptsächlich die Schlagfläche ab. MHV\_3 ist mit mehr Abstand frontal positioniert und stellt damit eine Variante zu MHV\_1 dar.

Abbildung 5.19 zeigt die beiden MHV-Positionen am Schlagzeug. Auf eine dritte Position wurde verzichtet, stattdessen wurden Bassdrum und Snaredrum jeweils mit einem einzelnen Mikrofon als Zusatzsignale für die Mischung abgenommen. MHV\_1 befindet sich an klassischer Overhead-Position über dem Schlagzeug, mit Ausrichtung auf die Snaredrum, während MHV\_2 zwischen Bassdrum, Snaredrum und Becken positioniert ist und von vorne auf die Snaredrum zeigt.



**Abbildung 5.19:** Positionen der MHV für Schlagzeug.



**Abbildung 5.20:** Panning der dekodierten MHV Signale LR und BT im MultiPanner.

## 5.7 Bereitstellung der Daten zur Evaluation

Von allen Takes der Tonaufzeichnung wurden mit DSLR-Kameras Videoaufzeichnungen aus den beiden Blickwinkeln frontal und seitlich erstellt. Damit stehen als Rohmaterial pro Instrument neun Audiospuren (drei MHV-Positionen mit je drei Mikrofonen) sowie zwei synchronisierte Videos, welche das Geschehen vor dem Mikrofon zeigen, zur Verfügung. Die Audiospuren wurden zur Evaluation und Weiterverwertung in Form mehrerer *Nuendo*-Sessions (DAW des Herstellers *Steinberg*) auf einem Datenträger zur Verfügung gestellt. Die Benennung der Daten und Spuren erfolgte aufgrund der vielfältigen Weiterwendungsmöglichkeiten auf englisch. Die aktuelle *Nuendo*-Version 10 bietet den Vorteil, dass mit dem MultiPanner eine native Lösung zur Positionierung virtueller Schallquellen in einer mehrdimensionalen kanalbasierten Mischung integriert ist und somit keine Drittanbieter-Plugins Verwendung finden müssen.

Für jedes Einzelinstrument wurde eine Session mit den jeweiligen MHV-Positionen erstellt und die beiden Videodateien eingebettet. Diese Sessions dienen zur vergleichenden Evaluation der Anordnung an den verschiedenen Positionen für das spezifische Instrument. Außerdem wurde eine Session mit allen aufgezeichneten Spuren und eine weitere, mit vorausgewählter MHV-Position je Instrument, erstellt. Letztere diente als Startpunkt für die Testmischung, deren Session ebenfalls hinterlegt ist. Die im Mischprozess gewonnenen subjektiven Eindrücke sind in Kapitel 6 geschildert. Die MHV Signale wurden in allen Sessions pro Instrument jeweils auf zwei Busse **LR** für *Left/Right* und **BT** für *Bottom/Top* dekodiert. Diese beiden Stereosignale können über den MultiPanner positioniert werden (siehe Abbildung 5.20).

## 6 Testmischung und Eindrücke

Auf Grundlage der in Abschnitt 5.7 beschriebenen Nuendo-Session wurde eine beispielhafte Mischung für das 4+7+0 System angefertigt. Dafür wurde pro Instrument eine MHV-Position ausgewählt und deren LR- bzw. BT Signal an der gewünschten Position platziert. Es wurde eine breitere Verteilung als bei der konventionellen Positionierung auf einer frontalen virtuellen Bühne gewählt, die sich seitlich bis zu den Lautsprechern M+135 bzw. M-135 erstreckt. Die Anordnung der Einzelbestandteile zueinander wurde fürs Erste bei der in Abbildung 5.20 dargestellten Rautenform belassen, die der natürlichen Orientierung des Instruments nahe kommen soll. An dieser Stelle bestehen zukünftig zahlreiche weitere Möglichkeiten, die unter Umständen auch eine Vergrößerung insbesondere der vertikalen Ausdehnung erreichen können. Beispielhaft sei eine um 90° gedrehte Position der Gitarre genannt, sodass höherfrequente Griffgeräusche eleviert positioniert sind. Die Auswahl der Positionen je Instrument erfolgten nach dem Prinzip, die Balance aus ausgewogenem Klang und stimmiger ausgedehnter Abbildung zu finden. Im Folgenden werden die subjektiven Eindrücke geschildert, welche im Laufe des Mischprozesses gewonnen wurden:

Der gewonnene Klangeindruck lässt sich im Allgemeinen als „natürlich“ bezeichnen. Je nach Instrument gestaltet sich die Wahrnehmung einer Ausdehnung stärker oder schwächer. Dabei kann die Tendenz, dass insbesondere eine Ausdehnung nach oben für Instrumente mit genügendem Anteil hoher Frequenzen besser funktioniert, bestätigt werden. Ohne den Vergleich zum reinen Monosignal kann beispielsweise bei Oboe, Klarinette und Saxofon nicht direkt auf eine ausgedehnte Darstellung geschlossen werden. Es wird vermutet, dass dies an den verschiedenen beschriebenen Schwierigkeiten und Abhängigkeiten in der vertikalen Lokalisation liegt. Vergleicht man allerdings mit dem Mono-Signal, wird die vorhandene Ausdehnung mit dem MHV-Verfahren deutlich und die Instrumente passen sich natürlicher in das Gesamtbild ein. In horizontaler Richtung sind die drei realen Instrumente kaum ausgedehnt. Entsprechend ergibt sich hier nur bei größeren seitlichen Bewegungen des Instruments ein hörbarer Effekt, welcher im ersten Eindruck eher ungewohnt erscheint - dies kann aber auch auf die Hörgewohnheiten nicht-ausgedehnter Instrumente in konventionellen Mischungen zurückzuführen sein. Bei entsprechend hoch gewähltem H-Pegel verbreitert sich allerdings auch bei statischem Instrument das Bild in horizontaler Richtung auf natürliche Weise.

Für die Querflöte ist entsprechend ihrer Orientierung die horizontale Richtung dominant, wobei aufgrund der guten Lokalisation durch Pegeldifferenzen in dieser Dimension ein ausgewogenes Bild mit definiertem Anblasgeräusch und natürlicher Ausdehnung entsteht.

Die vertikale Richtung bleibt auch bei Bewegungen weitestgehend ungenutzt. Für Oboe und Klarinette wurde die seitliche Position MHV\_1 ausgewählt, da diese ein ausgewogenes Klangbild liefert, ohne den vom Trichter ausgehenden, präsenten Charakter zu verlieren, welcher in MHV\_2 hingegen zu stark und in MHV\_3 zu schwach ausgeprägt erscheint. Für das Saxofon gestaltet sich die seitliche Variante zu dumpf. Aufgrund des nach oben gerichteten Trichters bietet die elevierte Variante MHV\_3 das ausgewogenste Bild. Für die Querflöte wurde ebenfalls die frontal elevierte MHV\_3 aufgrund der lautereren Anblasgeräusche ausgewählt.

Für den Gesang wurde jeweils MHV\_1, also die nahe Position seitlich unterhalb des Mundes, aufgrund klanglicher Präferenz ausgewählt. Was für die vertikal orientierten Holzblasinstrumente gilt, macht sich hier in beide Ausdehnungsrichtungen bemerkbar. Wird der Pegel von H und V erhöht, entsteht ein leicht ausgedehntes, weniger scharf lokalisierbares Klanggebilde, welches jedoch zusammen mit den anderen ausgedehnten Instrumenten stimmig erscheint. Die hörbaren Bewegungen in horizontaler Richtung wirken bei zu breiter Positionierung von L und R jedoch unnatürlich, in vertikaler Richtung sind sie kaum bemerkbar.

Die frontale Position MHV\_2 auf Höhe der Violine wurde aus klangästhetischen Gründen ausgewählt. Sie stellt aus diesem Winkel jedoch nur ein sehr kleines Bild des realen Instrumentes dar. Die wahrgenommene Ausdehnung gestaltet sich daher eher klein, jedoch ist eine Homogenisierung des Klanges im Vergleich zur monofonen Abnahme, dem Effekt der Directional Tone Color entsprechend, zu verzeichnen. Es bleibt zu überlegen, andere nahe Positionen an der Violine zu untersuchen, um ein breiteres Bild zu generieren.

Für den Kontrabass wurde die seitliche Position MHV\_1 ausgewählt, da sie die beste Balance zwischen vollem Klang des Korpus und ausreichend hörbaren Saitengeräuschen bietet. Es ergibt sich ein sehr natürliches, großes Bild des Instrumentes, welches merklich von der Ausdehnung in vertikaler Richtung, insbesondere durch die Saitengeräusche profitiert. Aber auch die horizontale Ausdehnung des Korpusklangs ist wahrnehmbar und unterstützt das homogene Gesamtbild.

Die beiden frontalen Positionen MHV\_1 und MHV\_2 erzeugen an der Gitarre jeweils eine harmonische ausgedehnte Abbildung, wobei in horizontaler Richtung Anschlaggeräusche und Griffgeräusche separat lokalisierbar sind. Die Ausdehnung in vertikaler Richtung ist erwartungsgemäß weniger stark ausgeprägt, aber dennoch wahrnehmbar.

Für den Shaker wurde die seitliche Position MHV\_1 ausgewählt, welche die Schüttel-Bewegungen horizontal abbildet. Das Gesamtbild erscheint im Vergleich mit den anderen Instrumenten stimmig, jedoch werden die schnellen horizontalen Bewegungen bei zu breiter Positionierung von LR als störend empfunden.

Zur Abbildung der Congas fand Position MHV\_1 Anwendung, welche frontal und nah an den Instrumenten, knapp über Schlagfellhöhe positioniert war. Diese Variante bietet das aufgrund des Pitch-Height-Effekts erwartete natürliche Bild aus eher oben befindlichem Anschlaggeräusch höherer Frequenzen und dem tiefen Nachklang der Congas eher unten. Die Abbildung wirkt sowohl klanglich als auch bezüglich Ausdehnung sehr natürlich und rund.

Am Schlagzeug wurde die zwischen Bassdrum, Snaredrum und Becken positionierte Anordnung MHV\_2 verwendet, welche vor allem direkteren, feiner aufgelösten Beckenklang bietet. Des Weiteren zeigt sich diese Positionierung im Bezug auf den Pitch-Height-Effekt günstig, da sich die Bestandteile des Schlagzeugs mit Dominanz tiefer Frequenzen unten und die hohen Frequenzen oben befinden. Das entstehende Bild zeigt eine deutliche, homogene Ausdehnung, wobei die oben lokalisierten Becken durch die konventionellen Hörerfahrungen ungewohnt erscheint.

Bezüglich der Positionierung der Schallquellen ergeben sich einige, teilweise erwartete Einschränkungen. Wie in Abschnitt 4.1 dargelegt, ist die Wahrnehmung horizontaler Phantomschallquellen in Richtung der interauralen Achse sehr unscharf und instabil. Entsprechend kann die horizontale Ausdehnung seitlich platzierter Instrumente kaum wahrgenommen werden. Bei statischer Kopfposition ist die horizontale Ausdehnung der Instrumente demnach abhängig von der Positionierung dieser. Jedoch steht es dem Hörer im Allgemeinen frei, den Kopf zu drehen, wodurch der Effekt an Relevanz verliert. Die in Abschnitt 4.2 erläuterten Konsequenzen des VBAP-Verfahrens bezüglich positionabhängiger Schallquellenausdehnung bleibt jedoch bestehen. Weiterführend wurde in Anknüpfung an die Erkenntnisse aus Abschnitt 5.1 festgestellt, dass die nahen MHV-Anordnungen anfällig für das seitliche Verschieben des Gesamtbilds sind, wenn die Nullposition des Instrumentes während des Vortrags durch Bewegungen des Musikers leicht versetzt wurde. Außerdem muss an dieser Stelle festgehalten werden, dass aufgrund der spektral unterschiedlichen Signalen, welche von den Mikrofonen aufgezeichnet werden, eine Pegeländerung der Achtermikrofone zum Anpassen der dargestellten Breite (siehe Abschnitt 4.3) klangliche Unterschiede mit sich bringt. Eine dynamische Veränderung der Ausdehnung während des Stücks ist bei annähernd konstanter Klangfarbe daher nur durch Panning-Automation zu erzielen.

## 7 Fazit und Ausblick

Im Rahmen der Arbeit wurde die Stütz-mikrofonanordnung MHV zur ausgedehnten Abbildung akustischer Instrumente in einer kanalbasierten, mehrdimensionalen Audiomischung konzipiert. Die Anordnung besteht aus einem Mittenmikrofon M sowie zwei Achtermikrofonen H und V, und basiert auf dem MS-Prinzip. Das Verfahren wurde unter Berücksichtigung psychoakustischer Aspekte zur Richtungslokalisation und Ausdehnungswahrnehmung sowie der Anforderungen bezüglich Praktikabilität und Flexibilität entwickelt. Insbesondere die Möglichkeit verzerrungsfreier Abwärtskompatibilität und die hohe Relevanz von ILD für die Wahrnehmung ausgedehnter Schallquellen führten zur Konzeption eines koinzidenten Verfahrens. Die nahe Positionierung der MHV-Anordnung vor einem Instrument soll, aufgrund der schallabstrahlenden Teilflächen desselben, zu de-korrelierten Signalen aus unterschiedlichen Richtungen führen. Dies resultiert - nach entsprechender Dekodierung und Positionierung im mehrdimensionalen Wiedergabesystem - in der Wahrnehmung einer sowohl horizontal, als auch vertikal ausgedehnten virtuellen Quelle. Zur Erprobung des Verfahrens wurde ein Musikstück für ausgewählte Instrumente arrangiert und mögliche Positionen der MHV-Anordnungen für das jeweilige Instrument spezifiziert. In Vorbereitung dazu wurden die zu berücksichtigenden akustischen Eigenschaften der Instrumente, insbesondere im Hinblick auf spektralen Besonderheiten und Abstrahlcharakteristika, herausgestellt. Die im reflexionsarmen Raum aufgezeichneten Audiospuren, sowie die dokumentierenden Bilder und Videoaufzeichnungen, wurden für die weitere Verwertung zu Forschungs- und Demonstrationszwecken aufbereitet und zur Verfügung gestellt.

Durch eine Testmischung, unter Verwendung einer MHV-Anordnung je Instrument, konnten subjektive Eindrücke gesammelt, Optimierungspotential erkannt und Felder zukünftiger Forschung ausgemacht werden. Beispielsweise wurde festgestellt, dass für einen reibungslosen Workflow bezüglich Processing und Panning einer solchen mehrdimensionalen Anordnung flexiblere Routingmöglichkeiten geschaffen und intuitivere Panning-Tools entwickelt werden müssten. Im vorliegenden Fall wäre ein vierkanaliger Bus für LR und BT wünschenswert, welcher weiterführend in Kombination mit einem Panning-Tool zur freien Positionierung und Rotation ausgestattet werden kann, welches Funktionen wie horizontale und vertikale Ausdehnung mit einem Regler bedienbar macht. Ebenfalls zu untersuchen sind Konzepte zur Generierung künstlicher Reflexionen und Hall aus dem Signal einer MHV-Anordnung.

Im Rahmen der subjektiven Beurteilung des Verfahrens wurde festgestellt, dass der Effekt einer wahrgenommenen Ausdehnung grundsätzlich eintritt und für die meisten untersuchten Instrumente einen natürlichen Höreindruck vermittelt. Insbesondere die ausgedehnte Wahrnehmung größerer Instrumente mit transienten, charakterisierenden Geräuschen, wie beispielsweise Beckenschläge oder Anschlagsgeräusche von Saiten, wird positiv bewertet. Die erwarteten Effekte durch Bewegungen der Musiker gestalten sich im Allgemeinen eher subtil bzw. sind diese nicht direkt hör- oder lokalisierbar, sondern drücken sich hauptsächlich im natürlichen Gesamtbild aus. Es wird jedoch vermutet, dass die Bewegungen im Kontext eines Konzerts größer ausfallen als in der wissenschaftlichen Umgebung des reflexionsarmen Raumes. An dieser Stelle sind weiterführende Untersuchungen denkbar, welche beispielsweise fixierten und bewegten Vortrag bei gleicher MHV Position vergleichen. Ebenso bleibt zu untersuchen, inwieweit der subjektiv störende Eindruck zu großer horizontaler Bewegungen, wie beispielsweise am Shaker, bestätigt werden kann, und ob dieser auch nach gewisser Eingewöhnungszeit an Audiomischungen mit ausgedehnten Einzelinstrumenten weiterhin besteht. Des Weiteren ist die Tauglichkeit der Anordnung beim praktischen Einsatz im Tonstudio oder in Konzertsälen zu evaluieren. In diesem Kontext wird beispielsweise ein Augenmerk auf die Richtcharakteristika der Mikrofone und die daraus resultierende Menge an Übersprechen zwischen mehreren Anordnungen zu legen sein. Ebenfalls sind Hörversuche zur Bewertung der konzipierten Anordnung, eventuell auch im Bezug auf verschiedene Wiedergabeverfahren, in Betracht zu ziehen. In diesem Zusammenhang sind außerdem mehrere Modifikationen der Signale denkbar, um die ausgedehnte Abbildung zu optimieren. Beispielfähig könnten Untersuchungen zur Änderung der Ausdehnung durch Hochpassfilterung des V-Signals durchgeführt (vgl. [Lee, 2016](#), [Kamekawa & Marui, 2020](#)) oder eine frequenzabhängige Positionierung zur breitbandigen Vereinheitlichung der wahrgenommenen Ausdehnung implementiert werden (vgl. [Franck & Fazi, 2019](#); [Hughes et al., 2018](#)).

Abschließend ist festzuhalten, dass das Stützmikrofonverfahren zur ausgedehnten Abbildung akustischer Instrumente in Form der koinzidenten MHV-Anordnung erfolgreich konzipiert und kritisch erprobt wurde. Die Grenzen der Möglichkeiten des Verfahrens durch Einschränkungen psychoakustischer oder technischer Natur wurden erkannt, und Ansatzpunkte zur Optimierung bestimmt. Die Arbeit dient somit als Basis zukünftiger Forschung im Bereich der Abbildung realer Schallquellen mit Ausdehnung in einem mehrdimensionalen Wiedergabesystem.

## Literaturverzeichnis

- Ackermann, D., Böhm, C., Brinkmann, F. & Weinzierl, S. (2019). The Acoustical Effect of Musicians' Movements During Musical Performances. *Acta Acustica united with Acustica*, 105 (2), 356–367. <https://doi.org/10.3813/AAA.919319>
- Albrecht, C. (2010). *Der Tonmeister: Mikrofonierung akustischer Instrumente in der Popmusik; Live- und Studiosetups*. Berlin: Schiele & Schön.
- Barbour, J. L. (2003). Elevation Perception: Phantom Images in the Vertical Hemisphere..
- Blauert, J. (1974). *Räumliches Hören*. Hirzel Stuttgart.
- Blauert, J. (1996). *Spatial Hearing: The Psychophysics of Human Sound Localization*. MIT Press, Revised Edition. <https://doi.org/10.7551/mitpress/6391.001.0001>
- Boren, B. & Roginska, A. (2011). The Effects of Headphones on Listener HRTF Preference. In *Audio engineering society convention 131*. <http://www.aes.org/e-lib/browse.cfm?elib=16063>
- Bork, I. (1990). Klang und Schallabstrahlung der Querflöte. *Beitrag Tonmeistertagung*.
- Bork, I. (1994). Analyse von Schallfeldern im Nahfeld mechanischer Schwinger. *PTB-Mitteilungen, Vol. 104*.
- Boueri, M. & Kyriakakis, C. (2004). Audio Signal Decorrelation Based on a Critical Band Approach. In *Audio Engineering Society Convention 117*. <http://www.aes.org/e-lib/browse.cfm?elib=12948>
- Brown, A. W. (2004). *Acoustical Studies on the Flat-backed and Round-backed Double Bass* (phdthesis). Universität für Musik und darstellende Kunst Wien.
- Cabrera, D. & Tilley, S. (2003). Vertical Localization and Image Size Effects in Loudspeaker Reproduction. In *Audio Engineering Society Conference: 24th International Conference: Multichannel Audio, The New Reality*. <http://www.aes.org/e-lib/browse.cfm?elib=12269>
- Curdt, O. (1998). *Übertragung der Abstrahleigenschaften von Musikinstrumenten in Raumsimulationsprogramme* (mathesis). Hochschule für Musik Detmold.
- Dickreiter, M., Dittel, V., Hoeg, W. & Wöhr, M. (2014). *Handbuch der Tonstudiotchnik*. Gruyter, de Saur. <https://doi.org/10.1515/9783110316506>
- Fletcher, N. H. & Rossing, T. D. (1998). *The Physics of Musical Instruments*. Springer New York. <https://doi.org/10.1007/978-0-387-21603-4>

- Franck, A. & Fazi, F. M. (2019). Towards Constant-Spread Panning for Multi-Loudspeaker Configurations. In *Audio Engineering Society Conference: 2019 AES International Conference on Immersive and Interactive Audio*. <http://www.aes.org/e-lib/browse.cfm?elib=20403>
- Friesecke, A. (2014). *Die Audio-Enzyklopädie*. Gruyter, de Saur. <https://doi.org/10.1515/9783110340181>
- Geluso, P. (2012). Capturing Height: The Addition of Z Microphones to Stereo and Surround Microphone Arrays. In *Audio Engineering Society Convention 132*. <http://www.aes.org/e-lib/browse.cfm?elib=16233>
- Görne, T. (2011). *Tontechnik*. München: Hanser. <https://doi.org/10.3139/9783446427402>
- Gribben, C. & Lee, H. (2014). The Perceptual Effects of Horizontal and Vertical Interchannel Decorrelation Using the Lauridsen Decorrelator. In *Audio Engineering Society Convention 136*. <http://www.aes.org/e-lib/browse.cfm?elib=17174>
- Gribben, C. & Lee, H. (2017). The Perceptual Effect of Vertical Interchannel Decorrelation on Vertical Image Spread at Different Azimuth Positions. In *Audio Engineering Society Convention 142*. <http://www.aes.org/e-lib/browse.cfm?elib=18623>
- Gribben, C. & Lee, H. (2019). Increasing the Vertical Image Spread of Natural Sound Sources Using Band-Limited Interchannel Decorrelation. In *Audio Engineering Society Conference: 2019 AES International Conference on Immersive and Interactive Audio*. <http://www.aes.org/e-lib/browse.cfm?elib=20397>
- Gribben, C. & Lee, H. (2020). The Perception of Band-Limited Decorrelation Between Vertically Oriented Loudspeakers. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 28, 876-888. <https://doi.org/10.1109/TASLP.2020.2969845> doi: 10.1109/TASLP.2020.2969845
- Herre, J., Hilpert, J., Kuntz, A. & Plogsties, J. (2015). MPEG-H 3D Audio: The New Standard for Coding of Immersive Spatial Audio. *IEEE Journal of Selected Topics in Signal Processing*, 9 (5), 770-779. <https://doi.org/10.1109/JSTSP.2015.2411578>
- Howard, D. & Angus, J. (2009). *Acoustics and psychoacoustics*. Amsterdam London: Focal. <https://doi.org/10.4324/9780080961873>
- Hughes, R. J., Franck, A., Cox, T. J., Shirley, B. G. & Fazi, F. M. (2018). Dual Frequency Band Amplitude Panning for Multichannel Audio Systems. In *Audio Engineering Society Conference: 2018 AES International Conference on Spatial Reproduction - Aesthetics and Science*. <http://www.aes.org/e-lib/browse.cfm?elib=19638>

- Huttunen, T. & Vanne, A. (2017). End-To-End Process for HRTF Personalization. In *Audio Engineering Society Convention 142*. <http://www.aes.org/e-lib/browse.cfm?elib=18723>
- ITU-R. (2012). *Recommendation ITU-R BS.775-3: Multichannel stereophonic sound system with and without accompanying picture* (Bericht). International Telecommunication Union. <https://www.itu.int/rec/R-REC-BS.775-3-201208-I/en>
- ITU-R. (2018). *Recommendation ITU-R BS.2051-2: Advanced sound system for programme production* (Bericht). International Telecommunication Union. <https://www.itu.int/rec/R-REC-BS.2051-2-201807-I/en>
- Jacques, R. (2005). *Untersuchungen von Verfahren zur naturgetreuen Aufnahme und Reproduktion der räumlich-akustischen Eigenschaften von Schallquellen* (mathesis). Technische Universität Ilmenau.
- Kamekawa, T. & Marui, A. (2020). Are full-range loudspeakers necessary for the top layer of Three-Dimensional audio? In *Audio Engineering Society Convention 148*. <http://www.aes.org/e-lib/browse.cfm?elib=20779>
- Katz, B. & d'Alessandro, C. (2007). Directivity measurements of the singing voice. In *19th international congress on acoustics* (S. 1-6). Madrid. [https://www.researchgate.net/publication/236213215\\_Directivity\\_measurements\\_of\\_the\\_singing\\_voice](https://www.researchgate.net/publication/236213215_Directivity_measurements_of_the_singing_voice)
- Kendall, G. S. (1995). The Decorrelation of Audio Signals and its Impact on Spatial Imagery. *Computer Music Journal*, 19 (4), 71. <https://doi.org/10.2307/3680992> doi: 10.2307/3680992
- Lee, H. (2011). The Relationship Between Interchannel Time and Level Differences in Vertical Sound Localization and Masking. In *Audio Engineering Society Convention 131*. <http://www.aes.org/e-lib/browse.cfm?elib=16082>
- Lee, H. (2016). Perceptual Band Allocation (PBA) for the Rendering of Vertical Image Spread with a Vertical 2D Loudspeaker Array. *Journal of the Audio Engineering Society*, 64 (12), 1003–1013. <https://doi.org/10.17743/jaes.2016.0052> doi: 10.17743/jaes.2016.0052
- Lee, H. (2017). Sound Source and Loudspeaker Base Angle Dependency of Phantom Image Elevation Effect. *Journal of the Audio Engineering Society*, 65, 733-748. <https://doi.org/10.17743/jaes.2017.0028>
- Lee, H. & Gribben, C. (2015). Effect of Vertical Microphone Layer Spacing for a 3D Microphone Array. *J. Audio Eng. Soc*, 62 (12), 870–884. <https://doi.org/10.17743/jaes.2014.0045>
- Lerch, R., Sessler, G. & Wolf, D. (2009). *Technische Akustik*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-540-49833-9>

- Melchior, F. (2012). The Reactive Source: A Reproduction Format Agnostic and Adaptive Spatial Audio Effect. In *Audio Engineering Society Convention 133*. <http://www.aes.org/e-lib/browse.cfm?elib=16618>
- Menzies, D. (2002). W-Panning and O-Format, Tools for Object Spatialization. In *Audio Engineering Society Conference: 22nd International Conference: Virtual, Synthetic, and Entertainment Audio*. <http://www.aes.org/e-lib/browse.cfm?elib=11151>
- Meyer, J. (2009). *Acoustics and the Performance of Music*. Springer-Verlag GmbH. <https://doi.org/10.1007/978-0-387-09517-2>
- Mironovs, M. & Lee, H. (2016). Vertical Amplitude Panning for Various Types of Sound Sources.. [https://www.researchgate.net/publication/308962002\\_Vertical\\_Amplitude\\_Panning\\_for\\_Various\\_Types\\_of\\_Sound\\_Sources](https://www.researchgate.net/publication/308962002_Vertical_Amplitude_Panning_for_Various_Types_of_Sound_Sources)
- Mironovs, M. & Lee, H. (2017). The Influence of Source Spectrum and Loudspeaker Azimuth on Vertical Amplitude Panning. In *Audio Engineering Society Convention 142*. <http://www.aes.org/e-lib/browse.cfm?elib=18658>
- Mironovs, M. & Lee, H. (2018). On the Accuracy and Consistency of Sound Localization at Various Azimuth and Elevation Angles. In *Audio Engineering Society Convention 144*. <http://www.aes.org/e-lib/browse.cfm?elib=19469>
- Potard, G. & Spille, J. (2003). Study of Sound Source Shape and Wideness in Virtual and Real Auditory Displays. In *Audio Engineering Society Convention 114*. <http://www.aes.org/e-lib/browse.cfm?elib=12499>
- Power, P., Davies, B. & Hirst, J. (2014). Investigation into the Impact of 3D Surround Systems on Envelopment. In *Audio Engineering Society Convention 137*. <http://www.aes.org/e-lib/browse.cfm?elib=17431>
- Pulkki, V. (1997). Virtual Sound Source Positioning Using Vector Base Amplitude Panning. *J. Audio Eng. Soc*, 45 (6), 456–466. <http://www.aes.org/e-lib/browse.cfm?elib=7853>
- Pulkki, V. (1999, Oct). Uniform Spreading of Amplitude Panned Virtual Sources. In *Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics* (S. 187-190). <https://doi.org/10.1109/ASPAA.1999.810881>
- Pätynen, J. & Lokki, T. (2010). Directivities of Symphony Orchestra Instruments. *Acta Acustica united with Acustica*, 96 (1), 138–167. <https://doi.org/10.3813/AAA.918265>
- Rindel, J. (2020). *Acoustical aspects of the transverse flute from 1700 to present*. [https://www.researchgate.net/publication/340051666\\_Acoustical\\_aspects\\_of\\_the\\_transverse\\_flute\\_from\\_1700\\_to\\_present](https://www.researchgate.net/publication/340051666_Acoustical_aspects_of_the_transverse_flute_from_1700_to_present)

- Roginska, A. & Geluso, P. (2018). *Immersive sound: The art and science of binaural and multi-channel audio*. New York: Routledge. <https://doi.org/10.4324/9781315707525>
- Rumsey, F. (2001). *Spatial audio*. Oxford Boston: Focal Press. <https://doi.org/10.1017/S1355771804220155>
- Santala, O. & Pulkki, V. (2011). Directional Perception of Distributed Sound Sources. *The Journal of the Acoustical Society of America*, 129 (3), 1522-1530. <https://doi.org/10.1121/1.3533727>
- Schmele, T. & Sayin, U. (2018). Controlling the Apparent Source Size in Ambisonics Using Decorrelation Filters. In *Audio Engineering Society Conference: 2018 AES International Conference on Spatial Reproduction - Aesthetics and Science*. <http://www.aes.org/e-lib/browse.cfm?elib=19612>
- Schneider, A. (2017). *Studies in Musical Acoustics and Psychoacoustics*. Springer-Verlag GmbH. <https://doi.org/10.1007/978-3-319-47292-8>
- SCHOEPS. (2020a). Manual CCM Compact Microphones [Software-Handbuch].
- SCHOEPS. (2020b). Manual Colette Modular System: MK Microphone Capsules and CMC Microphone Amplifiers [Software-Handbuch].
- SCHOEPS. (2020c). OCT Surround: Anleitung [Software-Handbuch].
- Theile, G. & Wittek, H. (2011). Principles in Surround Recordings with Height. In *Audio Engineering Society Convention 130*. <http://www.aes.org/e-lib/browse.cfm?elib=15870>
- Tregonning, A. & Martin, B. (2015). The Vertical Precedence Effect: Utilizing Delay Panning for Height Channel Mixing in 3D Audio. In *Audio Engineering Society Convention 139*. <http://www.aes.org/e-lib/browse.cfm?elib=18025>
- Wallis, R. & Lee, H. (2015). Directional Bands Revisited. In *Audio engineering society convention 138*. <http://www.aes.org/e-lib/browse.cfm?elib=17702>
- Wallis, R. & Lee, H. K. (2016). Vertical Stereophonic Localization in the Presence of Interchannel Crosstalk: The Analysis of Frequency-Dependent Localization Thresholds. *Journal of the Audio Engineering Society*, 64 (10), 762–770. <https://doi.org/10.17743/jaes.2016.0039>
- Weinreich, G. (1997). Directional tone color. *The Journal of the Acoustical Society of America*, 101 (4), 2338-2346. <https://doi.org/10.1121/1.418213>
- Weinzierl, S. (2008). *Handbuch der Audiotechnik*. Springer-Verlag GmbH. <https://doi.org/10.1007/978-3-540-34301-1>
- Zaar, J. (2010). *Vermessung von Außenohrübertragungsfunktionen mit reziproker Messmethode*.

## Anhang A: Bilder zur Dokumentation

### A.1 Weiblicher Gesang



(a) Azimut  $\phi \approx 30^\circ$



(b) Azimut  $\phi \approx -90^\circ$



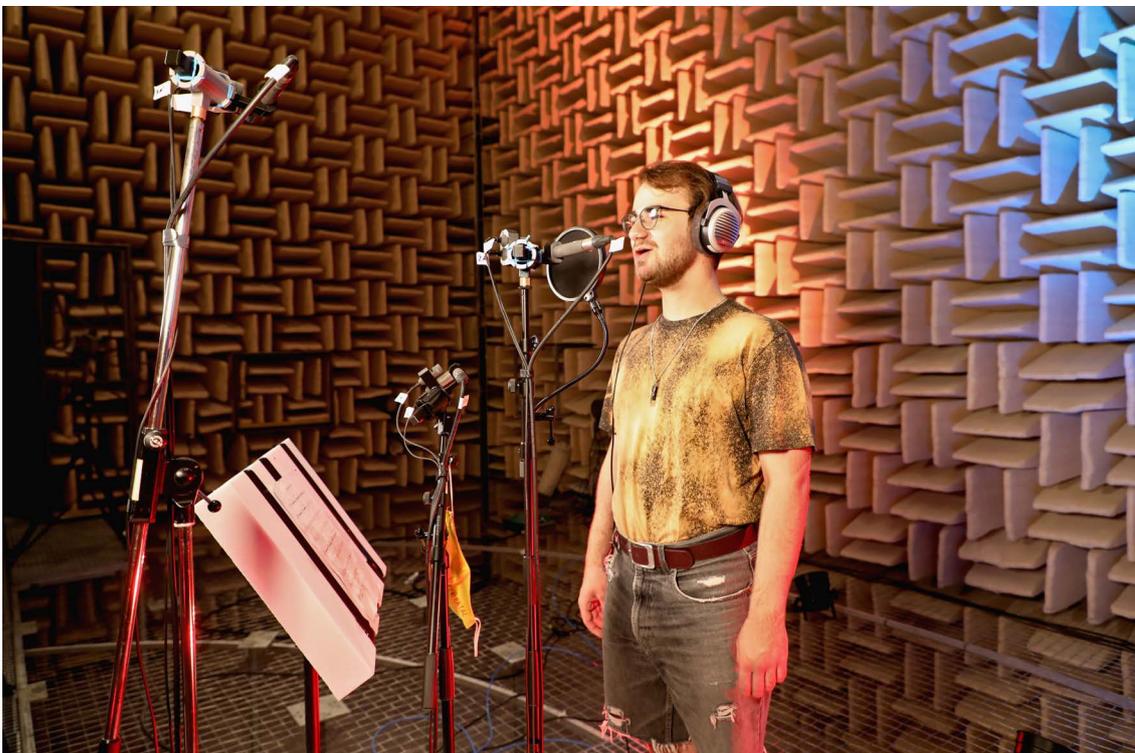
(c) Azimut  $\phi \approx -150^\circ$

**Abbildung A.1:** Bilder weiblicher Gesang.

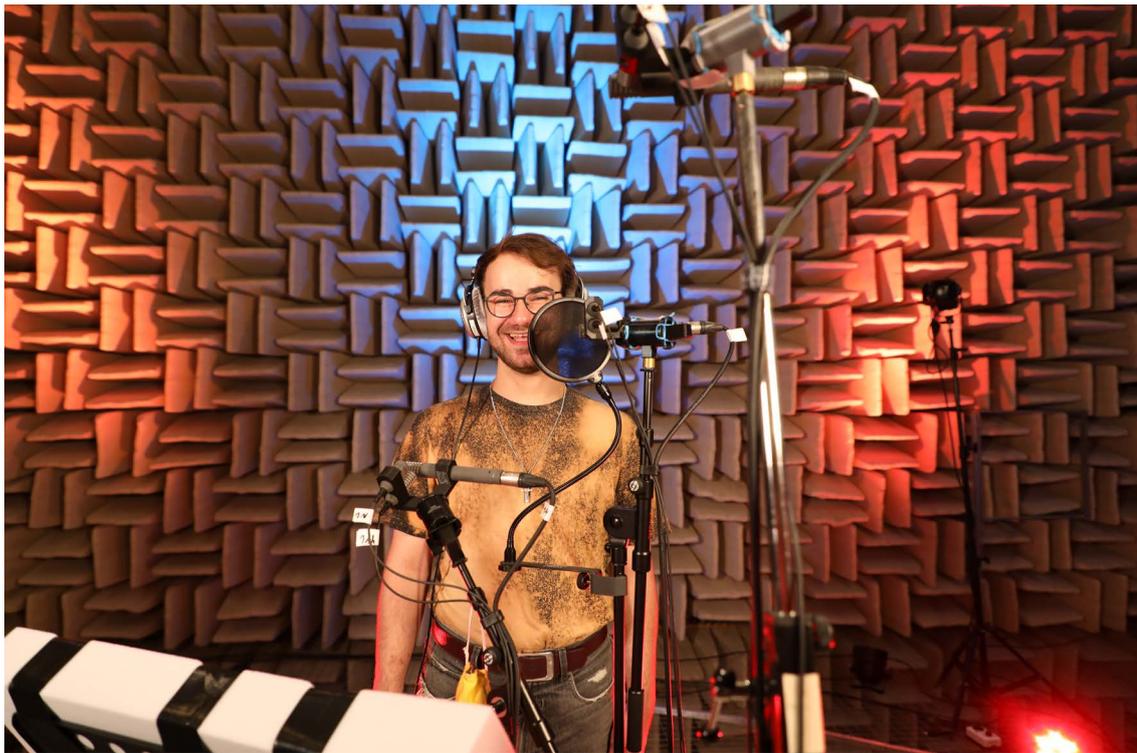
## A.2 Männlicher Gesang



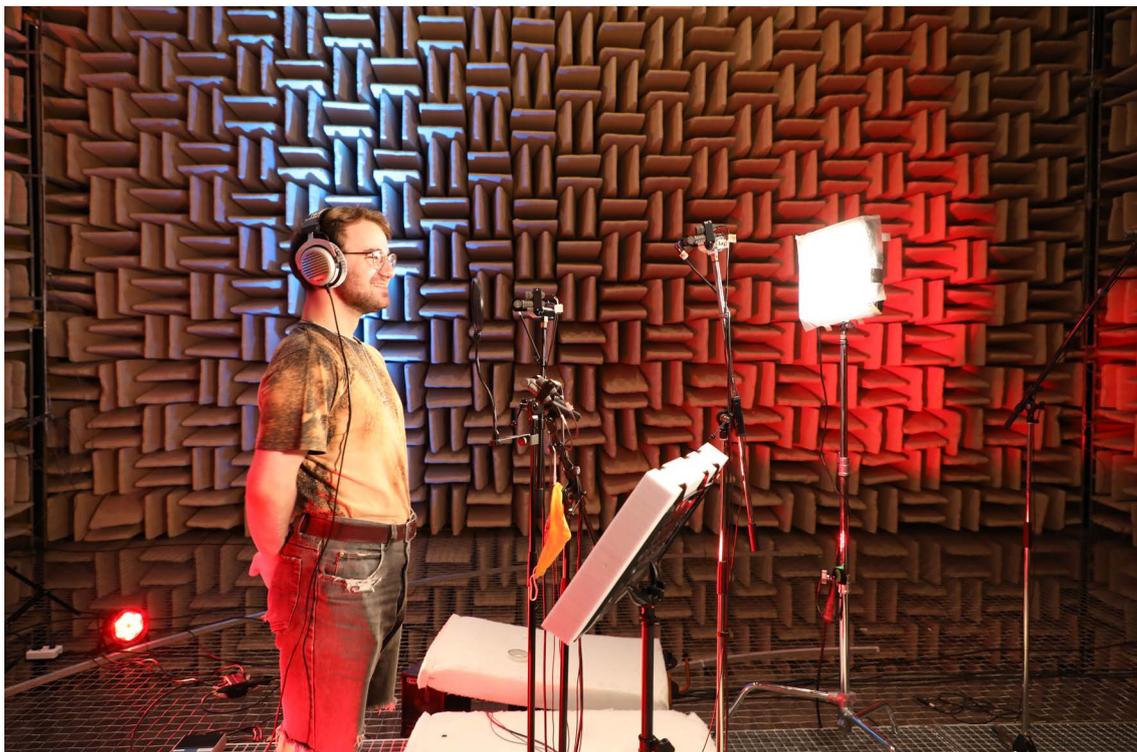
(a) Azimut  $\phi \approx 135^\circ$



(b) Azimut  $\phi \approx 70^\circ$



(c) Azimut  $\phi \approx 0^\circ$



(d) Azimut  $\phi \approx -90^\circ$

**Abbildung A.2:** Bilder männlicher Gesang.

### A.3 Querflöte



(a) Azimut  $\phi \approx 90^\circ$



(b) Azimut  $\phi \approx 45^\circ$



(c) Azimut  $\phi \approx -10^\circ$



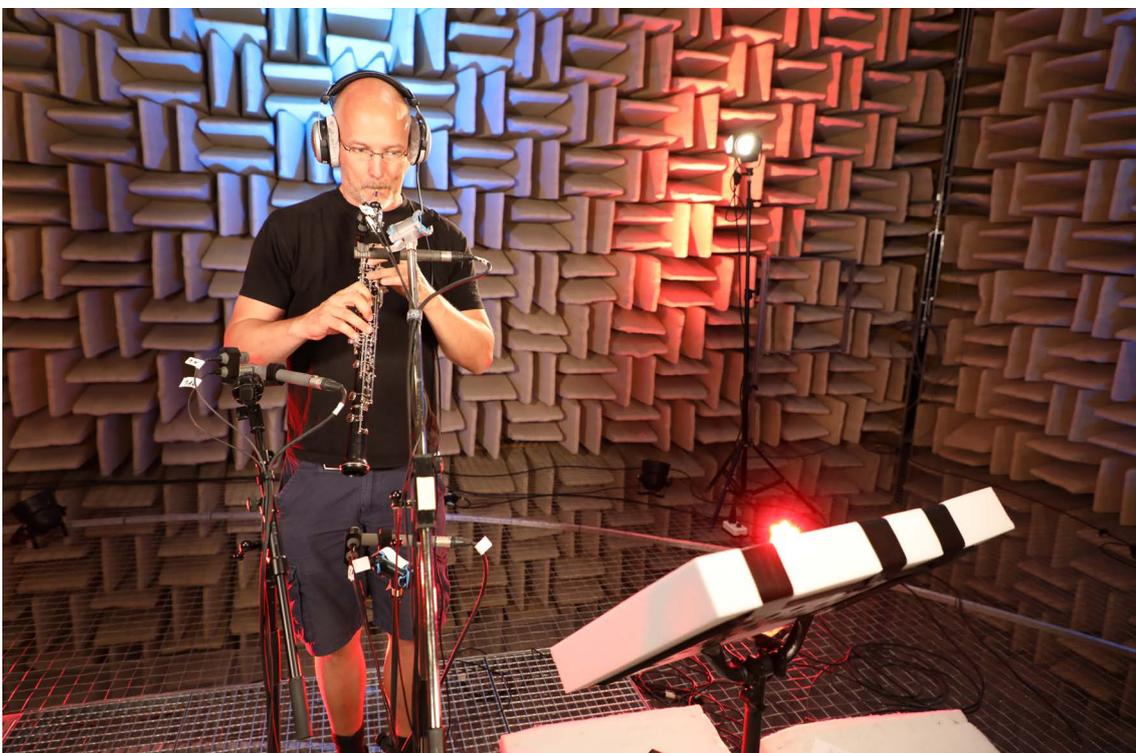
(d) Azimut  $\phi \approx -90^\circ$

**Abbildung A.3:** Bilder Querflöte.

## A.4 Oboe



(a) Azimut  $\phi \approx 45^\circ$



(b) Azimut  $\phi \approx 0^\circ$



(c) Azimut  $\phi \approx -90^\circ$



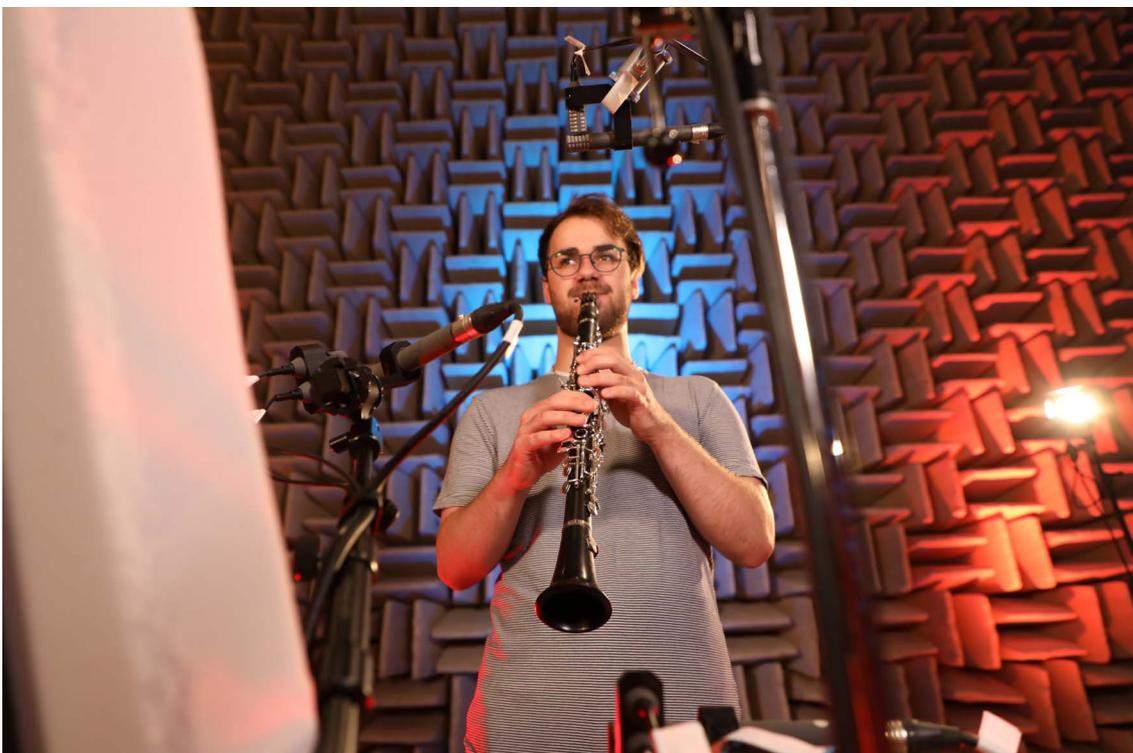
(d) Azimut  $\phi \approx -135^\circ$

**Abbildung A.4:** Bilder Oboe.

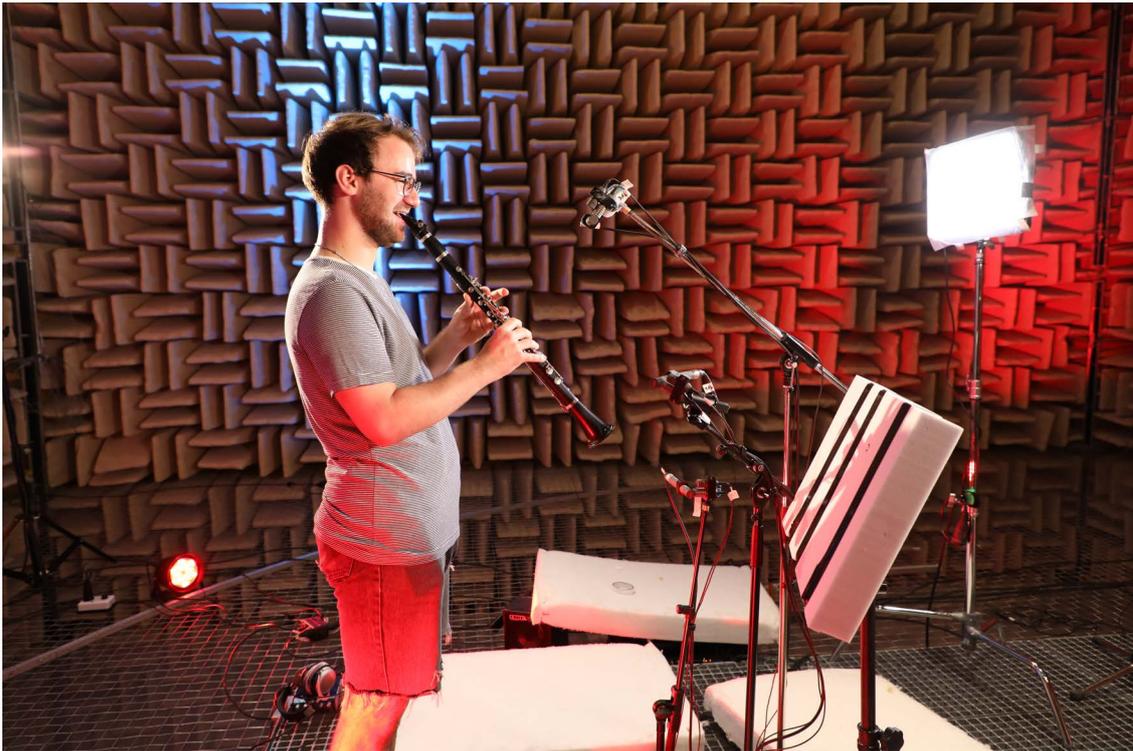
## A.5 Klarinette



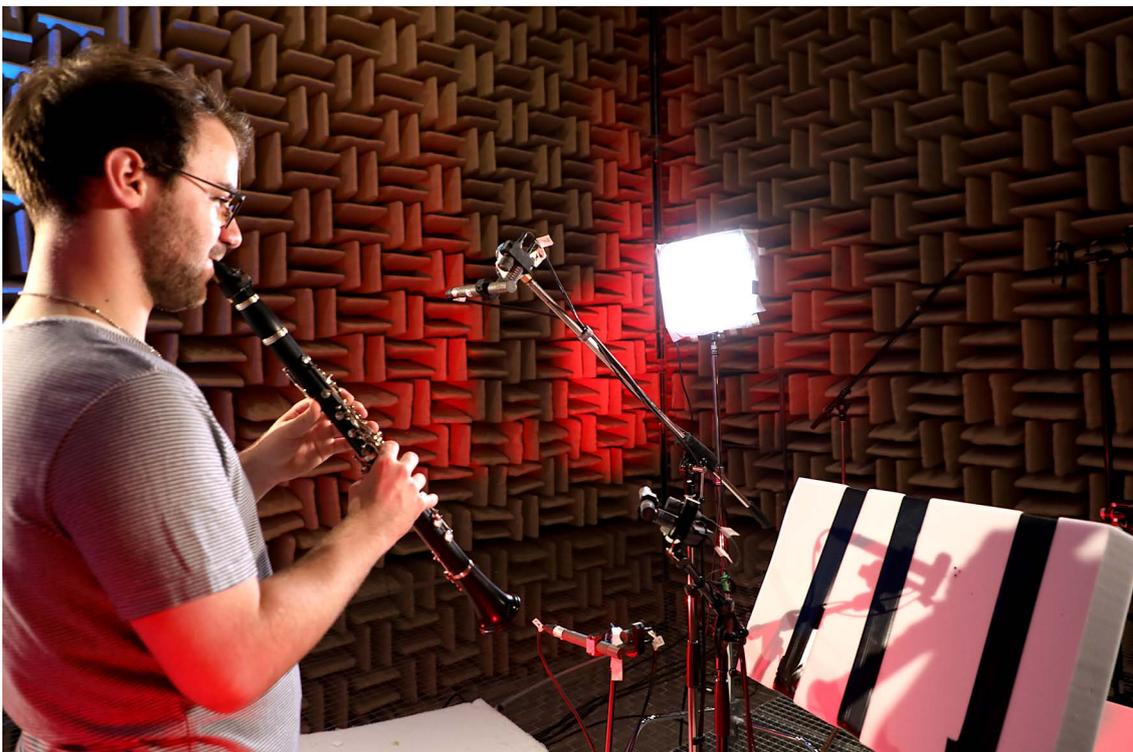
(a) Azimut  $\phi \approx 90^\circ$



(b) Azimut  $\phi \approx 0^\circ$



(c) Azimut  $\phi \approx -90^\circ$



(d) Azimut  $\phi \approx -120^\circ$

**Abbildung A.5:** Bilder Klarinette.

## A.6 Tenorsaxofon



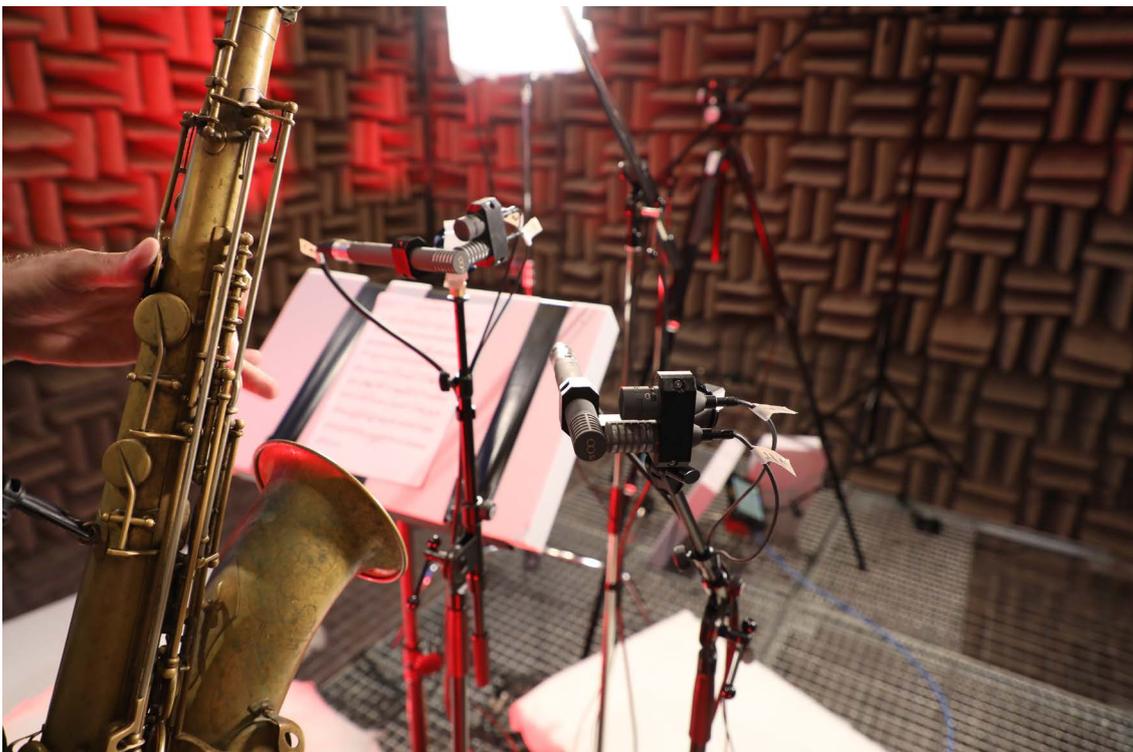
(a) Azimut  $\phi \approx 90^\circ$



(b) Azimut  $\phi \approx 0^\circ$



(c) Azimut  $\phi \approx -90^\circ$



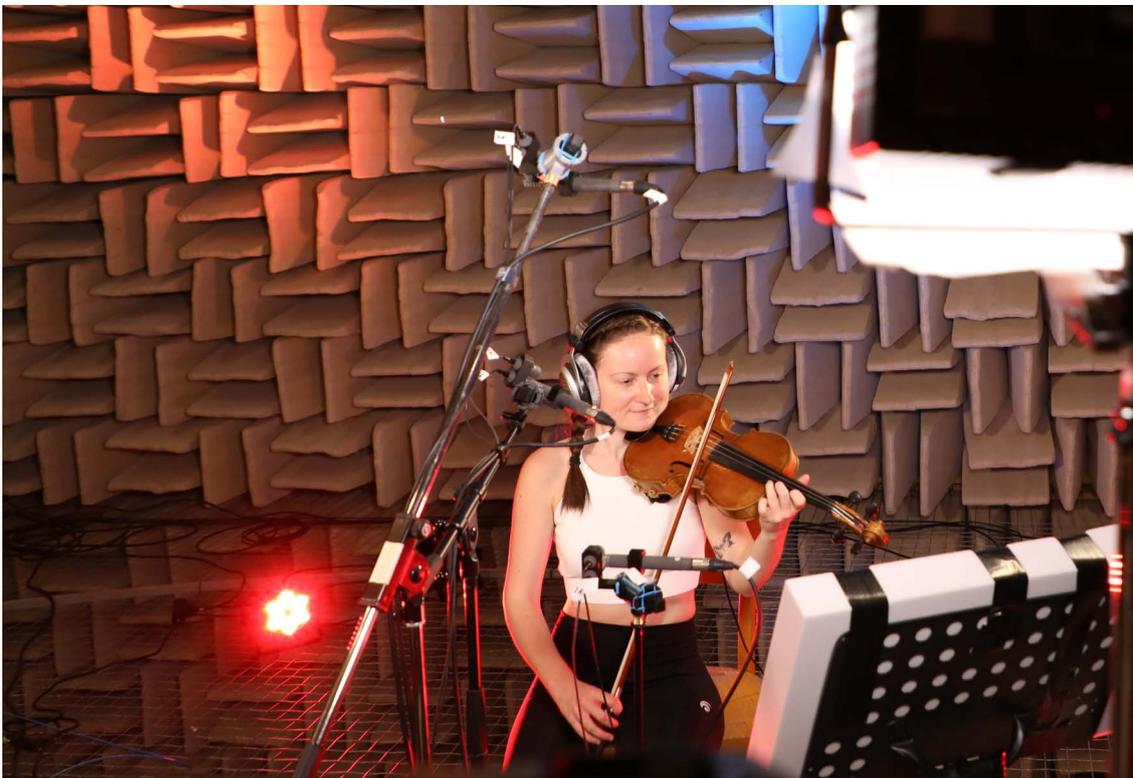
(d) Azimut  $\phi \approx -135^\circ$

**Abbildung A.6:** Bilder Tenorsaxofon.

## A.7 Violine



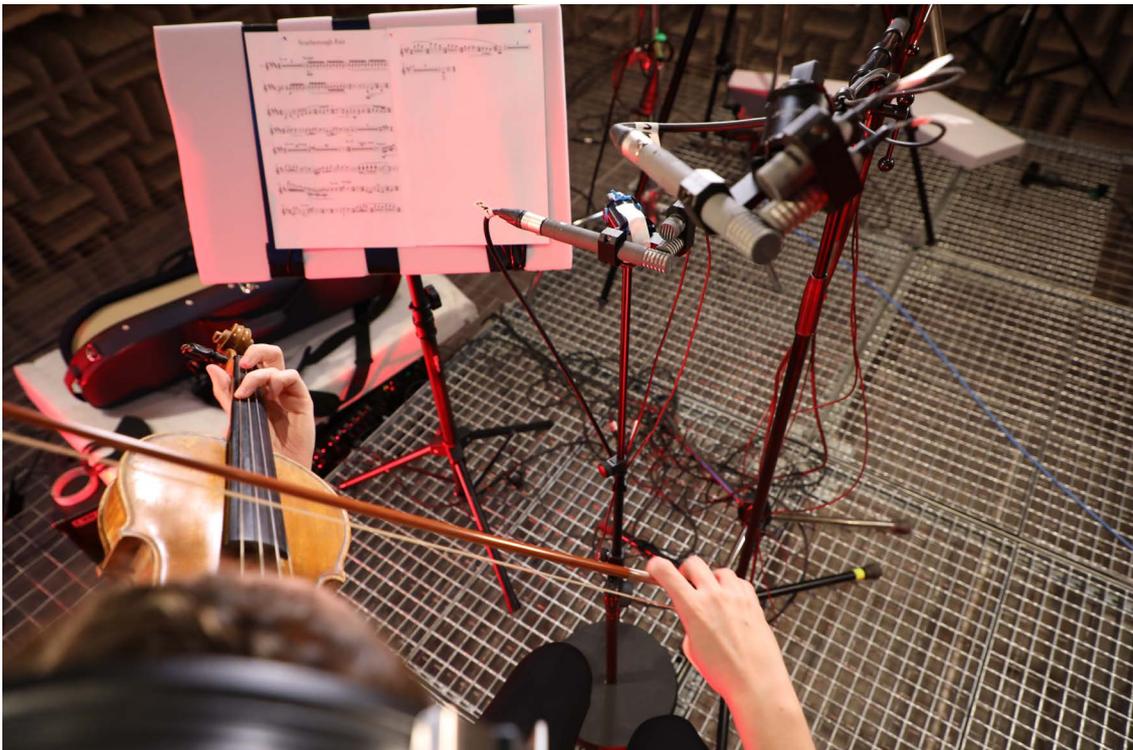
(a) Azimut  $\phi \approx 75^\circ$



(b) Azimut  $\phi \approx 20^\circ$



(c) Azimut  $\phi \approx -90^\circ$



(d) Azimut  $\phi \approx -150^\circ$

**Abbildung A.7:** Bilder Violine.

## A.8 Kontrabass



(a) Azimut  $\phi \approx 135^\circ$



(b) Azimut  $\phi \approx -45^\circ$



(c) Azimut  $\phi \approx -90^\circ$

**Abbildung A.8:** Bilder Kontrabass.

## A.9 Gitarre



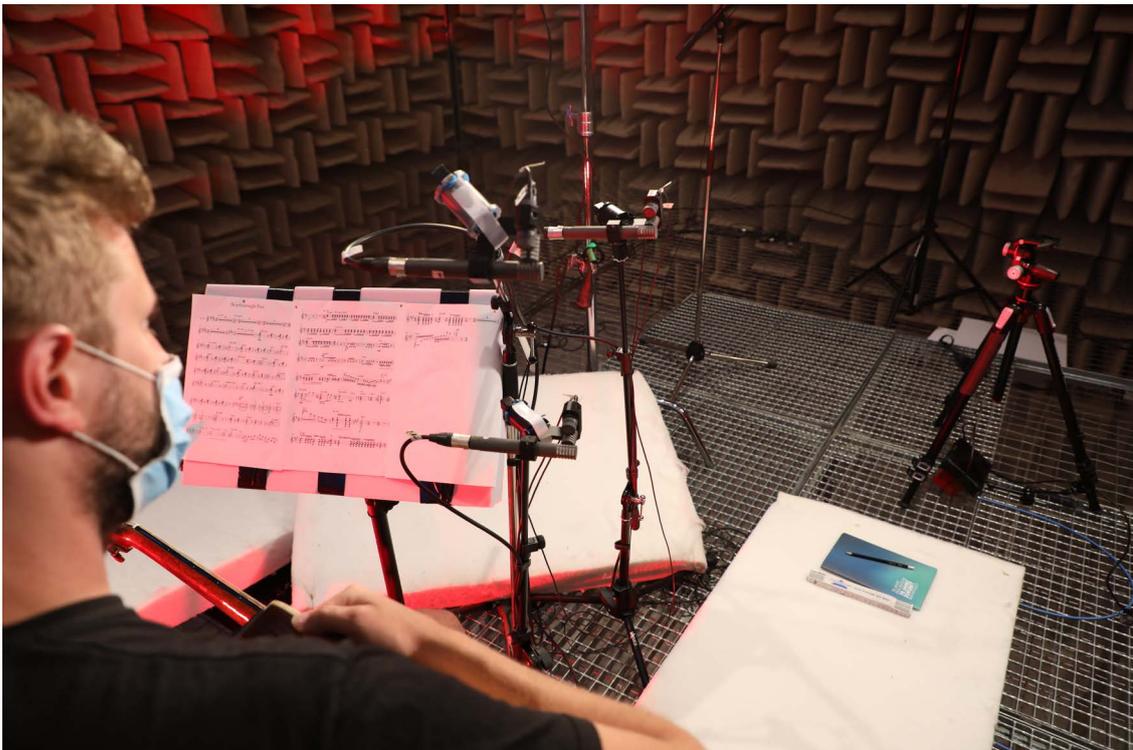
(a) Azimut  $\phi \approx 75^\circ$



(b) Azimut  $\phi \approx -45^\circ$



(c) Azimut  $\phi \approx -90^\circ$



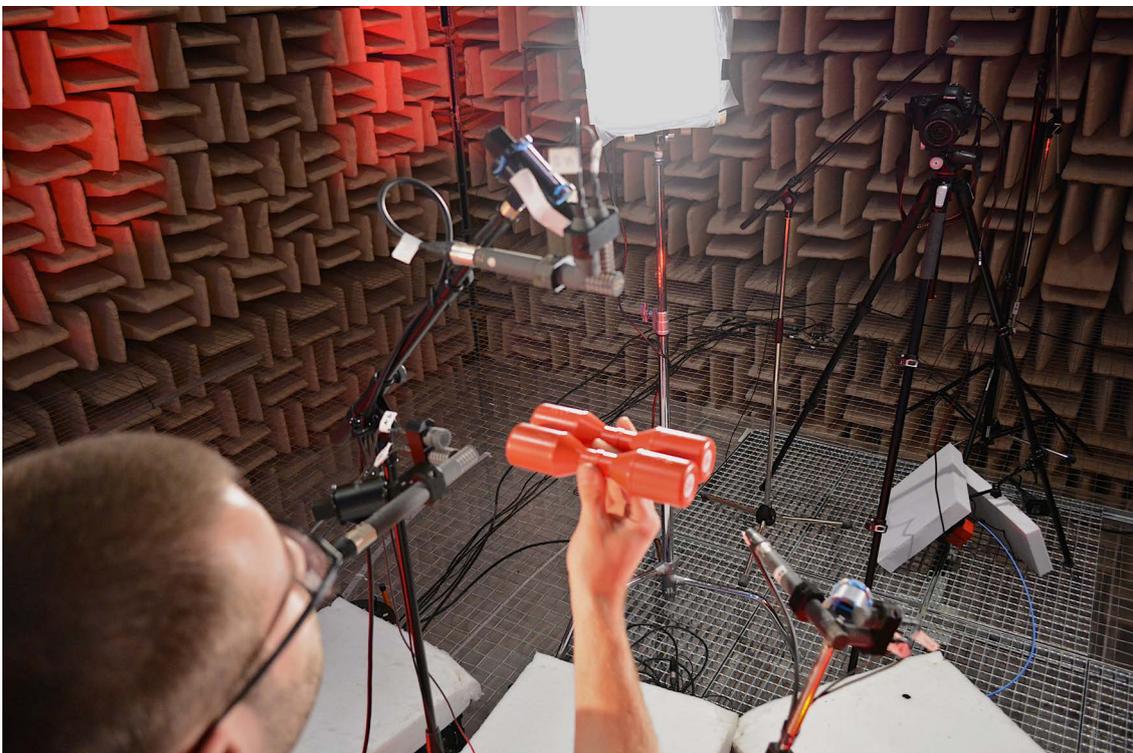
(d) Azimut  $\phi \approx -150^\circ$

**Abbildung A.9:** Bilder Gitarre.

## A.10 Shaker



(a) Azimut  $\phi \approx -45^\circ$



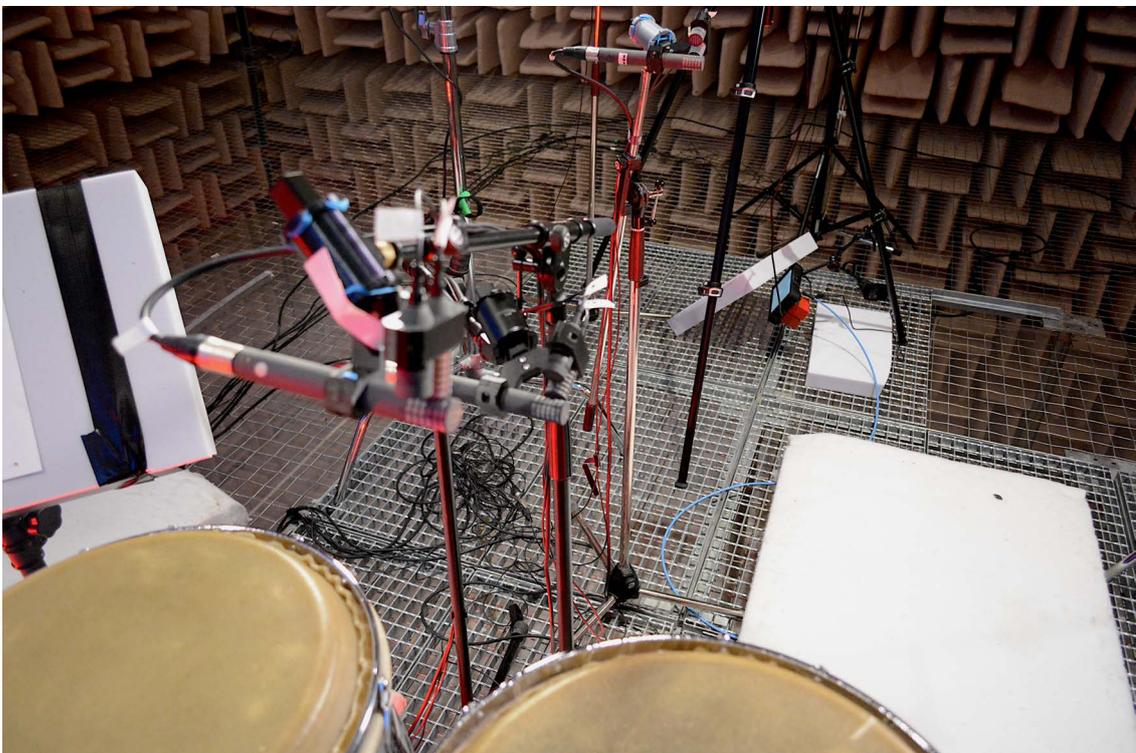
(b) Azimut  $\phi \approx -150^\circ$

**Abbildung A.10:** Bilder Shaker.

## A.11 Congas



(a) Azimut  $\phi \approx 0^\circ$



(b) Azimut  $\phi \approx -150^\circ$

Abbildung A.11: Bilder Congas.

## A.12 Schlagzeug



(a) Azimut  $\phi \approx 150^\circ$



(b) Azimut  $\phi \approx -20^\circ$

Abbildung A.12: Bilder Schlagzeug.

# Anhang B: Datenblätter

## B.1 Freifeldraum



### Akustischer Freifeldraum P22

<b>Messgröße</b>	Schalleistung (Hüllflächenverfahren), Schallintensität, Richtcharakteristik, Lokalisierung von Schallquellen
<b>Norm</b>	DIN EN ISO 3745
<b>Messobjekte</b>	Maschinen, Geräte und Anlagen, z.B. der Lüftungs- und Klimatechnik, sowie andere Geräuschquellen

#### Technische Daten

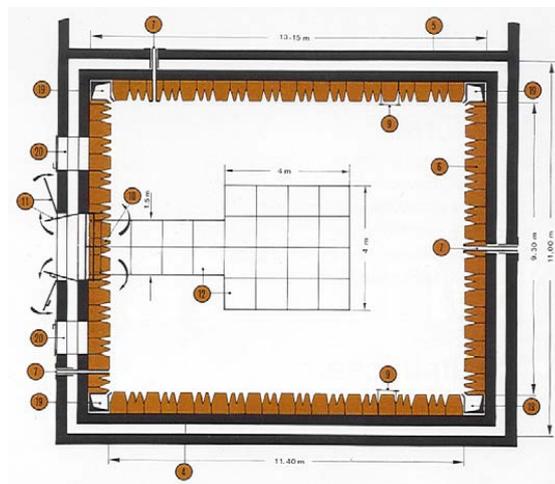
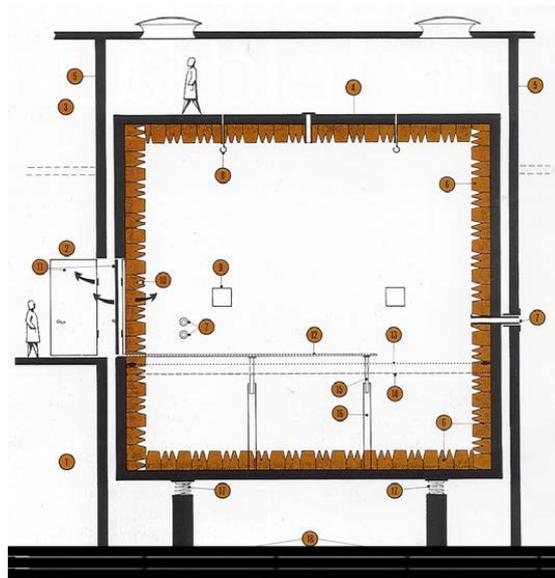
Freie Raumgröße (L x B x H)	9,3 m x 10,4 m x 10,3 m
Raumvolumen	1090 m <sup>3</sup>
Tür-Durchgangsmaße	2,34 m x 1,89 m
Demontierbares Montagegitter als Arbeitsebene	4,0 m x 4,0 m Fläche
Belastbarkeit des Montagegitters	5000 N/m <sup>2</sup>

#### Weitere Informationen

- Untere Grenzfrequenz 80 Hz
- Elastische Lagerung des Raumes und separates Fundament zur Schwingungsisolierung
- Raum-in-Raum Konstruktion mit sehr niedrigem Fremdgeräusch
- Extrem leise Lüftungsanlage (< 25 dB/Okt.)
- Stahlseilnetz und Gitterrost für den Einbau schwerer Gegenstände, für die Messungen demontierbar.
- Durchführungen für Mess- bzw. Versorgungsleitungen vorhanden
- Druckluft- und Elektroanschluss vorhanden

Der Freifeldraum ist durch seine Größe auch für große Schallquellen und große Messabstände besonders geeignet. Auch die niedrige Grenzfrequenz von 80 Hz sowie die sehr niedrigen Fremdgeräusche im Raum stellen eine Besonderheit des Prüfstands dar.

### Schnitt und Grundriss des Prüfstands (Maße in m)



1 Keller	5 Vorsatzwände	9 Befestigungsrahmen	13 Gehnetz	17 Stahlfedern
2 Erdgeschoss	6 Akustik-Keile	10 Auskleidung	14 Fangnetz	18 getrennte Fundamente
3 Obergeschoss	7 Durchführungen	11 Türen	15 Adapter f. Gitterrost	19 Lüftungskanäle
4 Baukörper	8 Deckenhaken	12 Gitterrost	16 Stützen f. Gitterrost	20 große Durchführungen

Das Prüflabor des Fraunhofer IBP ist nach DIN EN ISO/ IEC 17025 durch die Deutsche Akkreditierungsstelle GmbH (DAkkS) akkreditiert.

## B.2 Mikrofone

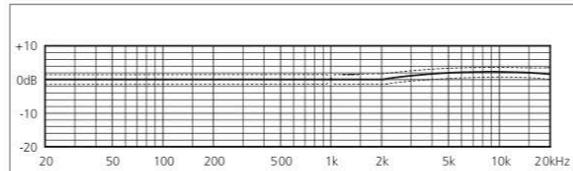


### CCM 5

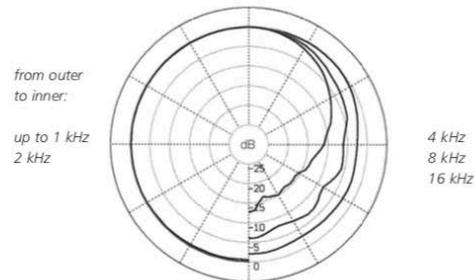
- mechanically switchable single-diaphragm microphone (omni/ cardioid)
- smoother, more extended high-frequency response than most other multi-pattern microphones (e.g. dual-diaphragm microphones of other manufacturers)
- slightly brighter than the CCM 2H (omni) or CCM 4 (cardioid)
- a pure pressure transducer when in the "omni" setting (flat, extended low-frequency response without proximity effect or undue sensitivity to wind or solid-borne sound)

#### Uses:

preferred uses similar to those of the CCM 2 or CCM 2S and the CCM 4:  
 In the cardioid setting: for use with singing or speaking voices or most instruments, as a spot microphone, and for stereo recording with coincident, ORTF or M/S microphone arrangements.  
 In the omnidirectional setting: for recording instruments, singers, etc. at relatively close range

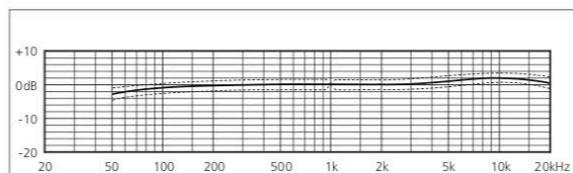


Frequency response curve CCM 5 "omni" position

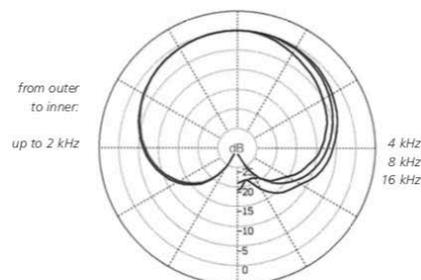


Polar diagram CCM 5

"omni" position



Frequency response curve CCM 5 "cardioid" position



Polar diagram CCM 5

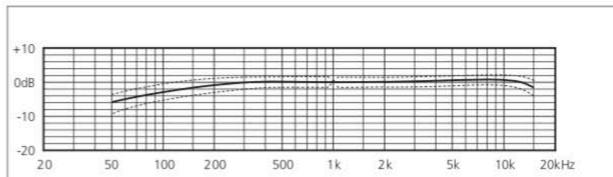
"cardioid" position



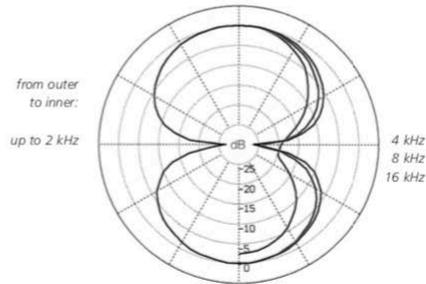
**CCM 8**

- figure-8 ("bidirectional") pattern
- clear sound quality, free of coloration
- microphone for M/S and Blumlein stereo
- highly consistent frequency and polar response
- response essentially free of off-axis peaks like a good ribbon microphone (but not as delicate physically)
- lateral pickup

Uses: optimal for M/S and Blumlein stereo recording



Frequency response curve CCM 8

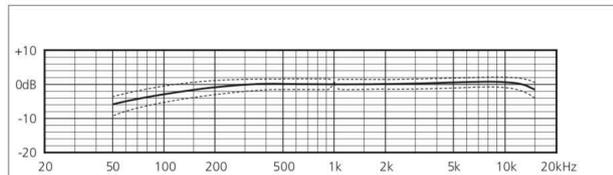


Polar diagram CCM 8

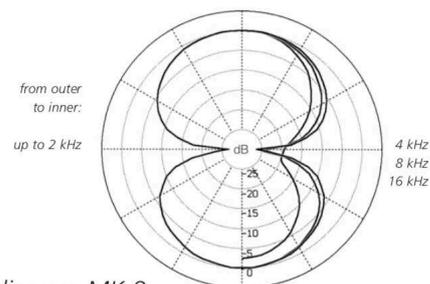


**MK 8**

- figure-8 ("bidirectional") pattern
- clear sound quality, free of coloration
- for M/S and Blumlein stereo
- highly consistent frequency and polar response
- response essentially free of off-axis peaks like a good ribbon microphone (but not as delicate physically)
- lateral pickup



Frequency response curve MK 8 + CMC 6



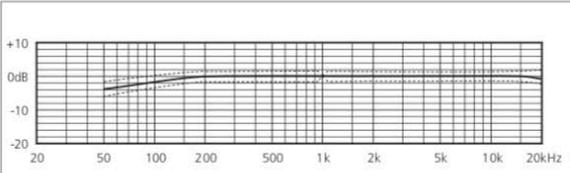
Polar diagram MK 8



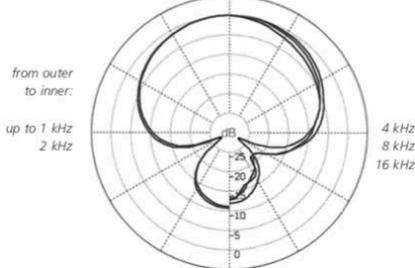
**CCM 41**

- all-purpose microphone for speech and music recording of all kinds
- well suited for use as the main microphones for stereo pickup and/or as "spot" microphones
- extended, smooth, well-balanced frequency response
- often used for film and video sound
- where it can be used, it has distinct sonic and practical advantages over most shotgun microphones
- highly consistent polar response
- 0° axis is at the microphone's tip

Uses: often preferred for use in film sound recording and as a spot microphone in orchestras



*Frequency response curve CCM 41*



*Polar diagram CCM 41*

microphone type	polar pattern	frequency range	sensitivity	equivalent noise level		signal-to-noise ratio A-weighted	max. SPL (0,5%THD)
				CCIR	A-weighted		
CCM 41	supercardioid	40 Hz – 20 kHz	14 mV/Pa	24 dB	15 dB	79 dB	132 dB
CCM 41V	supercardioid	40 Hz – 20 kHz	14mV/Pa	23 dB	14 dB	80 dB	132 dB
CCM 8	figure-8	40 Hz – 16 kHz	10 mV/Pa	26 dB	18 dB	76 dB	134 dB
CCM 5	omni	20 Hz – 20 kHz	10 mV/Pa	26 dB	14 dB	80 dB	133 dB
	cardioid	40 Hz – 20 kHz	13 mV/Pa	24 dB	15 dB	79 dB	132 dB
MK 8	figure-8	40 Hz – 16 kHz	10 mV/Pa	26 dB	18 dB	76 dB	134 dB

**Abbildung B.1:** Technische Daten Mikrofone.  
(SCHOEPS, 2020a, SCHOEPS, 2020b)

# Anhang C: Partitur Komposition

## Scarborough Fair

Arranged by Leon Hofmann

(♩ = 60) **Intro** Straight

Female Vocals

Male Vocals

Flute

Oboe

Clarinet

Tenor Sax.

Violin

Double Bass

Ac. Guitar

legato

pizz.

pp

pp

pp

ppp

pizz.

7

Voc. Fe.

Voc. Ma.

Fl. *legato*  
*p* *mf*

Ob. *p* *mf*

Cl. *p* *mf*

Sax. *p* *mf*

Vi. *p*

Bass

Git.

10

Voc. Fe. *mp*  
Are you going to Scar - bo-rough fair? Pars - ley, sa - ge, rose - ma - ry and

Voc. Ma.

Fl. *normal*  
*mp*

Ob.

Cl.

Sax.

Vi. *normal*  
*mp*

Bass *normal*  
*mp*

Git.

2

16

Voc. Fe. *thyme. Re - mem - ber me to the one who lives there For he once was*

Voc. Ma.

Fl.

Ob.

Cl.

Sax.

Vl. *mf*

Bass

Git.

---

21 **Verse 1** (♩ = 144) Swing

Voc. Fe. *for he once was a true love of mine*

Voc. Ma.

Fl.

Ob.

Cl.

Sax.

Vl.

Bass

Git.

3

27

Voc Fe.

Voc Ma.

Fl.

Ob.

Cl.

Sax.

VI.

Bass

Walking Bass

pizz.

*mf*

Git.

*mp*

C#m7 (9)

A6 (add9)

G#7 (#9)

C#m9

---

33

Voc Fe.

Voc Ma.

*mp*

Tell her \_\_\_\_\_ to make me a cam - br - ic shirt

Fl.

Ob.

Cl.

Sax.

VI.

Bass

Git.

F#9

G#7 (#9)

C#m9

A6 (add9)

F#/G#

4

39

Voc. Fe. \_\_\_\_\_

Voc. Ma. *Pars - ley, sa - ge, rose - ma - ry and thyme. With - out a - ny scam or*

Fl. \_\_\_\_\_

Ob. \_\_\_\_\_ *mp*

Cl. \_\_\_\_\_

Sax. \_\_\_\_\_

VI. \_\_\_\_\_

Bass \_\_\_\_\_

Git. *C#m9 F#9 G#7 (#9) C#m9 Emaj7*

---

45

Voc. Fe. \_\_\_\_\_

Voc. Ma. *ne - c - dle work And then she'll be *mf* a true love of \_\_\_\_\_*

Fl. \_\_\_\_\_

Ob. \_\_\_\_\_

Cl. \_\_\_\_\_

Sax. \_\_\_\_\_

VI. \_\_\_\_\_

Bass \_\_\_\_\_

Git. *B (maj7) Amaj7 (9) G#maj7 add11 F#m9 G#7 (#9)*

5

52 **Verse 2**

Voc Fe. *mf*  
Ask him to find me an a - cre of

Voc Ma. *mp*  
mine

Fl.

Ob.

Cl.

Sax.

Vi.

Bass *mp* *mf*

Git. *mp* *mf* *Bm7* *A*

---

59

Voc Fe. *mf*  
land Pars - ley, sa - ge, rose - ma - ry and thyme Be -

Voc Ma. *mp*  
sa -

Fl.

Ob.

Cl. *mf*

Sax.

Vi.

Bass

Git. *G* *Bm7* *E7* *F#7 (#9)*

6

66

Voc. Fe. *tween the salt - wa - ter and the sea - sand For then he'll be*

Voc. Ma.

Fl.

Ob.

Cl. *mp* *vib.*

Sax. *mp*

VI.

Bass

Git. *Bm9 Dmaj7 A (maj7) Gmaj7 (9) F#maj7 add11*

73 *f* *a true love of mine* *Solopart* *f*

Git. *Em9 F#7 (#9) Bm* *let ring...*

(♩ = 72)

7

79 Swing 16ths **Solo Part**

Voc. Fe.

Voc. Ma.

Fl.

Ob.

Cl.

Sax. *vib.* *f* *normal* *legato* 3 3

VI.

Bass

Git. *f* Gmaj7 Dmaj7

---

83

Voc. Fe.

Voc. Ma.

Fl. *Solopart* *mf*

Ob.

Cl.

Sax. *^* *^* *vib.* *Straight* *vib.*

VI.

Bass

Git. Gmaj7 (9) Dadd9

87

Voc. Fe.

Voc. Ma.

Fl.

Ob.

Cl.

Sax.

VI. *Solopart*

Bass *pizz.*

Git. *G Bm*

*mp* *p* *mf*

91

Voc. Fe.

Voc. Ma.

Fl.

Ob.

Cl.

Sax.

VI.

Bass

Git. *Gadd9 Bm9 Gmaj7/A*

*mf*

9

95 *rit.* (♩ = 69) **Straight Instrumental Theme**

Voc. Fe.

Voc. Ma.

Fl.

Ob.

Cl.

Sax.

VI.

Bass

Git. *Gadd9* *Aadd9* *Em* *Bm/D#*  
*let ring - - - - -*  
*mit den Fingern*  
*mp* *let ring - - - - -*

---

100 *rit.* *accel.* (♩ = 69)

Voc. Fe.

Voc. Ma.

Fl. *mp*

Ob. *mf*

Cl. *legato* *mf*

Sax.

VI.

Bass

Git. *D6 (add9)* *C#m7* *Cmaj7* *Bm7* *A sus2* *let ring - - - - -*  
*8va* *Gmaj7*

105 **Verse 3**

Voc. Fe. *f* Love im-po-ses im - pos - si - ble tasks Pars - ley sa - ge, rose - ma - ry and

Voc. Ma. *f* Love im-po-ses im - pos - si - ble tasks Pars - ley sa - ge, rose - ma - ry and

Fl. - - - - -

Ob. - - - - -

Cl. *ff* *mf*

Sax. *ff* *mf*

Vl. *f* legato

Bass *ff* normal

Git. *f* *F#m* normal *Dsus2* *Asus2* *B*

---

110

Voc. Fe. thyme But none more than a - ny heart would

Voc. Ma. thyme love im - po - ses some tasks heart would

Fl. *f*

Ob. - - - - -

Cl. *f* vib.

Sax. - - - - -

Vl. - - - - -

Bass - - - - -

Git. *F#m7 (add 11)* *F#m* *A* *B9/D#*

114

Voc. Fe. ask I must know *mp* oh I must know

Voc. Ma. ask, none more than a - ny heart would ask *mp* I must know I must know

Fl. *mp*

Ob. *mp*

Cl. *mp* vib.

Sax. *mp*

VI. *mp*

Bass. *mf*

Git. E

---

119

Voc. Fe. *rit.* *p* you're a true love of mine

Voc. Ma. *p* you're a true love of mine

Fl. *tr.*

Ob. *pizz.*

Cl. *pizz.*

Sax. *pizz.*

VI. *pizz.*

Bass. *pizz.*

Git. *p* *Fm11* let ring... *8va* *Fm*

12

## Anhang D: Datenträger

Der Arbeit liegt ein elektronischer Datenträger bei, dessen Inhalt im Folgenden kurz beschrieben wird. Die Benennung ist aufgrund der vielfältigen Weiterwendungsmöglichkeiten englisch.

- Vorliegende Masterarbeit als PDF-Dokument:  
*Masterthesis\_Leon\_Hofmann.pdf*
- Stereo-Playback Demo des arrangierten Musikstücks:  
*Playback\_Scarborough\_Fair.mp3*
- Verwendete Paper, Standards und Handbücher als PDF-Dokumente
- Dokumentierende Bilder zu Aufnahmeraum und Mikrofonpositionen
- Exportierte Videodaten aus Messungen mit akustischer Kamera in der Form:  
*Instrument\_Blickwinkel(\_Spieltechnik)\_Frequenzbereich(\_Geschwindigkeit).avi*