Technische Universität Berlin
Fakultät I: Institut für Sprache und Kommunikation
Fachgebiet Audiokommunikation

Magisterarbeit

Simulation einer musikalischen Aufführung durch die Auralisation von Musikinstrumenten mit korrekter Richtcharakteristik

Roman Schneider

[redacted]

Erstgutachter: Prof. Dr. Stefan Weinzierl

Zweitgutachter: Dr. Hans-Joachim Maempel

Die selbständige und eigenhändige Anfertigung dieser Arbeit versichere ich an Eides statt.

Berlin, den 29.09.2011

…….

Roman Schneider

Inhaltsverzeichnis

1 Einleitung 1

2 Die nachhallfreien Musikaufnahmen 5
2.1 „Contredanse“, Pascale Gautier Quartett 5
2.2 „River Man“, Nick Drake 6

3 Simulation der Richtcharakteristiken 9
3.1 Mittelung und Filterung 11
3.1.1 Kompensation der Quellverschiebung 14
3.2 Interpolation 21
3.2.1 Lautsprecherdaten in EASE 22
3.2.2 Interpolationsmethoden 23
3.2.3 Vergleich und Diskussion der Ergebnisse 36
3.2.4 Konvertierung der CATT-Daten des Flügels in das EASE-Format 38
3.3 Bestimmung der Bezugsrichtung 41
3.3.1 Ermittlung der Bezugsrichtung für zwei Musikaufnahmen 43
3.4 Bestimmung der Kennwerte und Importieren der Daten in EASE 52
3.4.1 Lautsprecherkenndaten 53
3.4.2 Richtfaktor 54
3.4.3 Wirkungsgrad 57
3.4.4 Messwerte 58
3.4.5 Bearbeitung der Lautsprecherdaten in EASE 59
3.5 Einfluss der Bezugsrichtung auf den Frequenzgang 60
3.6 Auswertung der Richtcharakteristiken 63
3.6.1 Richtcharakteristik der Singstimme 66
3.6.2 Allgemeine Fehlerbetrachtung und Diskussion 75

4 Auralisation 77
4.1 Der Raum 77
4.2 Hörerposition und Schallquellen – Einfluss der Richtcharakteristik auf den Hallabstand 80
4.3 Rendering 86
4.3.1 Schallfeldberechnung in EASE 86
4.3.2 Erzeugung der dynamischen BRIRs 90
4.4 Binaurale Mischung 91
4.5 Diskussion 92

5 Zusammenfassung 97

Literaturverzeichnis 99

Abbildungsverzeichnis 102

Tabellenverzeichnis 105

Anhang A: Inhalt der DVD 106

Anhang B: Patchtabelle 107

Anhang C: Quellcodes 108

Abkürzungen

ANSI	American National Standards Institute
AURA	Analysis Utility for Room Acoustics (EASE-Modul)
BRIR	Binaural Room Impulse Response
CAD	Computer Aided Design
CLF	Common Loudspeaker Format
DIN	Deutsches Institut für Normung
EASE	Enhanced Acoustic Simulator for Engineers (Akustische Simulationssoftware von Acoustic Design Ahnert)
Fachgebiet AK	Fachgebiet Audiokommunikation des Instituts für Sprache und Kommunikation an der Technischen Universität Berlin
HRTF	Head Related Transfer Function
IR	Impulsantwort
ITA	Institut für Technische Akustik
LF	Lateral Fraction
PTB	Phsyikalisch-Technische Bundesanstalt
RAR	Reflexionsarmer Raum
RIR	Room Impulse Response
RWTH Aachen	Rheinisch-Westfälische Technische Hochschule Aachen
S/R	Signal-Rauschabstand
TKK	Helsinki University of Technology
TU	Technische Universität Berlin
UDK	Universität der Künste Berlin
VP	Versuchsperson
WAV	Waveform Audio File Format

1 Einleitung

Die ständige Weiterentwicklung von Audiowiedergabetechnologien hat unter anderem zum Ziel, dem Hörer auf verbesserte Weise den Klangeindruck von realen akustischen Ereignissen zu vermitteln. Ihren derzeitigen Höhepunkt erreicht diese Entwicklung in den Verfahren der virtuellen Akustik wie der Binaural- und Wellenfeldsynthese. Das vom Fachgebiet Audiokommunikation durchgeführte Forschungsprojekt „Medium und Emotion“[1] untersucht den Einfluss der Rezeptionssituation auf die emotionale Wirkung von Musik. Dabei wird unter anderem die weithin verbreitete Vermutung untersucht, dass qualitativ hochwertige Übertragungsverfahren den Grad der Immersion in ein mediales Ereignis steigern und dadurch den emotionalen Gehalt einer musikalischen Werkaufführung erfahrbarer machen können. Es mag aber genauso möglich sein, dass bestimmte technische Wiedergabeverfahren, ganz losgelöst von der Frage einer „authentischen“ Reproduktion eines Live-Konzerts, einen spezifischen Beitrag zur Emotionalität der Medienrezeption liefern. Die dynamische Binauralsynthese macht es möglich, jenen Fragen auf den Grund zu gehen, indem traditionelle Wiedergabeverfahren der stereofonen Lautsprecher- und Kopfhörerwiedergabe mit den akustischen Rezeptionsbedingungen bei einer Live-Aufführung experimentell verglichen werden. In einem Hörversuch werden diese drei Wiedergabetechnologien auf ihre emotionale Wirkung untersucht. Für diesen Hörversuch wird in der vorliegenden Arbeit, die Teil dieses Forschungsprojekts ist, die Simulation einer Konzertaufführung durch die dynamische Binauralsynthese eines Konzertsaals durchgeführt. Um dem hohen Anspruch einer möglichst realgetreuen Reproduktion der akustischen Bedingungen einer Live-Aufführung gerecht zu werden, müssen sämtliche wahrnehmungsrelevanten Klangerzeuger- und Übertragungseigenschaften hinreichend simuliert werden. Der Ansatz meiner Arbeit ist, nicht nur die akustischen Raumeigenschaften zu simulieren, sondern auch die räumlichen Abstrahleigenschaften von Musikinstrumenten in einer dynamischen Binauralsynthese hörbar zu machen. Die Durchführung und Auswertung des Hörversuchs ist nicht mehr Bestandteil dieser Arbeit.

Die dynamische Binauralsynthese ist ein Wiedergabeverfahren der virtuellen Akustik. Binaurale Wiedergabesysteme basieren auf der Tatsache, dass alle wesentlichen Informationen über die akustische Umgebung in dem am Trommelfell beider Ohren anliegenden Schalldruck enthalten sind. Um die Schalldrucksignale am Trommelfell des Hörers realistisch nachzubilden, müssen alle relevanten Übertragungseigenschaften auf dem

[1] Untertitel: „Zur emotionalen Wirkung von Musik in verschiedenen medialen Rezeptionssituationen.“ Forschungsprojekt im Exzellenzcluster „Languages of Emotion“

Weg von der Quelle bis zum Ende des äußeren Gehörgangs in einem binauralen Signal enthalten sein. Die Übertragung eines Schallereignisses zum Hörer in einem Raum kann als ein lineares, zeitinvariantes System betrachtet werden. In der Impulsantwort eines solchen Systems sind sämtliche Übertragungseigenschaften enthalten. Eine reale akustische Umgebung wird mit der Binauraltechnik durch die Messung der Raumantwort eines impulsförmig angeregten Lautsprechers mit einem Kunstkopf simuliert. Wird ein Raum im Computermodell simuliert, ergibt sich die binaurale Impulsantwort aus den berechneten Übertragungseigenschaften des Modells mit Berücksichtigung der gemessenen Außenohrübertragungsfunktion (engl. head related transfer function, im Folgenden: HRTF) des Menschen. Die Faltung einer binauralen Raumimpulsantwort (engl. binaural room impulse response, im Folgenden: BRIR) mit trocken aufgenommenen Schallsignalen ermöglicht über geeignet entzerrte Kopfhörer die realistische Wiedergabe der Signale in der durch die BRIR charakterisierten akustischen Umgebung. Wissenschaftliche Untersuchungen haben ergeben, dass bei der Lokalisation von Schallquellen spontane Kopfbewegungen eine entscheidende Rolle spielen. Um dieser Tatsache gerecht zu werden, wurden Abhörsysteme entwickelt, die die horizontalen Kopfdrehungen des Hörers erfassen und die Faltung durch um die veränderte Außenohrübertragungsfunktion korrigierte BRIRs anpasst. Ein solches Wiedergabeverfahren wird als dynamische Binauralsynthese bezeichnet.

In dieser Arbeit wird die akustische Umgebung durch ein im Computer erzeugtes Raummodell simuliert. Anhand der modellierten Geometrie des Raumes und der Absorptions- sowie Diffusionseigenschaften seiner Begrenzungsflächen berechnet das Simulationsprogramm die spektrale Energie und den Eintreffwinkel jeder an einer Hörerposition eintreffenden Reflektion bei der Anregung einer im Raum positionierten Quelle mit einem Impuls. Mit der resultierenden Impulsantwort kann die virtuelle akustische Szene bestehend aus Quelle, Raum und Empfänger hörbar gemacht werden. Ein solcher Prozess wird als Auralisation bezeichnet. Die Erzeugung wahrnehmbarer Signale aus einem Computermodell wird Rendering genannt.

Ein Modell entspricht nie exakt der Realität. Die Simulation eines akustischen Ereignisses muss aber auch nur so genau sein, wie es vom Menschen wahrgenommen wird. Das Verhältnis zwischen akustischem Reiz und menschlichen Wahrnehmungsschwellen spielt dabei eine entscheidende Rolle. Mit den Erkenntnissen der Psychoakustik ist es möglich Vereinfachungen bei der Modellierung vorzunehmen, die die Berechnung eines virtuellen Schallfeldes mit der begrenzten Rechenleistung von Computern erst möglich machen.

Einen grundlegenden Teil ihrer räumlichen Wirkung verdanken Musikinstrumente ihrer Richtcharakteristik. Bei der Auralisation einer akustischen Umgebung ist es daher notwendig die Abstrahlcharakteristik einer Quelle in das Modell einzubeziehen (Vorländer 2008). Ihre

Ausrichtung wird durch den Hörer über richtungsabhängige Unterschiede in der Abstrahlung erfasst. Ein Sprecher klingt an der Rückseite beispielsweise wesentlich dumpfer, als wenn der Hörer vor ihm steht. Aus der Abstrahlung einer Quelle in einem Raum resultieren zwei sich überlagernde Schallfelder, das Diffus- und das Direktfeld. Der Abstand von der Quelle an dem die Energiedichte beider Felder gleich groß ist wird als Hallabstand bezeichnet. Der Anteil des Direktschalls in einer bestimmten Richtung wird unmittelbar durch die Richtwirkung der Quelle beeinflusst. In der Vorzugsrichtung weitet sich der Hallabstand, also der Bereich mit vorwiegend direktem Schallanteil, proportional zum Bündelungsrad aus. Die Richtwirkung von Instrumenten ist zudem stark frequenzabhängig und steigt meist mit der Frequenz an. Daraus resultiert eine charakteristische Frequenzabhängigkeit des Direkt-Diffus-Verhältnisses. Die Wiedergabe der Aufnahme eines Instrumentes ohne Klangfärbung ist im Diffusfeld nur dann möglich, wenn die frequenzabhängige Richtcharakteristik von der Quelle hinreichend korrekt nachgebildet wird (vgl. Weinzierl 2008, S. 502). Dieser Zusammenhang ergibt sich aus der Tatsache, dass sich der Schalldruck im Diffusfeld aus allen Richtungen annähernd gleich verteilt überlagert. Sowohl für ungerichtete als auch für gerichtete Quellen ist der Schalldruck im Diffusfeld demnach proportional zur Leistung der Quelle. Die Leistung ergibt sich aus der Energie, die die schwingenden Teile eines Klangerzeugers pro Zeiteinheit in das Medium Luft abgeben. Verteilt sich diese Leistung ungleichmäßig um die Quelle, ist die Gesamtleistung abhängig von der Richtungsverteilung der abgestrahlten Energie. Bei der Aufnahme eines Instrumentes mit einem Mikrofon enthält das Signal den Schalldruck in einer einzigen Richtung. Soll die Frequenzabhängigkeit der Gesamtleistung des Instrumentes korrekt reproduziert werden, muss das Wiedergabemedium die Richtcharakteristik entsprechend nachbilden.

Um den genannten Auswirkungen der Richtcharakteristik auf das Schallfeld in einem Raum Rechnung zu tragen, soll die Auralisation einer Konzertaufführung mit Quellmodellen durchgeführt werden, die eine dem realen Musikinstrument äquivalente Richtcharakteristik haben. Dazu werden in Abschnitt drei aus Messungen der TU Berlin Quellmodelle erstellt, die in eine akustische Simulationssoftware importiert werden können. Im vierten Abschnitt wird mit diesen Quellmodellen und der Raumakustiksoftware EASE eine virtuelle akustische Umgebung erstellt, die eine Konzertsituation für zwei Musikstücke simuliert, die das Forschungsprojekt „Medium und Emotion“ ausgewählt hat. Das Ergebnis dieser Simulation werden dynamische BRIRs sein, mit denen die Musikstücke so zu Gehör gebracht werden können, dass der Eindruck entsteht, als ob sich die Schallquellen in dem simulierten Raum befinden. Dabei kann die Wirkung der modellierten Richtcharakteristiken auf die Hörwahrnehmung in einer solchen simulierten Rezeptionssituation mit Hilfe von akustischen Parametern abgeschätzt werden.

Zunächst werden die Musikaufnahmen, mit denen die Konzertaufführung simuliert werden soll, im folgenden Abschnitt vorgestellt.

2 Die nachhallfreien Musikaufnahmen

Für den Hörversuch wurden vier Musikstücke nach dem Kriterium ausgewählt, möglichst gegenpolige Emotionen beim Hörer hervorzurufen. Zwei Stücke sind dem Genre Pop/Jazz zuzuordnen und die anderen beiden sind orchestrale Klassikwerke. Am Beispiel der beiden Pop/Jazz Stücke wird die Simulation der Konzertaufführung in dieser Arbeit vorgestellt. Dazu wird möglichst nachhallfreies Mehrspurmaterial der Stücke benötigt, dessen Beschaffung und Produktion im Folgenden kurz dargestellt wird.

2.1 „Contredanse“, Pascale Gautier Quartett

Bei dem Stück „Contredanse“[1] handelt es sich um eine aktuelle CD-Produktion von 2010. Auf eine Anfrage der Forschungsgruppe willigten die Künstler ein, ihre Aufnahmen zur Verfügung zu stellen. Das zugeschickte Material besteht aus 16 Einzelspuren im Format 16 Bit / 44 kHz. Ein vollständiger Spurenplan ist in Tabelle 1 dargestellt.

Instrument	Spurbezeichnung
Bass	DI
	KM 85
	RE 20
Schlagzeug	Kick
	SN bottom
	SN top
	T1
	T2
	T3
	Overhead L
	Overhead R
Flügel	Piano L
	Piano R
	Piano Room L
	Piano Room R

[1] Gautier 2010

Stimme	Lead

Tabelle 1 Spurbezeichnungen für das Stück „Contredanse“.

Die Besetzung besteht aus Stimme, Flügel, Bass und Schlagzeug. Für den Bass gibt es drei Spuren. Zwei davon wurden mit Mikrofonen an unterschiedlichen Positionen aufgenommen (KM85 und RE20). Die dritte Spur enthält das Signal eines Tonabnehmers (DI), das direkt über eine DI-Box aufgenommen wurde.

Das Schlagzeug wurde mit acht Spuren aufgenommen. Zwei davon tragen das Signal von zwei Overheadmikrofonen, die links und rechts vom Spieler aufgestellt waren. Die restlichen sechs Spuren wurden für einzelne Klangerzeuger des Schlagzeugs verwendet: Bassdrum (Kick), Snaredrum oben (SN top), Snaredrum unten (SN bottom), kleine Tom (T1), mittlere Tom (T2) und Standtom (T3).

Vier Spuren wurden für den Flügel gebraucht. Zwei Mikrofone wurden dicht an das Instrument gestellt (Piano L/R) und zwei waren mehr im Raum positioniert (Piano Room L/R). Bei beiden Positionen handelt es sich um eine A/B-Aufstellung.

Eine Spur wurde schließlich für die Stimme (Lead) aufgenommen.

Nach Anlegen der Spuren in einem Multitrackeditor konnte man sich einen Eindruck über die Qualität der unbearbeiteten Aufnahmen verschaffen. Der Anteil des Nachhalls wurde für alle Spuren bis auf die Raummikrofone des Flügels als ausreichend gering befunden.

2.2 „River Man“, Nick Drake

Der Song „River Man“[1] von Nick Drake wurde 1969 aufgenommen. Die Besetzung ist: Gesang, Gitarre, Bass und ein Ensemble aus 12 Streichern. Auf Nachfragen bei dem Produzenten Joe Boyd wurde in Erfahrung gebracht, dass alle Instrumente zusammen „live“ in einem Raum aufgenommen wurden. Es wurden insgesamt vier Spuren mitgeschnitten: Eine für Gesang, eine für Gitarre und eine Stereomischung der Streicher. Dabei gab es reichlich Übersprechen zwischen den Instrumentenmikrofonen[2]. Zudem schien es nach Aussagen von Joe Boyd schwierig, an die Originalaufnahmen heranzukommen. Daher wurde die Produktion eines so genannten „sound-alikes“ in Auftrag gegeben. Es wurde eine Aufnahme produziert, die sich in allen klanglichen und musikalischen Aspekten möglichst genau an die Vorlage hält. Dazu gehört, dass ein Sänger gesucht wurde, der eine ähnliche Klangfarbe wie Nick Drake hat, und dass sich die künstlerische Interpretation am „Original“

[1] Drake 1969

[2] Zitat aus einer E-Mail von Joe Boyd (2011): “All performers were close together, there was no separation between the instruments. Even Nick's vocal and guitar mics were close to the strings and his tracks had plenty of 'spill' from the strings and vice versa.”... “It was a 4-track recording - guitar, voice and stereo mix of the strings.”

orientiert. Bass, Gitarre und Gesang wurden separat aufgenommen. Die Aufnahme der Streicher wurde mit einem Streichquartett durchgeführt, das mehrere Overdubs[1] einspielte, die später zu einem größeren Streicherarrangement kombiniert werden können. Die Mikrofone wurden wie folgt ausgewählt und aufgestellt:

Der Sänger wurde mit einem Neumann U 67 in 20 bis 30 cm Abstand aufgenommen. Die Membran des Mikrofons befand sich auf Höhe der Mundöffnung und war parallel zu dieser ausgerichtet.

Die Gitarre wurde mit zwei Kugelmikrofonen der Firma Schoeps (MK 2H) aufgenommen. Die Basislänge der AB-Konfiguration betrug 20 cm. Die Mikros wurden mit einem Abstand von 50 bis 60 cm und einem Neigungswinkel von ca. 45° von schräg oben auf die Mitte des Instrumentes gerichtet. Die beiden Mikros waren dabei so ausgerichtet, dass sich das Schallloch ungefähr in ihrer Mitte befand.

Das Streichquartett saß in der üblichen Aufstellung: 1. Geige, 2. Geige, Bratsche und Cello (von links nach rechts). Die Stützmikrofone wurden in einem Abstand von 70 bis 80 cm aufgestellt. Um etwa 30 bis 40 Grad aus der Resonanzdeckennormale nach oben verschoben, zeigten die Mikrofone auf das Zentrum des Instrumentes. Für die erste und zweite Geige sowie für die Bratsche wurden Neumann U 89 verwendet.

Für das Cello kam ein Neumann U 87 zum Einsatz. Hier war das Mikrofon um ca. 20° von der Resonanzdeckennormale nach rechts verschoben und zeigte ebenfalls auf die Mitte des Instrumentes.

Für die Stereoversion wurden mehrere Stereofoniekonfigurationen in der Mitte vor den Streichern platziert (Abbildung 2.1):

1. AB (Basislänge = 50 cm) mit Brühl und Kjer 4006
2. AB mit einem Neigungswinkel von 45° nach außen (Basislänge = 40 cm), verwendet wurden Schoeps MK 21
3. XY (Neigungswinkel = 90°) mit Neumann KM 140
4. großes AB (Basislänge = 70 cm) mit Neumann KM 140

[1] Gängige Bezeichnung im Tontechnikjargon für die Tonaufnahme über eine bestehende Tonaufnahme.

Abbildung 2.1 Foto der Mikrofonaufstellung für das Streichquartett bei den Aufnahmen zu „River Man“. Die Streicher saßen von links nach rechts in der Reihenfolge: Violine 1, Violine 2, Bratsche, Violoncello. Quelle: Jonas Fehrenberg.

Die Streicher saßen circa 1,5 m bis 2 m voneinander entfernt. Stellen, an denen das Cello solo spielt, eignen sich dazu, die Übersprechdämpfung abzuschätzen. Hier werden zwischen erster Violine und Cello circa 15 dB erreicht. Bei der zweiten Violine und der Bratsche ist der Pegelabstand zum Cello ungefähr gleich groß und beträgt circa 11 dB.

Der Bass wurde in Form von Mididaten aufgenommen. Als Klangerzeuger dienten Pizzicatosamples der Library LA Scoring Strings[1].

Als Aufnahmemischpult wurde für die Streicher ein Yamaha DM 2000 und für Gitarre und Gesang ein Ramsa WRT 820 verwendet.

Für die beiden Stimulusversionen im konventionellen Wiedergabeverfahren wurde eine Stereomischung angefertigt. Dabei wurden drei Overdubs der Streicher verwendet.

[1] http://audiobro.com/la-scoring-strings/

3 Simulation der Richtcharakteristiken

Die Messung der richtungsabhängigen Abstrahlung von Lautsprechern wird üblicherweise mit einem stationären Schallsignal durchgeführt. Die Richtcharakteristiken können durch eine Abfolge von Messungen in einer hohen räumlichen Auflösung bestimmt werden. Da beim Spielen eines Musikinstrumentes die Dauer eines Tones meistens begrenzt ist, muss die Messung hier simultan durchgeführt werden. Das Institut für Technische Akustik (im Folgenden: ITA) der RWTH Aachen hat zu diesem Zweck ein Kugelflächenarray aus 32 Mikrofonen entwickelt (Pollow und Masiero 2009). Mit diesem Mikrofonarray wurden vom Fachgebiet AK Messungen zur Abstrahlung von 41 historischen und modernen Musikinstrumenten im reflexionsarmen Raum (im Folgenden: RAR) der TU Berlin durchgeführt (Krämer et al. 2010), (Detzner et al. 2010). Da in EASE keine natürlichen Quellen als Modelle zur Verfügung stehen, sollen die Messdaten der 19 modernen Instrumente für die Erstellung einer Lautsprecherbibliothek mit den Richtcharakteristiken von Musikinstrumenten ausgewertet werden.

Bereits bestehende Lautsprechermodelle von Musikinstrumenten werden im Folgenden kurz erläutert, um die Motivation für eine Erstellung von neuen Quellmodellen herzuleiten:

Von der Physikalisch-Technischen Bundesanstalt (im Folgenden: PTB) wurden Lautsprechermodelle im Format SD1[1] des Raumsimulationsprogramms CATT-Acoustic[2] zur Verfügung gestellt[3]. Diese Daten basieren auf Messungen, die von Meyer (1995) durchgeführt wurden. Die meisten der Instrumente wurden mit einem Schwingerreger angeregt, was zur Folge hat, dass die akustischen Abschattungseffekte des Spielers nicht berücksichtigt werden. Die Pegelwerte werden in Oktaven ausgewertet. In EASE werden die raumakustischen Berechnungen in Terzbändern durchgeführt. Es gibt keine Möglichkeit diese höhere Auflösung der Lautsprecherdaten aus den Werten der PTB-Daten zu extrapolieren. Das Koordinatengitter des SD1-Formates hat eine Winkelbreite von 10°. Um die Werte in das Format von EASE zu konvertieren, müsste eine Interpolation der Pegelwerte in ein 5°-Netz durchgeführt werden. Eine solche Formatkonvertierung wird in Abschnitt 3.2.4 für die Quelle des Flügels durchgeführt, da dieses Instrument nicht Bestandteil der von der TU Berlin durchgeführten Messungen war.

Weitere Messungen zur Abstrahlung von Musikinstrumenten wurden an der Helsinki University of Technology (im Folgenden: TKK) von Pätynen (2010) durchgeführt. Auch diese

[1] Lautsprecherformat der Software CATT-Acoustic, http://www.catt.se/srcdir_mod.htm
[2] Dalenbäck 2011, http://www.catt.se/
[3] http://www.ptb.de/de/org/1/16/163/richtchar.htm

Richtcharakteristiken stehen im Internet zur Verfügung[1]. Hier wurden die Messdaten in das Lautsprecherformat CLF1[2] konvertiert. Wie bei den Daten der PTB werden die Pegelwerte in Oktaven angegeben. Die Messwerte wurden ebenfalls auf ein Koordinatengitter mit 10° Winkelbreite interpoliert. Im Vergleich zu den Richtcharakteristiken der PTB gibt es bei Pätynen (2008) ausführliche Informationen darüber, wie die Messungen durchgeführt wurden. Die Anzahl der Messmikrofone, die auf einer Kugelfläche um das Instrument verteilt waren, lag bei 22. Das sind rund 30 % weniger Messproben, die für die Interpolation der Richtcharakteristik zur Verfügung standen, als bei den Messungen der TU Berlin.
Als Abschluss des Kapitels über die Simulation von Richtcharakteristiken wird das Lautsprechermodell der PTB und der TKK für die Singstimme mit der hier erstellten Richtcharakteristik verglichen.

Bei den Messungen der TU Berlin kam ein 32-kanaliges Mikrofonarray der RWTH Aachen zum Einsatz. Die 32 Messmikrofone des Arrays befinden sich an gleichmäßig verteilten Punkten einer das Instrument umgebenden Kugelfläche. Der Durchmesser des Arrays beträgt 4,17 m.
Der RAR der TU Berlin ist nach der DIN EN ISO 3745 bis zu der unteren Grenzfrequenz von 63 Hz nachhallfrei[3]. Der gemessene Schalldruck wird für die Richtcharakteristiken in Terzbändern ausgewertet. Da die untere Frequenzgrenze des tiefsten Terzbandes von 100 Hz bei 89 Hz und damit über der kritischen Grenzfrequenz des RAR liegt, ist im gesamten auszuwertenden Frequenzbereich von einer korrekten Messung auszugehen.
Es wurden alle Halbtöne im Ambitus des jeweiligen Instrumentes und in zwei Dynamikstufen, Pianissimo und Fortissimo, erfasst. Um eine generelle Information über die Abstrahlung des jeweiligen Instrumentes zu erhalten, werden die Schallpegel an den 32 Messpunkten über den gemessenen Dynamik- und Tonumfang gemittelt. Änderungen der Richtcharakteristik in Abhängigkeit von der Anspielstärke und der Tonhöhe werden bei der Simulation nicht berücksichtigt.
Informationen über die Phasenlage werden hier nicht erfasst.
In der akustischen Simulationssoftware EASE werden die Pegelwerte der Richtcharakteristiken in einem 5°-Koordinatennetz ausgewertet. Die Daten der 32 Messpunkte müssen daher auf einer Kugeloberfläche in dieses feinmaschigere Koordinatennetz interpoliert werden. Unterschiedliche Methoden der Interpolation und Koordinatentransformation werden dabei ausgewertet.
Richtcharakteristiken werden als winkelabhängige Pegeldifferenzen zu einer Bezugsrichtung definiert. Die interpolierten Schallpegelwerte müssen dafür auf eine Bezugsrichtung normiert

[1] http://auralization.tkk.fi/node/44
[2] Common Loudspeaker Format, http://www.clfgroup.org/
[3] http://www.akustik.tu-berlin.de/fileadmin/fg23/Sonstiges/Technische_Einrichtungen/Anechoic_Room.pdf

werden, deren Bestimmung in Abhängigkeit von der Verwendung der Richtcharakteristik diskutiert wird.

Die Abstrahlcharakteristiken werden in EASE als Lautsprechermodelle importiert. Daher wird es notwendig sein elektroakustische Kennwerte zu bestimmen und die Daten in einem von EASE lesbaren Dateiformat zu exportieren.

Die zur Simulation der Richtcharakteristiken erforderlichen Berechnungen werden in Matlab[1] durchgeführt. Die dafür erstellten Quellcodes sind im Anhang C: dokumentiert. Alle notwendigen Routinen vom Einlesen der Daten bis zum Export der Ergebnisse sind in Abbildung 3.1 als Blockdiagramm zusammengefasst.

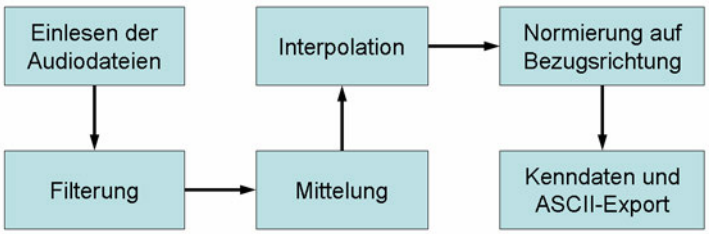

Abbildung 3.1 Blockdiagramm der Verarbeitungsschritte zur Berechnung der Richtcharakteristiken

3.1 Mittelung und Filterung

Die Daten liegen in Form von 32-kanaligen Audiodateien[2] vor. Diese werden einzeln in Matlab eingelesen und ausgewertet.

Zunächst wird ein Filter angewendet, welches den Frequenzgang der einzelnen Messmikrofone kompensiert.

Die richtungsabhängige Pegeldifferenz der Quellabstrahlung wird von EASE in 21 Terzbändern mit den Mittenfrequenzen von 100 Hz bis 10 kHz erfasst. Die Signale jeder einzelnen Mikrofonposition werden zunächst mithilfe einer Terzbank gefiltert. Anschließend werden deren Effektivwertquadrate berechnet und über die Menge der eingespielten Einzeltöne gemittelt. Als Ergebnis dieser Reihe von Prozessen erhält man für jeden der 32 Kanäle 21 gemittelte Terzbandpegel. Eine grafische Übersicht der notwendigen Berechnungsschritte ist in Abbildung 3.2 als Blockdiagramm dargestellt.

[1] The MathWorks 2009
[2] im Broadcast Wave Format

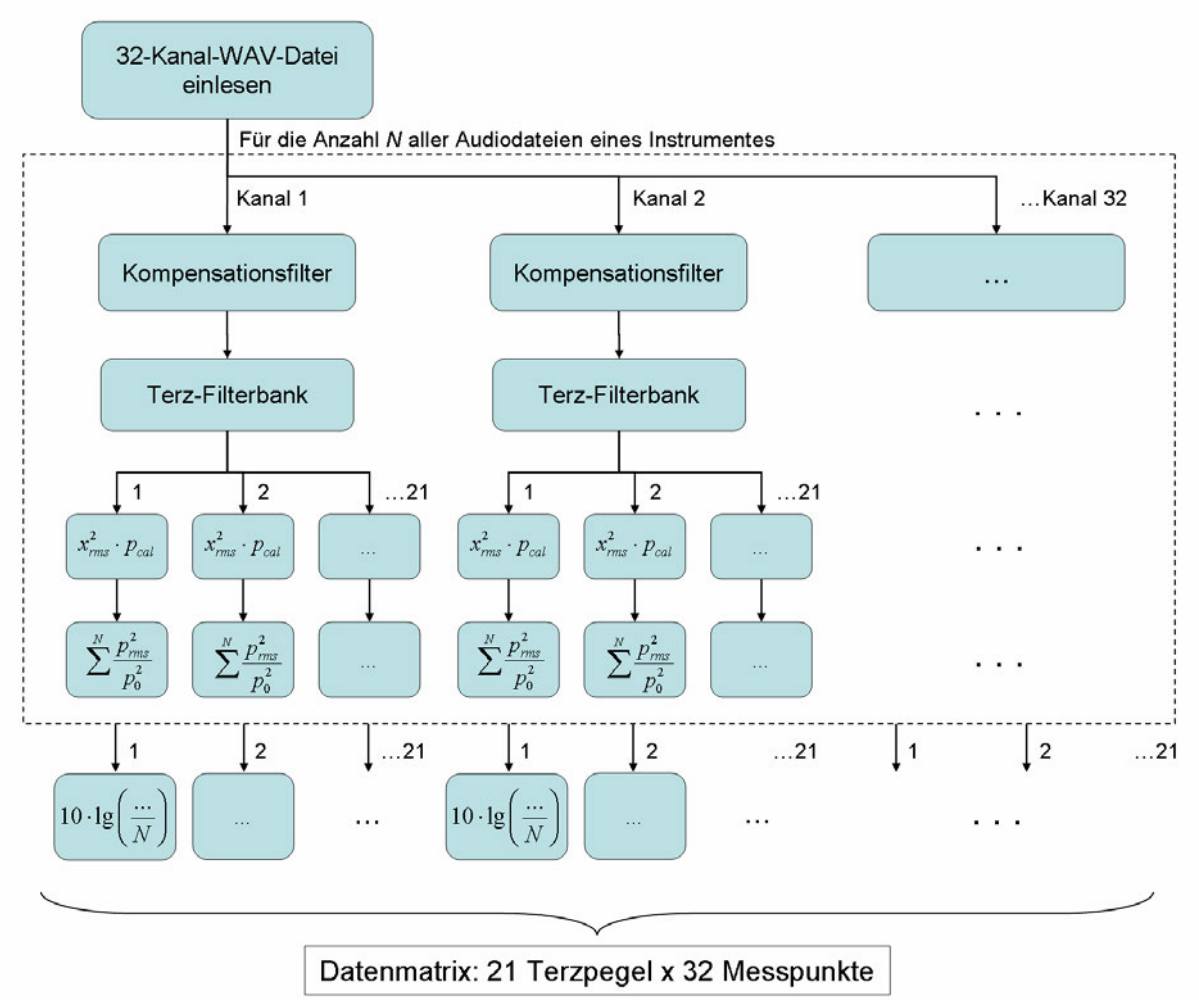

Abbildung 3.2 Blockdiagramm für die Filterung und Mittelung der Schalldruckpegel.

Die Kompensation des Frequenzganges der Messmikrofone erfolgt durch ein FIR-Filter, das mit der Matlab-Funktion *fftfilt* realisiert wird. Die dafür verwendeten Impulsantworten wurden zusammen mit den WAV-Datensätzen vom Fachgebiet AK zur Verfügung gestellt.

Für das Design der Filterbank wurde auf eine Arbeit von Couvreur (1998) zurückgegriffen. Nach den Vorgaben des ANSI-Standards für Terzbandfilter[1] wird die Filterbank durch eine Reihe von Butterworthfiltern dritter Ordnung realisiert. In seiner Arbeit beschreibt Couvreur das Design für die Mittenfrequenzen von 100 Hz bis 5 kHz. Dieser Ansatz muss auf eine oberste Mittenfrequenz von 10 kHz erweitert werden.
Die Mittenfrequenzen werden gemäß DIN 61260 nach der Formel

$$f_m = 2^{x/3} \cdot 1000\ Hz$$

(x: Index des Frequenzbandes) berechnet.

[1] ANSI S1.11-2004

Digitale Filter haben die Eigenschaft einerseits bei sehr tiefen Frequenzen und schmalen Bandbreiten und andererseits bei sehr hohen Frequenzen instabil zu werden. Die Filterfrequenzen digitaler Filter sind im Verhältnis zur Samplefrequenz definiert. Die Bandpässe werden mithilfe der Matlab-Funktion *butter* entworfen. Um ein korrektes Filterdesign mit dieser Funktion zu gewährleisten, muss die Mittenfrequenz f_m in Abhängigkeit von der Samplefrequenz f_s im Bereich

$$\frac{f_s}{200} < f_m < \frac{f_s}{5} \tag{1.1}$$

liegen (Couvreur 1998). Die zu filternden Audiodaten wurden mit einer Samplefrequenz von 44,1 kHz aufgenommen. Die Mittenfrequenzen der Bandpassfilter sollten also den Bereich

$$220,5Hz < f_m < 8820Hz$$

weder über- noch unterschreiten.

Die Samplefrequenz des Signals wird daher für die Filterung bei 10 kHz verdoppelt. Die Interpolation der Samplewerte wird mithilfe der Funktion *resample* durchgeführt. Für Gleichung (1.1) ergibt sich jetzt

$$441Hz < f_m < 17640\,.$$

Mit einer Samplefrequenz von 88,2 kHz kann also jetzt unter Einhaltung des Kriteriums (1.1) ein Terzfilter bei 10 kHz erstellt werden.

Für die Frequenzen von 1,6 kHz bis 8 kHz werden die Filterkoeffizienten mit f_s = 44,1 kHz direkt aus der Funktion *butter* ermittelt.

Für die Frequenzen von 100 Hz bis 1,25 kHz wird ein so genanntes *multirate filter* implementiert. Man macht sich dabei die besonderen Eigenschaften einer Filterbank mit konstanter relativer Bandbreite zunutze: Da Mittenfrequenz und Bandbreite eines digitalen Filters in Relation zur Samplefrequenz definiert sind, ergeben die für eine bestimmte Frequenz und Bandbreite errechneten Filterkoeffizienten für ein Signal mit halber Samplefrequenz einen Bandpass, welcher bei genau der halben Frequenz und Bandbreite filtert. Die Filterkoeffizienten der drei Terzfilter einer Oktave ergeben also nach Halbierung der Samplefrequenz die Koeffizienten der drei Terzfilter, die eine Oktave tiefer liegen. Durch Iteration dieses Prozesses können die tieferen Terzfilter der Filterbank mit dem Vorteil rechnerischer Effizienz und Einhaltung des Kriteriums (1.1) berechnet werden. Die Halbierung der Samplefrequenz durch Tiefpassfilterung und Eliminierung jedes zweiten Samplewertes erfolgt mit der Matlab-Funktion *decimate.*

Die Frequenzgänge der Terzfilter sind in Abbildung 3.3 dargestellt.

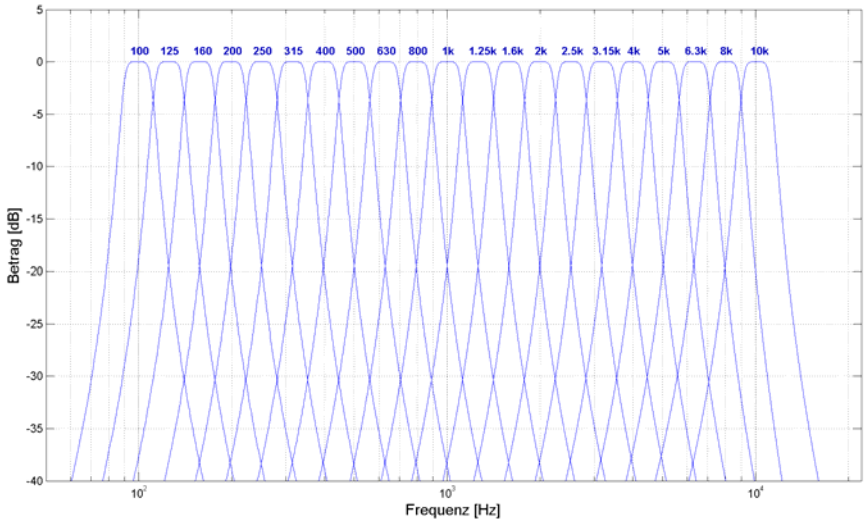

Abbildung 3.3 Frequenzgänge der 21 Terzbandfilter. Die Frequenzbezeichnungen entsprechen den Normfrequenzen nach der DIN EN ISO 266 und nicht den exakt berechneten Frequenzwerten.

Die Funktion *filter* erzeugt schließlich mithilfe der berechneten Terzbandkoeffizienten das gefilterte Signal, aus dem dann durch Quadrierung und Mittelung über alle Samplewerte das Quadrat des Effektivwertes x_{rms}^2 berechnet wird.

Durch Multiplikation mit einem Kalibrierungswert p_{cal}, der für jede Mikrofonposition mit einem Pistonfon bei der Messung ermittelt und mit den Audiodaten bereitgestellt wurde, wird aus dem numerischen Wert der tatsächliche Schalldruck als Verhältnis zum Schalldruck bei der Hörschwelle

$$\frac{p^2}{p_0^2}$$

($p_0 = 20 \cdot 10^{-6}\ N/m^2$) bestimmt.

Durch Aufsummieren der Schalldruckquadrate und anschließendes Dividieren durch die Anzahl der einzelnen Signale nach der Formel für die „Mittelung von Einzelereignispegeln“ (DIN 45641) erhält man schließlich den gemittelten Schalldruckpegel:

$$L_m = 10 \cdot \lg\left(\frac{1}{N} \cdot \sum_{i=1}^{N} \frac{p_i^2}{p_0^2}\right) \tag{1.2}$$

3.1.1 Kompensation der Quellverschiebung

Nach der Interpolation und Auswertung der Daten für den Kontrabass kamen Zweifel auf, ob die ermittelte Richtcharakteristik korrekt sein kann. Bei allen Terzbändern hatte die

Abstrahlung des Instruments eine Vorzugsrichtung nach vorne/unten und konzentrierte sich bei circa 0° Azimut und 135° Polarwinkel. Dieser Ef fekt steht im Kontrast zu den Auswertungen bei Meyer (1995) und Pätynen (2010). Zur Veranschaulichung sind in Abbildung 3.4 die Plots für auszugsweise sieben Terzbänder dargestellt.

Im Vergleich wäre nach Meyer (1995) bei 100 Hz „mit einer allseitig gleichmäßigen Schallabstrahlung“ zu rechnen. Als ungerichtet werden Abstrahlcharakteristiken bezeichnet, deren Pegel bei keinem Winkel um mehr als 3 dB vom Maximalwert abfällt (ebd., S. 108). Die Messdaten für den Kontrabass weisen für 100 Hz in der Richtung vorne/unten bereits einen Pegelunterschied von 5 dB zur seitlichen und vertikalen Abstrahlung und zur rückwärtigen Richtung sogar einen Unterschied von 10dB auf.

Der maximale Pegelunterschied für den gesamten Messbereich liegt im Terzband von 2 kHz und beträgt 22,7 dB. Der statistische Richtfaktor[1] liegt in dieser Richtung bei 3,4. Bei Meyer wird ein maximaler Richtfaktor von nur 2,6 ermittelt[2]. Da das Maximum in der verdächtigen Richtung liegt, wird vermutet, dass diese Abweichung auf eine inkorrekte Quellpositionierung zurückzuführen ist.

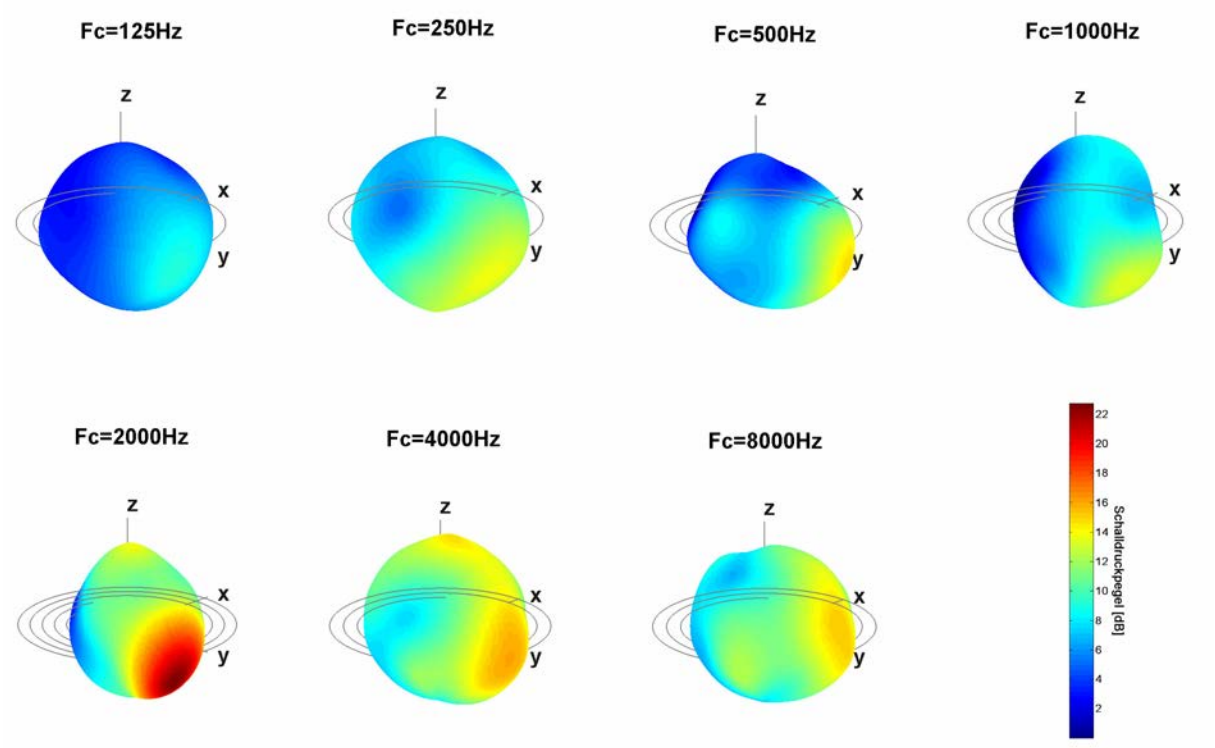

Abbildung 3.4 Richtcharakteristiken des Kontrabasses für sieben exemplarische Terzbänder im Oktavabstand aus dem Blickwinkel vorne/rechts/unten (Azimut=315° und Polarwinkel=110°). Der Pegel ist auf den jeweiligen Minimalwert der Abstrahlcharakteristiken normiert.

[1] Zur Bedeutung des Richtfaktors siehe Abschnitt 3.4.2

[2] Bei Meyer wird das Maximum des Richtfaktors bei 3 kHz erreicht. Für welche Bandbreite dieser Wert bestimmt ist, wird nicht angegeben.

Aus praktischen Gründen konnte der Kontrabass bei der Messung nicht exakt im Zentrum des Mikrofonarrays positioniert werden. Wie man auf dem Foto, das bei der Messung gemacht wurde, sehen kann, war er etwas weiter nach vorne und nach unten aufgestellt, als man es erwarten würde (Abbildung 3.5). Er war also genau in die Richtung verschoben, in der sich der Pegel bei allen Terzbändern konzentriert.

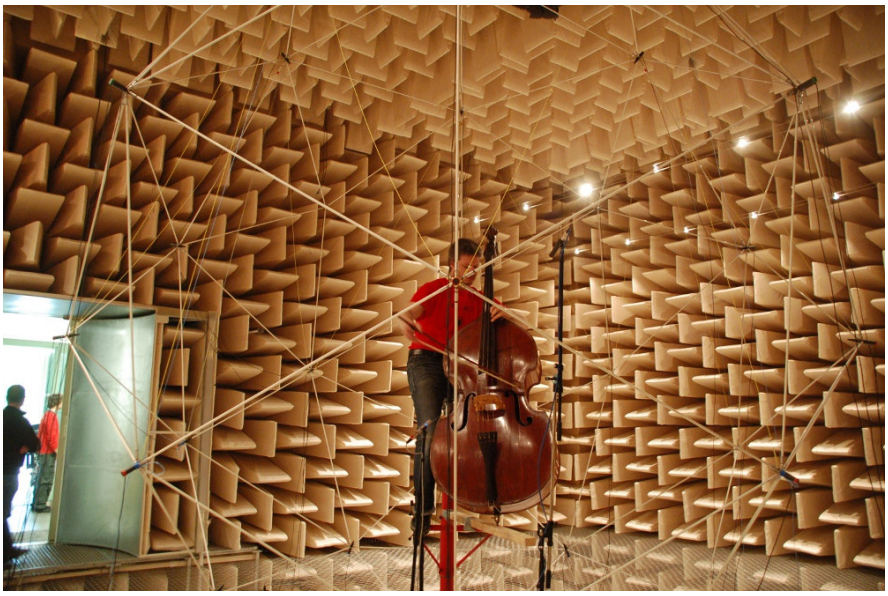

Abbildung 3.5 Frontalansicht des Kontrabasses bei der Messung im reflektionsarmen Raum der TU Berlin. Die Verlängerung der Längsachse des Messmikrofons in der Mitte des Bildes ist ein guter Anhaltspunkt für den Blick auf das Zentrum des Kugelflächenarrays. Quelle: Fachgebiet Audiokommunikation, TU Berlin

Wenn sich das akustische Zentrum des Instruments nicht in der Mitte des Kugelarrays befindet, kommt es nach Pollow und Masiero (2009) bei einer Interpolation der Daten mithilfe der Zerlegung in Kugelflächenfunktionen zu folgenden Effekten: Aufgrund von Laufzeitunterschieden entstehen zwischen den Messmikrofonen Phasenverschiebungen, die bei einer Interpolation des komplexen Schalldrucks zu Richtcharakteristiken führen, deren Verlauf wesentlich komplexer erscheint, als zu erwarten ist. Dieser mit der Frequenz ansteigende Effekt ist auf räumliches Aliasing zurückzuführen und hängt mit der begrenzten Anzahl von 32 Messproben zusammen. Wertet man nur den Betrag des Schalldrucks aus, kommt es zu keinen Aliasingeffekten. Die Verschiebung der Schallquelle wirkt sich nur insofern aus, als der Schalldruck aufgrund der unterschiedlich langen Wege unterschiedlich stark abnimmt, bis er die Mikrofone erreicht.

Da hier nur der Betrag ausgewertet wird, können die durch eine Quellverschiebung verursachten Pegelunterschiede mit dem für Punktquellen geltenden 1/r-Gesetz näherungsweise berechnet werden. Diese approximative Pegelkorrektur wird sowohl von Pollow und Masiero (2009) als auch von Pätynen (2010) vorgeschlagen und im Folgenden für den Kontrabass angewendet.

Durch jede beliebige, die Schallquelle umfassende Hüllfläche fließt im Freifeld die gleiche akustische Leistung. Für den Schalldruckpegel L_p im Abstand *r* auf einer umhüllenden Kugelfläche gilt im Fernfeld einer ungerichteten Punktquelle:

$$L_p = L_w - 20\lg\frac{r}{m} - 11dB \tag{1.3}$$

(L_w: Schallleistungspegel) (Möser 2009, S. 76). Vor und nach der Quellverschiebung fließt die gleiche Leistung durch die gedachte Kugelfläche des Mikrofonarrays:

$$L_{W_1} = L_{W_2}$$

Aus diesem Ansatz und (1.3) ergibt sich:

$$L_{p_1} + 20\lg\frac{r_1}{m} + 11dB = L_{p_2} + 20\lg\frac{r_2}{m} + 11dB$$

$$L_{p_2} = L_{p_1} + 20\lg\frac{r_1}{r_2} = 10\lg\left(\frac{p_1}{p_0}\right)^2\left(\frac{r_1}{r_2}\right)^2$$

Der durch die Quellverschiebung korrigierte Schallpegel L_{p_2} unterscheidet sich von dem unkorrigierten Schallpegel L_{p_1} durch den Faktor des quadrierten Abstandsverhältnisses. Der Algorithmus zur Berechnung der Mittelungspegel muss also einfach durch das Korrekturelement

$$\left(\frac{r_1}{r_2}\right)^2$$

an geeigneter Stelle erweitert werden (siehe Abbildung 3.7).

Nachdem die unterschiedlichen Mikrofonabstände der unkorrigierten Quellposition ermittelt wurden (siehe 3.1.1.1 Berechnung der Quellabstände), können die korrekten Richtcharakteristiken ausgewertet werden. Abbildung 3.6 stellt die Richtcharakteristiken der korrigierten Quellposition für die gleichen Terzbänder wie in Abbildung 3.4 dar. Man beachte dabei den veränderten Maßstab der Farbskalierung für den Pegel.

Das Maximum der Kompensation liegt bei einer Abschwächung des Pegels um 4,5 dB und das quadratische Mittel der Pegelunterschiede beträgt 2 dB. Bei Pätynen (2010) wurde eine Verschiebung des Kontrabasses um 0,5 m und 0,2 m kompensiert. Der größte Pegelunterschied, den dort die Kompensation ausmacht liegt bei 2,3 dB.

Die Pegeldifferenzen für das Terzband bei 100 Hz betragen nach der Korrektur maximal 4 dB. Das entspricht einer eher ungerichteten Abstrahlung, die nach Meyer (1995) bei dieser Frequenz zu erwarten wäre. Die maximale Abweichung für die gesamte Messbandbreite beträgt 16,5 dB und liegt bei 2kHz. Das entspricht einem maximalen Richtfaktor von 2,3 und liegt damit um 0,3 unter dem maximalen Wert bei Meyer.

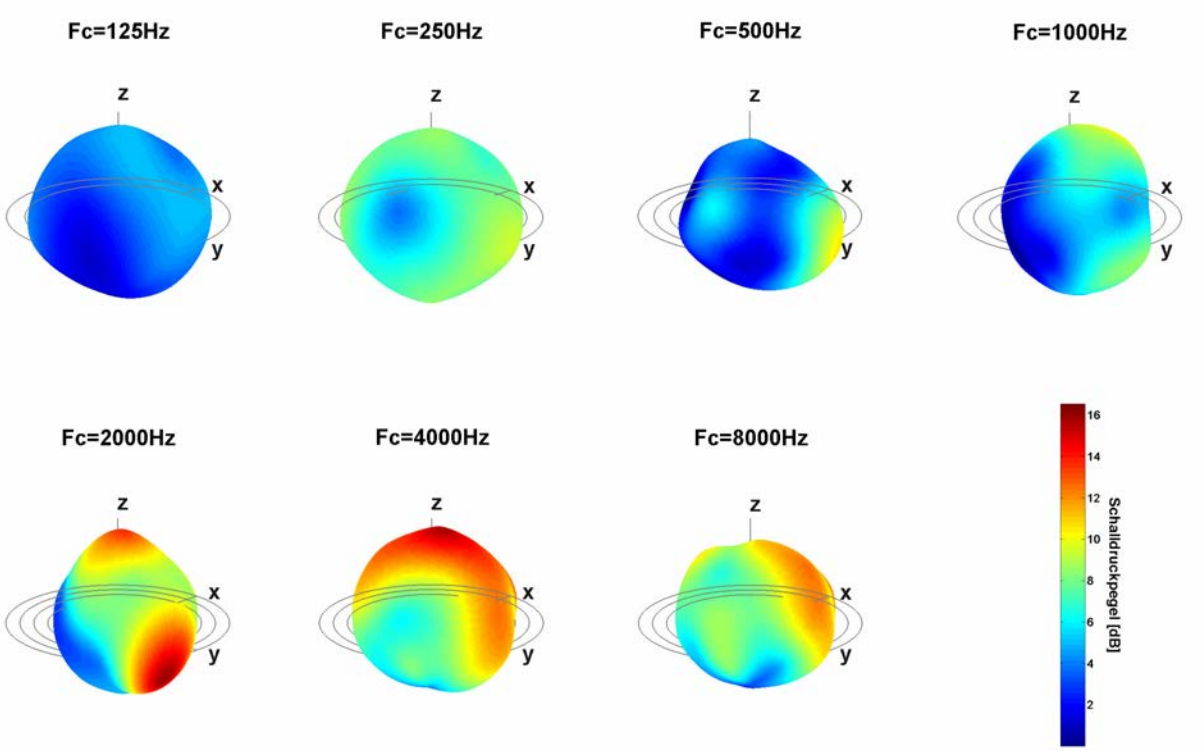

Abbildung 3.6 Richtcharakteristiken des Kontrabasses nach Korrektur der Quellposition für sieben exemplarische Terzbänder im Oktavabstand aus dem Blickwinkel vorne/rechts/unten (Azimut=315° und Polarwinkel=110°). Der Pegel ist auf den jeweiligen Minimalwert der Abstrahlcharakteristiken normiert.

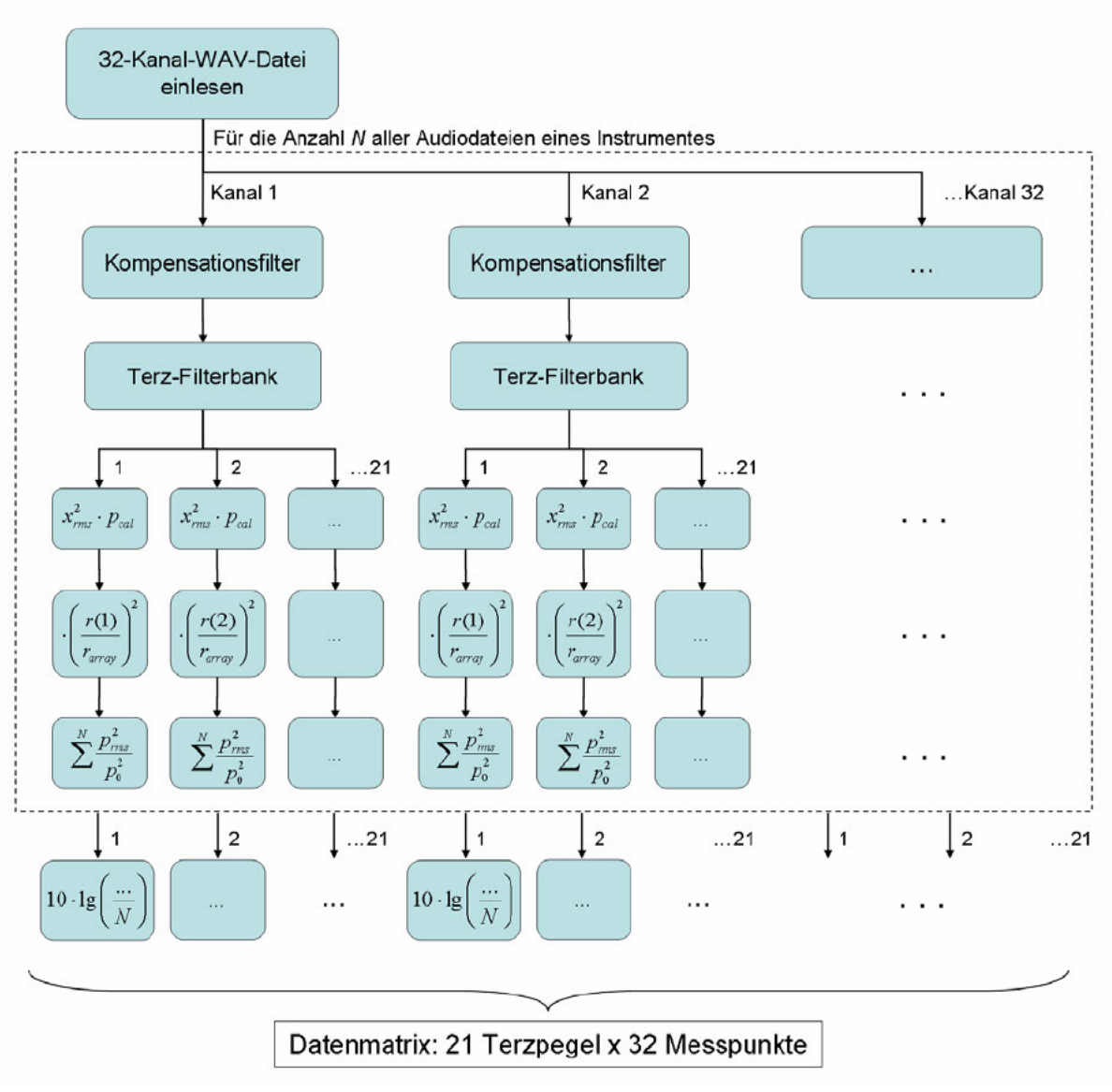

Abbildung 3.7 Blockdiagramm für die Filterung und Mittelung der Schalldruckpegel mit Kompensation der Quellposition.

3.1.1.1 Berechnung der Quellabstände

Der korrekte Quellabstand r_2 ist der Radius des Mikrofonarrays und beträgt $r_{array} = 2{,}085m$.

Die unterschiedlich langen Wege r_1 zwischen ursprünglicher Quellposition und Messmikrofonen können aus den kartesischen Koordinaten über die Gleichung für den euklidischen Abstand zwischen zwei Punkten *a* und *b*

$$r(a,b) = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2 + (z_a - z_b)^2} \tag{1.4}$$

berechnet werden.

Die kartesischen Koordinaten (x, y, z) der Messpunkte ergeben sich durch Koordinatentransformation aus den zu den Messdaten bereitgestellten Kugelkoordinaten. Z entspricht der vertikalen, y der transversalen Achse und x der Medianachse des Arrays.

Bei der Messung wurde die Verschiebung des Instrumentes aus dem Mittelpunkt nicht dokumentiert. Daher muss die Position des Kontrabasses aus Fotografien, die vor der Messung gemacht wurden, abgeschätzt werden. Abbildung 3.5 zeigt die frontale Ansicht (Blick auf die y/z-Ebene). Das Mikrofon in der Mitte des Bildes zeigt ungefähr auf das Zentrum des Arrays und in diesem Fall zugleich auf den Kopf des Spielers. Das Instrument war demnach etwas nach unten verschoben. Die seitliche Verschiebung ist so gering, dass sie außer Betracht gelassen werden kann.

Ein zweites Foto zeigt den senkrechten Blick auf die x/z-Ebene. Nach den Angaben von Pollow und Masiero (2009) und eines Mitarbeiters des Fachgebiets AK lag das Zentrum des Arrays ungefähr auf der vertikalen Achse des Stuhls und auf der Höhe des Kopfes (Abbildung 3.8).

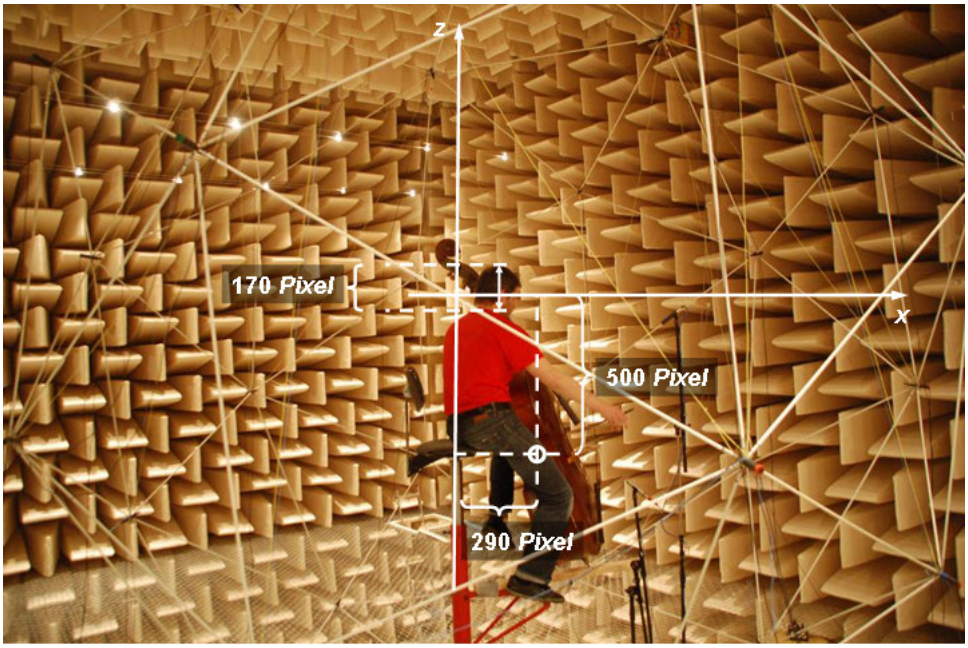

Abbildung 3.8 Grafisches Ausmessen der Quellposition für den Kontrabass mithilfe der seitlichen Ansicht.

Die Schallabstrahlung erfolgt beim Kontrabass durch die Schwingungen der Decke und des Bodens. Es ist daher schwer das akustische Zentrum festzulegen, das hier durch das geometrische Zentrum des Korpus approximiert wird. Die Kreismarkierung in Abbildung 3.8 stellt den Punkt dar, der dafür angenommen wird. Durch grafisches Ausmessen der Achsenlängen erhält man die Anzahl der Pixel: $x = 290$ und $z = 500$. Das Verhältnis von Pixelanzahl zu den realen Ausmaßen kann man herleiten, indem man ein Objekt ausmisst, von dem die reale Größe bekannt ist. Die Strecke von den Schultern bis zur Oberkante des Kopfes beträgt circa 170 Pixel. Der Kopf befindet sich ungefähr in gleichem Abstand wie das Instrument und wird daher als Maßstab für die Bestimmung der x- und z-Abstände benutzt.

Wenn man bei der der Strecke zwischen Schultern und Oberkante des Kopfes von 25 cm ausgeht, ergibt sich:

$$1Pixel = 1,5 \cdot 10^{-3} m .$$

Daraus lassen sich die Koordinaten der Quelle berechnen. Es ergibt sich ein Verschiebung um:

$$x = 0,43m$$
$$y = 0m$$
$$z = -0,74m$$

Mit diesem Ergebnis und (1.4) erhält man die Abstände $r_1(n)$ für die $n = 32$ Messpunkte aus der Gleichung:

$$r_1(n) = \sqrt{\left(x_{quelle} - x_{mik}(n)\right)^2 + \left(y_{quelle} - y_{mik}(n)\right)^2 + \left(z_{quelle} - z_{mik}(n)\right)^2}$$

Die Fotos von Cello und Gitarre lassen annehmen, dass auch diese Instrumente etwas aus dem Zentrum des Kugelarrays verschoben waren. Die Ermittlung der Quellverschiebung über das Auszählen von Pixel anhand einer Fotografie ist eine grobe Abschätzung der tatsächlichen Verschiebung. Da die Werte der Richtcharakteristiken von der Gitarre und dem Cello keine so eindeutigen Verzerrungseffekte aufzeigen wie beim Kontrabass, wird bei diesen Instrumenten von einer Kompensation der Quellposition abgesehen.

3.2 Interpolation

Um aus den gemittelten Terzpegeln der 32 Messpunkte die Abstrahlcharakteristik eines Quellmodells zu erzeugen, müssen aus diesen Werten die Pegel an den Punkten des in EASE definierten Koordinatennetzes interpoliert werden. Da die Winkelangaben in EASE einem anderen System unterliegen als bei der Messung, ist zusätzlich eine Koordinatentransformation notwendig. Es werden im Folgenden zunächst die Positionen der Datenpunkte in den beiden Koordinatensystemen erläutert. Anschließend folgt eine Auswertung von zwei möglichen Transformations- und Interpolationsmethoden.
Da ein Flügel nicht Bestandteil der Messungen der TU Berlin war, wird es notwendig sein die für dieses Instrument von der PTB zur Verfügung gestellten Lautsprecherdaten in das System von EASE zu konvertieren. Auch wird eine Koordinatentransformation und eine Interpolation notwendig sein, wofür eine der beiden vorgestellten Methoden in abgewandelter Form zur Anwendung gebracht wird.

3.2.1 Lautsprecherdaten in EASE

In EASE werden die Lautsprecherdaten horizontal und vertikal in 5°-Schritten ausgewertet. Zur Definition der Koordinaten wird nicht das übliche System für Kugelkoordinaten mit Azimut und Polarwinkel benutzt, sondern die Punkte werden durch Winkel auf einem Halbkreisbogen bestimmt, der sich von der Forder- zur Rückseite zieht und sich 360° um den Lautsprecher dreht (Olson und Feistel Stefan 2002).

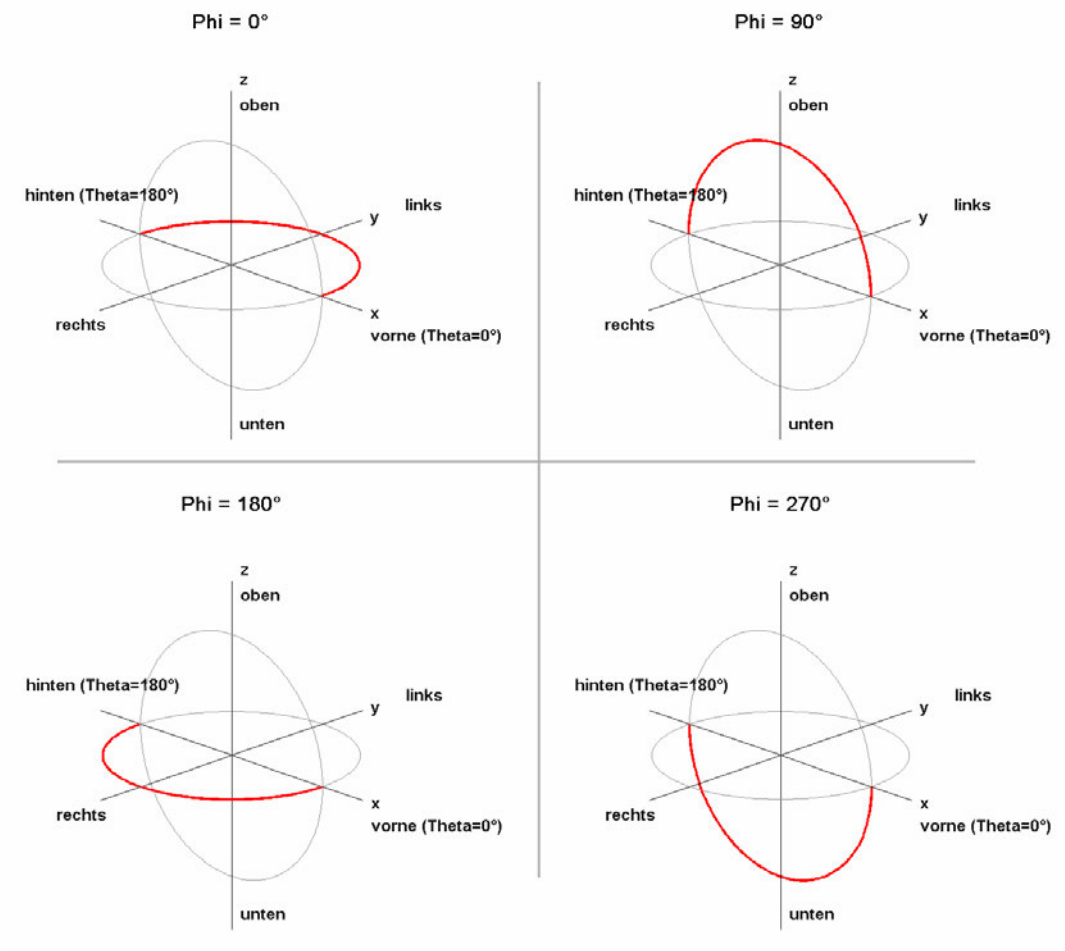

Abbildung 3.9 Das Koordinatensystem von EASE. Die vier Positionen des Halbkreisbogens, der sich von der Front zur Rückseite der Quelle erstreckt, sind für die Phiwinkel 0°, 90°, 180° und 270° jeweil s rot markiert.

In Abbildung 3.9 ist der prinzipielle Aufbau des Koordinatensystems von EASE dargestellt. Von der Vorderseite zur Rückseite zählt der Halbkreisbogen 0° bis 180° in 5°-Schritten. Das ergibt 37 Winkel, die mit Theta bezeichnet werden. Wenn man aus Sicht der Quelle nach vorne schaut, dreht sich der Halbkreisbogen mit dem Uhrzeigersinn von 0° nach 355° in 5°-Schritten um die Medianachse[1] der Quelle. Das ergibt 72 Winkel, die mit Phi bezeichnet

[1] Die Bezeichnung der drei Raumachsen und –ebenen stützt sich im Folgenden auf die in der Biologie zur Beschreibung der menschlichen Anatomie benutzten Begriffe. Die Richtung der Achsen ist wie

werden. Phi=0° entspricht dem Halbkreisbogen auf der Horizontalebene links von der Quelle. 90° ist demzufolge oben, 180° rechts und 270° unten . Die Lautsprecherdaten werden also in EASE auf einem Netz von 72 mal 37 Winkeln erfasst.

3.2.2 Interpolationsmethoden

Die Koordinaten der 32 äquidistanten Messmikrofone liegen im konventionellen Kugelkoordinatensystem vor. Phi entspricht dem Azimut, welcher sich bei Blick in die negative Richtung der Vertikalachse auf der Horizontalebene entgegen dem Uhrzeigersinn um 360° dreht und auf der positiven Seite der Medianachse zu zählen beginnt. Theta entspricht dem Polarwinkel, der von oben nach unten 180° zählt. Bei einer Transformation der Messpunkte in das kartesische Koordinatensystem entspricht die x-y-Ebene der Horizontalebene, wobei die positive Richtung der x-Achse nach vorne zeigt (Abbildung 3.10). Die positive y-Achse zeigt nach links. Die z-Achse entspricht der Vertikalachse und verläuft in positiver Richtung nach oben.

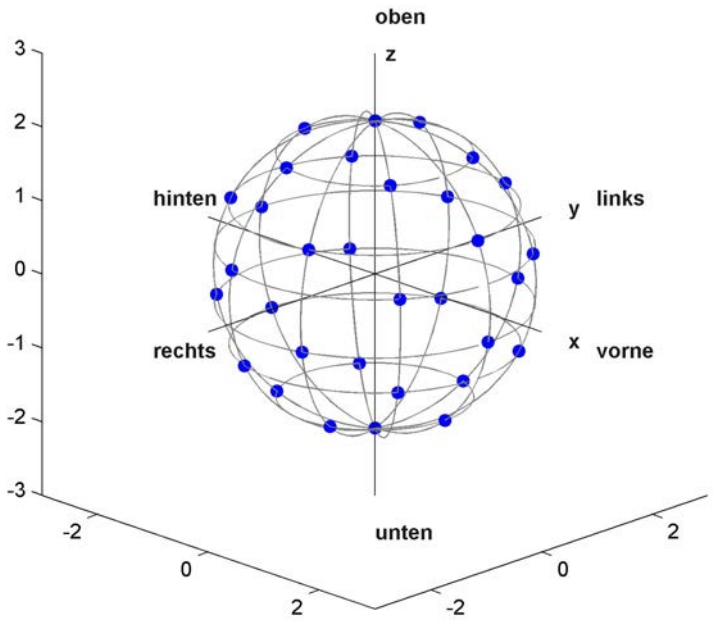

Abbildung 3.10 Positionen der 32 gleichmäßig auf einer Kugeloberfläche verteilten Messmikrofone. Die Bezeichnung x, y oder z einer Achse zeigt ihre positive Richtung an.

Die Koordinaten der Messmikrofone sind in Tabelle 2 zusammengefasst. Es gibt einen Messpunkt am Nordpol und einen am Südpol. Dazwischen befinden sich Messmikrofone auf jeweils drei Höhen der Nordhemisphäre und der Südhemisphäre. Der Azimut verläuft in 36°-Schritten. Das ergibt neun von oben nach unten verlaufende Halbkreisbögen, auf denen sich

folgt definiert: Die Medianachse verläuft von hinten nach vorne, die Frontalachse von rechts nach links und die vertikale Achse von unten nach oben.

jeweils zwei Messpunkte befinden. Die Nummerierung der Mikrofone stimmt nicht mit der Kanalnummer der aufgenommenen Multikanalaudiodateien überein. Eine Zuordnungsliste der beiden Nummerierungen ist in Anhang B zu finden.

Mik. Nr.	Polarwinkel [Grad]	Azimut [Grad]	
1 - 5	37,38	36; 108; 180; 252; 324	nördliche Hemisphäre
6 - 10	79,19	36; 108; 180; 252; 324	"
11 - 15	100,81	72; 144; 216; 288; 0	südliche Hemisphäre
16 - 20	142,62	72; 144; 216; 288; 0	"
21	0	0	Nordpol
22 - 26	63,43	0; 72; 144; 216; 288	nördliche Hemisphäre
27 - 31	116,57	108; 180; 252; 324; 36	südliche Hemisphäre
32	180	0	Südpol

Tabelle 2 Positionen der Messmikrofone um die Schallquelle.

Um die Messwerte auf den Koordinaten von EASE zu interpolieren, muss zunächst eine Transformation des Koordinatensystems vorgenommen werden. Diese Transformation kann durch zwei Drehungen des Kugelkoordinatensystems erzielt werden. Die Punkte der Messdaten sind in einem rechtshändigen Koordinatensystem definiert. Das heißt: Wenn man in die negative Richtung einer Achse schaut, verläuft die positive Drehrichtung entgegen dem Uhrzeigersinn.

Zunächst wird das System mit 90° um die y-Achse nach vorne gedreht. Das ergibt ein Koordinatensystem, bei dem Theta einen Halbkreisbogen beschreibt, der mit Phi im Uhrzeigersinn um die x-Achse rotiert. Die 0°-Stellung dieser Rotation befindet sich auf der unteren Hälfte der vertikalen Ebene. Das entspricht fast dem Koordinatensystem von EASE. Damit die Halbkreisrotation aus Sicht der Quelle links beginnt und nicht unten, muss das Koordinatensystem ein weiteres Mal gedreht werden, in diesem Fall 90° um die x-Achse.

Um aus den verteilten Messpunkten des Arrays eine kontinuierliche Abstrahlcharakteristik zu gewinnen, gibt es unterschiedliche Methoden. Eine Rekonstruktion der Abstrahlcharakteristik durch die Überlagerung von Kugelflächenfunktionen wurde von Zotter (2007) durchgeführt. Dabei muss angenommen werden, dass es sich bei der Quelle um einen Punktstrahler handelt. Ein weiteres Verfahren ist die Multipolsynthese (Ochmann 1990). Hier wird die Abstrahlung aus einer Kombination von Monopolquellen „re-synthetisiert“. Dieser Ansatz eignet sich auch für ausgedehnte Quellen. Eine eher konventionelle Methode wurde von Pätynen (2009) vorgestellt. Er hat eine Konvertierung von Messdaten eines Kugelflächenarrays durch eine Interpolation in der 2D-Ebene durchgeführt. Es wurden dabei zwei Vorgehensweisen vorgeschlagen, welche im Folgenden für den ähnlichen Fall einer Konvertierung der Messdaten des 32-Mikrofonarrays in das Lautsprecherformat von EASE ausgewertet werden.

3.2.2.1 Methode der Koordinatenabbildung

Im ersten Ansatz werden die Messwerte zunächst im sphärischen Koordinatensystem auf der Fläche einer Kugel interpoliert. Anschließend werden die Koordinaten der EASE-Lautsprecherdaten auf der Kugeloberfläche bestimmt, um diesen die interpolierten Werte zuzuordnen.

Die Interpolation wird in Matlab mit der Funktion *griddata* durchgeführt, mit der man auf einer Fläche innerhalb der Grenzen von verteilten Messpunkten auf ein definiertes feinmaschigeres Netz von Messpunkten interpolieren kann.

Die Werte an den EASE-Koordinaten werden ermittelt, indem diesen der geometrisch nächstgelegene Punkt auf der interpolierten Kugeloberfläche zugeordnet wird. Die Rasterung der Oberfläche muss daher kleiner sein als das 5°-Gitter der EASE-Koordinaten. Um für diesen Zweck eine hinreichende Auflösung zu gewährleisten und zugleich den Rechenaufwand im Auge zu behalten, wird mit Winkelabständen von 2° interpoliert.

Bei der Interpolation mit *griddata* wird die Kugeloberfläche in eine zweidimensionale Ebene gelegt, in der die x-Achse den Azimut und die y-Achse den Polarwinkel darstellt. In Abbildung 3.11 ist die Interpolation des gemittelten Schalldruckpegels der Stimme für die Terz mit der Mittenfrequenz $f_m = 4kHz$ dargestellt. Man sieht, dass die Funktion *griddata* nur Daten innerhalb der Fläche, die sich aus den vorhandenen Koordinaten ergibt, interpoliert. Die blauen Kreise markieren die Messwerte an den Mikrofonpositionen und die farbige Fläche stellt den daraus interpolierten Bereich dar. Die Ziffern an den Messwerten entsprechen den Kanalnummern der eingelesenen WAV-Dateien.

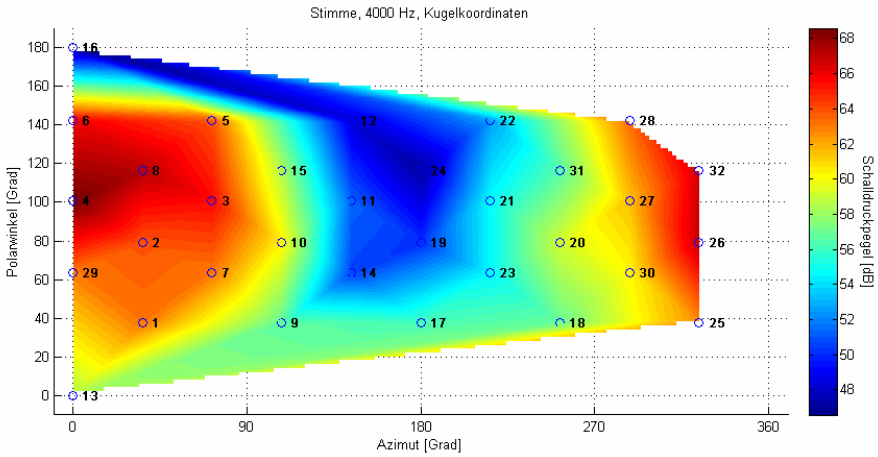

Abbildung 3.11 Interpolation des gemittelten Schalldruckpegels der Stimme in Kugelkoordinaten bei einer Terz von f_m = 4 kHz. Die blauen Kreise mit den Kanalnummern stellen die Werte der Messmikrofone dar.

Da durch die Interpolation auf einer zweidimensionalen Ebene die Kugelfläche auf der Linie zwischen den Polstellen bei Phi=0° aufgetrennt wird, entsteht eine Unterbrechung der Oberfläche. Die Periodizität der Kugeloberfläche kann wieder hergestellt werden, indem die Interpolationsfläche durch Kopieren der Datenpunkte um 360° in positiver Richtung des Azimuts erweitert wird (Abbildung 3.12).

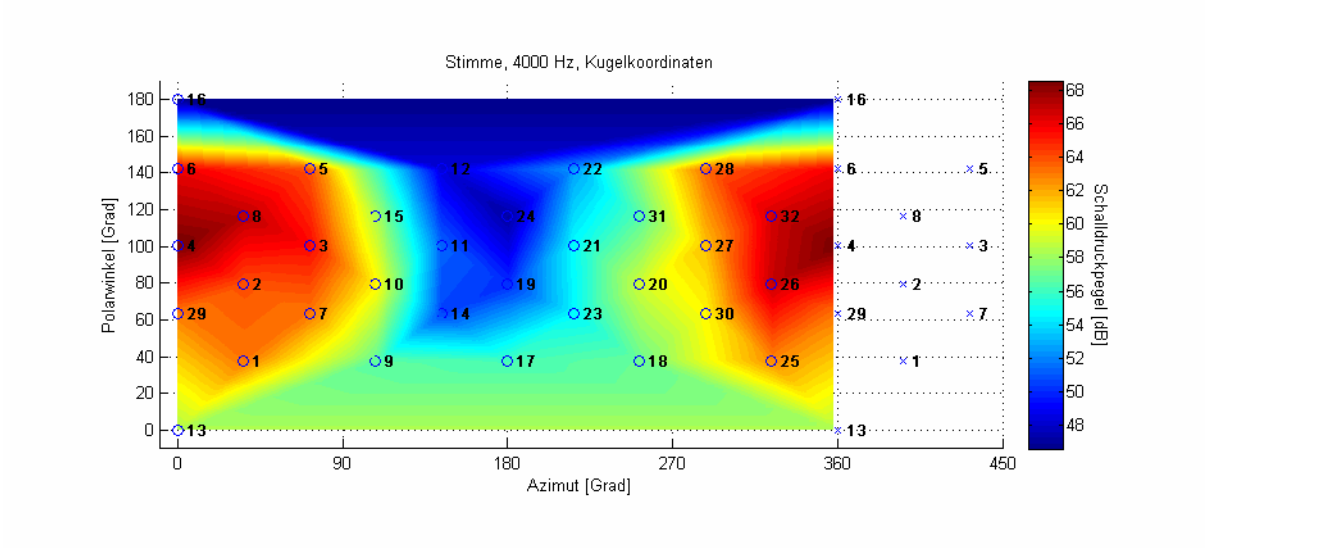

Abbildung 3.12 Interpolation des gemittelten Schalldruckpegels der Stimme in Kugelkoordinaten bei einer Terz von f_m = 4 kHz. Die blauen Kreuze mit den Kanalnummern stellen die kopierten Messwerte dar.

Für die Interpolation auf einer Kugeloberfläche müssen die Werte bei Annäherung aus allen Richtungen an einem bestimmten Punkt konvergieren. Um bei Phi=0° eine Interpolation durch Werte auf beiden Seiten der Trennlinie herzustellen, müssen die Messwerte auch um eine Periodenlänge auf die negative Seite des Azimuts kopiert werden.

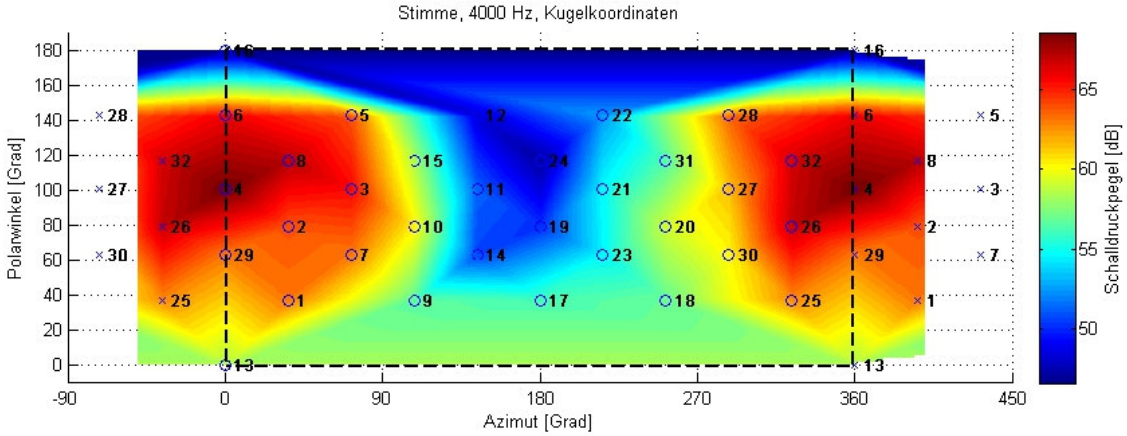

Abbildung 3.13 Interpolation des gemittelten Schalldruckpegels der Stimme in Kugelkoordinaten bei einer Terz von f_m = 4 kHz. Die blauen Kreuze mit den Kanalnummern stellen die kopierten Messwerte dar. Die Interpolation wurde zur Veranschaulichung der Periodizität in negativer und positiver Richtung um 45° einer Kreisumdrehung erweitert. Das gestrichelte Rechteck entspricht dem herausgeschnittenen Bereich einer vollständigen Kugelfläche.

In Abbildung 3.13 wurde zur Veranschaulichung des periodischen Verlaufs auch ein Teil des kopierten Bereichs interpoliert. Der Abschnitt einer Kugelumdrehung wird aus der Fläche der kopierten Datenpunkte, wie durch die gestrichelte Linie dargestellt, herausgeschnitten. Der homogene Verlauf der Interpolation auf einer Kugeloberfläche ist jetzt in der zweidimensionalen Darstellung über alle Winkel einer Kreisumdrehung des Azimuts wieder hergestellt.

Vom Auftrennen der Kugeloberfläche sind auch die Polstellen durch eine Unterbrechung der Homogenität betroffen. Für jeden Bogen, den der Polarwinkel von oben nach unten durchschreitet, werden die Polstellen aufgerastert. Aus den Polstellenpunkten im dreidimensionalen Raum wird in der zweidimensionalen Ebene eine Polstellenlinie entlang des Azimuts bei Theta=0° (Nordpol) und Theta=180° (Südpol). Die interpolierten Werte müssen bei Annäherung aus allen Richtungen an den Polstellen konvergieren. Dies wird erreicht, indem die Messwerte der Pole (Kanal 16 und 13) auf die Rasterpunkte der Polstellenlinien kopiert werden (Abbildung 3.14).

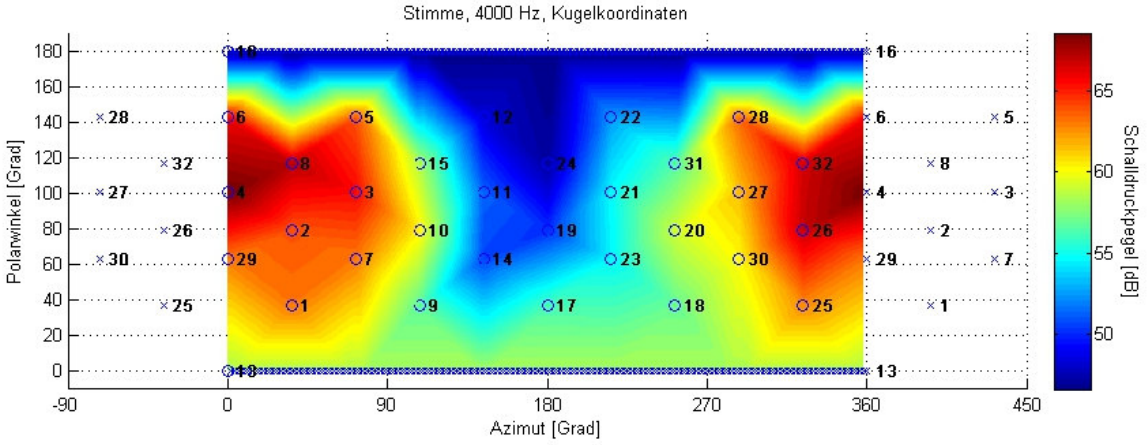

Abbildung 3.14 Interpolation des gemittelten Schalldruckpegels der Stimme in Kugelkoordinaten bei einer Terz von f_m = 4 kHz. Die blauen Kreuze stellen die kopierten Messwerte dar. Die Messwerte der Polstellen wurden in 2°-Abständen entlang des Azimu ts kopiert.

Nachdem durch das Kopieren der Datenpunkte eine korrekte Interpolationsfläche zwischen Azimut gleich 0° bis 360° hergestellt wurde, kann d iese aus der erweiterten Fläche von Messwerten herausgeschnitten und als 91x180-Datenmatrix abgespeichert werden.

Die Funktion *griddata* stellt vier Methoden der Interpolation zur Verfügung. Eine grafische Auswertung ist in Abbildung 3.15 dargestellt.

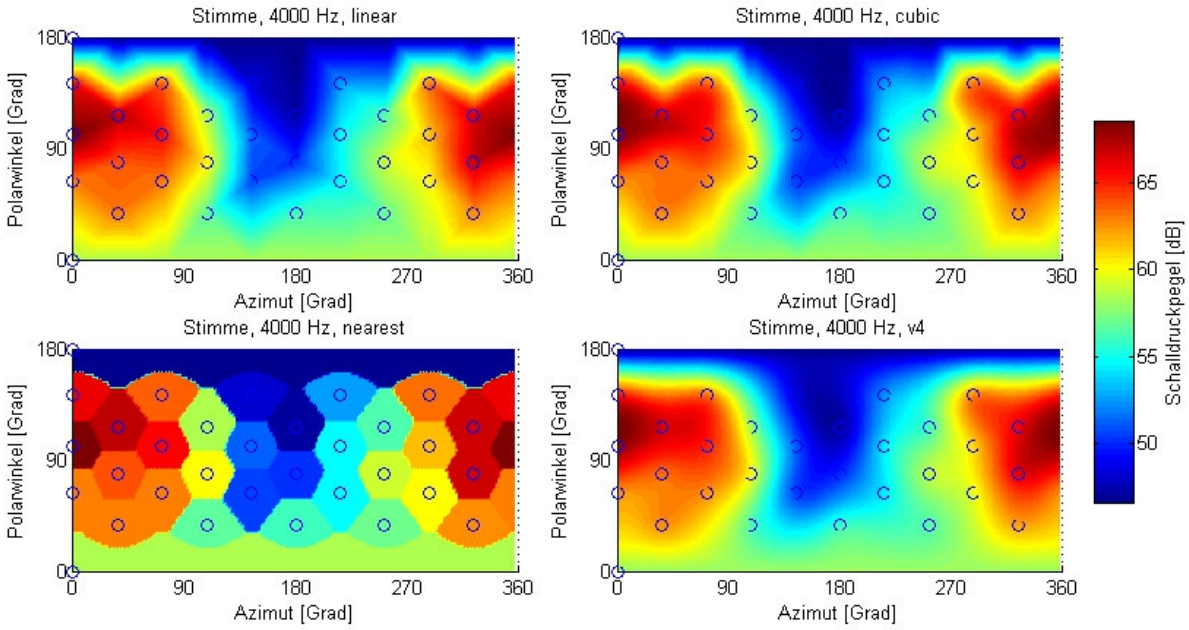

Abbildung 3.15 Interpolation des gemittelten Schalldruckpegels der Stimme in Kugelkoordinaten bei einer Terz von f_m = 4 kHz. Darstellung der Interpolationsergebnisse für die vier von *griddata* zur Verfügung gestellten Methoden.

Bis hierhin wurde die Standardmethode *linear* benutzt. Sie basiert auf einer linearen Delaunay-Triangulation. C*ubic* interpoliert mit einer quadratischen Delaunay-Triangulation. *Nearest* steht für eine nearest-neighbor-Interpolation und ist für diese Zwecke ungeeignet. Mit der Methode *v4* wird eine biharmonische Spline-Interpolation angewendet, ein Algorithmus, der zur Interpolation von Höhenmessungen von Geosatelliten entwickelt wurde (Sandwell 1987). Während sich die interpolierten Flächen für *linear* und *cubic* kaum unterscheiden, ergibt die Spline-Interpolation einen wesentlich glatteren Verlauf. In Tabelle 3 sind die Werte für die durchschnittliche und die maximale Abweichung zwischen den Interpolationsmethoden für die Stimme bei einem Terzband von 4 kHz dargestellt.

Methoden	Quadratischer Mittelwert der Abweichung [dB]	Maximale Abweichung [dB]
linear – cubic	0,56	1,89
linear – v4	0,96	5,02
cubic – v4	0,64	3,62

Tabelle 3 Abweichungen der Interpolationsmethoden von griddata für den Mittelungspegel der Stimme bei f_m = 4 kHz.

Bei einer Abstrahlcharakteristik, die den Schall in unterschiedliche Richtungen des Raumes streut, sind die Abweichungen erwartungsgemäß kleiner. Für die Bratsche bei 1 kHz ist der

maximale Unterschied zwischen *linear* und *v4* nur 1,7 dB. Aber auch für eine Abstrahlcharakteristik mit stärkerer Richtwirkung wie beispielsweise der Trompete bei 4 kHz fallen die Abweichungen zwischen den Methoden nicht stärker aus. Zwischen *linear* und *v4* ergibt sich hier nur ein Unterschied von maximal 2,78 dB.

Da die Spline-Methode (*v4*) die homogenste Interpolation bietet, wird im Folgenden diese Methode bei Berechnungen mit *griddata* verwendet. Abbildung 3.16 zeigt die aus dem beschriebenen Verfahren resultierenden dreidimensionalen Abstrahlcharakteristiken. Die Geometrie der Hüllfläche entspricht durch ihren Abstand vom Ursprung des Koordinatensystems dem Verlauf des Pegels.

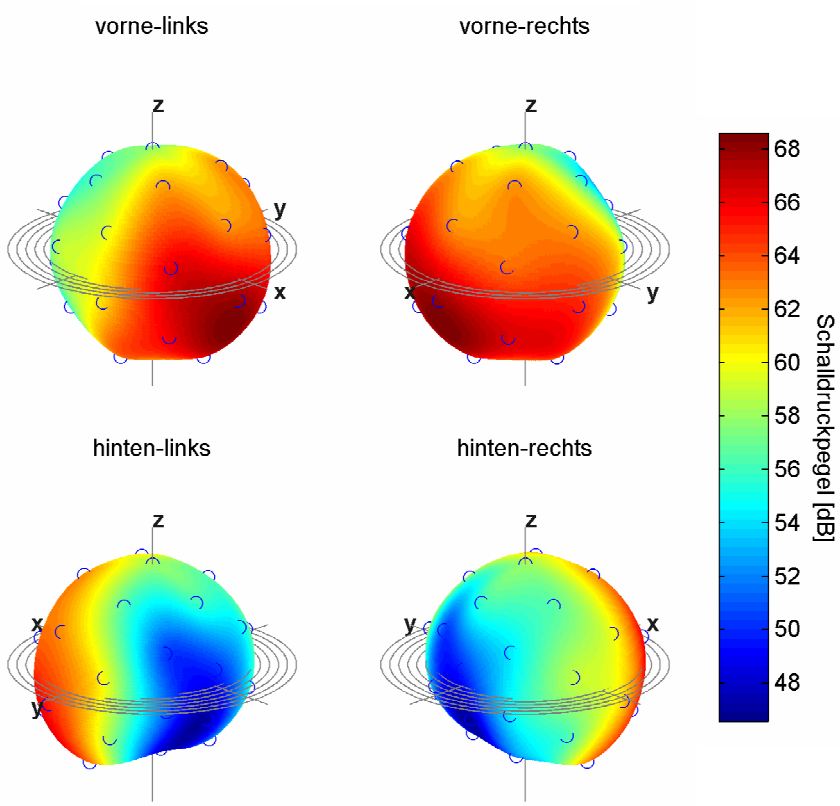

Abbildung 3.16 Interpolation des gemittelten Schalldruckpegels der Stimme in Kugelkoordinaten bei einer Terz von f_m = 4 kHz. Der Abstand der Kreislinien beträgt 5 dB. Die kleinen blauen Kreise zeigen die Position und die Pegelwerte der Messmikrofone an.

Um die interpolierten Werte auf die Koordinaten von EASE abzubilden, muss zunächst ein entsprechendes Koordinatensystem erzeugt und auf die Kugeloberfläche gelegt werden. Mit der Funktion *meshgrid* wird im sphärischen Koordinatensystem ein Gitter erzeugt, das in 5°-Schritten den Halbkreisbogen des Polarwinkels mit fortschreitendem Azimut über die Kugeloberfläche zieht (Abbildung 3.17).

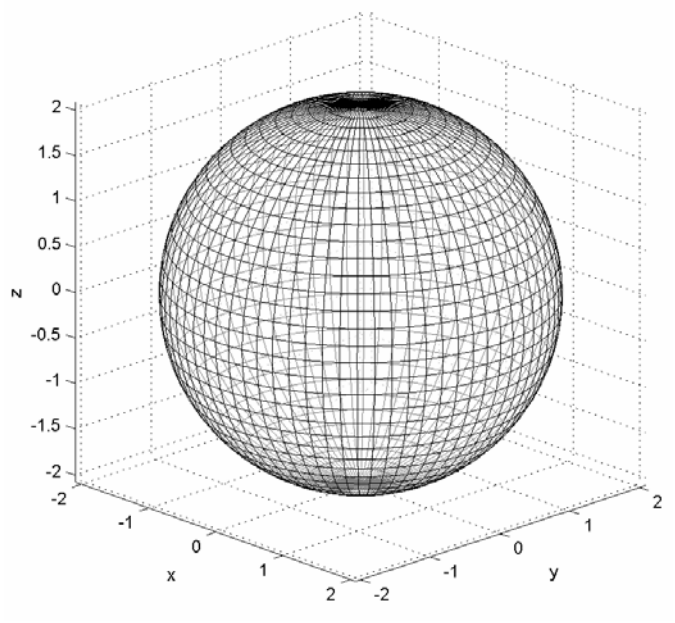

Abbildung 3.17 5°-Koordinatennetz vor der Drehung

Anschließend wird dieses Koordinatennetz wie oben beschrieben durch Drehungen in das System von EASE übertragen:

Die Transformation wird im kartesischen System durchgeführt und durch eine Drehmatrix beschrieben. Ein Punkt $\vec{p} = (x, y, z)$ wird durch Multiplikation mit der Drehmatrix

$$\begin{pmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{pmatrix} \tag{1.5}$$

mit dem Winkel α um die y-Achse gedreht. Für $\alpha = 90°$ ergibt sich daraus für die Koordinaten des gedrehten Punktes $\vec{p}\,'$

$$\vec{p}\,' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} z \\ y \\ -x \end{pmatrix}$$

Nach der Transformation der Punkte des Koordinatennetzes in das kartesische System und anschließender Multiplikation mit der Drehmatrix müssen die Datenpunkte noch einmal 90° um die x-Achse gedreht werden. Für die Koordinaten eines zweimal gedrehten Punktes $\vec{p}\,''$ ergibt sich

$$\vec{p}\,'' = \begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} = \vec{p}\,' \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} x' \\ -z' \\ y' \end{pmatrix}$$

Mit dem Koordinatennetz, das jetzt die korrekten Punkte der EASE-Lautsprecherdaten enthält, werden nun die am nächsten liegenden Koordinaten auf der interpolierten Kugeloberfläche gesucht (Abbildung 3.18).

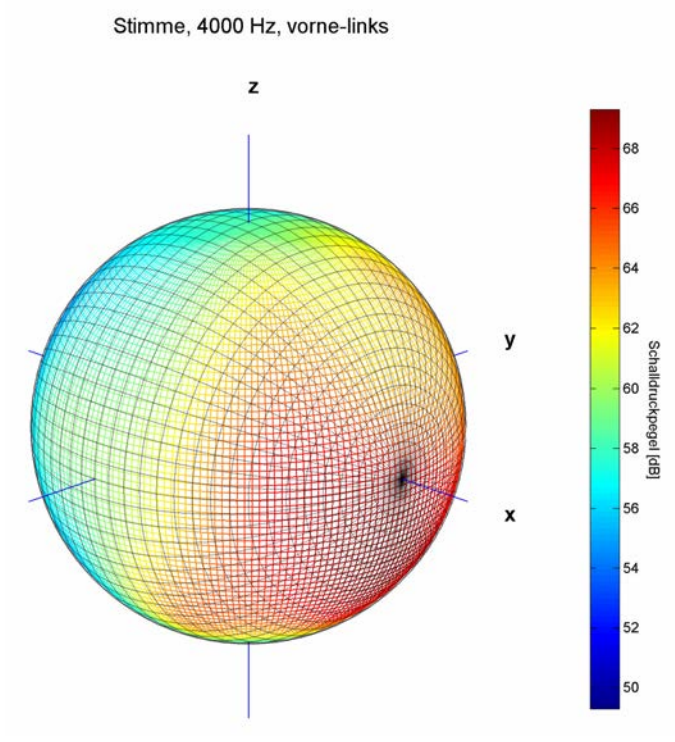

Abbildung 3.18 Interpolation des gemittelten Schalldruckpegels der Stimme bei einer Terz von f_m = 4 kHz. Das 5°-Koordinatennetz von EASE wurde auf das interpolierte 2°-Gitter gelegt.

Mithilfe der Matlab-Funktion *dsearchn* werden die Indizes der am nächsten liegenden Punkte zwischen den zwei Koordinatenmatrizen gefunden. Der Fehler bei der Zuordnung von einem 2°- auf ein 5°-Koordinatennetz liegt maximal bei ei nem Winkelabstand von 1,39° und im Durchschnitt bei 0,65°.

Mit Hilfe der Indizes werden die entsprechenden Werte der Schalldruckmatrix den EASE-Koordinaten zugeordnet. Man erhält eine 37x72-Matrix, die die korrekten Pegel für alle Werte von Phi und Theta enthält. Das grafische Ergebnis der Interpolation und Transformation in das EASE-Format ist in Abbildung 3.19 dargestellt. Die quadratische Musterung des Farbverlaufs entspricht der 5°-Winkelauflösung der EASE-Koordinaten.

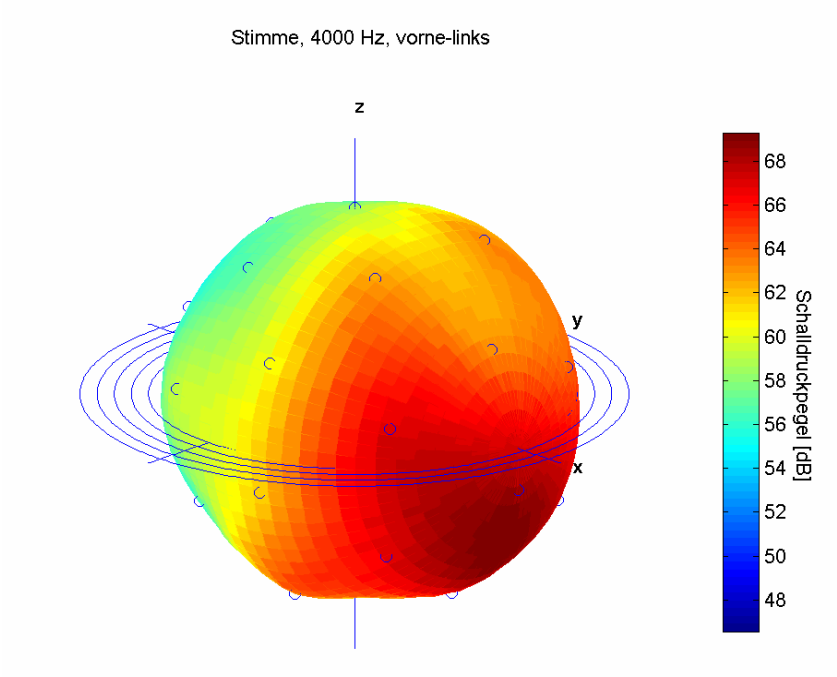

Abbildung 3.19 Interpolation des gemittelten Schalldruckpegels der Stimme bei einer Terz von f_m = 4 kHz im EASE-Lautsprecherformat nach der Methode der Koordinatenabbildung.

3.2.2.2 Methode der direkten Koordinatentransformation

Der zweite Ansatz basiert darauf, dass die Transformation des Koordinatensystems vor der Interpolation vorgenommen wird. Das Koordinatensystem muss wie oben beschrieben einmal $90°$ um die Frontalachse und dann noch einmal $90°$ um die Medianachse der Quelle gedreht werden. Die Rotation eines Koordinatensystems entspricht einer entgegengesetzten Drehung der Koordinatenpunkte. Die kartesischen Koordinaten der Messmikrofone $\vec{p}_{mik} = (x, y, z)$ werden also nach (1.5) zunächst mit $\alpha = -90°$ um die y-Achse gedreht:

$$\vec{p}'_{mik} = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \vec{p}_{mik} \begin{pmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -z \\ y \\ x \end{pmatrix}$$

Bei einer Rotation des Koordinatensystems dreht sich mit den Koordinaten die Orientierung im Raum. Das bedeutet, dass sich die Ausrichtung der x-, y- und z-Achse ändert. In Abbildung 3.20 wurde der Blickwinkel in Bezug auf die Achsenbezeichnungen zunächst beibehalten.

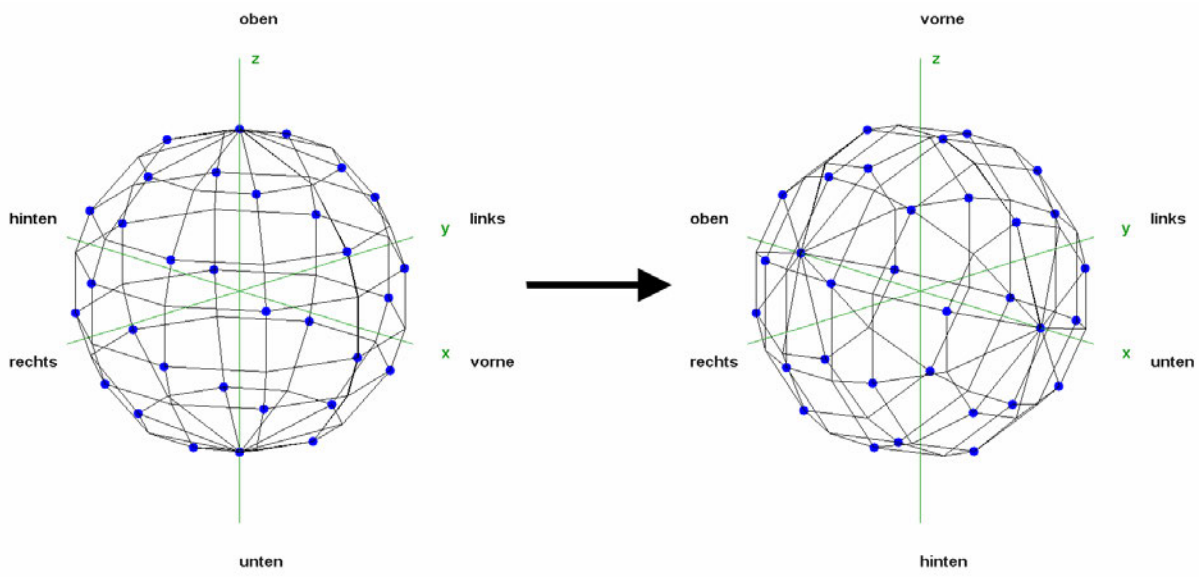

Abbildung 3.20 Drehung der Koordinaten der Messmikrofone mit $\alpha = -90°$ um die y-Achse

Das Koordinatensystem beschreibt jetzt durch den ehemaligen Polarwinkel Halbkreisbögen, die mit dem ehemaligen Azimut um die Medianachse der Quelle rotieren. Die Rotation der Bögen beginnt mit 0° auf der unteren Seite der Quelle. Da die Rotation im Koordinatensystem von EASE aus Sicht der Quelle links beginnt, müssen die Koordinaten ein weiteres Mal mit $\alpha = -90°$ um die z-Achse gedreht werden. Mit der entsprechenden Drehmatrix für eine Rotation um die z-Achse werden die neuen Koordinaten der Messmikrofone berechnet:

$$\vec{p}''_{mik} = \begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} = \vec{p}'_{mik} \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} y' \\ -x' \\ z' \end{pmatrix}$$

Man beachte, dass sich der dreidimensionale Raum ein weiteres Mal mit den Koordinaten dreht.

Um zur gleichen Ansicht des Arrays wie in der ursprünglichen Ausrichtung des Koordinatensystems zu gelangen, muss der Blickwinkel entsprechend korrigiert werden (Abbildung 3.21). In der rechten Darstellung von Abbildung 3.21 kann man gut erkennen, inwiefern sich das Koordinatensystem gedreht hat: Die z-Achse beschreibt nun die Medianachse, y entspricht der vertikalen und x der frontalen Achse.

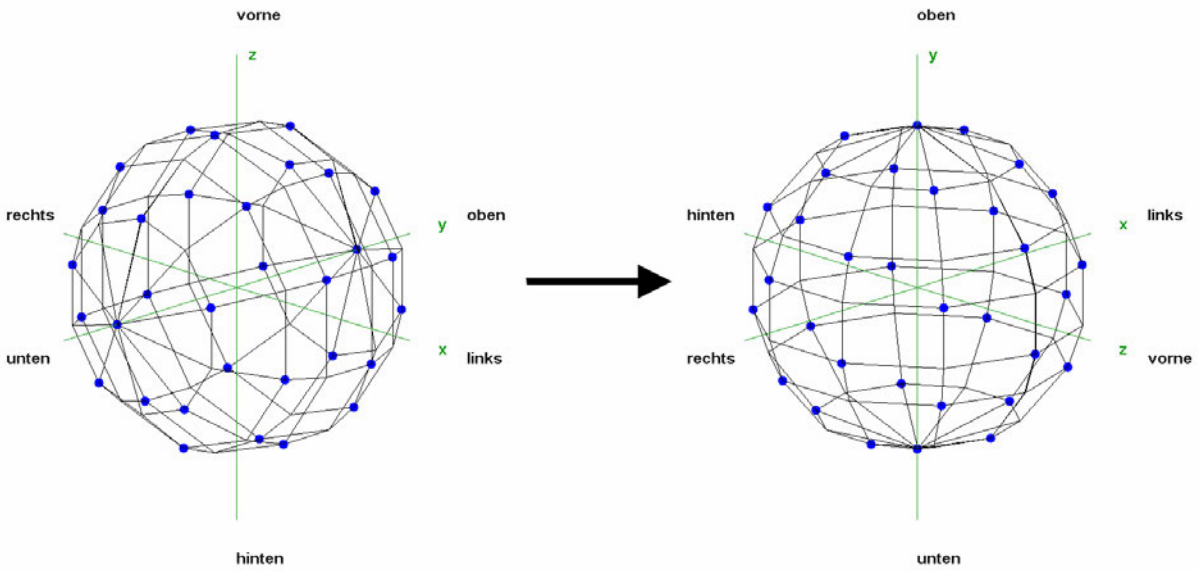

Abbildung 3.21 Mikrofonarray nach zweimaliger Drehung (links) und Änderung des Blickwinkels zur Korrektur in die ursprüngliche Ansicht (rechts).

Im gedrehten Koordinatensystem kann nun ein 5°-Netz erzeugt werden, welches sich über die Hüllfläche des Mikrofonarrays zieht und direkt den Koordinaten von EASE entspricht. Die Hüllfläche wird mit der Funktion *griddata* in eine zweidimensionale Ebene gelegt und an den Punkten des Koordinatennetzes interpoliert (Abbildung 3.22). Die x-Achse entspricht jetzt dem Rotationswinkel und die y-Achse dem Winkel auf dem rotierenden Halbkreisbogen.

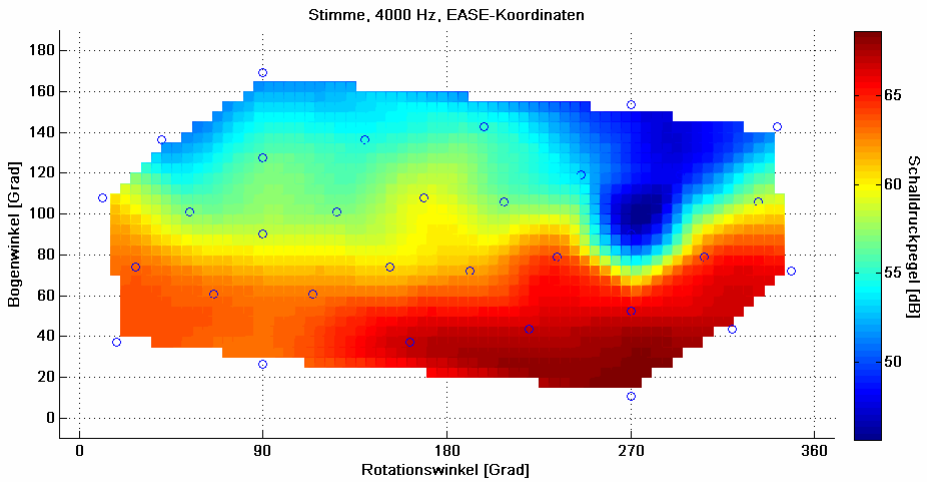

Abbildung 3.22 Verlauf der Interpolation des gemittelten Schalldruckpegels der Stimme in EASE-Koordinaten bei einer Terz von f_m = 4 kHz. Die blauen Kreise stellen die Werte der Messmikrofone dar.

Da *griddata* nur innerhalb der Grenzen der gegebenen Koordinaten interpoliert, müssen die Messwerte auch hier um eine Periodenlänge in positiver und negativer Richtung des

Rotationswinkels kopiert und der Bereich für eine Umdrehung des Rotationswinkels herausgeschnitten werden (Abbildung 3.23).

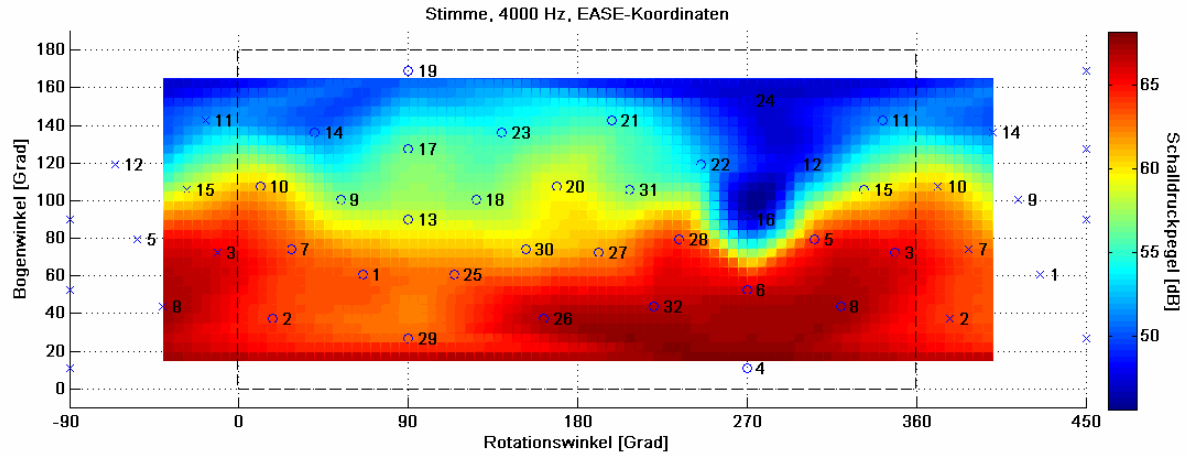

Abbildung 3.23 Interpolation des gemittelten Schalldruckpegels der Stimme in EASE-Koordinaten bei einer Terz von f_m = 4 kHz. Die blauen Kreuze stellen die kopierten Messwerte dar. Die Interpolation wurde zur Veranschaulichung der Periodizität in negativer und positiver Richtung des Rotationswinkels um 45° erweitert. Das gestrichelte Rechteck entspricht dem herausgeschnittenen Bereich einer vollständigen Kugelfläche.

In der Repräsentation des Mikrofonarrays in den Koordinaten von EASE existieren an den Polstellen vorne und hinten an der Quelle keine Messwerte. Wie man in Abbildung 3.23 sieht, führt das zu einer unvollständigen Interpolationsfläche. Das gleiche Problem trat bei dem Mikrofonarray von der TKK auf (Pätynen 2009). Pätynen hat dieses Problem gelöst, indem er die Werte der Polstellen aus den ihnen am nächsten liegenden Punkten linear interpolierte. Hier müsste man die Frontpolstelle aus den Werten von Kanal 2, 29 und 4 interpolieren. Kanal 19, 24 und 11 wären die Werte, aus denen man den rückwärtigen Pol ermitteln könnte. Bei einer linearen Interpolation werden die unterschiedlichen Entfernungen der Werte zu den jeweiligen Polen nicht berücksichtigt. Es ist daher ein gewisser Fehler zu erwarten. Da es aus dem Interpolationsverfahren der Koordinatenabbildung bereits einen gut interpolierten Wert für die Front- und Rückpolstelle gibt, liegt es nahe diese Werte mit der linearen Interpolation am Beispiel der Stimmenrichtcharakteristik zu vergleichen: Die Frontpolstelle hat in Kugelkoordinaten die Winkelwerte Phi_{Front} = 0° und $Theta_{Front}$ = 90°. Der Wert der 2°-Spline-Interpolation beträgt an dieser Stelle L_{Front} = 67,25 dB. Der Mittelwert aus Kanal 2, 29 und 4 beträgt 65,21 dB. Das ergibt eine Abweichung von rund 2dB. Für die Rückpolstelle mit den Kugelkoordinaten $Phi_{Rück}$ = 180° und $Theta_{Rück}$ = 90° beträgt der Pegel der Spline-Interpolation $L_{Rück}$ = 49,21 dB. Der Mittelwert aus Kanal 19, 24 und 11 berechnet sich zu 49,50 dB. Die Abweichung ist mit 0,3 dB wesentlich geringer.

Zur Bewahrung der Konvergenz werden die manuell interpolierten Werte der Polstellen, wie im Verfahren der Koordinatenabbildung, für alle Winkel der Halbkreisrotation und des

jeweiligen Thetawinkels für die Front- und Rückpolstellen kopiert. Abbildung 3.24 zeigt das vollständig interpolierte 5°-Koordinatennetz, das direkt dem Format von EASE entspricht.

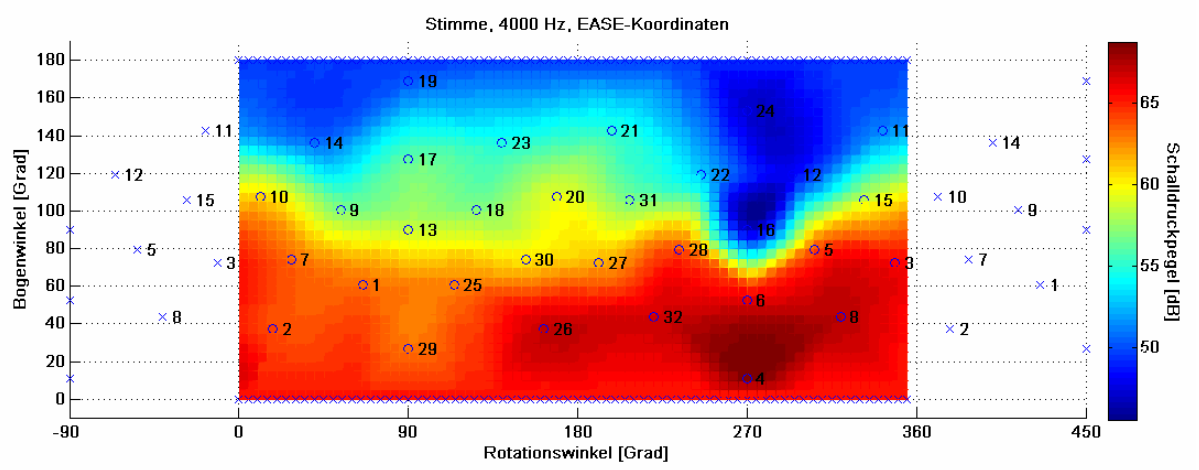

Abbildung 3.24 Verlauf der Interpolation des gemittelten Schalldruckpegels der Stimme in EASE-Koordinaten bei einer Terz von f_m = 4 kHz. Die blauen Kreuze stellen die kopierten Messwerte dar.

3.2.3 Vergleich und Diskussion der Ergebnisse

Zum Vergleich sind die Richtcharakteristiken beider Interpolationsverfahren in Abbildung 3.25 gegenübergestellt. Es ist sichtbar, dass sich an der Vorderseite und der unteren Rückseite der Quelle unterschiedliche Kurvenverläufe ergeben.

Bei den vorgestellten Methoden wird die Oberfläche einer Kugel in einer zweidimensionalen Ebene interpoliert. Die Transformation einer dreidimensionalen Fläche in eine Ebene ist nicht ohne eine gewisse Deformation möglich. Eine Fläche die sich aus zwei Winkelabständen für Phi und Theta ergibt wird abhängig von Theta zu den Polstellen immer kleiner während sie am Äquator am größten ist. Bei der zweidimensionalen Repräsentation der Kugeloberfläche ist der Abstand zwischen zwei Punkten $A(\varphi,\theta_1)$ und $B(\varphi,\theta_2)$ immer gleich. Dies hat zur Folge, dass sich die Abstände zwischen den Messmikrofonen verändern und die flächenmäßig gleichmäßige Verteilung auf die Kugeloberfläche nicht ganz bewahrt wird. Auch die zu interpolierenden Punkte erfahren Verzerrungen in ihren Abständen, die sich an den Polstellen, wo die Fläche aufgetrennt wurde, am stärksten auswirken. Daher ist die Interpolation dort am ungenauesten und es entstehen, bei einer Rückführung der interpolierten Werte auf eine Kugeloberfläche, Inhomogenitäten im Kurvenverlauf. Der Fehler wirkt sich bei beiden Interpolationsverfahren unterschiedlich stark auf bestimmte Richtungen aus. Da jeweils die Polstellen am meisten von der Deformation betroffen sind, kann man sagen, dass sich der Fehler bei der Interpolation in EASE-Koordinaten vor allem auf die frontale und rückwärtige Abstrahlung auswirkt, während die Abstrahlung nach oben und

unten recht gut interpoliert wird. Bei der Interpolation in Kugelkoordinaten wirkt sich der Fehler vor allem auf die vertikale Abstrahlung aus.

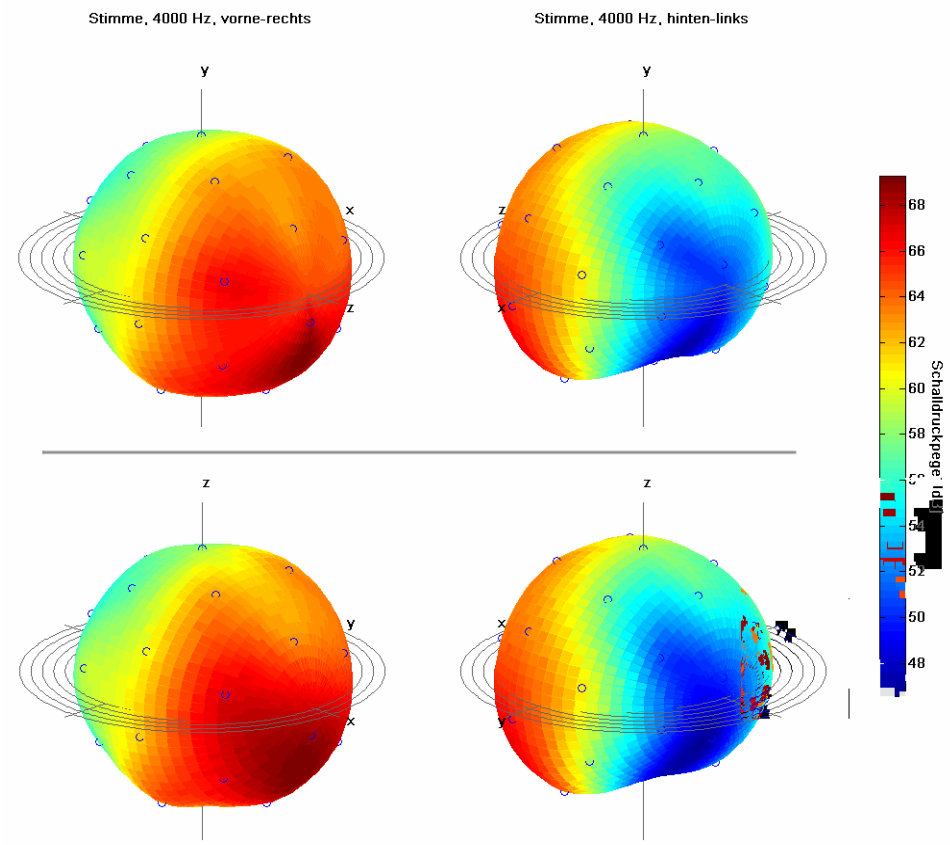

Abbildung 3.25 Interpolation des gemittelten Schalldruckpegels der Stimme bei einer Terz von f_m = 4 kHz. Oben: Direkte Interpolation im Koordinatensystem von EASE, Unten: Interpolation mit der Methode der Koordinatenabbildung.

Die Differenz beider Verfahren liegt für die Stimme bei einem Terzband von 4 kHz im Durchschnitt bei 0,74 dB und maximal bei 2,33 dB. Für den Zweck der Auralisation kann man den Unterschied wohl eher als gering betrachten (Pätynen 2009). Die Verteilung der Differenz auf die Kugelfläche ist in Abbildung 3.26 dargestellt. Man kann hier gut erkennen, wie sich der Unterschied beider Verfahren auf die Gegenden an den Polstellen, bei Theta gleich 0° (vorne) und 180° (hinten) sowie bei Phi g leich 90° (oben) und 270° (unten), auswirkt. Vor allem an den Stellen, wo der Pegel der Richtcharakteristik maximal bzw. minimal wird, also an den Stellen des Frontpols und in der Gegend des Südpols Richtung Rückachse, wirkt sich der Unterschied am stärksten aus.

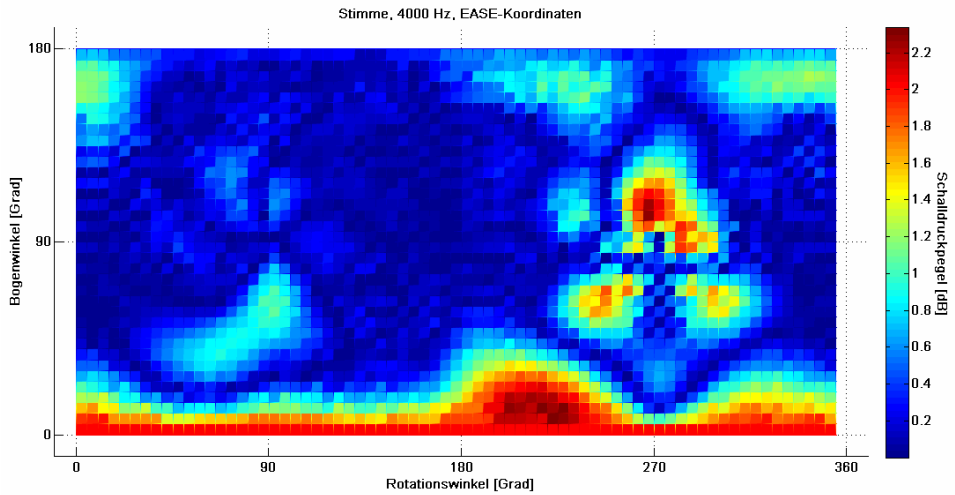

Abbildung 3.26 Betrag der Abweichung zwischen den Interpolationsverfahren der Koordinatenabbildung und der direkten Interpolation in EASE-Koordinaten für die Stimme bei einem Terzband von 4 kHz

Aus zwei Gründen fällt die Wahl für die weitere Auswertung der Daten auf die etwas komplexere Interpolation in den konventionellen Kugelkoordinaten mit anschließender Koordinatenabbildung. Erstens ist die horizontale Abstrahlung für die meisten Instrumente mehr von Bedeutung als die vertikale Abstrahlung und sollte demzufolge möglichst frei von Ungenauigkeiten sein. Zweitens müssen bei der Interpolation in EASE-Koordinaten die Front- und Rückpolstellen von Hand linear gemittelt werden. Dies führt zu mehr Ungenauigkeiten an diesen Stellen, als bei dieser Methode sowieso schon in Kauf genommen werden müssen.

Da in den gängigen Formaten[1] für Lautsprechermodelle in akustischen Simulationsprogrammen ein ähnliches System von um die Medianachse rotierenden Halbkreisbögen verwendet wird, könnte man für zukünftige Messungen mit Kugelflächenarrays in Betracht ziehen, an der Vorder- und Rückachse zusätzliche Messmikrofone aufzustellen. Sind die Werte der Rück- und Frontpolstellen bei der Messung ermittelt worden, wäre das direkte Interpolationsverfahren in den Koordinaten des rotierenden Halbkreisbogens schon eher in Betracht zu ziehen, da hier wiederum die Ungenauigkeiten der Koordinatenabbildung nicht in Kauf genommen werden müssen und der Rechenaufwand etwas sparsamer wäre.

3.2.4 Konvertierung der CATT-Daten des Flügels in das EASE-Format

Da bei den Messungen an der TU Berlin keine Daten für einen Flügel erhoben wurden, musste für die Auralisation dieses Instrumentes auf bereits bestehende Daten der PTB

[1] Z.B. CATT (siehe auch Kapitel 3.2.4), CLF, EASE

zurückgegriffen werden. Um die Werte in das Format von EASE zu konvertieren, muss eine Interpolation auf das 5°-Datennetz durchgeführt werden. Das Koordinatensystem für die Raumwinkel basiert in CATT ebenfalls auf einem Halbkreisbogen, der vor der Quelle mit 0° beginnt und hinten mit 180° abschließt. Die Rotation verläuft im Gegensatz zum EASE-Format, wenn man aus Sicht des Instrumentes nach vorne schaut, entgegen dem Uhrzeigersinn und beginnt senkrecht über der Quelle (Abbildung 3.27).

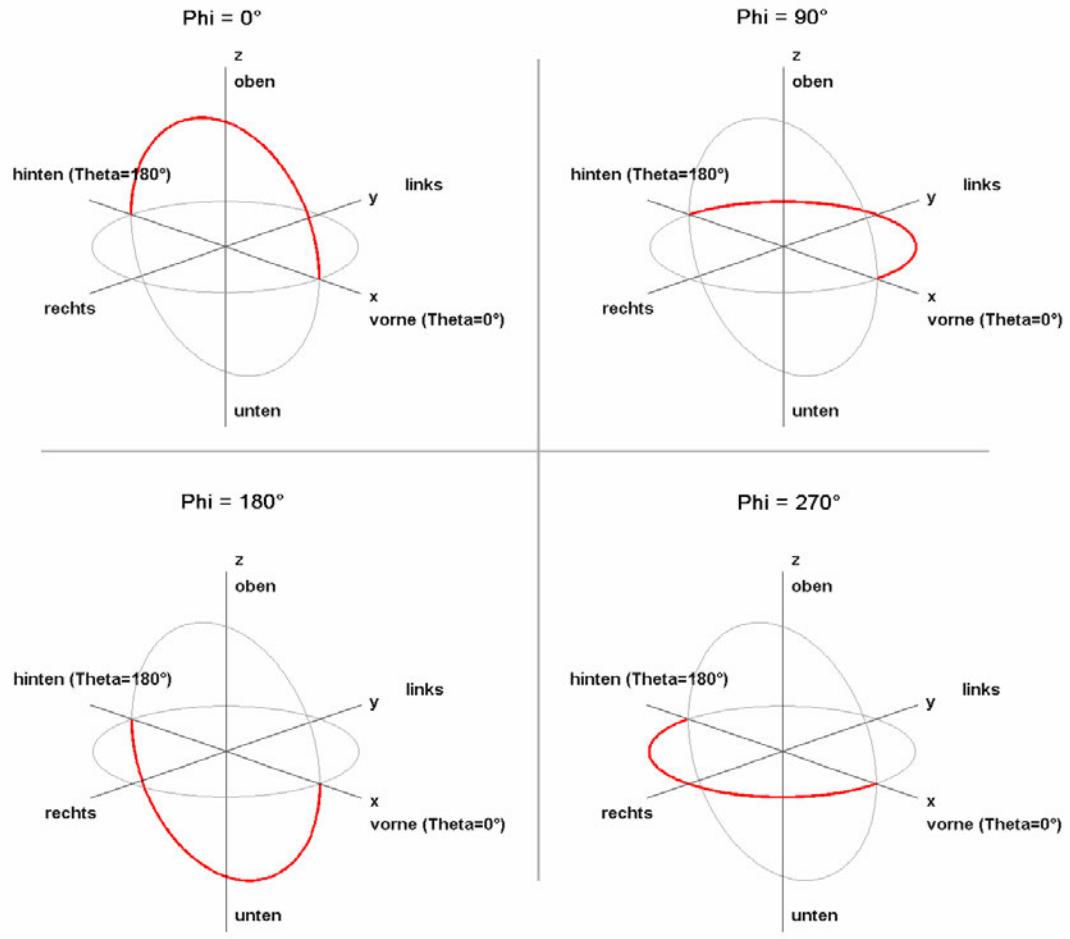

Abbildung 3.27 Das Koordinatensystem von CATT. Die Positionen des Halbkreisbogens sind für einen Rotationswinkel von 0°, 90°, 180° und 270° jeweils rot markiert.

Die Konvertierung der Daten wird so verlaufen, dass die Koordinaten zunächst transformiert und anschließend die Werte interpoliert werden. Der frequenzmäßig höheren Terzbandauflösung von EASE kann nur entsprochen werden, indem lediglich die resultierenden Messwerte der Oktaven für jedes der in ihr enthaltenen drei Terzbänder kopiert werden. Eine Extrapolation wird hier nicht durchgeführt, da dies aufgrund der nicht ohne weiteres vorhersagbaren Komplexität der Abstrahlung eines Flügels bei unterschiedlichen Frequenzen nicht plausibel erscheint.

Die PTB stellt ihre Daten als Textdatei und in dem mit CATT erstellten SD1-Format zur Verfügung. Eine Untersuchung beider Dateien ergab, dass sich ihr Frequenzumfang

unterscheidet. Während in der Textdatei nur Werte für die Oktaven von 1 kHz, 2kHz und 4kHz angegeben werden, findet man beim Öffnen der SD1-Datei in CATT auch Werte für die Oktaven von 125Hz, 250 Hz und 500Hz, so wie es auch in dem Infotext der Webseite beschrieben wird. Warum die unteren Frequenzen im Textfile fehlen, ist unbekannt. Zum Importieren der Daten in Matlab ist nur das Format als Textdatei brauchbar. Die Werte der untersten drei Oktaven werden in CATT abgelesen und von Hand in Matlab importiert.

Da jeder Messpunkt der SD1-Daten auf einem Winkel der Kugelfläche liegt, an dem auch Messwerte in EASE definiert sind, kann die Transformation der Koordinaten durch eine einfache Abbildung durchgeführt werden: Beide Systeme werden übereinander gelegt, um anschließend die Werte auf die Koordinaten von EASE zu kopieren.

Die Interpolation wird wieder mit der Matlab-Funktion *griddata* durchgeführt. Die Daten werden in eine zweidimensionale Ebene gelegt, in der die x-Achse dem Winkel der Rotation (Phi) und die y-Achse dem Winkel auf dem Halbkreisbogen (Theta) entspricht (Abbildung 3.28). Als Methode wird für *griddata* die biharmonische Spline-Interpolation (*v4*) gewählt. Das 10°-Raster der Ausgangswerte muss durch Kopieren von Messpunkten gemäß dem periodischen Verlauf von Kugelumdrehungen um 360° in negativer und positiver Richtung des Rotationswinkels erweitert werden, um auf der Trennlinie der Kugelfläche (bei Phi=0° bzw. Phi=360°) eine Interpolation der Werte aus allen Richtungen zu gewährleisten. Die Werte der Frontpol- und Rückpolstelle werden in der zweidimensionalen Darstellung für die Thetawinkel der Pole, 0 Grad und 180 Grad, auf 10°-Schritte des Rotationswinkels aufgerastert. Um eine Konvergenz der Messwerte bei Annäherung an die Polstelle aus allen Richtungen zu gewährleisten, werden ihre Werte auf die 5°-Schritte des Rotationswinkels kopiert.

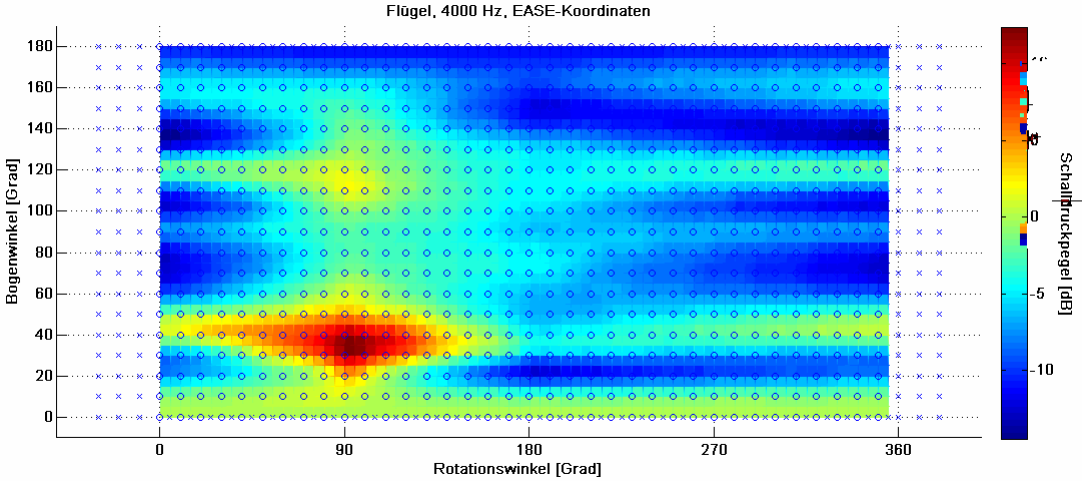

Abbildung 3.28 Interpolation des Schalldruckpegels in EASE-Koordinaten für den Flügel bei einer Oktave von f_m = 4 kHz. Die blauen Kreise stellen die Pegelwerte der CATT-Daten und die blauen Kreuze stellen die kopierten Werte dar. Der Pegel ist auf die 0°-Achse normiert.

Das Ergebnis ist für das Oktavband bei 4 kHz in Abbildung 3.28 dargestellt. Der Pegel der CATT-Daten war bereits auf die Bezugsrichtung bei Phi und Theta gleich null Grad normiert. Das heißt alle von null abweichenden Pegelwerte sind als Differenz zur Bezugsrichtung zu verstehen.

Eine Grafik, die von der PTB mit dem Quellmodell zur Verfügung gestellt wird, veranschaulicht wie das Koordinatensystem bei der Messung ausgerichtet war (Abbildung 3.29).

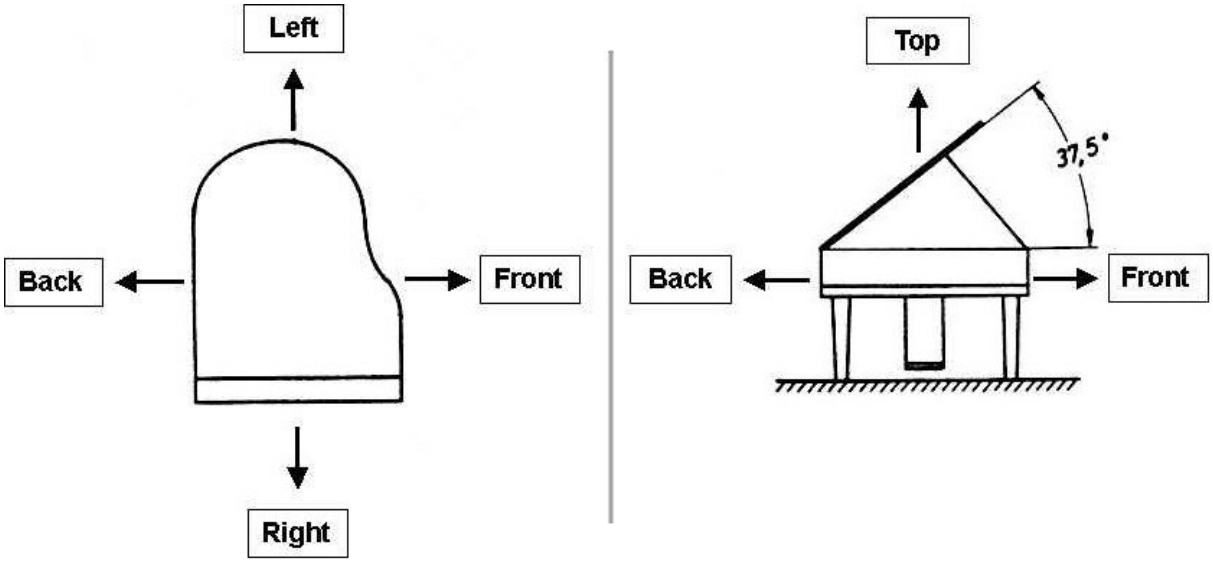

Abbildung 3.29 Ausrichtung des Koordinatensystems bei der Messung des Flügels. Quelle: Homepage der Physikalisch-Technischen Bundesanstalt.

Wissenswert ist weiterhin, dass die Messung mit einem Schwingerreger durchgeführt wurde. Das bedeutet, der raumakustische Einfluss des Spielers ist nicht in der Richtcharakteristik des Flügels enthalten.

3.3 Bestimmung der Bezugsrichtung

Lautsprecherdaten werden in akustischen Simulationsprogrammen allgemein so ausgewertet, dass die Hauptabstrahlrichtung des Lautsprechers die Bezugsrichtung ist. Wird die Pegeldifferenz als Verhältnis der beiden Schalldrücke ausgedrückt, bezeichnet man dieses Verhältnis als Richtungsfaktor. In EASE werden die Pegeldifferenzen zur Bezugsrichtung als so genannte „attenuation“-Werte, also als Dämpfungswerte bestimmt. Das heißt, ein positiver Pegelwert steht für eine Abschwächung und ein negativer Wert für eine Verstärkung des Pegels in Bezug zur Hauptabstrahlrichtung.

Bei Lautsprechern ist die Definition der Hauptabstrahlrichtung ziemlich eindeutig. Da diese in den allermeisten Fällen nach vorne gerichtet ist, wird diese Richtung, welche auch als 0°-

Achse bezeichnet wird, als Bezugsrichtung deklariert. Bei den meisten Lautsprechern verhält sich die Abstrahlung symmetrisch zur Hauptabstrahlrichtung. Das hat für Simulationsprogramme den Vorteil der Einsparung rechnerischen Aufwandes.

Um die Abstrahlung von Musikinstrumenten korrekt auszuwerten, sollte das Instrument so ausgerichtet sein, dass seine Hauptabstrahlrichtung in die für die Messanordnung definierte Bezugsrichtung zeigt. Dies kann jedoch nicht für alle Instrumente gewährleistet werden, da einerseits die Richtcharakteristik mitunter recht komplex wird, und andererseits die Hauptabstrahlrichtung bei manchen Instrumenten nicht für alle Frequenzen dieselbe ist. Das betrifft vor allem die Holzbläser, aufgrund ihrer dipolähnlichen Abstrahlung, und die Streichinstrumente, aufgrund des komplexen Abstrahlverhaltens von schwingenden Flächen. Bei den Blechblasinstrumenten ist die Situation ziemlich einfach. Hier wird die Richtwirkung durch das Verhältnis der Wellenlänge zu den Abmessungen des Schalltrichters bestimmt. Mit steigender Frequenz wird der Schall hauptsächlich in die Verlängerung der Längsachse des Instrumentes abgegeben. Da bei den Instrumenten mit komplexerer Abstrahlung keine eindeutige Hauptabstrahlrichtung festgelegt werden kann, muss man sich ein anderes Konzept zur Bestimmung der Bezugsrichtung überlegen, als es bei Lautsprechern der Fall ist.

Für den Gegenstand dieser Arbeit ist das Vorhaben der räumlichen Auralisation von quasi nachhallfreien Musikaufnahmen für die Bestimmung der Bezugsrichtung von großer Bedeutung. Um das Abstrahlverhalten einer Quelle in Bezug auf die nachhallfreie Aufnahme korrekt wiederzugeben, muss nämlich die Bezugsrichtung mit der Position des Mikrofons, in der das Instrument aufgenommen wurde, übereinstimmen (Bork 2010). In der Richtung, in der das Mikrofon bei der Aufnahme stand, muss der Pegel für alle Frequenzen unverändert bleiben. In allen anderen Richtungen wird der Pegel der Aufnahme im Verhältnis zu dem Schalldruck an der Position des Mikrofons der Charakteristik des Instrumentes entsprechend abgeschwächt oder verstärkt. Ähnlich wie die Bedingung, dass das Zentrum der Klangerzeugung des Instrumentes bei der Messung genau in der Mitte des Mikrofonarrays stehen muss, sollte das Mikrofon bei der Aufnahme auf das Zentrum der Abstrahlung gerichtet sein, damit die Pegelbewertungen korrekt vorgenommen werden.

Zur Bestimmung der Koordinaten für die Bezugsrichtung muss zunächst bekannt sein, wie das Instrument bei der Messung ausgerichtet war. Weiß man zudem, wo das Mikrofon bei der Aufnahme stand, kann man den entsprechenden Winkel auf der umhüllenden Kugeloberfläche finden. Zur Berechnung der Dämpfungspegel wird der Schalldruckverlauf der Richtcharakteristik ins Verhältnis zu dem Schalldruck der Bezugsrichtung gesetzt und mit minus eins multipliziert:

$$L_{Dämpfung}(\varphi,\theta) = \left[L(\varphi,\theta) - L_{Bezug}(\varphi_{Bezug},\theta_{Bezug}) \right] \cdot (-1) \,.$$

Alle für diese Arbeit relevanten Instrumente, bis auf den Flügel, wurden so gemessen, dass der Blick des Spielers in die positive Richtung der Medianachse zeigt. Eine Tabelle, die mit den Messdaten erstellt wurde, enthält Angaben über Zielpunkte auf der Arrayoberfläche, die durch den Spieler angepeilt wurden. Er hatte dabei die Anweisung, eine für die Abstrahlung markante Achse des Instrumentes auf ein bestimmtes Messmikrofon zu richten. Diese zwei Anhaltspunkte müssen zusammen mit den Fotos, die bei der Messung gemacht wurden, ausreichen, um die Koordinaten einer beliebigen Abstrahlrichtung zu finden.
Informationen über die Position der Mikrofone bei den Aufnahmen wurden recherchiert und stehen in Form von Fotografien, Skizzen und verbalen Aussagen der Instrumentalisten und Toningenieure zur Verfügung. Da die Aufnahmesituation für jeden Stimulus eine andere war, mussten für jede Kombination von Instrument und Musiktitel neue Ermittlungen zur Bezugsrichtung angestellt werden.

3.3.1 Ermittlung der Bezugsrichtung für zwei Musikaufnahmen

„Contredanse“

Für den Titel „Contredanse“ konnten für das Schlagzeug, den Bass und den Flügel von der Band Fotos zur Verfügung gestellt werden, die während der Aufnahmen gemacht wurden.
Für die Stimme gab es kein Foto. Nach Aussagen des Pianisten stand das Mikrofon in der üblichen Position im Abstand von 20 bis 30 cm vor der Sängerin und war senkrecht auf die Mundöffnung gerichtet. Bei der Messung hatte die Sängerin die Vorgabe in Richtung Mikrofon 15 zu singen, welches sich auf der Medianebene und etwas unterhalb der Medianachse befindet (Abbildung 3.30).

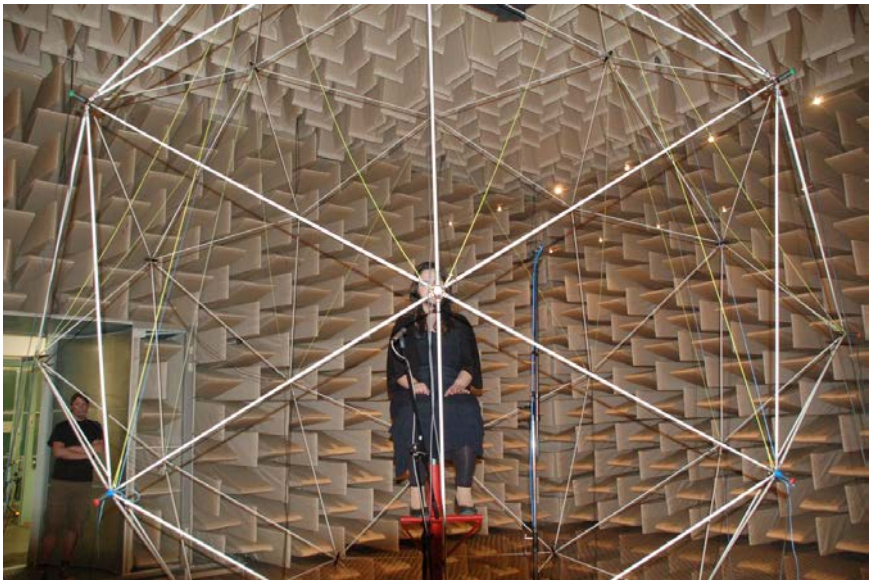

Abbildung 3.30 Frontalansicht der Sängerin bei der Messung. Das Arraymikrofon Nummer 15 befindet sich in der Mitte des Bildes und zeigt in Richtung Mundöffnung.

Diese Richtung dürfte ungefähr mit der beschriebenen Position des Mikrofons bei „Contredanse“ übereinstimmen. Die Kugelkoordinaten von Mikrofon 15 betragen Phi=0° und Theta=100,8°. Um zu diesen Koordinaten den am nächsten liegenden Punkt auf dem Koordinatennetz von EASE zu finden, wurde eine Matlab-Funktion erstellt. Ähnlich wie in Abschnitt 3.2.2.1 wird zunächst mit der Funktion *meshgrid* ein Datennetz erzeugt und anschließend so rotiert, dass es dem Koordinatengitter von EASE entspricht. Die Koordinaten des Rasters und die Kugelkoordinaten des Messmikrofons werden in das kartesische System transformiert, um dann mit der Funktion *dsearchn* den am nächsten liegenden Punkt auf dem Koordinatennetz zu finden. Die Rotations- und Bogenwinkel der Punkte auf dem Gitter können mithilfe der gefundenen Indizes aus dem ursprünglich erstellten Datennetz ermittelt werden. Dieser Algorithmus wird für alle weiteren Fälle, in denen zu einem beliebigen Punkt auf der Kugeloberfläche die adäquaten EASE-Koordinaten gefunden werden sollen, angewendet.
Mit den gefundenen Indizes ergibt sich in EASE für die Winkel der Bezugsrichtung Phi = 270° und Theta = 10°. Der Pegel dieser Koordinaten wird von allen Werten auf dem Datennetz subtrahiert. Werden die Differenzpegel noch mit minus eins multipliziert, erhält man den für EASE gesuchten Dämpfungspegel.

Das Foto des Flügels zeigt zwei Mikrofone, die ziemlich dicht am Instrument aufgestellt waren (Abbildung 3.31). Es gibt zusätzlich eine Aufnahme von einem Mikrofonpaar, das weiter im Raum stand. Dieses kann jedoch nicht verwendet werden, da dieses Signal zuviel Raumanteil enthält.

Abbildung 3.31 Foto des Flügels während der Aufnahme von „Contredanse“. Quelle: Thierry Gautier; Pianist des Pascale Gautier Quartet

Das eine der beiden nah aufgestellten Mikrofone zeigt auf die hohen Saiten, das andere auf die tiefen Saiten. Die Mikrofonierung war für den Zweck einer stereophonen Studioproduktion ausgewählt und diente dazu, den Frequenzanteil der hohen und der tiefen Saiten auf dem Stereopanorama zusammenzumischen. Ein Flügel ist eine Quelle mit einem relativ stark ausgedehnten Zentrum der Schallabstrahlung. Daher werden beide Mikrofonsignal benötigt, um alle klanglichen Aspekte dieses Instrumentes abzubilden. Ein punktuelles Zentrum der Klangerzeugung, welches zur Ermittlung der Bezugsrichtung herangezogen werden könnte, gibt es nicht. Man kann sich die beiden Bereiche der hohen und tiefen Saiten als zwei getrennte Schallquellen ein und desselben Instrumentes vorstellen, die später in der Simulation quasi räumlich zusammengemischt werden. Als Bezugsrichtung wird für beide Positionen die Richtung frontaler Abstrahlung gewählt, welche aus Abbildung 3.29 zu entnehmen ist. Die Beträge der Winkel in den Lautsprecherkoordinaten betragen also jeweils null Grad.

Wie in Abbildung 3.32 zu sehen ist, wurde der Bass auch mit zwei Mikrofonen aufgenommen. Das Signal des schräg von oben kommenden Mikros (im Folgenden: Position 1) enthält sehr viele Geräuschanteile und gibt wenig von den tiefen Frequenzen des Instruments wieder. Dies ist unter anderem darauf zurückzuführen, dass hier ein KM 85 von Neumann verwendet wurde, aus dessen Betriebsanleitung hervorgeht, dass der Frequenzgang ab 300 Hz um circa 4 dB pro Oktave zu den tiefen Frequenzen abfällt. Das schräg von unten auf den Bass gerichtete Mikrofon (im Folgenden: Position 2) nahm vor allem den tieffrequenten Anteil des Instrumentes auf. Es handelt sich hier um ein RE 20 von dem Hersteller Electro-Voice. Das Konzept der zweifachen Mikrofonaufstellung ist auf die

klangästhetischen Zielstellungen einer Studioproduktion zurückzuführen. Beide Signale werden in der Mischung kombiniert und geben dem Tonmeister die Möglichkeit, das Verhältnis der verschiedenen spektralen Anteile des Basses zu beeinflussen und an die Gesamtmischung des Musikstücks anzupassen.

Abbildung 3.32 Fotografie des Bassspielers während der Aufnahmen von „Contredanse“. Quelle: Thierry Gautier.

Da keines der beiden Mikrofone eine neutrale Aufnahmeposition einnahm, werden Richtcharakteristiken für beide Abstrahlwinkel normalisiert. Um dem Mischtonmeister noch eine dritte Option zur Verfügung zu stellen, wird auch ein Modell für die 0°-Achse erstellt. Das entspricht der Abstrahlung in die Medianachse des Instrumentes. Die Idee ist hierbei, beide Signale für diese „neutralere“ Bezugsrichtung zusammenzumischen. Der Bass wurde zudem noch mit einem Tonabnehmer aufgenommen, so dass dieses Signal für die neutrale Bezugsrichtung hinzugemischt werden kann.

Die Mikrofonpositionen werden schätzungsweise in die fotografische Darstellung des gemessenen Basses eingetragen und die Winkel mit einem Grafikprogramm ausgemessen. Aus Abbildung 3.32 geht hervor, dass beide Mikrofone von vorne auf das Instrument zeigen. Für den Azimutwinkel ergeben sich also 0° (vgl. Abbildung 3.5).

Abbildung 3.33 Grafisches Ausmessen des Polarwinkels der zwei Mikrofonpositionen für den Bass.

Die Position des Basses wurde in Abschnitt 3.1.1 korrigiert. Man kann sich daher das Instrument in der seitlichen Darstellung als Zentrum des Kugelarrays vorstellen (Abbildung 3.33). Von der Mitte des Korpus ausgehend, werden die Längsachsen der beiden Mikrofone nach Augenmaß eingezeichnet. Für Mikrofon 1 ergibt sich ein Polarwinkel von 20°. Für Mikrofon 2 beträgt der Polarwinkel 105°. Das Ergebnis für die EASE-Koordinaten ist in Tabelle 4 zusammengefasst.

	Azimut	Polarwinkel	Rotationswinkel	Bogenwinkel
Mikrofon 1	0°	20°	90°	70°
Mikrofon 2	0°	105°	270°	15°

Tabelle 4 Die Koordinaten der beiden Mikrofone für den Bass. Mikrofon 1 war von schräg oben und Mikrofon 2 von schräg unten auf den Bass gerichtet.

Ein Schlagzeug war nicht Bestandteil der Messungen, die vom Fachgebiet AK durchgeführt wurden. Auch in den Daten der PTB und in den Daten der TKK findet sich keine Richtcharakteristik für dieses Instrument. Das Schlagzeug wird daher durch zwei Kugelschallquellen simuliert. Der Schall wird für alle Frequenzen in alle Richtungen gleichmäßig abgestrahlt.

„River Man“

Der Sänger von „River Man“ wurde wie bei „Contredanse“ senkrecht vor der Mundöffnung aufgenommen. Es kann daher für die Stimme die gleiche Richtcharakteristik verwendet werden.

Der Klang des Basses kommt von einer Samplelibrary. Wo das Mikrofon bei der Aufnahme der Samples stand, ist nicht bekannt. Aus dem Höreindruck ergibt sich ein relativ ausgewogener Klang. Es liegt daher nahe, das Quellmodell von „Contredanse“ zu benutzen, welches auf die 0°-Achse normiert wurde.

Die Streicher für „River Man“ wurden mit einem Quartett aus zwei Geigen, einer Bratsche und einem Cello aufgenommen. Die Instrumentalisten saßen in einem für ein Streicherquartett typischen Halbkreis. Leider gibt es von der Aufstellung der Mikrofone nur ein Foto ohne die Instrumentalisten (vgl. Abbildung 2.1). Dafür stehen aber im Gegensatz zu „Contredanse“ ausführliche Aussagen des Toningenieurs zur Verfügung.
Nach der Beschreibung des Toningenieurs war das Mikrofon für das Cello um circa 20° von der Medianebene aus Sicht des Spielers nach links verschoben und zeigte auf die Mitte des Instruments.

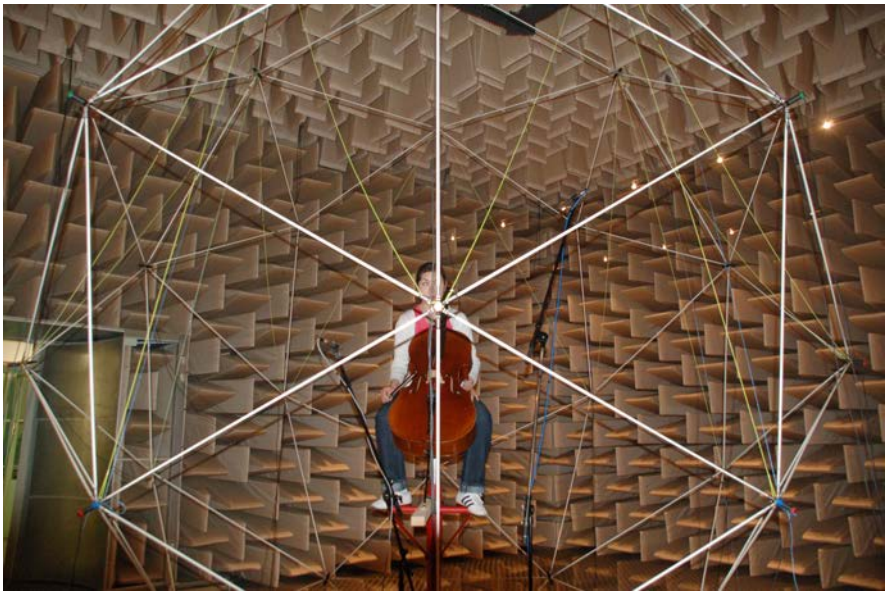

Abbildung 3.34 Frontalansicht des Cellos bei der Messung.

Wie man in Abbildung 3.34 sehen kann, war das Instrument bei der Messung so ausgerichtet, dass die Medianebene des Cellos und die Medianebene des Kugelarrays ungefähr in einer Fläche lagen. Der Versatz von 20° aus der Medianachse nach links entspricht demnach einem Azimutwinkel des Arrays von ebenfalls 20°. Aus der seitlichen Abbildung des gemessenen Cellos lässt sich abschätzen, dass die Höhe des Mikrofons ungefähr einem Polarwinkel von 90° entspricht. Die entsprechenden Werte der EASE-Koordinaten sind in Tabelle 5 eingetragen.

	Azimut	Polarwinkel	Rotationswinkel	Bogenwinkel
Mikrofon Cello	20°	90°	0°	20°

Tabelle 5 Die Koordinaten des Mikrofons bei der Aufnahme des Cellos.

Die Bratsche wurde so aufgenommen, dass das Mikrofon von oben nahezu senkrecht nach unten auf den Spieler herabschaute. Durch die etwas schräge Haltung der Bratsche war das Mikrofon um circa 20 bis 30 Grad aus der Resonanzdeckennormale verschoben.

Für die Ausrichtung der Bratsche bei der Messung wird angegeben, dass die Resonanzbodennormale auf einen Punkt etwas rechts neben dem Nordpol zielte. Das entspricht der natürlichen Haltung des Instruments, wenn der Spieler sich so ausrichtet, dass er in die Medianachse des Arrays schaut (vgl. Abbildung 3.35).

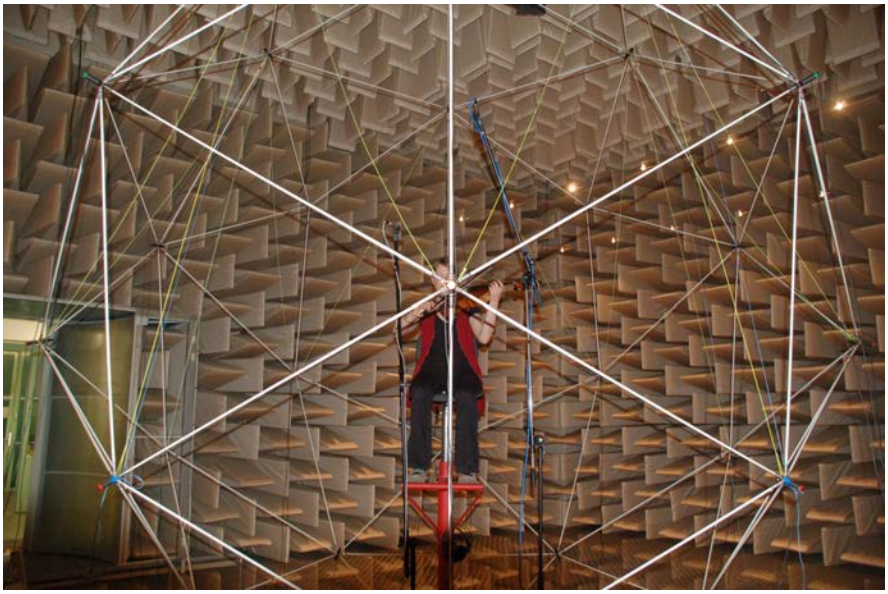

Abbildung 3.35 Frontalansicht der Bratsche bei der Messung

Der Nordpol entspricht ungefähr dem gleichen Punkt wie das senkrecht nach unten ausgerichtete Mikrofon bei der Aufnahme. Die Werte der entsprechenden Winkelkoordinaten sind in Tabelle 6 eingetragen.

	Azimut	Polarwinkel	Rotationswinkel	Bogenwinkel
Mikrofon Bratsche	0°	0°	90°	90°

Tabelle 6 Die Koordinaten des Mikrofons für die Aufnahme der Bratsche.

Ähnlich wie bei der Bratsche stand das Mikrofon bei der Violine nahezu senkrecht über dem Spieler. Es war durch die schräge Haltung des Instruments um circa 30° aus der Resonanzdeckennormale verschoben und war auf die Mitte der Quelle gerichtet.

Bei der Messung war die Resonanzbodennormale der Violine auf Mikrofon 5 gerichtet, für das der Azimut 324° und der Polarwinkel 37,4° beträ gt. Die gesuchte Mikrofonposition entspricht einer Drehung dieses Punktes mit α=-30° um die Längsachse des Instruments. Diese Drehung kann in kartesischen Koordinaten durchgeführt werden. Verläuft die Achse, um die gedreht werden soll, durch den Koordinatenursprung, kann die Drehung eines Punktes $\vec{p}$ mit dem Winkel α um den Einheitsvektor der Drehachse $\vec{n}$ mit der Gleichung

$$\vec{p}' = \vec{p}\cos\alpha + \vec{n}\left(\vec{n}\cdot\vec{p}\right)\left(1-\cos\alpha\right) + \left(\vec{p}\times\vec{n}\right)\sin\alpha \qquad (1.6)$$

berechnet werden (Goldstein 1987).

Dafür müssen die Koordinaten der Längsachse bekannt sein. Nach Abbildung 3.36 kann man davon ausgehen, dass die Längsachse ungefähr in der horizontalen Ebene liegt. Die Drehachse geht also durch den Koordinatenursprung und für den gesuchten Einheitsvektor beträgt der Wert für z gleich null.

Abbildung 3.36 Seitliche Ansicht der Violine. Der Koordinatenursprung wurde etwas nach unten verschoben, um darzustellen, dass die Längsachse des Instruments annäherungsweise auf den Koordinatenursprung trifft.

Der Vektor der Resonanzbodennormale $\vec{r}$ steht senkrecht auf der Längsachse. Das Skalarprodukt beider Vektoren ist also gleich null. Die Werte der Resonanzbodennormalen ergeben sich aus den kartesischen Koordinaten für Mikrofon 5. Daraus ergibt sich:

$$\vec{n} \cdot \vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -0,7 \\ 1,7 \end{pmatrix} = x - 0,7y + 1,7z = 0$$

und mit z=0 ist

$$x = 0,7y\,. \tag{1.7}$$

Aus der Gleichung für den Betrag des Einheitsvektors der Längsachse folgt

$$\sqrt{x^2 + y^2 + z^2} = 1$$

und mit z=0 ergibt sich

$$x^2 + y^2 = 1\,. \tag{1.8}$$

Durch gegenseitiges Einsetzen von (1.7) und (1.8) erhält man die gesuchten Werte für x und y. Die Koordinaten des Einheitsvektors der Längsachse betragen:

$$\vec{n} = \begin{pmatrix} 0,5734 \\ 0,8192 \\ 0 \end{pmatrix}.$$

Mit (1.6) können jetzt die kartesischen Koordinaten des gedrehten Punktes $\vec{p}'$ berechnet werden. Die entsprechenden Winkelwerte sind in Tabelle 7 eingetragen.

	Azimut	Polarwinkel	Rotationswinkel	Bogenwinkel
Mikrofon Violine	320,2°	7,4°	95°	85°

Tabelle 7 Die Koordinaten des Mikrofons für die Aufnahme der Violine

Die Gitarre wurde im Abstand von 50 bis 60 cm mit zwei Mikrofonen in einer AB-Aufstellung aufgenommen. Die beiden Mikrofone zeigten mit einem Neigungswinkel von circa 45° schräg von oben auf das Zentrum des Korpus. Das Schallloch befand sich dabei ungefähr in ihrer Mitte. Eine getrennte Betrachtung der beiden Mikrofonpositionen wird nicht vorgenommen, da ihr Abstand voneinander recht klein und der spektrale Unterschied zwischen den beiden Signalen relativ gering ist. Welches der beiden Signale verwendet wird, kann später bei der Mischung entschieden werden. Gesucht wird also ein Mittel aus beiden Mikrofonpositionen. Betrachtet man Spieler und Instrument als eine Einheit und definiert die Medianachse als Senkrechte auf der Resonanzdecke, entspricht das einem Punkt, der sich 45° oberhalb der Medianachse befindet.

Aus der Frontalansicht der Gitarre geht hervor, dass das Instrument bei der Messung so gehalten wurde, dass das Schallloch senkrecht in die Medianachse des Arrays zeigte. Der Azimut für die gesuchte Mikrofonposition beträgt also 0°.

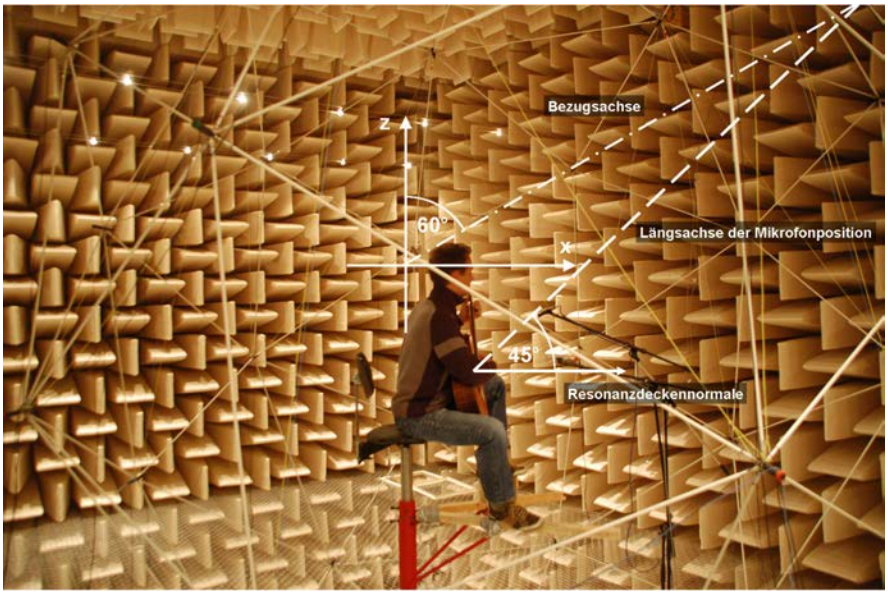

Abbildung 3.37 Seitlicher Blick auf die Gitarre bei der Messung.

In der seitlichen Betrachtung (Abbildung 3.37) fällt auf, dass die Gitarre aus dem Zentrum nach unten verschoben war. Die Position der Quelle wurde bei der Pegelauswertung nicht korrigiert. Der Winkel, um den das Mikrofon aus der Resonanzdeckennormale verschoben ist, stimmt daher nicht mit dem Polarwinkel des Arrays überein. Verlängert man die Längsachse des Aufnahmemikrofons bis zum Erreichen der Grenze des Kugelarrays, landet

man bei dem gesuchten Punkt auf der Hüllfläche. Verbindet man diesen Punkt mit dem Zentrum des Arrays, erhält man den Polarwinkel der Bezugsachse. Dass das „virtuelle“ Mikrofon an diesem Punkt nicht auf das Zentrum der Gitarre zeigt, ist zu vernachlässigen, da die Messmikrofone nicht gerichtet waren[1]. Die Winkelwerte sind in Tabelle 8 zusammengefasst.

	Azimut	Polarwinkel	Rotationswinkel	Bogenwinkel
Mikrofon Gitarre	0°	60°	90°	30°

Tabelle 8 Die Koordinaten der Mikrofonposition für die Aufnahme der Gitarre

3.4 Bestimmung der Kennwerte und Importieren der Daten in EASE

Um die Daten der Richtcharakteristiken in EASE zu importieren, werden diese mithilfe der Matlab-Funktion *fprintf* in ein Textfile geschrieben und dann über das EASE-Modul *Speaker Base* importiert. Das Textfile muss dem von EASE festgelegten Standard für Lautsprecherdaten im ASCII-Format entsprechen. Jede ASCII-Datei beginnt mit drei Headerzeilen, in denen drei Variablen festgelegt werden, die den Datentyp, die Formatversion und die Längeneinheit bestimmen:

```
;
"FileType","Speaker Types"
"Format",4.0
"LengthUnit","meters"
;
```

Für jedes Frequenzband gibt es einen „frequency header“, in dem Informationen zur Herkunft des Lautsprechers, die Mittenfrequenz und die elektroakustischen Kenndaten bestimmt werden:

```
;
"SpeakerName","Measured Loudspeaker"
“Manufacturer”,”Loudspeakers, Inc.”
"Frequency",100 ,"Hz"
"Sensitivity",96,"dB"
"Impedance",4,"ohms"
“Q“,1.6
“Efficiency”,3.1,“%“
"MaxPower",1200,"W"
"DataGood",1
```

Es folgt eine nähere Erläuterung der elektroakustischen Kenndaten.

[1] Die Messmikrofone des Arrays sind vom Typ Sennheiser KE4-211-2.

3.4.1 Lautsprecherkenndaten

3.4.1.1 Kennschalldruck

Die erste Headerzeile ist die Bestimmung der Variable "Sensitivity". Sie bezeichnet den Kennschalldruck, also den Schalldruck des betreffenden Frequenzbandes, den der Lautsprecher in einem Meter Entfernung bei einer Anregung mit einem Watt elektrischer Leistung erreicht. Als Anregungssignal ist ein rosa Rauschen definiert. Die Leistung der Anregung bezieht sich auf das entsprechende Terzband und nicht auf die Gesamtleistung des Rauschens. Der Kennschalldruck bezieht sich zudem auf alle Richtungen, in denen die Richtcharakteristik den Wert null hat. In allen anderen Richtungen ist der Kennschalldruck dem Dämpfungswert entsprechend höher oder niedriger. Bei der Simulation von Musikinstrumenten für die Auralisation von Mikrofonaufnahmen ist nicht von Interesse wieviel Leistung das Quellmodell absolut abstrahlt, sondern nur wie sich die abgestrahlte Leistung in Abhängigkeit der Frequenz auf die verschiedenen Richtungen verteilt. Das Verhältnis der Abstrahlungsleistung zwischen den Quellen wird später in der Mischung durch die Lautstärkeeinstellung der einzelnen Instrumente bestimmt. In der Bezugsrichtung, wo das Mikrofon bei der Aufnahme stand, soll das Quellmodell einen linearen Frequenzgang aufweisen. Daher wird der Kennschalldruck für alle Frequenzbänder auf einen konstanten Wert gesetzt. Die Wahl fällt hier willkürlich auf den Wert 100 dB.
Wenn man in EASE mit dem Mapping-Modul akustische Messungen zur frequenzabhängigen Schallpegelverteilung im Raum durchführen will, bei denen die Quelle mit einem Rauschsignal angeregt wird, wäre es sinnvoll über den Kennschalldruck der Terzbänder den verallgemeinerten Frequenzgang der entsprechenden Quelle wiederzugeben. Für den Fall der Auralisation von Musikinstrumenten dagegen ist der Frequenzgang ja quasi in den Musikaufnahmen schon enthalten. Nur die Pegelunterschiede in von der Bezugsrichtung abweichenden Richtungen sollen von dem Quellmodell wiedergegeben werden.

3.4.1.2 Maximale Leistungsaufnahme

Die Größe "MaxPower" gibt die maximale Leistungsaufnahme des entsprechenden Frequenzbandes bei Anregung mit rosa Rauschen an. Zusammen mit dem Kennschalldruck wird somit zugleich der maximale Ausgangspegel eines Frequenzbandes bestimmt. Nach neuen Pegelkonventionen in EASE 4.3 darf zudem der Wert der maximalen Leistungsaufnahme eines beliebigen Passbandes nicht von der anregenden Leistungssumme der enthaltenen Terzbänder überschritten werden.

Um mindestens den Kennschalldruck erreichen zu können, muss der Wert für die maximale Leistungsaufnahme 21 Watt betragen. Er kann aber auch ohne weitere Auswirkungen auf die Simulation höher liegen und wird hier für alle Terzbänder auf 200 Watt festgelegt.
Für den Zweck der Messung von akustischen Parametern mit Rauschsignal angeregten Instrumentenmodellen wäre es denkbar über den Wert der maximalen Leistungsaufnahme die Quellen auf die tatsächlich erreichbare Schallleistung der gemessenen Instrumente zu beschränken. Maximale Schallleistungspegel von Einzeltönen der gemessenen Instrumente wurden bereits für die Streichinstrumente sowie die Holz- und Blechblasinstrumente ausgewertet (Detzner et al. 2010), (Krämer et al. 2010). Für die Bestimmung der maximalen Leistungspegel in den Terzbändern wären allerdings neue Berechnungen nötig. Die Schallpegel der Messpunkte müssten zunächst gefiltert und über die Einzeltöne der Fortissimodarbietungen gemittelt werden, um dann einen verallgemeinerten Leistungspegel in Terzbändern zu berechnen.

3.4.1.3 Impedanz

Die Bestimmung des Wertes „Impedance“ hat eher informativen Charakter. Die anregende elektrische Energie wird in EASE über einen Leistungswert bestimmt, der für jede Quelle und jedes Terzband eingegeben werden kann. Da die Impedanz schon in der elektrischen Leistung enthalten ist, spielt sie hier keine Rolle. Würde man die Quellanregung über eine Spannung bestimmen, wäre die Impedanz von großer Bedeutung, da die aufgenommene elektrische Leistung von der Impedanz abhängig ist. Für den Fall der Modellierung von Instrumenten wird die Impedanz hier willkürlich für alle Frequenzen auf acht Ohm festgelegt.

3.4.2 Richtfaktor

Der Parameter „Q“ bestimmt die Richtwirkung der Quelle auf der 0°-Achse. Die Bestimmung des Richtfaktors kann in der ASCII-Datei auch ausgelassen werden und stattdessen im Modul „Speaker Base“ über das Menü „Speaker Data“ von EASE berechnet werden. Dabei fällt auf, dass sich die von EASE ermittelten Werte von den eigenen Berechnungen unterscheiden. Um den Grund dieser Unstimmigkeiten zu erläutern, soll zunächst eine Herleitung des statistischen Richtfaktors folgen:
Nach Meyer (1995) bestimmt der statistische Richtfaktor das Verhältnis des Schalldrucks in einer beliebigen Richtung p(φ,θ) zu dem Schalldruck, den eine Kugelschallquelle mit einer gleichen Gesamtleistung in derselben Entfernung hervorrufen würde:

$$\Gamma_{st} = \frac{p(\varphi,\theta)}{p_{Kugel}} .$$

Wenn die Leistung der Quelle bekannt ist, kann man den Schalldruck einer Kugelquelle berechnen aus

$$p_{Kugel} = \sqrt{\frac{P\rho_0 c}{4\pi r^2}} \tag{1.9}$$

(Möser 2009, S. 49) (P: Leistung; ρ_0: Ruhedichte der Luft; c: Schallgeschwindigkeit; r: Abstand von der Quelle). Die Leistung der gerichteten Quelle ergibt sich aus dem Integral der Intensität über die umhüllende Fläche der Quelle:

$$P = \int I\, dS \tag{1.10}$$

.

(Möser 2009, S. 47). Wenn man bei dem im Abstand r gemessenen Schalldruck von einer ebenen fortschreitenden Welle ausgeht, ist die Intensität

$$I = \frac{p^2}{\rho_0 c}. \tag{1.11}$$

.

Für das Flächenelement einer Kugeloberfläche gilt:

$$dS = r^2 \sin\theta\ d\varphi\ d\theta$$

($\theta : Bogenwinkel,\ \varphi : Rotationswinkel$) (Papula 2008, S. 126). Die Schallpegelwerte der Richtcharakteristik stehen für eine Fläche, deren Größe sich aus 5°-Bogenwinkel mal 5°-Rotationswinkel ergibt. Man geht also von den Koordinaten eines Punktes einmal 2,5° in die eine und 2,5° in die andere Richtung. Daraus folgt für das Oberflächenintegral einer Teilfläche:

$$S = \int_{\varphi-a}^{\varphi+a} \int_{\theta-a}^{\theta+a} r^2 \sin\theta\, d\theta\, d\varphi$$

($a = 2{,}5° \frac{\pi}{180°} = \frac{1}{72}\pi$). Durch die Lösung dieses Integrals erhält man die Größe einer Teilfläche:

$$S_i = r^2 2a\left(\cos[\theta_i - a] - \cos[\theta_i + a]\right)$$

Für die Messung von diskreten Schalldruckwerten geht das Integral aus (1.10) in eine Summe der Intensitäten auf den N Teilabschnitten der Kugelfläche über:

$$P = \sum_{i=1}^{N} \frac{p_i^2}{\rho_0 c} r^2 2a(\cos[\theta_i - a] - \cos[\theta_i + a]) \tag{1.12}$$

Durch Einsetzen von (1.12) in (1.9) ergibt sich für den Schalldruck einer leistungsäquivalenten Kugelquelle:

$$p_{Kugel} = \sqrt{\frac{\sum_{i=1}^{N} p_i^2 \, \frac{1}{36}\pi(\cos[\theta_i - \frac{1}{72}\pi] - \cos[\theta_i + \frac{1}{72}\pi])}{4\pi}} \, .$$

Das entspricht einer Mittelung des Schallpegels über die umhüllende Kugelfläche.

In der Bedienungsanleitung von EASE wird der Parameter „Q“ bedeutungsgleich als Verhältnis des Schalldrucks in der 0°-Achse zu dem über die gesamte Kugelfläche gemittelten Schalldruck bezeichnet (Ahnert 2009, S. 540).

Zur Pegelbetrachtung ist es nach Meyer (Meyer 1995, S. 109) sinnvoll den statistischen Richtfaktor in ein Richtwirkungsmaß umzurechnen:

$$L_{Richtwirkungsmaß} = 20\log\left(\frac{p(\varphi,\theta)}{p_{Kugel}}\right) \tag{1.13}$$

Das Richtwirkungsmaß gibt an, um wie viel dB sich der Pegel in der betroffenen Richtung vom Pegel einer kugelförmigen Abstrahlung unterscheidet.

In EASE wird das Richtwirkungsmaß („directivity index“) laut Bedienungsanleitung mit $10*\log(R)$ bestimmt. Es kommt die Vermutung auf, dass die Angaben für den Richtfaktor nicht ganz korrekt sind und eigentlich der Quotient der Schalldruckquadrate gemeint ist. Dann wäre nämlich das Richtwirkungsmaß

$$L_{directivity\ index} = 10\log(\Gamma_{st}^2) = 20\log(\Gamma_{st})$$

genau wie bei Meyer bestimmt. Ein Vergleich des aus den Messdaten berechneten Richtungsfaktors mit den Werten von EASE zeigt, dass der quadrierte Richtungsfaktor näher an den von EASE berechneten Werten liegt. Allerdings treten auch dann noch Differenzen auf, die von Terzband zu Terzband unterschiedlich groß sind. Erstaunlicherweise ändert sich der Richtfaktor in EASE, wenn der Pegel der Richtcharakteristik auf eine andere Bezugsrichtung normalisiert wird. Dies ist nach der Definition des Richtfaktors eigentlich nicht möglich, da sich die Pegelverhältnisse durch Änderung der Bezugsrichtung nicht ändern. Berechnet man den quadrierten Richtfaktor für eine Richtcharakteristik mit Bezug auf die 0°-Achse, stimmen die Ergebnisse mit den vo n EASE ermittelten Werten überein. Der Grund für diese Anomalität liegt darin, dass EASE davon ausgeht, dass die Werte der Richtcharakteristik stets auf die 0°-Achse normalis iert sind. Anstatt also den tatsächlichen Pegelwert in dieser Richtung für die Berechnung einzubeziehen, wird immer mit einem Pegel von 0 dB gerechnet.

Für den Fall auf Mikrofonposition bezogener Richtcharakteristiken von Musikinstrumenten berechnet EASE demnach keine korrekten Werte für den Richtfaktor. Dieser wird daher im Vorhinein berechnet und über die ASCII-Dateien in EASE importiert.

Abschließend sei noch erwähnt, dass die Bedeutung des statistischen Richtfaktors in der 0°-Achse für den Fall der Musikinstrumente eine andere ist als bei Lautsprechern. Bei

Lautsprechern liegt die Hauptabstrahlrichtung immer in der 0°-Achse. Daher ergibt der statistische Richtfaktor in dieser Richtung einen aussagekräftigen Wert über die Richtwirkung der Quelle. Im Allgemeinen wird der Richtfaktor auf der Hauptachse eines Lautsprechers als Bündelungsgrad bezeichnet. Bei der Simulation der Richtcharakteristiken befindet sich die Hauptabstrahlrichtung nur bei einigen Instrumenten auf der 0°-Achse. Daher kann dieser Wert keine Aussage über die generelle Richtwirkung des Instrumentes treffen. Allerdings enthält dieser Wert eine Information darüber, wie stark der Schall in die Blickrichtung des Spielers gebündelt wird. Da diese Richtung oft die dem Publikum zugewandte Seite des Spielers ist, erhält man somit eine Aussage darüber, wie stark das Instrument seine Richtwirkung auf das Publikum entfaltet.

3.4.3 Wirkungsgrad

Die Größe „Efficiency“ bestimmt den über ein Terzband gemittelten Wirkungsgrad des Lautsprechers. Das ist das Verhältnis von der abgestrahlten Schallleistung zu der zugeführten elektrischen Leistung in Prozent ausgedrückt. Dieser Wert kann aus dem Kennschalldruck und den Dämpfungswerten der Richtcharakteristik bestimmt werden. Der Kennschalldruck einer leistungsäquivalenten Kugelquelle berechnet sich aus dem über die Kugelfläche gemittelten Dämpfungspegel, wie es für den Richtfaktor hergeleitet wurde, und dem Kennschalldruck der gerichteten Quelle:

$$L_{Kenn_Kugel} = L_{Kenn} - \overline{L}_{Dämpfung}$$

Aus der Formel für die Leistung einer ungerichteten Schallquelle

$$P = \frac{p_{eff}^2}{\rho_0 c} 4\pi r^2$$

(Möser 2009, S. 76) (p_{eff} : Effektivwert des Schalldrucks, ρ_0 : Ruhedichte der Luft, c: Schallgeschwindigkeit, r: Abstand von der Quelle) erhält man (mit r=1m und P_{zu}=1W, da sich der Kennschalldruck auf 1m Entfernung bei 1W anregender Leistung bezieht) den Wirkungsgrad:

$$\frac{P_{ab}}{P_{zu}} = \frac{\frac{10^{L_{Kenn_Kugel}/10}\, p_0^2}{\rho_0 c} 4\pi}{1W} \tag{1.14}$$

Da sich mit der Änderung der Bezugsrichtung die Skalierung der Dämpfungspegel und damit auch deren Mittelungswert ändert, erhält man auch für jede Bezugsrichtung einen anderen Wirkungsgrad. Der Wert für den Wirkungsgrad ist nur dann wahr, wenn sich der Kennschalldruck auf den tatsächlich gemessenen Wert in der Bezugsrichtung bezieht. Da das Quellmodell hier nicht die Übertragungseigenschaften eines Lautsprechers simulieren,

sondern die richtungsabhängige Entzerrung eines Mikrofonsignals bewirken soll, muss der Kennschalldruck über alle Frequenzen konstant sein. Der Absolutbetrag wäre ohnehin schwer zu ermitteln, da die vom Spieler in das Instrument eingeführte Leistung nicht bekannt ist. Der Wirkungsgrad ist aber in anderer Hinsicht eine aussagekräftige Größe, da er als Funktion der Frequenz wiedergibt inwiefern die Leistung des eingespeisten Mikrofonsignals durch die simulierte Quelle beeinflusst wird (siehe Abschnitt 3.5.).
Der Wirkungsgrad hat für die Lautsprecherdaten in EASE einen rein informativen Charakter und wirkt sich nicht auf Schallfeldberechnungen aus. Seine Bestimmung wird daher aus den exportierten ASCII-Dateien weggelassen.

3.4.4 Messwerte

Mit der Variable "DataGood" wird bestimmt, ob die Quelle in dem entsprechenden Frequenzband für die Schallfeldberechnungen einbezogen wird. Ein Hochtonlautsprecher, dessen nominaler Frequenzgang erst bei 1 kHz beginnt, würde beispielsweise für alle Terzbänder unterhalb 1 kHz nicht mit in die Berechnung einbezogen werden. Das heißt, der Wert für "DataGood" wäre bei diesen Frequenzbändern gleich null. Da für alle Instrumentenmodelle die Abstrahlung auf der ganzen Bandbreite von 100 Hz bis 10 kHz simuliert werden soll, wird für alle Frequenzbänder der Wert auf eins gesetzt.

Zusammenfassend kann man sagen, dass sich die Frequenzheader für alle exportierten Instrumente nur in der Lautsprecherbezeichnung und dem statistischen Richtfaktor für die 0°-Achse („Q“) unterscheiden und wie folgt aussehen:

```
;
"SpeakerName","cello"
"Frequency",100  ,"Hz"
"Sensitivity",100,"dB"
"Impedance",8,"ohms"
“Q“,1.6
"MaxPower",200,"W"
"DataGood",1
;
```

3.4.4.1 Pegelwerte

Auf den Frequenzheader folgt die Angabe der Pegelwerte für die Richtcharakteristik. Jede neue Zeile entspricht dem Fortschreiten des Rotationswinkels um 5°. Die 37 Werte eines Halbkreisbogens werden durch Komma getrennt in eine Zeile geschrieben:

```
" 0 °", -0.63, -0.55, -0.37, …
" 5 °", -0.63, -0.50, -0.35, …
…
…
" 355°", -0.63, -0.55, -0.37, …
```

3.4.4.2 Dämpfungskurven

In den drei darauf folgenden Zeilen können die Pegelwerte für die als „Isobaren“ bezeichneten -3dB-, -6dB-, und -9dB-Kurven angegeben werden. Wie bei Lautsprechern üblich wird hier von einer Hauptabstrahlung in die Richtung der Medianachse ausgegangen. Für jede Kurve gibt es eine Zeile, die 72 Werte für Winkelangaben des Halbkreisbogens enthält. Jeder neue Wert entspricht dabei einem Fortschreiten des Rotationswinkels um 5°. Es wird dabei eine nach vorne gerichtete Keule gebildet, innerhalb derer, die Pegelwerte den genannten Wert der Kurve nicht unterschreiten. Dies macht jedoch nur für Quellen mit genau einer Hauptabstrahlrichtung Sinn. Da Musikinstrumente mitunter mehrere bevorzugte Abstrahlrichtungen haben, wird hier auf die Angabe der -3dB-, -6dB-, und -9dB-Kurven verzichtet. Für Instrumente mit nur einer in die vordere Hemisphäre gerichteten Hauptabstrahlrichtung können diese Kurven im Nachhinein mit dem Modul „Speaker Base“ von EASE berechnet werden.

3.4.4.3 Phasenwerte

Es ist weiterhin möglich Daten für die Phasenlage der Richtcharakteristik in der ASCII-Datei anzugeben. Da hier keine Daten zur Phasenlage ermittelt wurden, werden diese Angaben weggelassen.

3.4.5 Bearbeitung der Lautsprecherdaten in EASE

Nach dem Import der ASCII-Dateien können mithilfe des Moduls „Speaker Base“, insofern das für das entsprechende Instrument Sinn macht, über das Menü „Speaker Data“ die „isobaren“ Dämpfungskurven berechnet werden. Weiterhin kann man über das Menü „Speaker Model“ einstellen, ob die Richtcharakteristik symmetrisch ist und ob die Daten zur Phasenlage für Berechnungen einbezogen werden. Als symmetrisch werden Richtcharakteristiken angesehen, deren Daten links und rechts der Medianebene identisch sind. Dies trifft für keine der gemessenen Quellen zu, auch wenn dies bei einigen Instrumenten theoretisch zu erwarten wäre. Die Option wird abgewählt, da die Pegelwerte

ansonsten an der Medianebene gespiegelt würden. Die Option „use phase data“ wird ebenfalls nicht angewählt. Für alle Richtcharakteristiken wird von einer gleichphasigen Abstrahlung ausgegangen.

3.5 Einfluss der Bezugsrichtung auf den Frequenzgang

Die Wahl der Bezugsrichtung ist mit großer Sorgfalt zu treffen, da eine andere Bezugsrichtung auch immer eine Veränderung des Frequenzgangs des simulierten Instrumentes mit sich bringt. Liegt der Pegel in der Bezugsrichtung für ein Frequenzband unter den über alle Richtungen gemittelten Schalldruckdämpfungs- und verstärkungswerten, wird der mittlere Pegel insgesamt angehoben und umgekehrt. Auf diese Weise entzerrt das Quellmodell das Signal, das in die Richtung der Mikrofonposition abgestrahlt wurde, auf die Abstrahlung in alle Richtungen hin. Das bedeutet, dass sich die Leistung der Quelle mit Änderung der Bezugsrichtung ändert: Der Schallpegel in einer bestimmten Richtung ergibt sich aus dem Dämpfungswert und dem für die Bezugsrichtung bestimmten Kennschalldruck:

$$L(\varphi,\theta) = \Delta L(\varphi,\theta) + L_{Kenn}$$

Der Kennschalldruck bezieht sich auf 1 Watt anregende Leistung und einen Meter Abstand. Die abgestrahlte Leistung der Quelle bei 1 W anregender Leistung ergibt sich aus der Summe der Teilintensitäten auf den N Teilflächen der quellumhüllenden Kugelfläche mit einem Radius von einem Meter:

$$P = \sum_{i=1}^{N} \frac{10^{L(\varphi_i,\theta_i)/10}\, p_0^2}{\rho_0 c} S(\varphi_i,\theta_i)$$

(ρ_0: Ruhedichte der Luft; c: Schallgeschwindigkeit; $p_0 = 2\cdot 10^{-5} N/m^2$). Teilt man diese Leistung durch die 1W zugeführter Leistung, entspricht das dem Wirkungsgrad der Quelle. Aus der Leistung lässt sich der Leistungspegel berechnen:

$$L_w = 10\log\frac{P}{P_0} = 10\log\sum_{i=1}^{N} 10^{L(\varphi_i,\theta_i)/10} S(\varphi_i,\theta_i) = 10\log\sum_{i=1}^{N} p_i^2(\varphi_i,\theta_i)\, p_{Kenn}^2 S(\varphi_i,\theta_i)$$

($P_0 = \frac{p_0^2}{\rho_0 c} 1m^2$). Der Kennschalldruckpegel wird für alle Frequenzen konstant gehalten und beeinflusst nur den Absolutbetrag des Leistungspegels. Bezieht man den Leistungspegel auf die Abstrahlung einer ungerichteten Quelle mit gleichem Kennschalldruck, erhält man die Pegeldifferenz, die die Richtcharakteristik auf den Frequenzgang der abgestrahlten Leistung ausübt. Der Leistungspegel einer Kugelquelle mit gleichem Kennschalldruck ist:

$$L_{w-Kugel} = 10\log p_{Kenn}^2\, 4\pi r^2$$

(mit r=1m). Daraus ergibt sich für den normierten Leistungspegel:

$$L_w - L_{w-Kugel} = 10\log\sum_{i=1}^{N}\frac{p_i^2(\varphi_i,\theta_i)S(\varphi_i,\theta_i)}{4\pi r^2} \tag{1.15}$$

Das entspricht einer Mittelung der Dämpfungspegel $\Delta L(\varphi,\theta)$ über die gesamte Kugelfläche. Der Frequenzgang gibt wieder, inwieweit die Richtcharakteristik des Quellmodells das Verhältnis von eingespeister und abgestrahlter Leistung beeinflusst. Mit der Bedingung, dass der Kennschalldruck über alle Terzbänder konstant gehalten wird, entspricht dieser Frequenzgang dem Leistungsfrequenzgang der Richtcharakteristik. Er ist die Invertierung des Frequenzganges des Richtwirkungsmaßes in der Bezugsrichtung[1]. Da sich der eintreffende Schall im Diffusfeld über alle Richtungen addiert, ist dort der Frequenzgang der Energiedichte auch für gerichtete Quellen proportional zum Leistungsfrequenzgang der Quelle (Goertz und Laws 1998). Dieser Zusammenhang ergibt sich auch aus der Leistungsbestimmung nach dem Hallraumverfahren (DIN EN ISO 3741). Für die Auralisation eines Konzertsaals wird die Hörerposition üblicherweise bei einem Abstand von der Quelle außerhalb des Hallradius gewählt. Daher ist hier der Diffusfeldfrequenzgang von großer Bedeutung.

Für die beiden Mikrofonpositionen des Kontrabasses in „Contredanse“ ist in Abbildung 3.38 der normierte Leistungsfrequenzgang der beiden Richtcharakteristiken dargestellt. Die beiden Kurven für den Bass liegen bei einer maximalen Verstärkung von 3 dB und maximalen Dämpfung von 5 dB. Der Bass strahlt hohe Frequenzen bevorzugt nach vorne ab. Beide Richtcharakteristiken beziehen sich auf eine Mikrofonposition in der vorderen Hemisphäre. Folglich wird dieser Frequenzbereich bei den Richtcharakteristiken eher gedämpft. Die maximale Differenz beider Bezugsrichtungen liegt bei 5 dB.

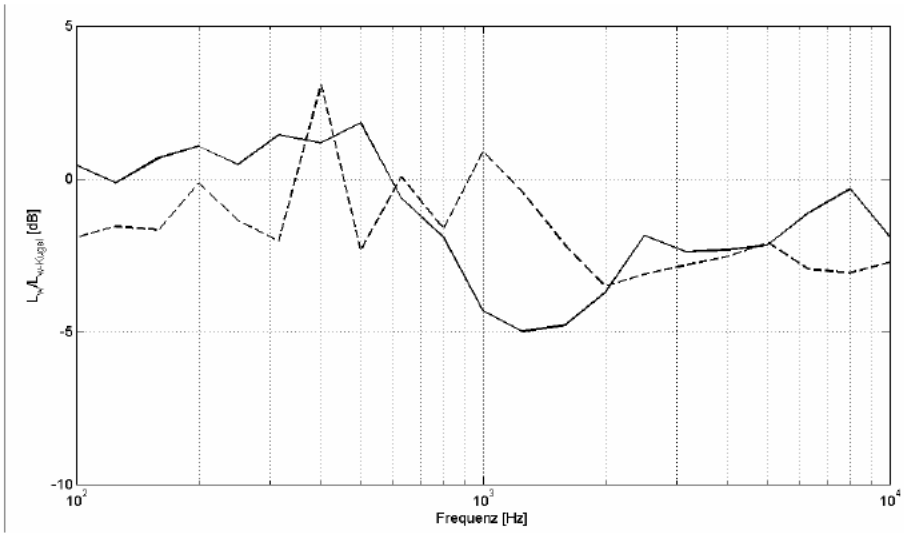

Abbildung 3.38 Normierter Leistungsfrequenzgang des Kontrabasses für zwei Bezugsrichtungen. Durchgehend: Mikrofonposition 1 (Azimut=0°, Polarwinkel=20°); Gestrichelt: Mikrofonposition 2 (Azimut=0°, Polarwinkel=105°)

[1] Gl. (1.13) ist mit p(φ,θ)=1 (Schalldruck in der Bezugsrichtung) das Reziproke von Gl. (1.15).

Auch der Freifeldfrequenzgang der Quelle wird durch die Wahl der Bezugsrichtung beeinflusst, da die Pegeldifferenz zwischen zwei Bezugsrichtungen nicht für alle Frequenzen konstant ist:

$$\Delta L_1(\varphi,\theta,f) = L(\varphi,\theta,f) - L_{Bezug_1}(\varphi_{Bezug_1},\theta_{Bezug_1},f)$$

$$\Delta L_2(\varphi,\theta,f) = L(\varphi,\theta,f) - L_{Bezug_2}(\varphi_{Bezug_2},\theta_{Bezug_2},f)$$

$$\Delta L_2 - \Delta L_1 = L_{Bezug_1}(\varphi_{Bezug_2},\theta_{Bezug_2},f) - L_{Bezug_2}(\varphi_{Bezug_2},\theta_{Bezug_2},f) \neq const.$$

In Abbildung 3.39 ist der Frequenzgang des Basses in der frontalen Richtung für die Richtcharakteristiken mit Bezug auf die zwei unterschiedlichen Mikrofonpositionen dargestellt. Die Kurven geben eine Verstärkung des Kennschalldrucks von maximal 3,5 dB und eine Dämpfung um bis zu 6 dB wieder. Sie gehen maximal um 5 dB auseinander.

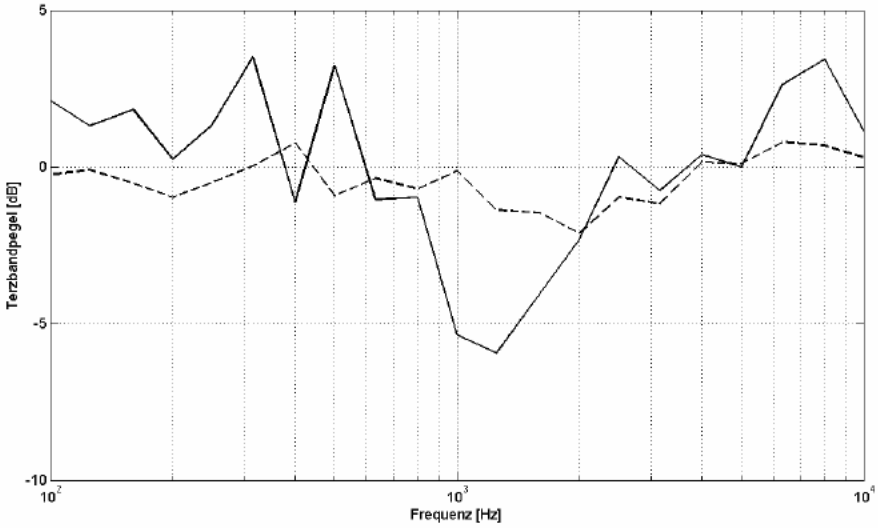

Abbildung 3.39 Verlauf des Schalldruckpegels für die frontale Richtung bei zwei verschiedenen Normierungen. Durchgehend: Mikrofonposition 1 (Azimut=0°, Polarwinkel=20°); Gestrichelt: Mikrofonposition 2 (Azimut=0°, Polarwinkel=105°).

Für die Empfängerposition in einem simulierten Konzertsaal werden sich Diffus- und Direktfeld immer überlagern. Welcher der beiden Frequenzgänge überwiegt, hängt dabei von der Entfernung zur Schallquelle und von der frequenzabhängigen Richtwirkung in Bezug auf die Hörerposition ab. Außerhalb des Hallradius wird der Klangeindruck beim Hörer „überwiegend durch die Frequenzabhängigkeit der abgestrahlten und absorbierten Schallleistung“ bestimmt (Weinzierl 2008, S. 502). Da der Bündelungsgrad von Instrumenten und der Stimme mit der Frequenz ansteigt, ist zu den hohen Frequenzen in der Hauptabstrahlrichtung der Quelle mit einem stärkeren Einfluss des Direktfeldes zu rechnen. Zudem ist der Absorptionsgrad von Räumen bei den hohen Frequenzen meist stärker, was die Bedeutung des Direktfeldes für die hohen Frequenzen bestärkt.

3.6 Auswertung der Richtcharakteristiken

Es wurden von 19 gemessenen Instrumenten Lautsprechermodelle im EASE Format erstellt und auf einer DVD beigefügt. Da eine Auswertung aller Quellmodelle den Umfang dieser Arbeit übersteigen würde, wird hier beispielhaft eine vergleichende Betrachtung der Singstimme vorgenommen.

Zum Vergleich werden die Ergebnisse von Pätynen und Meyer hinzugezogen. Für beide Erhebungen stehen Informationen in beschreibender Textform[1] und als Lautsprecherdaten zur Verfügung. Um eine gemeinsame grafische Auswertung zu ermöglichen, werden die Daten von Pätynen und Meyer in Matlab importiert. In beiden Fällen wurden diese in Oktaven ausgewertet. Zum Vergleich ist es sinnvoll, die Ergebnisse der TU Berlin ebenfalls in Oktaven zur Verfügung zu haben. Die Schalldruckpegel der drei Terzbänder einer Oktave werden dafür jeweils addiert.

Es sei hier darauf hingewiesen, dass bei dem Vergleich der Daten keine Messwerte sondern aus Messwerten interpolierte Daten verglichen werden. Inwiefern sich die Daten zueinander verhalten ist also nicht allein auf Bedingungen der Messungen sondern auch auf unterschiedliche Interpolationsverfahren zurückzuführen. Hier werden eben keine reinen Messungen sondern Quellmodelle verglichen.

Zur grafischen Darstellung der Richtcharakteristiken wird der Pegelverlauf wie bei der Interpolation der Richtcharakteristiken in einer zweidimensionalen Ebene abgebildet. Da die zum Vergleich herangezogenen Daten von der TKK und der PTB in den für Lautsprecherformate üblichen Koordinaten des um die Medianachse der Quelle kreisenden Halbkreisbogens zur Verfügung stehen, werden die gemessenen Richtcharakteristiken der TU Berlin ebenfalls in diesem System dargestellt. Die Darstellung der Pegelverläufe in Abbildung 3.40 kann man so verstehen, dass die Kugelfläche bei der Draufsicht von oben einmal zur Hälfte nach links und einmal zur Hälfte nach rechts ausgeklappt wurde.

[1] Für Pätynen: Pätynen und Lokki 2010 und für Meyer: Meyer 1995, Marshall und Meyer 1985

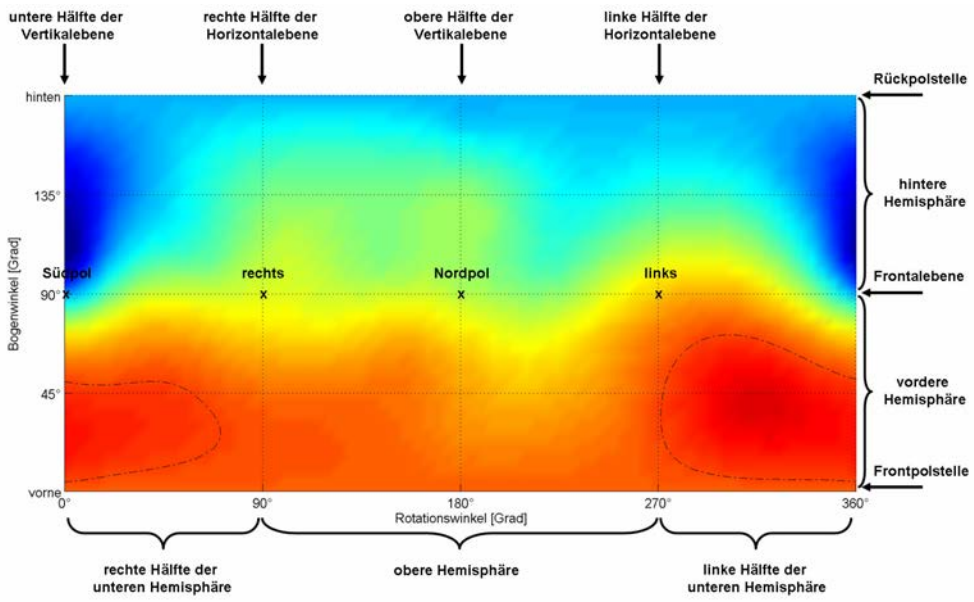

Abbildung 3.40 Bezeichnung der Achsen, Ebenen und Polstellen in der zweidimensionalen Darstellung der Richtcharakteristik in den Koordinaten des um die Medianachse rotierenden Halbkreisbogens.

Für die Rotationsrichtung und den Start des Halbkreisbogens wird eine Drehung entgegen dem Uhrzeigersinn bei Blick nach vorne mit Beginn des Rotationswinkels unterhalb der Quelle gewählt. Der Halbkreisbogen verläuft bei einem Rotationswinkel von 0° auf der unteren Hälfte der Vertikalebene, bei 180° auf der oberen Hälfte der Vertikalebene, bei 90° auf der rechten Hälfte der Horizontalebene und bei 270° auf der linken Hälfte der Horizontalebene. Die 90°-Linie des Halbkreisbogens entspricht der Frontalebene. Bei dieser Linie teilt sich die obere und untere Hälfte der Darstellung in die hintere und vordere Hemisphäre. Alle Winkel des Halbkreisbogens zwischen 0° und 90° entsprechen der vorderen Halbkugel und alle Winkel zwischen 90° und 180° entsprechen der hinteren Halbkugel. Zwar hätte die Darstellung in Polarkoordinaten den Vorteil die Horizontalebene ohne Unterbrechung im Blick zu haben. Eine Transformation der Modelle von der PTB und der TKK in das Polarkoordinatensystem hätte aber gewisse Verzerrungen zur Folge, die hier nicht in Kauf genommen werden sollen.

Die Pegelwerte werden auf den über die gesamte Kugelfläche gemittelten Wert normalisiert. Man kann also direkt das Richtwirkungsmaß ablesen. Zusätzlich wird die Halbwertsbreite der Maxima also die Bereiche, welche bis -3 dB unter dem Maximum liegen, als gestrichelte Kontur dargestellt.

Sowohl für die vertikalen als auch für die horizontalen Polardiagramme entspricht 0° der Vorderseite der Quelle. Von dort aus bewegt sich der Winkel im vertikalen Polardiagramm nach oben und im horizontalen Polardiagramm nach links. Die Pegelkurven beziehen sich auf den Betrag ihres Maximums. Sie stellen also die Pegeldifferenz zur stärksten Hauptabstrahlrichtung dar. Es werden nur Pegelwerte oberhalb -35 dB angezeigt. Fällt die

Kurve unter diesen Wert, entsteht eine Lücke in diesem Bereich. Neben den Pegelkreisen im Abstand von -10 dB gibt es einen Strichpunktkreis bei -3 dB. Der Abschnitt einer Pegelkurve, der sich über diesem Grenzkreis befindet, stellt den Bereich der Halbwertsbreite der Hauptabstrahlrichtung dar.

Neben der Betrachtung von Polardiagrammen bieten sich zur Bewertung von Richtcharakteristiken nach Meyer (1995) folgende Kennwerte an:

- Die „Dynamik“ der Richtcharakteristik beschreibt den Pegelunterschied zwischen dem kleinsten und größten Wert eines Diagramms. Dieser Wert hat nach Meyer für die räumliche Wirkung eine verhältnismäßig geringe Bedeutung, da sich die kleinsten Werte aus den häufig sehr schmalen Einbrüchen ergeben. Für die Auswertung der Messergebnisse des Mikrofonarrays ist allerdings zu beachten, dass sehr schmale Einkerbungen aufgrund der begrenzten Anzahl von Messpunkten nicht erfasst werden können. Für die Auswertung der „Dynamik“ wird dieser Umstand zu berücksichtigen sein.
- Für die Einschätzung der richtungsabhängigen Auswirkung der Hauptabstrahlungskeulen wird die so genannte *Halbwertsbreite* definiert. Diese ist vergleichbar mit der in EASE angewandten -3dB-Kurve[1]. Das ist der Bereich in dem der Pegelwert um nicht mehr als 3 dB unter das Maximum sinkt. Der Begriff Halbwertsbreite leitet sich aus der Tatsache her, dass 3 dB eine Halbierung der abgestrahlten Energie bedeutet, was dem Lautstärkeunterschied bei einer Halbierung der Anzahl der Spieler entspricht.
- Das *Vor-Rück-Verhältnis* bezeichnet den Pegelunterschied zwischen dem nach vorne und nach hinten abgestrahlten Schalldruck. Es bietet sich an, das Verhältnis aus den beiden Mittelwerten des vorderen und hinteren ±10°-Bereichs zu bestimmen. Die *Seitendämpfung* beschreibt den Pegelunterschied zwischen vorderer und seitlicher Abstrahlung. Dabei gilt: $L_{\text{Dämpfung}} = L_{\text{seitlich/hinten}} - L_{\text{vorne}}$.
- Zur Einschätzung der Richtwirkung einer gefragten Richtung gibt es schließlich den *statistischen Richtfaktor* und das aus ihm hergeleitete logarithmische *Richtwirkungsmaß*. Beide Kenngrößen wurden bereits in Abschnitt 3.4.2 hergeleitet.

3.6.1 Richtcharakteristik der Singstimme

[1] In EASE ist es allerdings nur möglich *eine* Hauptabstrahlrichtung in Form einer -3dB-Keule zu bestimmen.

Die Messsignale der Stimme wurden von einer Sopranistin als Einzeltöne in chromatischem Tonabstand eingesungen. Der Tonumfang erstreckt sich von G3 bis G#5[1]. Jeder Ton wurde einmal im Fortissimo und einmal im Pianissimo präsentiert.[2] Damit ergeben sich 52 Einzelsignale, über die der Schalldruckpegel gemittelt wird. Das phonetische Material der Stimme bestand aus dem Vokal „a“.

Nach Meyer (ebd.) ergeben sich für verschiedene Vokale geringfügig unterschiedliche Abstrahlcharakteristiken. Der Unterschied der Abstrahlung liegt an der Trichterwirkung des Mundes, der je nach Vokal eine andere Stellung einnimmt. Dadurch wirken sich Bündelungs- und Phasenauslöschungseffekte der Mundöffnung unterschiedlich aus (Marshall und Meyer 1985, S. 134). Von Marshall und Meyer (ebd.) wurden die Vokale „a“, „e“ und „o“ verglichen. Die größten Unterschiede ergeben sich für die seitliche Abstrahlung und im Frequenzbereich zwischen 500 Hz bis 4 kHz. Diese Differenzen liegen in einer Größenordnung von 5 dB. In Abbildung 3.41b ist die Seitendämpfung der drei Vokale als Funktion der Frequenz dargestellt. Der Vokal „e“ hat den schwächsten Seitenpegel, während dieser für „o“ am stärksten ausgeprägt ist. Der Kurvenverlauf für den Vokal „a“ liegt dazwischen und tendiert dabei mehr zum „o“.

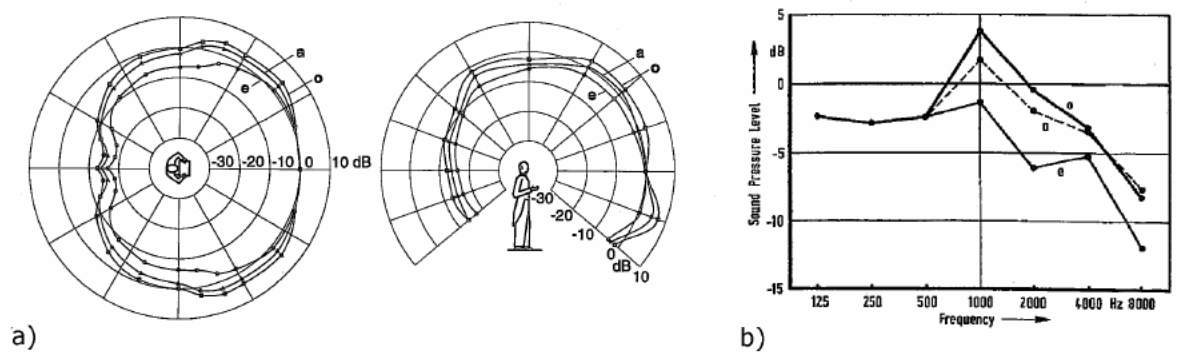

Abbildung 3.41 a: Horizontales und Vertikales Polardiagramm für drei Vokale. b: Pegeldifferenz zwischen vorderer und seitlicher Abstrahlung für dieselben drei Vokale als Funktion der Frequenz. Quelle: Marshall und Meyer (1985).

Die horizontale und vertikale Abstrahlung der drei Vokale ist für die Oktave bei 2 kHz in Abbildung 3.41a dargestellt. Charakteristisch ist, dass sich für die Vokale „a“ und „o“ ein Maximum der Abstrahlung bei 30° unterhalb der Horizontalebene an der Vorderseite des Sängers ergibt, welches für den Vokal „e“ nicht so stark ausgeprägt ist. In der horizontalen Ebene bilden sich bei 40° seitlich der Medianachse Maxima aus, die auch hier für „e“ am schwächsten ausgeprägt sind.

[1] Für die gesamte Arbeit wird die wissenschaftliche Notenbezeichnung nach der *Acoustical Society of America* verwendet Young 1939.

[2] Die Sängerin hatte dabei die Anweisung einmal „so laut wie möglich, ohne dass der Ton unschön klingt“ und einmal „so leise wie möglich, ohne dass der Ton abreißt“, zu singen. Detzner et al. 2010

Die Unterschiede zwischen Frauen- und Männerstimmen werden als gering eingestuft (ebd.). Insofern ist es unbedenklich, dass, wie bei dem Stück „River Man“, die gemessene Charakteristik eines Soprans für die Simulation einer Männerstimme eingesetzt wird. Allerdings ist zu beachten, dass der tiefste eingesungene Ton G3 einer Grundtonfrequenz von 196 Hz entspricht. Daher werden die untersten drei Terzbänder von 100 Hz, 125 Hz und 160 Hz nicht mit in die Berechnung der Richtcharakteristik einbezogen. EASE simuliert für Frequenzen, die außerhalb des für eine Quelle definierten Frequenzbereiches liegen überhaupt keine Abstrahlung, was einem Hochpassfilter gleich kommt. Da der Frequenzgang der nachhallfreien Musikaufnahmen nicht auf diese Weise beeinflusst werden soll, werden die unteren drei Frequenzbänder durch die Richtcharakteristik der PTB ersetzt. Dort wurde die Abstrahlung mit einem Bariton gemessen.

In Abbildung 3.42 ist mit der blauen Kurve der Maximal- und der Durchschnittspegel (gestrichelt) für alle Terzbänder dargestellt. Bei 200 Hz ist wegen des Einsetzens der tiefsten Grundtonfrequenz erwartungsgemäß ein Sprung zu beobachten. Der Pegel steigert sich bis zum Frequenzband bei 800 Hz. Die obere Grenze des Tonumfangs ist G#5 und entspricht einer Grundtonfrequenz von 830 Hz. Der Pegel sinkt von 1000 Hz an wieder ab weil ab hier nur noch Obertöne angeregt werden. In dem Frequenzbereich bei 3 kHz gibt es ein weiteres Maximum, welches auf den Sängerformanten zurückzuführen ist (vgl. ebd.).

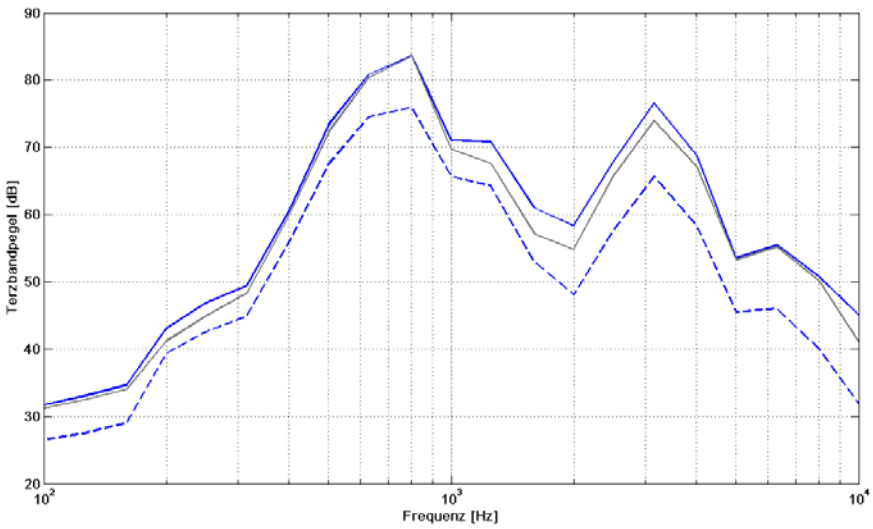

Abbildung 3.42 Terzbandpegel für die Stimme. Blau: Maximum, Blau gestrichelt: Mittelung über alle Richtungen, Grau: Pegel bei 45° unterhalb der Media nachse.

In Abbildung 3.43a sind die zweidimensionalen Richtcharakteristiken der TU Berlin für fünf Oktaven dargestellt. Die Pegelverläufe zeigen, dass es für alle Frequenzen in der vorderen Hemisphäre bei circa 45° unterhalb der Horizontaleb ene ein Maximum gibt. Die graue Kurve in Abbildung 3.42 beschreibt den Pegelverlauf bei 45° unterhalb der Medianachse. Mit Ausnahme der Abweichungen im Frequenzbereich zwischen 1 kHz und 3 kHz (maximal 5

dB) stimmt diese Kurve mit dem maximalen Schalldruckpegelverlauf nahezu überein. Die Differenzen der beiden Kurven bei den mittleren Frequenzen ergeben sich aus den von der Medianebene seitlich abweichenden Maxima der Pegelverläufe, die bei der Auswertung in Terzbändern zum Teil deutlicher hervortreten als bei den Oktaven.

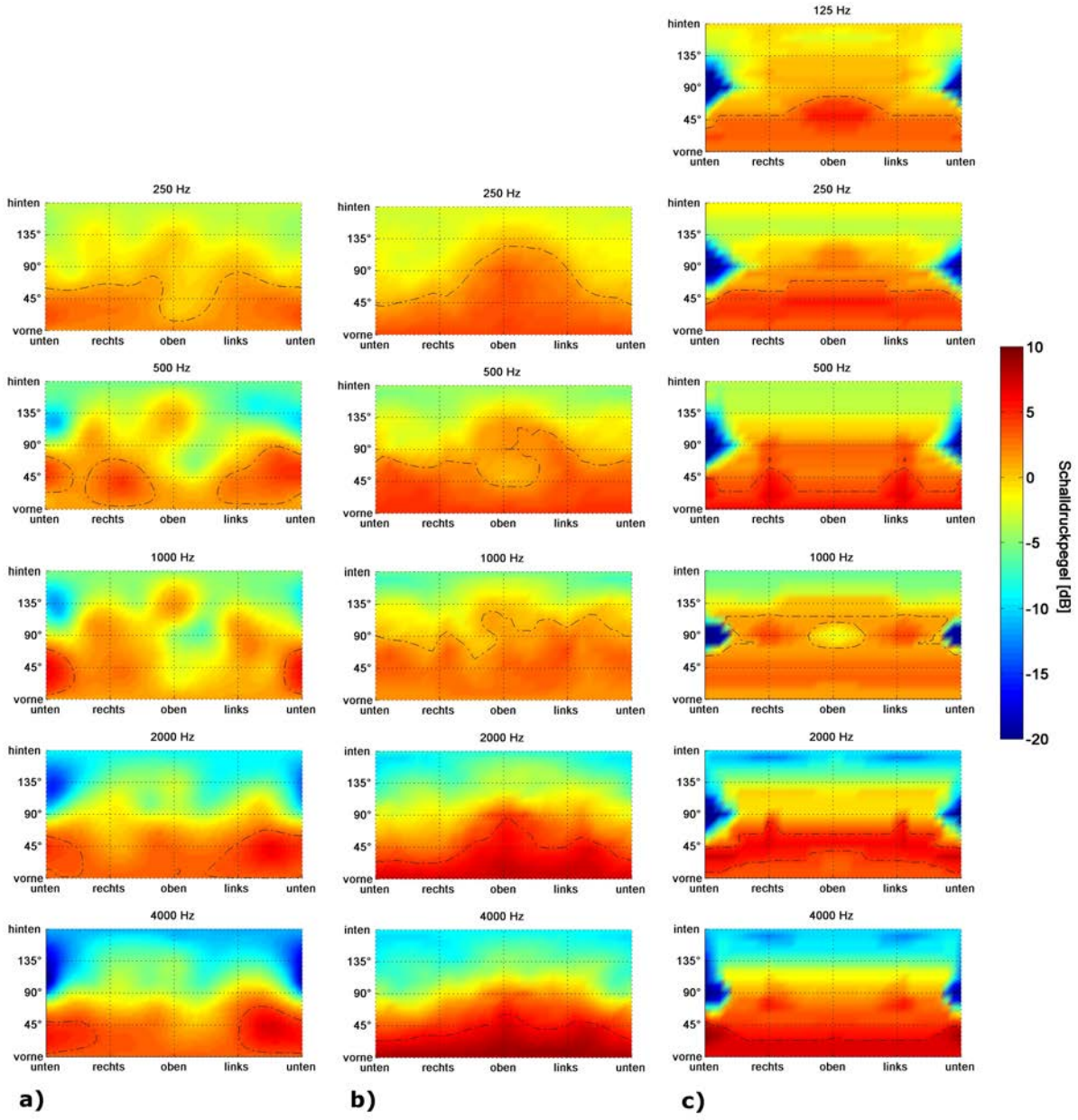

Abbildung 3.43 Richtcharakteristiken der Stimme für sechs Oktaven. (x-Achse: Rotationswinkel; y-Achse: Bogenwinkel). a: TU Berlin, b: TKK Helsinki, c: PTB. Die Richtcharakteristiken bei 125 Hz für TU und TKK fehlen, da hier mit einem Sopran gemessen wurde.

Im Vergleich der drei Richtcharakteristiken in Abbildung 3.43 fällt zunächst auf, dass die Daten der PTB zur Medianebene symmetrisch sind. Sie basieren auf Messungen von Marshall und Meyer (1985). Hier wurde in nur drei Ebenen gemessen: der Horizontal-, der Vertikalebene und in einer nach links verschobenen Vertikalebene, deren Messwerte

aufgrund der theoretisch vorhandenen Symmetrie der Abstrahlung der Stimme auf die rechte Seite gespiegelt wurden.

Des Weiteren fällt auf, dass vor allem bei den tiefen Frequenzen die Richtcharakteristik der PTB eine wesentlich größere Dynamik hat. Bei 4 kHz ist die Dynamik für alle drei Messungen am höchsten. Während hier bei der PTB ein Pegelunterschied von 44 dB erreicht wird[1], liegt die Dynamik bei der TKK mit 21 dB nur bei der Hälfte. Die Daten der TU liegen mit einem Pegelabstand von 28 dB dazwischen und näher an denen der TKK. Die hohe Dynamik bei der PTB kommt dadurch zustande, dass an der Unterseite der Quelle für alle Frequenzen ein starker Pegelabfall herrscht. Dies ist auch gut anhand des vertikalen Polardiagramms in Abbildung 3.44 zu erkennen. Erstaunlicherweise wird bei Marshall und Meyer keine Dynamik in dieser Größenordnung erreicht. Dort findet man Werte zwischen 6 dB bei 125 Hz und 23,5 dB bei 4 kHz. Da bei Marshall und Meyer nur bis zu einem Höhenwinkel von 40° unterhalb der Horizontalebene g emessen wurde, wird vermutet, dass die starken Dämpfungen an der Unterseite der Quelle für die Interpolation der Messdaten hinzugefügt wurden. Ob diese Werte aus der Theorie oder anderen Messungen stammen, konnte nicht ermittelt werden. Betrachtet man die Daten der PTB nur bis zu einem Winkel von 40° unterhalb der Horizontalachse kommt man auf Dynamiken in der gleichen Größenordnung wie bei der TU und der TKK. Die Dämpfungen an der Rückseite sind bis 40° unterhalb der Horizontalebene für alle drei Messungen recht ähnlich, außer bei 500 Hz und 1 kHz, wo einmal die Kurve der TKK und einmal die der TU von den jeweils anderen beiden Messungen um 5 bis maximal 10 dB abweicht.

[1] Die volle Dynamik der Daten der PTB ist in Abbildung 3.43 nicht sichtbar, da der Farbverlauf für die Intensität auf eine Dynamik von 30 dB beschränkt wurde.

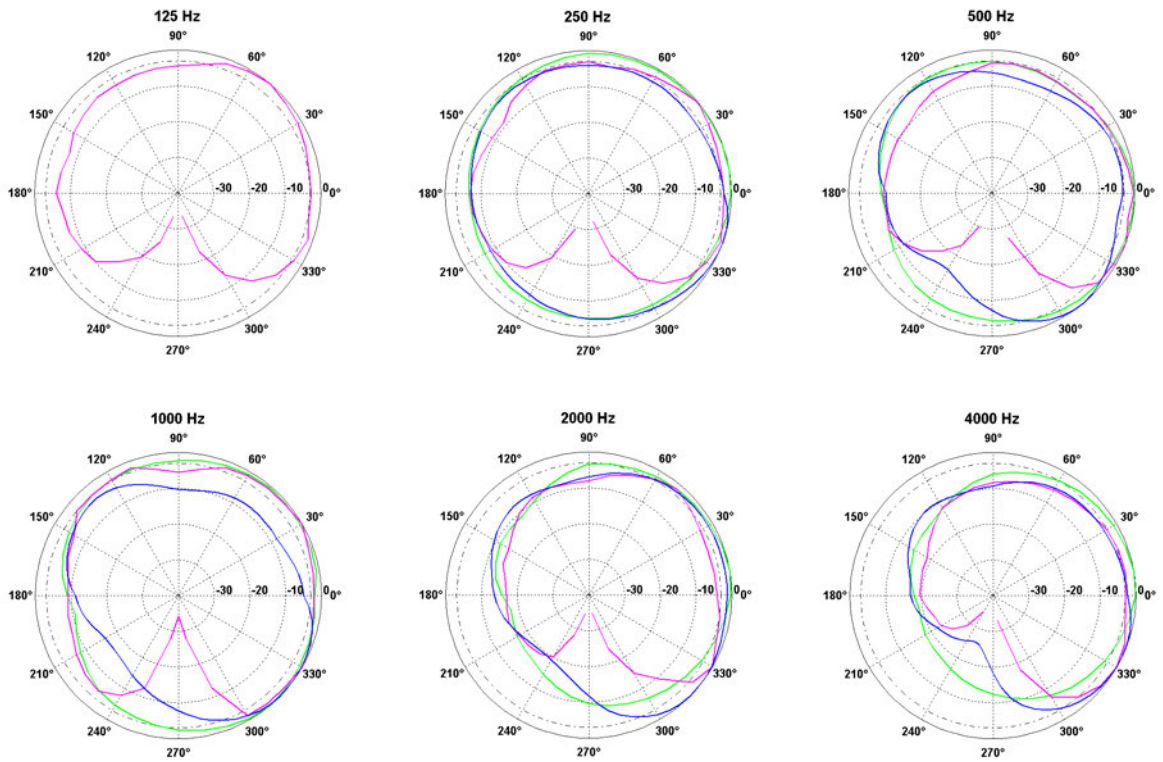

Abbildung 3.44 Vertikales Polardiagramm der Stimme für sechs Oktaven. Blau: TU Berlin, Grün: TKK Helsinki, Magenta: PTB.

Bei den Daten der PTB wurden die Richtcharakteristiken über die drei gesungenen Vokale „a“, „e“ und „o“ gemittelt (vgl. ebd.). Wie bei der TU weist der vertikale Pegelverlauf an der Vorderseite der Quelle unterhalb der Horizontalebene für Frequenzen ab 250 Hz teilweise eine stärkere Abstrahlung als bei 0° auf (Abbildung 3.44). Besonders bei 2 und 4 kHz ist dieser Effekt ausgeprägt. Die lokalen Maxima in der unteren Hemisphäre sind bei der PTB schmaler. Ihre vertikalen Halbwertsbreiten liegen im Bereich von 20° bis 40°, während die Maxima der TU eine Breite von 40° bis 120° erreiche n. Bei den Daten der TKK ist dieses Phänomen nicht zu beobachten. Bei 2 kHz und 4 kHz weitet sich das frontale Maximum eher nach oben und bei 500 Hz ein bisschen nach unten aus. Die Messungen wurden hier nicht mit vokalisierten Einzeltönen vorgenommen. Die Sopranistin hatte hier einen Ausschnitt einer Arie aus Mozarts „Don Giovanni“ zu singen. Die Richtcharakteristik ergibt sich aus einer Mittelung über eine Vielzahl von gesungenen Vokalen und Konsonanten. Da die Bevorzugung der Richtung unterhalb der Medianachse nach Marshall und Meyer vorrangig ein Effekt der „a“- und „o“-Vokale ist, ergeben sich die Abweichungen der vertikalen Höhe der Hauptabstrahlrichtung bei der TKK möglicherweise aus der Tatsache, dass anderes Phonemmaterial dargeboten wurde.

Bei 1 kHz weitet sich der Schall für TKK und PTB an der Vorderseite auch nach schräg oben relativ stark aus, was bei der TU nicht der Fall ist. Eine Gemeinsamkeit zwischen TU und PTB ist ein weiteres lokales Maximum bei dieser Frequenz auf der hinteren Hemisphäre bei

circa 30° bis 60° über der Horizontalebene. Allen d rei Modellen ist eine Verbreiterung der horizontalen Abstrahlung in die Seiten bei 1 kHz gemeinsam (Abbildung 3.45). Bei den Frequenzbändern unter 1 kHz ist der Verlauf der horizontalen Abstrahlung für alle drei Richtcharakteristiken sehr ähnlich. Es ergeben sich moderate Seitendämpfungen bis zu 4 dB. Bei 1 und 2 kHz ergibt sich für die PTB bei 30 bis 40 Grad seitlich ein 3 dB stärkerer Pegel als für die frontale Abstrahlung. Dieser Effekt tritt bei den anderen beiden Modellen nicht auf, findet sich jedoch bei der TU in abgeschwächter Form bei 125 und 500 Hertz wieder. Die theoretisch vorhandene Symmetrie der horizontalen Abstrahlung wird vor allem bei 2 und 4 kHz von der TU und der TKK nicht so genau gemessen. Es ergibt sich in der Horizontalebene bei beiden eine Abweichung bis zu 3 dB zwischen linker und rechter Hemisphäre.

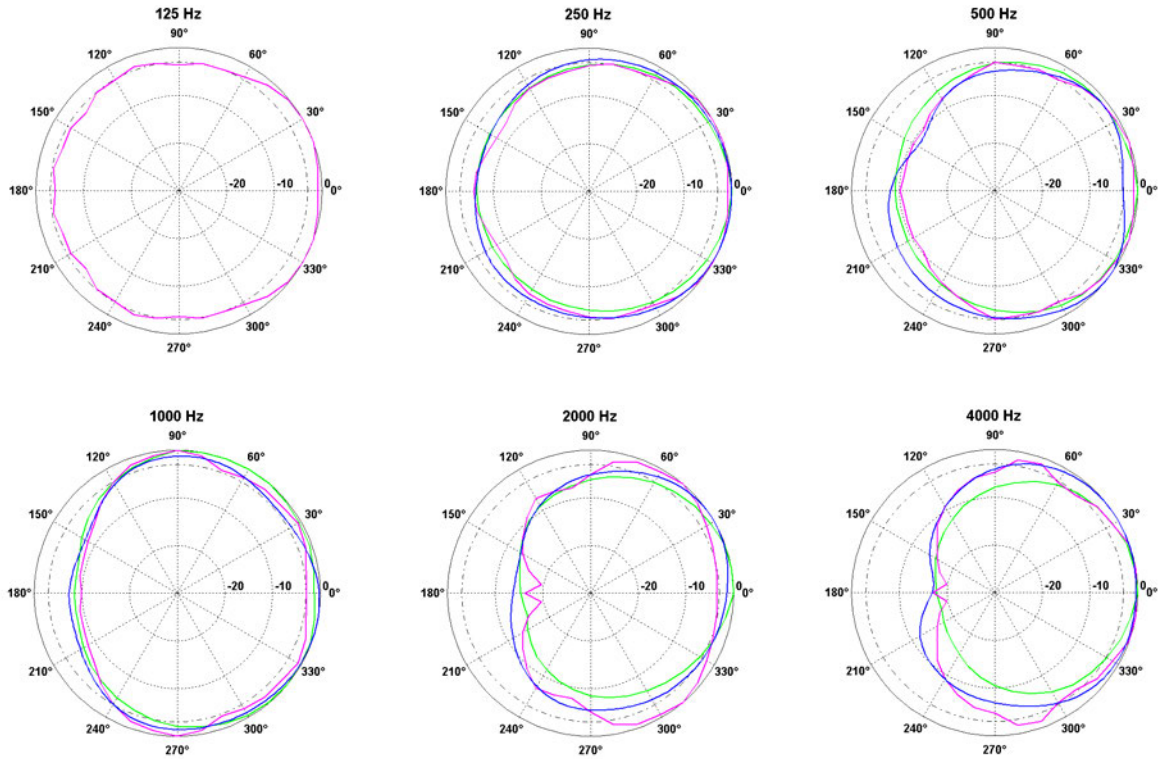

Abbildung 3.45 Horizontales Polardiagramm der Stimme für sechs Oktaven. Blau: TU Berlin, Grün: TKK Helsinki, Magenta: PTB.

In Abbildung 3.46a ist der Pegelunterschied zwischen seitlicher und frontaler Abstrahlung als Funktion der Frequenz dargestellt. Diese Seitendämpfung wurde mit dem Mittelwert aus linkem und rechtem Seitenpegel berechnet. Theoretisch müsste der Verlauf der blauen Kurve für die TU ähnlich wie die Kurve für den Vokal „a“ in Abbildung 3.41b sein. Sie liegen beide in der gleichen Größenordnung und es ergibt sich in beiden Fällen ein Maximum des seitlich abgestrahlten Schalls bei 1 kHz, wonach der seitliche Pegel mit steigender Frequenz abfällt. Für die Messung der TU Berlin ist das Maximum bei 1 kHz nicht so stark ausgeprägt wie bei der PTB. Dies könnte darauf zurückzuführen sein, dass die Seitendämpfung bei der

einen Messung aus einer Frauenstimme und bei der anderen aus einer Männerstimme hergeleitet wurde. Bei Marshall und Meyer (1985) ergeben sich nämlich für Frauenstimmen bei 1 kHz schwächere Seitenpegel als bei Männerstimmen. Der gleiche Effekt tritt auf, wenn die Stimme im piano dargeboten wird (ebd.). Die Kurve in Abbildung 3.41b wurde aus einem im forte singenden Bariton ermittelt. Da bei der Messung der TU aus beiden Dynamikstufen (Pianissimo und Fortissimo) gemittelt wurde, könnte hierin ein weiterer Grund für das schwächere Maximum bei 1 kHz liegen.

Das Verhältnis von vorderer zu hinterer Abstrahlung ist bei allen drei Messungen sehr ähnlich (Abbildung 3.46b). Die dafür verantwortliche Abschattung des Kopfes wird zu hohen Frequenzen wesentlich effektiver und ist hier circa 10 dB stärker als bei den tiefen Frequenzen. Bei 500 Hz gibt es für die TU eine Abweichung, da sich hier der Schall etwas mehr nach schräg unten und zur Seite verteilt, als bei den anderen beiden Richtcharakteristiken.

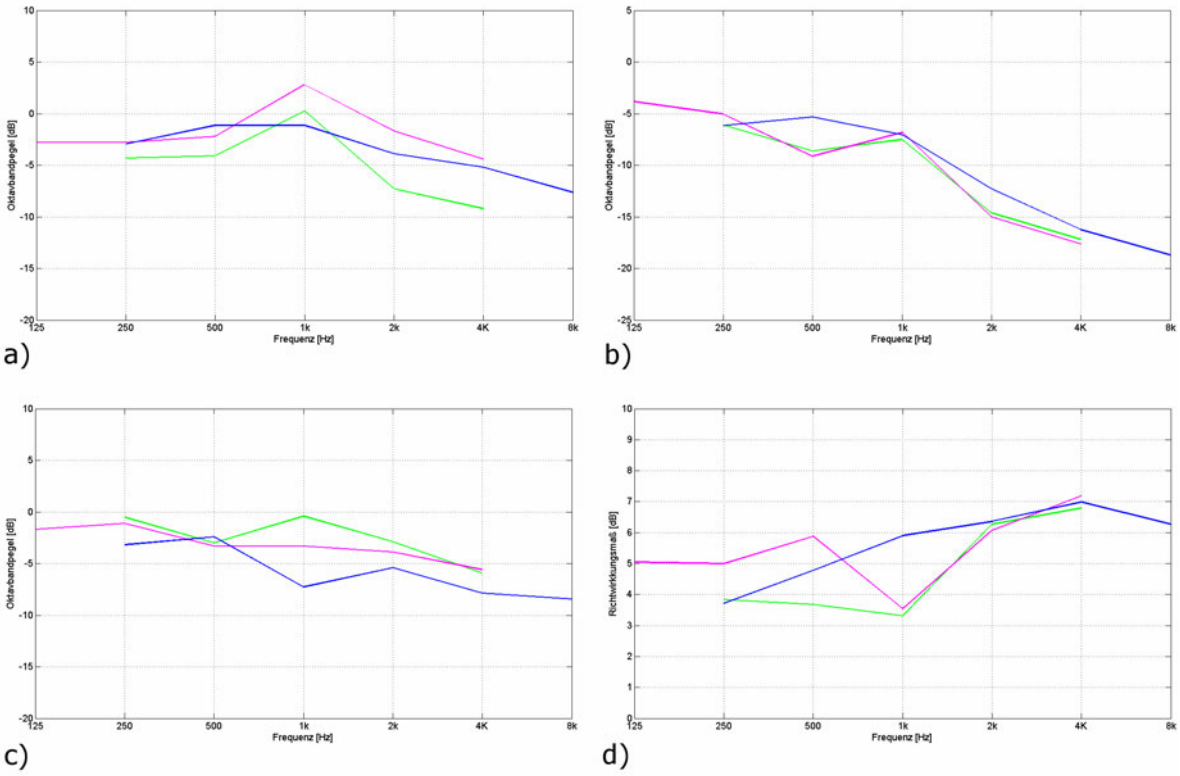

Abbildung 3.46 a: Pegelunterschied zwischen frontaler und seitlicher Abstrahlung. b: Pegelunterschied zwischen frontaler und rückwärtiger Abstrahlung. c: Pegelunterschied zwischen frontaler Abstrahlung und Abstrahlung nach oben. d: Richtwirkungsmaß für die Richtung maximaler Abstrahlung. Blau: TU Berlin, Grün: TKK Helsinki, Magenta: PTB.

Frequenzband [Hz]	TU hor. / vert. [°]		TKK hor. / ve rt. [°]		PTB hor. / vert. [°]	
125					0	50
250	-5	-15	0	10	0	40
500	5	-50	0	0	±20	0
1000	0	-50	70	10	±90	10

2000	30	-25	0	10	±15	-25
4000	35	-25	0	10	0	-30
8000	-5	-30				

Tabelle 9 Lage der Maxima. Die Winkel geben horizontale und vertikale Abweichungen aus der Medianachse an.

Generell kann man für die Stimme sagen, dass die Richtwirkung mit der Frequenz ansteigt. Das Richtungsmaß startet für alle drei Abstrahlcharakteristiken bei den unteren Frequenzen mit 3,5 – 5 dB und erreicht sein Maximum von circa 7 dB bei 4 kHz (Abbildung 3.46d). Eine Abweichung dieses Verlaufs liegt bei 1 kHz. Hier breitet sich der Schall bei der PTB und der TKK nahezu gleichmäßig in der gesamten vorderen Hemisphäre aus, während bei der TU ein klares Maximum unter der Medianachse herrscht, welches ein Richtungsmaß von 6 dB hat. Bei der PTB und TKK sinkt der Wert bei 1 kHz auf circa 3 dB ab.

Da die frontale Abstrahlung in allen Frequenzen bis 40° seitlich kaum unter 3 dB absinkt, kann man für alle drei Richtcharakteristiken sagen, dass in diesem Bereich mit keiner großen Veränderung der Klangfarbe zu rechnen ist.

Für die Halbwertsbreite in Richtung oben und unten ergeben sich eher Abweichungen zwischen den Messungen und den Frequenzbändern. Bei der TU existiert von 0° bis 45° nach unten über alle Frequenzen ein Hauptabstrahlungsgebiet. Für die Dämpfung des Schalls über dem Sänger ergeben sich bei allen drei Richtcharakteristiken Werte zwischen -3 und 0 dB bei den tiefen und -7 bis -5 dB bei den hohen Frequenzen (Abbildung 3.46c). Bei 1 kHz weichen die Kurven relativ stark voneinander ab.

Um den Einfluss der Quellmodelle auf den Diffusfeldfrequenzgang abzuschätzen, eignet sich eine Betrachtung der normierten Leistungsfrequenzgangs (vgl. Abschnitt 3.5.), welcher in Abbildung 3.47 dargestellt ist. Die Kurven liegen in der gleichen Größenordnung, haben eine Dynamik von 5 bis 8 dB und weichen um maximal 4 dB voneinander ab. Zusätzlich ist der Verlauf des Leistungsfrequenzgangs für das Modell der TU als Terzspektrum dargestellt (graue Kurve). Die Oktavpegel der blauen Kurve ergeben sich nicht aus der Addition der Terzpegel. Eine Normierung der Richtcharakteristik in Oktaven hat also einen anderen Diffusfeldfrequenzgang zur Folge als bei der Normierung mit Terzpegeln.

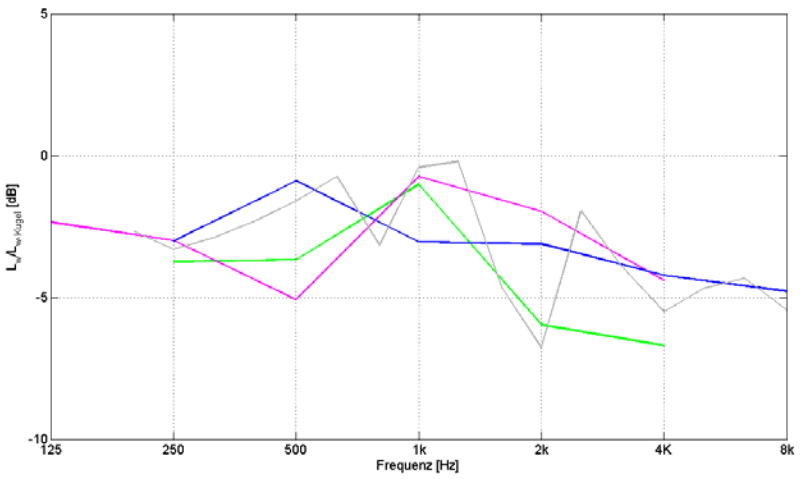

Abbildung 3.47 Normierter Leistungsfrequenzgang. Blau: TU Berlin, Grün: TKK Helsinki, Magenta: PTB, Grau: Terzbandspektrum der TU Berlin.

Dass das Charakteristikum der nahezu gleichmäßigen Verteilung des Schalls in der vorderen Hemisphäre bei 1 kHz in der Richtcharakteristik der TU nicht so ausgeprägt ist, liegt an dem kräftigen Abstrahlungsmaximum, welches bei Betrachtung der Terzbandpegel eher ein Phänomen im Bereich von 800 Hz ist (Abbildung 3.48). Bei dem Terzband von 1 kHz ergibt sich eine ähnliche Ausweitung der Hauptabstrahlung wie bei TKK und PTB. Da der absolute Durchschnittspegel bei 800 Hz um circa 10 dB höher als bei 1 kHz ist (siehe Abbildung 3.42), setzt sich diese Charakteristik bei der Ermittelung der Oktaven eher durch. An diesem Beispiel wird erstens deutlich wie sehr sich die Abstrahlung zwischen den Terzbändern unterscheiden kann. Bei einer Auswertung in Oktaven gehen gewisse charakteristische Details der Abstrahlung für die Simulation verloren. Zweitens zeigt sich, wie lautere Anteile des Messsignals die Richtcharakteristik dominieren (vgl. Pollow 2009, S. 1472), was bei einer Auswertung in Oktaven zu ungenaueren Ergebnissen führt als bei Terzbändern.

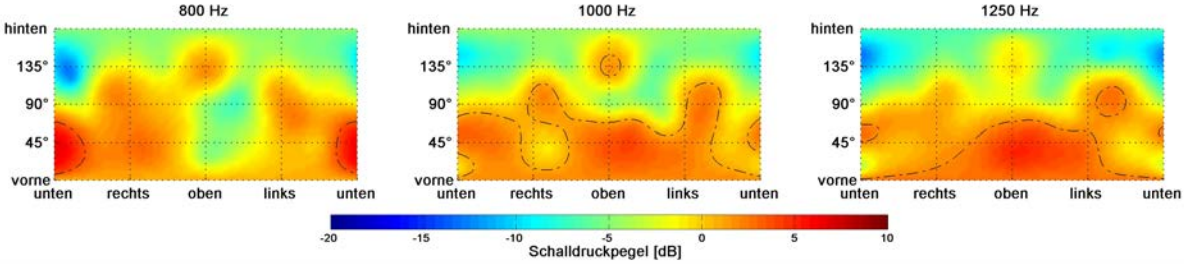

Abbildung 3.48 Richtcharakteristiken der Stimme für drei Terzbänder (x-Achse: Rotationswinkel; y-Achse: Bogenwinkel).

3.6.2 Allgemeine Fehlerbetrachtung und Diskussion

Wie eben gezeigt, geht eine Erfassung der Richtcharakteristiken in Oktavbändern mit negativen Effekten gegenüber der Terzbandauswertung einher. Dies hat Verzerrungen des Frequenzgangs der Quelle zur Folge (vgl. Abbildung 3.47). Zudem wertet das menschliche Gehör die Lautheit und die Klangfarbe von akustischen Ereignissen in Frequenzgruppen aus, die (ab 500 Hz) ungefähr eine Terzbreite haben (Fastl und Zwicker 2006). Um der menschlichen Wahrnehmung zu entsprechen, ist eine Bestimmung in Terzbändern für die Simulation von Abstrahlcharakteristiken hinreichend.

Bei der Auswertung der Stimme wurde bereits deutlich wie Unterschiede in der Darbietung die Richtcharakteristik beeinflussen können. Da nicht jedes Instrument baugleich ist, treten auch hier Unterschiede in den Abstrahlungscharakteristiken auf (Meyer 1995). Es ist jedoch anzunehmen, dass derartige Abweichungen in einer Größenordnung liegen, die für eine Auralisation als unerheblich angesehen werden kann.

Ein weiteres Problem ist die Positionierung der Instrumente im Zentrum des Arrays. Insbesondere der Kontrabass und das Cello waren deutlich aus der Mitte verschoben. Sich hieraus ergebende Veränderungen des Schallpegelverlaufs und deren näherungsweise Kompensation wurden bereits in Abschnitt 3.1.1 betrachtet.

Bei einer Anzahl von 32 Messpunkten ist eine Erfassung von Details der Abstrahlungscharakteristiken nur bis zu einem gewissen Grad möglich. Der verfälschende Effekt dieser Ungenauigkeit wird umso größer je komplexer die Richtcharakteristik wird. Dies betrifft vor allem die Abstrahlung von Holzblasinstrumenten, bei denen es zu mehrfachen schmalen Keulen und Einkerbungen in der Richtcharakteristik kommen kann, die durch die begrenzte Anzahl der Messmikrofone nur unzureichend erfasst werden. Analysiert man die Abstrahlcharakteristik durch eine Zerlegung in Kugelfunktionen, macht sich der Fehler, den eine begrenzte Anzahl von Messproben einbringt, als Aliasingeffekt bemerkbar, der ab einer bestimmten Ordnungszahl der Kugelharmonischen auftritt. Untersuchungen zum Zusammenhang von Ordnungszahl der Kugelharmonischen und Aliasing bei Mikrofonarrays wurden von Rafaely (2005) durchgeführt. Bei einfachen Charakteristiken wie der Stimme und den Blechbläsern, ist eine gute Abbildung der Abstrahleigenschaften zu erwarten. Es ist anzunehmen, dass das beschriebene Verfahren den Verlauf des Schalldrucks bei diesen Quellen sehr gut interpoliert.

Bei den Holzbläsern kommt das Problem hinzu, dass der Schall je nach gespielter Tonhöhe von unterschiedlichen Orten des Instrumentes abgestrahlt wird. Dadurch ändert sich die Richtcharakteristik von Ton zu Ton. Dieses Abstrahlverhalten kann durch Quellmodelle mit statischer Richtcharakteristik nicht simuliert werden. Otondo und Rindel (2005) haben zur Lösung dieses Problems eine Methode der Simulation von Multikanalaufnahmen vorgestellt. Dabei konnten Vorteile in der Reproduktion der Klangfarbe festgestellt werden.

4 Auralisation

Zur Auralisation der simulierten Instrumentenquellen werden diese als Lautsprechermodelle in das CAD-Modell eines Konzertsaals eingefügt. Für jede Quelle-Empfängerkonstellation muss eine Raumimpulsantwort gerendert werden. Anschließend werden den monauralen RIRs mithilfe von gemessenen HRTFs die Eigenschaften des menschlichen binauralen Hörens aufgeprägt. Dabei spielt die Übertragungsfunktion des Kopfes eine entscheidende Rolle. Je nach Kopfdrehung ergibt sich eine andere Laufzeit zwischen linkem und rechtem Ohr und eine andere Filterung des Signals. Um ein realistisches Anpassen der Außenohrübertragungsfunktion an die Kopfbewegung zu gewährleisten, ist eine Auflösung von einer BRIR pro 1°-Kopfdrehung hinreichend. Der resultierende Datensatz aus dynamischen BRIRs wird für die Simulation der Konzertaufführung im Rahmen des Hörversuchs in die Echtzeit-Faltungssoftware Fwonder geladen. Die Kopfdrehung des Rezipienten wird mit einem Headtracker erfasst und an Fwonder weitergeleitet, sodass die Faltung der trockenen Zuspielsignale mit den BRIRs dynamisch angepasst werden kann.

4.1 Der Raum

Als Raummodell wird ein bereits bestehender Konzertsaal von der Firma ADA[1] zur Verfügung gestellt. Die Werte für Größe und Nachhallzeit entsprechen denen üblicher Räume für Kammermusikaufführungen[2]. Das Volumen beträgt 6426 m^3. Mit Publikum liegt die mittlere Nachhallzeit nach der Sabine-Formel bei 1,8s.

Eine grafische Darstellung des Konzertsaals ist in Abbildung 4.1 zu sehen. Die längere Ausdehnung des Raumes verläuft entlang der y-Achse und die Breite des Raumes wird von der x-Achse bestimmt. Die z-Achse entspricht der Höhe. Mit Blick auf die Bühne verläuft die positive Richtung der Achsen für x nach rechts, für y in die Blickrichtung und für z nach oben. Die vordere Bühnenkante befindet sich bei x=y=0.

[1] Acoustic Design Ahnert, Hersteller von EASE.

[2] zum Vergleich: Kammermusiksaal der Philharmonie Berlin: V=11000m^3; RT=1,8s. Kammermusiksaal des Konzerthauses Berlin: V=2100m^3; RT=1,3s.

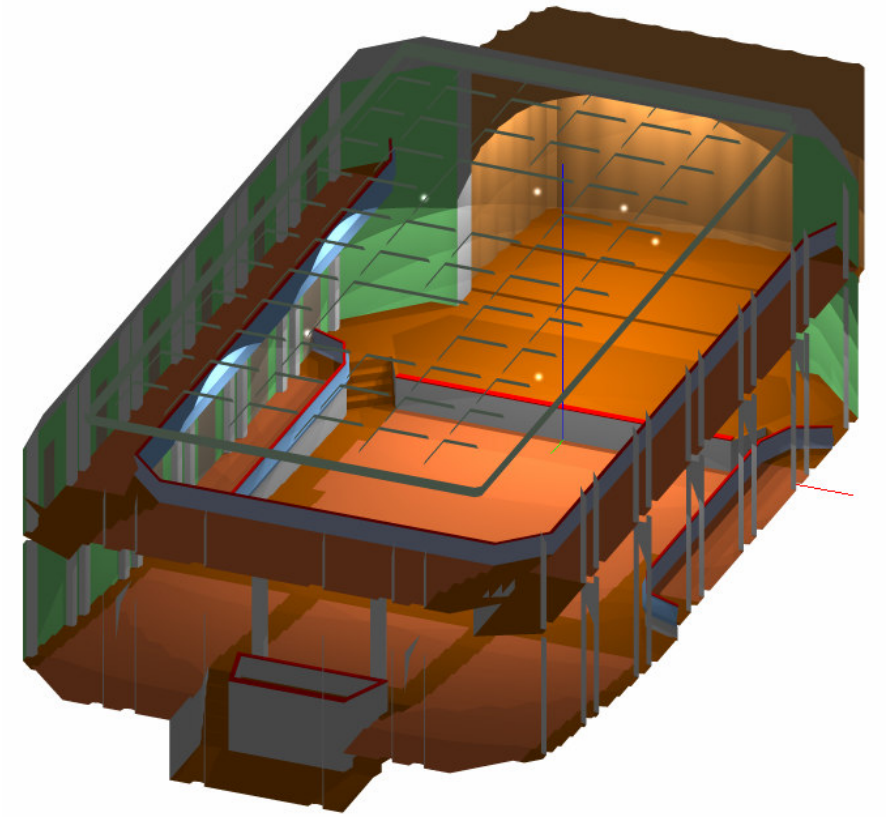

Abbildung 4.1 3D-Darstellung des Konzertsaals. Der Koordinatenursprung des Modells ist durch ein rot-grün-gelbes Koordinatenkreuz dargestellt. Er befindet sich auf der Mitte der Vorderkante der Bühne.

Die Zuschauerflächen bestehen aus dem Parkett, einem Rang und zwei länglichen Balkonen links und rechts vom Parkett. Die Bühne ist ohne Hinterbühne etwas mehr als sechs Meter breit und viereinhalb Meter tief. Dahinter schließt sich das Bühnenhaus an, welches circa sieben Meter tief und elf Meter breit ist. Links und rechts der Bühne ist diese durch eine Treppe mit dem Zuschauerraum verbunden. Die kassettenartig strukturierte Decke hat vom tiefsten Punkt des Parketts eine Höhe von ungefähr 11 m. Das Parkett verläuft zu den hinteren Reihen nach oben und erreicht am Ende des Saales einen Höhenunterschied von 2 m. Die Länge des Zuschauerraums beträgt von der Vorderkante der Bühne bis zur letzten Reihe circa 26 m. Die Oberfläche des Raumes ist in manchen Teilen sehr klein strukturiert. Die durchschnittliche Flächengröße beträgt 1,7 m^2, die durchschnittliche freie Weglänge liegt bei 8,1 m.

Die mittlere Nachhallzeit von 1,8 s liegt nach DIN 18041 für das Raumvolumen von 6426 m^3 innerhalb der für Musikdarbietungen empfohlenen Grenzen von 1,4 s bis 2,2 s. Bis auf die Überhöhung bei 250 Hz liegt der frequenzabhängige Verlauf der Nachhallzeit ebenfalls innerhalb des Tolleranzbereichs. Bei 250 Hz liegt die Nachhallzeit mit 2,3 s leicht über dem dort empfohlenen Maximum von 2,2 s.

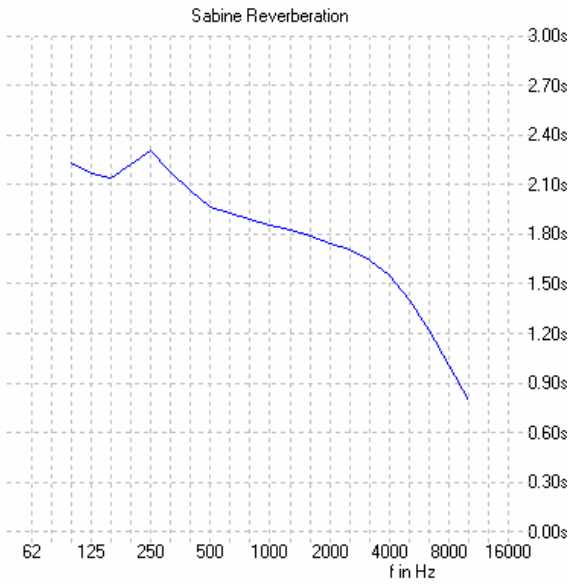

Abbildung 4.2 Verlauf der Nachhallzeit für den gewählten Konzertsaal.

Des weiteren war dem Höreindruck zufolge die Nachhallzeit für die beiden Pop/Jazz-Stücke, vor allem für das schnelle „Contredanse“ etwas zu lang und der Anstieg bei den tiefen Frequenzen zu steil. Daher wurde die Nachhallzeit mithilfe des EASE-Werkzeugs „Optimize RT“ bei den mittleren und tiefen Frequenzen herabgesetzt. Nach Anpassung der Absorptionskoeffizienten geeigneter Wandflächenmaterialien ergibt sich ein wesentlich flacherer Verlauf der Nachhallzeit (Abbildung 4.3). Der über alle Frequenzen gemittelte Durchschnitt liegt bei 1,5 s. Aus Volumen und Nachhallzeit lässt sich der Hallradius herleiten. Er beträgt im Mittel 3,7 m.

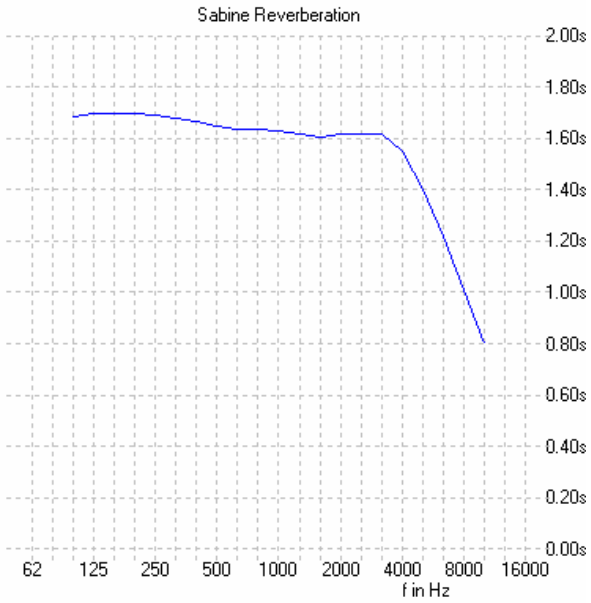

Abbildung 4.3 Verlauf der Nachhallzeit nach der Optimierung.

4.2 Hörerposition und Schallquellen – Einfluss der Richtcharakteristik auf den Hallabstand

Die Hörerposition befindet sich 3,9 m vor der Bühne auf der Symmetrieachse des Raumes (x = 0 m, y = -3,9 m). Der Abstand vom Boden entspricht der Ohrhöhe eines sitzenden Menschen und beträgt 1,30 m. Für die absolute Höhe kommen durch das Ansteigen des Parketts 5 cm dazu. Die Bühne hat eine Höhe von 1,12 m und liegt 23 cm unter der Empfängerposition. Die 0°-Achse des Hörers ist zunächst auf die Bühnenmitte ausgerichtet. Die Bestimmung der binauralen Impulsantworten für nach links und rechts abweichende Kopfdrehungen wird später mit gemessenen HRTFs außerhalb von EASE durchgeführt.

Bei der Positionierung der Schallquellen wurde von einer üblichen Konzertsituation ausgegangen. Abbildung 4.4 zeigt die Aufstellung der Instrumente für das Stück „Contredanse“ aus der Vogelperspektive.

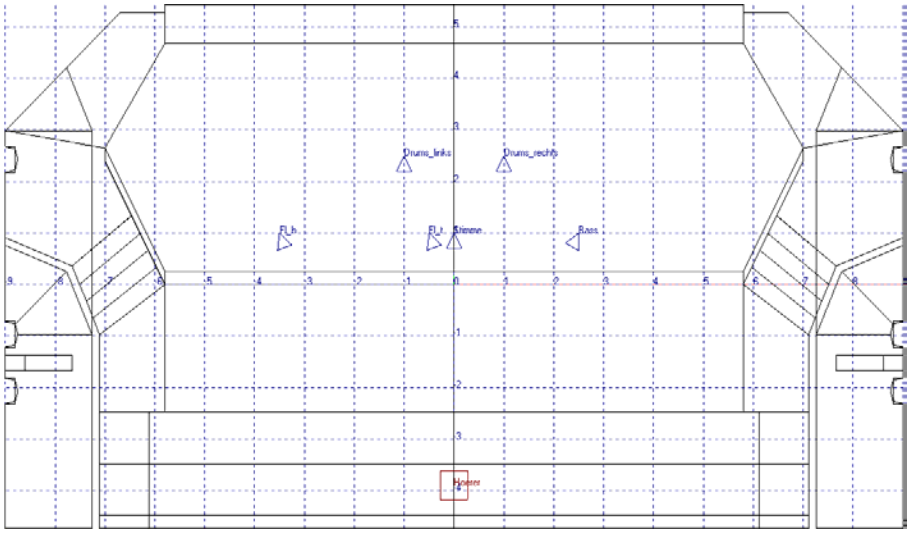

Abbildung 4.4 Aufstellung der Schallquellen für das Stück „Contredanse“. (Fl_h: Quelle für die hohen Saiten des Flügels; Fl_t: Quelle für die tiefen Saiten des Flügels)

Die Stimme befindet sich mit einem Meter Abstand zur Bühnenkante genau auf der Symmetrieachse der Bühne (x=0). Horizontal ist die Quelle auf die Hörerposition gerichtet. Die vertikale Ausrichtung wird bei allen Quellen bei 0° belassen, sodass der Spieler immer in die Horizontalebene schaut.

Der Bass steht 2,5 m rechts von der Stimme, hat den gleichen Abstand zur Bühnenkante und zeigt ebenfalls Richtung Hörer. Keines der Quellmodelle, die für die beiden Mikrofonpositionen erstellt wurden, gab in Zusammenhang mit den einzelnen Aufnahmen ein ausgewogenes Klangbild des Basses wieder. Mit der auf die 0°-Achse normalisierten Richtcharakteristik und einer Mischung aus den drei Signalen von Mikrofonen und Tonabnehmer wurden zufrieden stellende Klangergebnisse erzielt. An diesem Beispiel zeigt

sich, dass die verschiedenen klanglichen Anteile eines Instrumentes bei Aufnahmen für eine Simulation möglichst ausgewogen sein sollten. Spektrale Anteile, die in dem Signal nicht enthalten sind, können kaum „rekonstruiert“ werden.

Die beiden Schallquellen für den Flügel werden links von der Sängerin positioniert und haben einen Abstand von 3 m. Da die Mikrofonsignale nicht zwei völlig inkohärente Quellen wiedergeben, sondern ein gewisses „Übersprechen“ der hohen und tiefen Saitenanteile besteht, werden die beiden Quellen in einem größeren Abstand positioniert, als die Mikrofone bei der Aufnahme hatten. In Relation zur Hörerposition beträgt der Öffnungswinkel beider Quellen rund 30°. Für inkohärente Signale nahmen in einem Hörversuch von Lindau (2008) ab einem Öffnungswinkel von 15° fünfundneunzig Prozent der Versuchspersonen räumlich getrennte Schallquellen wahr. Es ist anzunehmen, dass bei einem doppelt so hohen Öffnungswinkel auch für Signale mit „Übersprechen“ eine räumliche Trennung wahrgenommen wird und man der relativ großen Ausdehnung des Flügels so gerecht werden kann. Die 0°-Achse des Flügels wird so ausgerichtet, dass die Öffnung des Deckels mit einer Anschrägung von 20° in Richtung Sänger auf das Publikum zeigt.

Auch das Schlagzeug wird mit zwei räumlich getrennten Schallquellen simuliert. Bei den Aufnahmen wurden zwei Overheadmikrofone aufgestellt, die links und rechts aus circa zwei Meter Höhe nach unten auf das Instrument gerichtet waren (Abbildung 4.5), sodass erstens die einzelnen Klangerzeuger des Schlagzeugs unterschiedlich stark in beiden Signalen vertreten sind und es zweitens zu Laufzeitunterschieden zwischen den Mikrofonsignalen kommt. Um die relativ große räumliche Ausbreitung des Schlagzeuges wiederzugeben, werden zwei Kugelquellenmodelle im Abstand von 2 m positioniert, um die Signale der zwei Overheadmikrofone getrennt abzustrahlen. Das entspricht einem Öffnungswinkel von 18°. Die Signale der Stützmikrofone für Bassdrum, Snare und die drei Toms werden zu den Overheadsignalen hinzugemischt. Die zwei Quellen des Schlagzeugs bekommen hier die Funktion so genannter „panning spots“. Ähnlich wie bei einer Stereomischung werden die Quellsignale über einen Panoramaregler auf einen linken und einen rechten Ort der Abstrahlung aufgeteilt, sodass auf der Achse zwischen den Lautsprechern Phantomschallquellen entstehen. Die beiden Schlagzeugquellen stehen symmetrisch hinter dem Sänger und haben einen Abstand zur Bühnenkante von 2,5 m.

Abbildung 4.5 Aufnahme des Schlagzeugs für das Stück „Contredanse".

Der Abstand der Schallquellen vom Boden wurde so gewählt, dass er ungefähr da liegt, wo sich in der Realität das akustische Zentrum der Abstrahlung des jeweiligen Instrumentes befindet (Tabelle 10). Um für die Reflektionen am Bühnenboden die Absorptions- und Diffusionseigenschaften der Musiker und Bühnenelemente zu berücksichtigen, wurden für die Bühnenoberfläche Werte gewählt, die den Eigenschaften einer mit Menschen besetzten Fläche entsprechen.

	Stimme	Bass	Flügel	Schlagzeug
Abstand vom Boden [m]	1,55	0,80	0,80	0,70

Tabelle 10 Abstand der Schallquellen vom Bühnenboden für das Stück „Contredanse".

Der Anteil des Diffusfeldes wird für das Schlagzeug am höchsten sein, da es den größten Abstand zum Hörer hat und zudem mit ungerichteten Quellen simuliert wird. Der Abstand zwischen Hörer und Instrument beträgt 6,5 m. Das entspricht dem 1,8-fachen des mittleren Hallradius. Die dichteste und zugleich am stärksten auf die Hörerposition gerichtete Quelle ist die Stimme. Der Höhenunterschied zwischen Stimmen- und Hörerposition beträgt 15°. Der Hörer sitzt also genau im Bereich des Hauptabstrahlungsgebietes, welches für die Stimme bei 0° bis 45° unterhalb der Medianachse lie gt. Vor allem die hohen Frequenzen werden relativ stark gebündelt. Mit dem statistischen Richtfaktor (Γ_{st}) lässt sich der Hallabstand nach der Formel

$$r_H = 0,057 \cdot \Gamma_{st} \cdot \sqrt{\frac{V}{T}}$$

(Meyer 1995, S. 153) (V: Raumvolumen [m^3]; T: Nachhallzeit [s]) berechnen. Bei der Oktave von 250 Hz liegt er bei 5,3 m und steigert sich bis zu 9,1 m bei 8 kHz. Die Weglänge

zwischen Stimme und Empfängerposition beträgt 5 m. Das Verhältnis zum Hallabstand liegt demnach bei einem Faktor von 0,6 bei den hohen, bis 1 bei den tiefen Frequenzen. Zum Vergleich: Für eine ungerichtete Quelle läge der Verhältnisfaktor zwischen 1 für hohe und 1,4 für tiefe Frequenzen. Zum einen folgt daraus, dass für die Stimme ein wesentlich trockener Klangeindruck zu erwarten ist als beim Schlagzeug. Zum anderen zeigt dieses Beispiel den Einfluss der Richtcharakteristik auf das Verhältnis von Direkt- und Diffusschall. Das Verhältnis von Quellabstand zu Hallabstand ist für die gerichtete Quelle bei den hohen Frequenzen circa zweimal so klein wie für eine ungerichtete Quelle. Da

$$r \sim \frac{p_{diff.}}{p_{direkt.}}$$

(r: Abstand zur Quelle; p: Schalldruck), lässt sich daraus ableiten, dass sich das Verhältnis von Diffusschallpegel zu Direktschallpegel für die Hörerposition durch die Richtcharakteristik um bis zu -6 dB von der Abstrahlung einer ungerichteten Quelle unterscheidet.

Bei dem Stück „River Man“ ist die Position der Stimme dieselbe wie bei „Contredanse“. Der Kontrabass steht 2 m rechts und die Gitarre 2 m links neben der Stimme. Beide Instrumente haben denselben Abstand zur Bühnenkante wie die Stimme und sind jeweils um 23° in Richtung Sänger eingedreht, sodass ihre 0°-Achse in der Horizontalebene auf den Hörer zeigt. Der Höhenwinkel liegt für alle Instrumente bei 0°.
Die Streicher werden in einem Halbkreis und in der gleichen Reihenfolge wie bei der Aufnahme im Studio aufgestellt. In der Stereomischung von „River Man“ wurden drei Overdubs des Streichquartetts zusammengemischt. Alle drei Overdubs werden auch in der Simulation wiedergeben. Hier besteht der Vorteil zum Stereo, dass die Einzelsignale räumlich aufgeteilt werden können. Jede Stimme bekommt sozusagen drei Spieler, die im üblichen Abstand von circa 1,5 m auf der Bühne positioniert werden (Abbildung 4.6). Im Verhältnis zum Hörer haben die Stimmen einen mittleren Abstand von 14° bis 18°. Der Öffnungswinkel zwischen dem äußersten linken und rechten Spieler beträgt 64°.

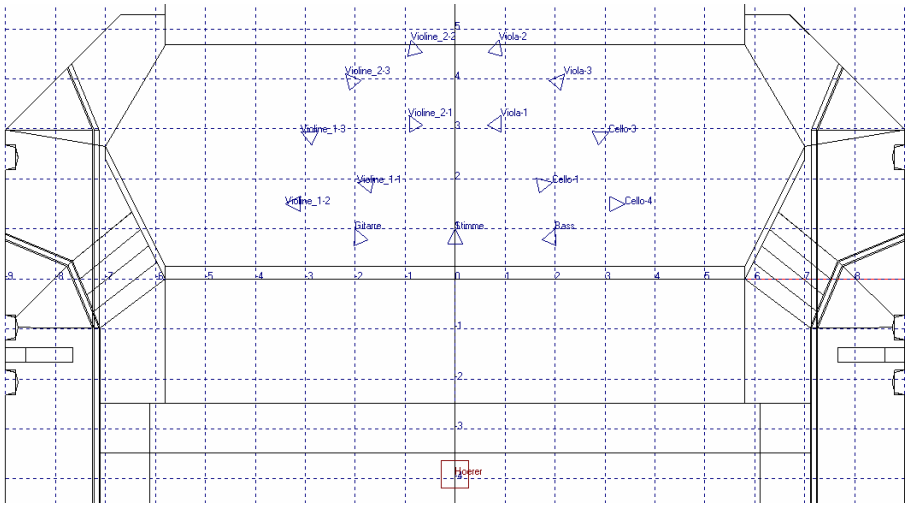

Abbildung 4.6 Aufstellung der Schallquellen für das Stück „River Man“.

Für die Höhe der Schallquellen wurde angenommen, dass die Streicher sitzen und der Gitarrenspieler steht. Es ergeben sich folgende Abstände vom Bühnenboden:

	Stimme	Gitarre	Bass	Violine/Bratsche	Cello
Abstand vom Boden [m]	1,55	1,10	0,80	1,10	0,40

Tabelle 11 Abstand der Schallquellen vom Bühnenboden für das Stück „River Man“.

Durch den Höhenunterschied zwischen Bühne und Empfängerposition erreicht den Hörer der Direktschall, der von den Quellen etwas unterhalb der Horizontalebene abgestrahlt wird. Zur Gitarre beträgt der Höhenunterschied circa 10°, zur Stimme 15° und zum Bass 5°. In Abbildung 4.7 ist das Richtwirkungsmaß auf die Hörerposition für diese drei Instrumente als Funktion der Frequenz dargestellt. Das Verhältnis von Direktschallpegel zu Diffusschallpegel in der Entfernung r von einer gerichteten Schallquelle berechnet sich nach der Formel:

$$L_{direkt} - L_{diffus} = 20\log\left(\Gamma_{st}\right) + 10\log\left(0{,}163\frac{V}{T}\right) - 20\log\left(\frac{R}{1m}\right) - 17dB \qquad (1.16)$$

(Weinzierl 2008, S. 502) ($20\log\left(\Gamma_{st}\right)$: Richtwirkungsmaß). Aus der Proportionalität

$$20\log\left(\Gamma_{st}\right) \sim L_{direkt} - L_{diffus}$$

folgt, dass das Richtwirkungsmaß angibt, um wie viel dB der Direktschall im Verhältnis zum Diffusschall lauter ist als bei einer ungerichteten Quelle. Der Kontrabass strahlt insgesamt weniger gerichtet auf den Hörer ab als Gitarre und Stimme. Das Richtwirkungsmaß bewegt sich zwischen -2 dB und 4 dB. Bei der Stimme sind wesentlich größere Schwankungen vorhanden und die Bündelung steigt bei den hohen Frequenzen relativ stark an. Die Gitarre hat bei den tiefen bis mittleren Frequenzen ein Richtwirkungsmaß von 2 dB bis 6 dB. Ab

circa 2 kHz ist die Richtwirkung auf den Hörer nicht so stark ausgeprägt. Sie bewegt sich hier zwischen 0 und 4 dB.

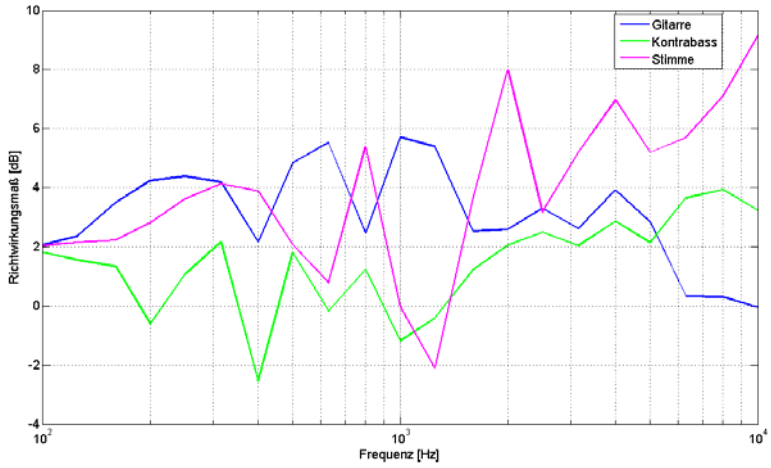

Abbildung 4.7 Richtwirkungsmaß bei der Abstrahlung auf den Hörer für drei Instrumente.

Um das absolute Verhältnis von Direkt- und Diffusschallpegel an der Hörerposition zu berechnen, spielen nach (1.16) die frequenzabhängige Nachhallzeit und der Abstand eine entscheidende Rolle. Für die gegebene Szene aus Schallquellen und Empfänger ist das für den Raumeindruck entscheidende Verhältnis zwischen Diffus- und Direktschall in Abbildung 4.8 wiedergegeben. Der Abstand der drei Instrumente zum Hörer ist ähnlich (Stimme: 5m; Gitarre: 5,36 m; Bass: 5,32 m). Bei 0 dB entspricht die Entfernung zwischen Quelle und Hörer dem Hallabstand. Für die Stimme ist bei den hohen Frequenzen ein eher trockener Klangeindruck zu erwarten. Für die Gitarre ist der Anteil von Diffus- und Direktschall über das gesamte Spektrum ungefähr gleich stark. Beim Bass ist in den tiefen und mittleren Frequenzen ein etwas „halligerer“ Klangeindruck zu erwarten als bei der Gitarre.

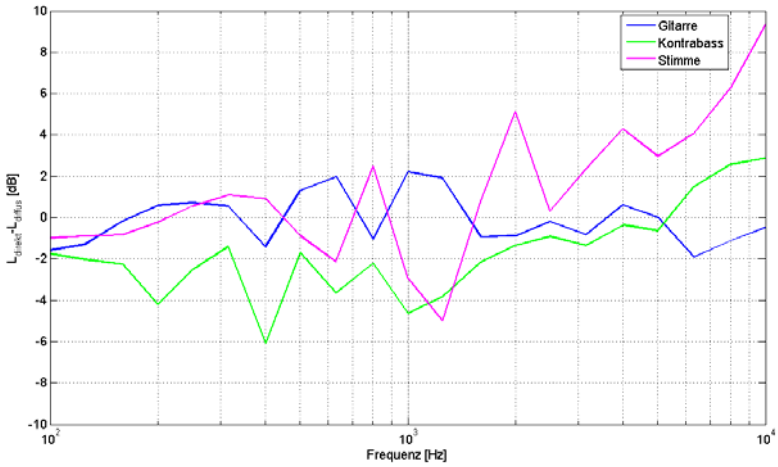

Abbildung 4.8 Verhältnis von Direkt- und Diffusschall an der Hörerposition.

Für die Streicher wird sich insgesamt ein diffuserer Klangeindruck ergeben, da diese Quellen den Schall weniger gerichtet abgeben und zudem durch die Variierung der Ausrichtung gleicher Instrumententypen in unterschiedliche Richtungen abstrahlen. Weiterhin stehen die Streicher weiter weg. Die hintersten Positionen von Violine und Bratsche haben einen 1,8-mal größeren Abstand zum Hörer als die Stimme. Das entspricht dem 2,4-fachen des frequenzgemittelten Hallradius.

4.3 Rendering

4.3.1 Schallfeldberechnung in EASE

Alle Schallquellen werden pro Terzband mit 1 W angeregt. Die anregende Leistung spielt für die resultierenden Lautstärkeverhältnisse der Quellen keine Rolle, da das Verhältnis der tatsächlich zugeführten Leistung durch die Pegelanpassung in der Mischung reguliert wird. Die Einstellung hat jedoch einen Einfluss auf den Dynamikumfang der resultierenden Reflektogramme, welche zur Erstellung der dynamischen BRIRs aus EASE exportiert werden. Diese Reflektogramme enthalten in Form einer Textdatei Informationen über die Ankunftszeit, die Terzbandpegel und die Richtung jeder einzelnen Reflektion für eine bestimmte Empfängerposition und -ausrichtung. 0 dB entspricht der simulierten Hörschwelle. Schalldruckpegelwerte unter 0 dB werden nicht exportiert. Der Nachhall bricht bei diesem Wert sozusagen ab. Das heißt, der Dynamikumfang ergibt sich aus dem Pegelwert des lautesten an der Hörerposition ankommenden Signals, welches der Direktschall ist. Bei einem Kennschalldruck von 100 dB pro Terzband und einer anregenden Leistung von einem Watt ergibt sich für die Quellen mit der größten Entfernung von der Hörerposition (ca. 7 m) ein Dynamikumfang von 83 dB pro Terzband. Für alle Terzbänder zusammen ergibt sich ein Dynamikumfang von 96 dB.

Die Berechnung der Reflektogramme wird mit dem EASE-Modul *AURA Response* durchgeführt. Dieses Modul verwendet die von der RWTH Aachen entwickelten CAESAR-Algorithmen. Sie basieren auf einer Kombination aus Spiegelquellen-Methode und Raytracingverfahren. Der Vorteil der CAESAR-Algorithmen liegt in der Verringerung der erforderlichen Berechnungszeit gegenüber den EASE-Algorithmen, welche unter anderem durch die Ausnutzung der Mehrkernausstattung handelsüblicher PCs erreicht wird. Dadurch ist es möglich Reflektogramme über die gesamte Dauer des Nachhalls in einer annehmbaren Zeit zu berechnen.

Zunächst werden in Aura Response die Quellen und Empfänger festgelegt, für die das Schallfeld berechnet werden soll. Für jedes Musikstück werden alle Quellen in einem Durchgang berechnet. Für alle Stücke wird die gleiche Hörerposition verwendet. Aura Response ermöglicht das Einstellen bestimmter Berechnungsparameter:

- *Particles* legt die Anzahl der Strahlen[1] pro Quelle fest, welche zur Berechnung des Schallfeldes verfolgt werden. Je größer die Anzahl, desto dichter und genauer wird das berechnete Reflektogramm. In Abhängigkeit von der mittleren freien Pfadlänge des Raumes werden nach Angabe einer Genauigkeitsstufe durch ein Auswahlmenü verschiedene Größen der Strahlenanzahl vorgeschlagen. Zum Zweck einer Auralisation wird die Option „high resolution“ oder „very high resolution“ empfohlen (Ahnert 2009). Bei der Berechnung *einer* Quelle ergibt die Option „very high resolution“ 7371000 Strahlen. Werden mehrere Quellen zugleich berechnet, muss diese Anzahl per Hand eingegeben werden (Option: „user-defined“), da sich die Option der Genauigkeitsstufe auf die Empfängerseite bezieht. Das bedeutet, dass die Auflösung für das Rendering einer Quelle durch die Anzahl der Quellen geteilt wird und dementsprechend weniger Strahlen pro Quelle vorgeschlagen werden. Nach der Betrachtung des resultierenden Reflektogramms und einer Hörprobe ergab sich, dass die vorgeschlagene Partikelanzahl noch nicht ausreicht, um den späten Anteil des Nachhalls artefaktefrei zu simulieren. In Abbildung 4.9 ist das an der Abnahme der Dichte und kleineren Spikes ab circa 1,5 s zu erkennen. Daher wird der Wert auf 10 Millionen erhöht.
- *Length* bestimmt die Länge des Reflektogramms. Diese sollte im Bereich der statistischen Nachhallzeit liegen. Hier gibt es ebenfalls ein Auswahlmenü, das die vorgeschlagene Länge des Reflektogramms in adaptiven Genauigkeitsklassen anhand der statistischen Eyring-Nachhallzeit festlegt (ebd.). Nach Blickensdorff (2010) wird die Option „Long“ mit einem Aufschlag von 30 % empfohlen. Im Gegensatz zu Blickensdorff wird hier das Reflektogramm zur Erzeugung der binauralen Impulsantworten in Form einer Textdatei aus EASE exportiert, was die bereits erwähnte Beschränkung des Dynamikumfangs zur Folge hat. Die effektive Länge des Reflektogramms wird unmittelbar durch den Dynamikumfang eingeschränkt. Die Option „Long“ ergibt eine Antwortlänge von 1990 ms. Bei plus 30% ergeben sich für die empfohlene Länge des Reflektogramms 2587 ms. Nach der statistischen Vorausberechnung mit der Sabine-Formel liegt die längste Nachhallzeit von 1,69 s zwischen 125 und 250 Hz (siehe Abbildung 4.3). Ein Blick auf das Reflektogramm zeigt, dass der Schall in diesem Frequenzbereich nach circa 1900 ms

[1] in EASE auch als Partikel bezeichnet

unter 0 dB fällt (Abbildung 4.9). Daher ist eine Festlegung der Antwortlänge auf 2000 ms ausreichend.

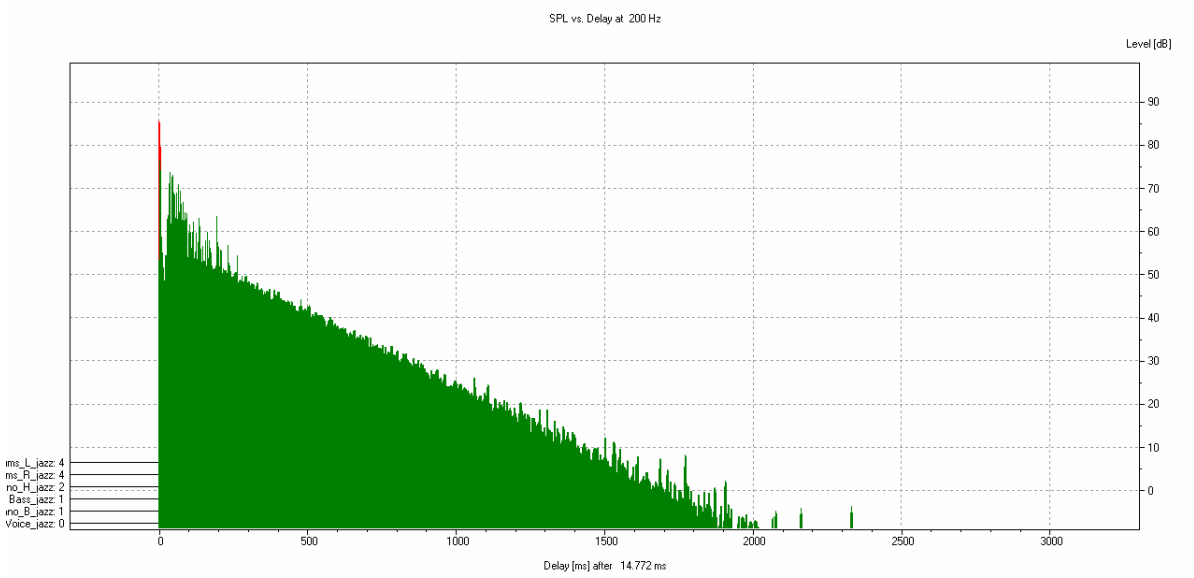

Abbildung 4.9 Reflektogramm aller Schallquellen von „Contredanse“ in der Form Schalldruckpegel als Funktion der Zeit. (Partikelanzahl: 7371000)

- Die Option *Scattering* ermöglicht es einen breitbandigen Streugrad für Wandoberflächen zu bestimmen, für die kein Streugrad definiert ist. In dem vorhandenen Raummodell ist dieser bei einigen Flächen definiert, bei anderen nicht. Da das Raummodell nur in vereinfachter Form die zur Streuung beitragenden Details eines Raumes enthält, ist das Hinzufügen einer gewissen Streuung von Vorteil. Nach Erfahrung von Mitarbeitern des Fachgebiets AK hat sich ein Streugrad von 60% bewährt.
- Aura berechnet das Schallfeld in zwei Teilen: Für frühe Reflektionsordnungen wird eine Spiegelquellen-Methode angewendet, die mit einem deterministischen ray-tracing-Algorithmus den Raum nach sichtbaren Spiegelquellen absucht, sodass die frühen Reflektionen aus einer ungleichmäßigen Verteilung der Abstrahlungsrichtung berechnet werden. Ab einer Ordnungszahl, die über den Parameter *Cut Off Order* festgelegt wird, wird nach einem statistischen Verfahren der Verlauf von in alle Richtungen gleich verteilten Schallstrahlen berechnet. Es können Werte zwischen 1 und 30 eingegeben werden. Die Bestimmung folgt nach der Abschätzung, ob starke einzelne Echos oder Flatterechos auftreten können, was vor allem bei einem niedrigen Streugrad vorkommt. Ist die Wahrscheinlichkeit von Flatterechos hoch, sollte ein großer Wert bestimmt werden. Ansonsten sind mittlere Werte hinreichend. In Anlehnung an Blickensdorff (2010) wird ein Wert von 16 gewählt.
- Die Parameter *Density Factor* und *Tail Resolution* bestimmen die Auflösung des späten Nachhallanteils. *Density Factor* gibt die Anzahl von Impulsen pro ms wieder

und *Tail Resolution* bestimmt die Zeitfensterbreite zur Bestimmung der Energieverteilung. Hier werden die Werte 16 und 5 ms verwendet.

- Mit der Option *Max. Diameter after 1s* wird die verwendete Anzahl der Strahlen für den ersten Teil der Berechnung (die Spiegelquellen-Methode) bestimmt. Dieser Wert ist maßgebend für die Auflösung der frühen Reflektionen. Bei einem Raum mit vielen unterschiedlichen Wandflächen sollte eher ein kleiner Wert verwendet werden. Als Richtwert wird der Durchmesser der kleinsten Wandfläche empfohlen (Renkus-Heinz Inc. 2009, S. 249). Der Wert sollte jedoch nicht unter 1 m liegen. Der Durchmesser der kleinsten Fläche im vorhandenen Modell beträgt 0,14 m.
- Die Option *Diffuse Rain* wird nicht angewählt, da von dieser Methode keine Vorteile erwartet werden (Blickensdorff 2010, S. 69).

In Tabelle 12 sind die für AURA-Response verwendeten Parameterwerte noch einmal zusammengefasst.

Particles:	10 Mio.
Length [ms]:	2000
Scattering [%]:	60
Cut Off Order:	16
Density Factor:	16
Tail Resolution [ms]:	5
Max. Diameter after 1s [m]:	1
Diffuse Rain:	uncheck

Tabelle 12 Liste der verwendeten Parameter für AURA Response.

Abhängig von der Anzahl der Quellen ergeben sich unterschiedlich lange Berechnungszeiten für die beiden Aufführungsszenen. Mit einem 4-Kern-Rechner liegen sie in der Größenordnung von 4 Stunden pro Quelle.

Nach der Berechnung mit AURA-Response wird die gerenderte Datei in das EASE-Modul *Probe* geladen. Mit diesem Programm kann die erstellte Raumantwort grafisch dargestellt und mit unterschiedlichen Analysemethoden untersucht werden. Das Verfahren Schroeder-RT ermittelt durch eine Rückintegrationsmethode die Nachhallzeit des Reflektogramms. Diese Methode ist genauer als die statistischen Berechnungen nach Eyring oder Sabine (Renkus-Heinz Inc. 2009, S. 64). Wie man in Abbildung 4.10 sehen kann, liegt die Schroeder-RT im Mittel circa 100 ms unter der nach Sabine berechneten Nachhallzeit.

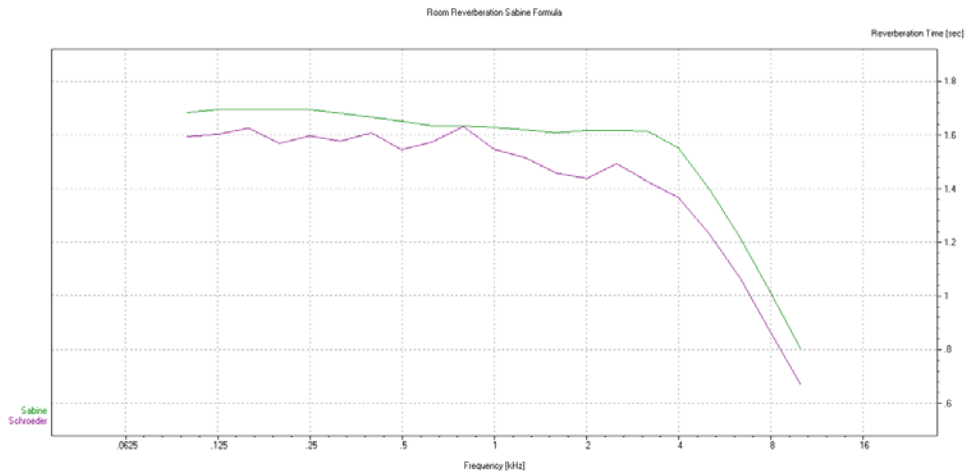

Abbildung 4.10 Vergleich der Nachhallzeitberechnung nach Schroeder und Sabine.

Probe ermöglicht es die gesamten Reflektionen einzelner Quellen zu deaktivieren, und über die Funktion „save as post file“ kann eine Textdatei des Reflektogramms gespeichert werden. Auf diese Weise wird für jede Quelle eine Raumantwort exportiert.

4.3.2 Erzeugung der dynamischen BRIRs

Die Konvertierung der Reflektogramme in dynamische BRIRs erfolgt mit vom Fachgebiet Audiokommunikation erstellten MATLAB-Skripten. Zunächst wird aus den Terzbandpegeln jeder einzelnen Reflektion ein FIR-Filter generiert und als Impulsantwort gespeichert[1]. Innerhalb der Bandbreite der Reflektogramme (100 Hz – 10 kHz) wird dazu ein kontinuierliches Spektrum aus den Terzbandpegeln interpoliert. Außerhalb dieser Bandgrenzen wird das Spektrum aus den Werten des tiefsten und höchsten Terzbandes extrapoliert. Bei den tiefen Frequenzen wird ab 20 Hz eine Flanke von -24 dB pro Oktave in Richtung tiefere Frequenzen angesetzt. Die hohen Frequenzen werden ab 10 kHz mit einem Pegelabfall von -3 dB pro Oktave extrapoliert.

Anschließend wird für jede Reflektion die Richtung des Eintreffens aus dem Reflektogramm extrahiert. Je nach Kopfdrehung wird der Azimutwinkel entsprechend korrigiert. Mit dieser Information wird aus einer Bibliothek von gemessenen HRTFs die entsprechende binaurale Impulsantwort gewählt und mit der Reflektion gefaltet[2]. Die HRTFs stammen aus einer Messung von Moldrzyk (2004). Dort wurden in einem RAR 52000 Übertragungsfunktionen eines Kunstkopfes gemessen (Winkelabstand horizontal: 0,5°, vertikal 5°). Die binauralen Impulsantworten der einzelnen Reflektionen werden schließlich ihrer Ankunftszeit entsprechend zu einer vollständigen BRIR zusammenaddiert. Dieser Prozess wird pro 1°

[1] MATLAB Skript erstellt von Frank Schultz, TU Berlin, Fachgebiet AK (2010)
[2] MATLAB Skript erstellt von Alexander Lindau, TU Berlin, Fachgebiet AK (2010)

einer Kopfdrehung durchgeführt. Bei einer Begrenzung der Kopfbewegung auf ± 70° in der Horizontalebene ergeben sich 141 dynamische BRIRs pro Quelle.

Das diffuse Schallfeld eines Raumes hat die Eigenschaft, dass die Richtungen der eintreffenden Reflektionen annähernd gleich verteilt sind. Aus diesem Grund ist die Wahrnehmung des stochastischen Nachhalls einer Raumimpulsantwort ab einer gewissen Zeit nach Eintreffen des Direktsignals unabhängig von der Kopfdrehung. Dieser Zeitpunkt beschreibt physikalisch den Übergang des Abschnitts der frühen Reflektionen in den diffusen Anteil und wird als mixing time (t_m) bezeichnet. Der Faltungsalgorithmus Fwonder, welcher zur Wiedergabe der dynamischen Binauralsynthese verwendet wird, ist in der Lage einen von der Kopfdrehung abhängigen Faltungsprozess für die frühen Reflektionen mit einer statischen Faltung des stochastischen Nachhalls zu kombinieren. Auf diese Weise wird es in Bezug auf die Rechenleistung möglich, eine hohe Anzahl an Quellen gleichzeitig zu synthetisieren. Es wird daher nur für die 0°-Ausrichtung des Kopfes eine vollständige BRIR berechnet. Bis zu welchem Zeitpunkt der frühe Anteil der Impulsantworten in Abhängigkeit von der Kopfdrehung gerendert werden muss, ist von der wahrgenommenen mixing time (t_{mp}) abhängig. Mittels Hörversuchen wurden unterschiedliche Abschätzungsverfahren für die t_{mp} ermittelt (Lindau et al. 2010). Das 95%-Perzentil der t_{mp} kann mit folgender Gleichung abgeschätzt werden:

$$t_{mp95\%} = 0,0117 \cdot V + 50,1 \qquad [ms]$$

(V: Volumen des Raumes). Für den hier simulierten Konzertsaal ergeben sich 125 ms.

4.4 Binaurale Mischung

Die Mischung der einzelnen Zuspielsignale erfolgt mithilfe des Multitrackeditors Logic Audio und wird als statische Simulation durchgeführt[1]. Zur Faltung der trockenen Musiksignale mit den BRIRs wird das Plugin[2] Space Designer als Inserteffekt in jede Einzelspur geladen. Bei der Kombination von Zuspielsignalen für eine Quelle werden diese auf eine Gruppe zusammengemischt und das Faltungsplugin als Gruppeninsert eingefügt. Die Mischung wird mit einem Kopfhörer Stax SRM 252-II[3] durchgeführt. Die Eigenschaften des Übertragungsweges vom Kopfhörer zum Trommelfell sind in den BRIRs bereits enthalten. Daher muss ein Korrekturfilter angewendet werden, das die Übertragungsfunktion des Gehörgangs und des Kopfhörers aus dem Signal entfernt (Blauert 1997). Die Übertragungsfunktion zwischen Stax Kopfhörer und Trommelfell wurde mithilfe eines Kunstkopfes vom Fachgebiet AK gemessen und als invertierter Frequenzgang in Form einer

[1] BRIRs für die 0°-Ausrichtung des Hörers
[2] Gängige Bezeichnung für einen Audioprozessor innerhalb eines Multitrackeditors
[3] mit Vorverstärker Stax SR 202

Impulsantwort zur Verfügung gestellt. Die Mischung wird von einem Tonmeister entsprechend seinen klangästhetischen Erwartungen von einer realen Konzertaufführung realisiert. Die Einstellung der trockenen Einzel- und Gruppensignale wird schließlich auf die Anzahl der Quellen zusammengemischt und als Multikanal-Audiodatei exportiert. Bei der dynamischen Simulation mit Fwonder kann diese Multikanal-Datei in das als Zuspielsoftware fungierende Programm Pure Data geladen und durch Fwonder mit den dynamischen BRIRs in Echtzeit gefaltet werden.

4.5 Diskussion

Wie bereits gezeigt wurde, bringen die Richtcharakteristiken eine Frequenzabhängigkeit des Diffus- und Direktschalldruckverhältnisses in die Simulation ein (vgl. Abbildung 4.8). Bei Auswertung des Schallfeldes mit ungerichteten Quellen wird eine Frequenzabhängigkeit nur dadurch charakterisiert, dass das Diffusfeld ab 4 kHz zu den höheren Frequenzen kontinuierlich abnimmt. Die tatsächlich durch die Richtcharakteristik hervorgerufene Klangfarbenveränderung ergibt sich an der Hörerposition durch die Überlagerung des Diffus- und Direktfeldfrequenzgangs. In Abbildung 4.11 ist der absolute Schalldruckpegel für die Stimme (gerichtet/ungerichtet) an der Hörerposition dargestellt. Dieser Frequenzgang stellt die Übertragungsfunktion der aus Raum und Quelle bestehenden akustischen Umgebung dar. Er hat Ähnlichkeit mit dem Leistungsfrequenzgang der Richtcharakteristik (vgl. graue Kurve in Abbildung 3.47). Die typische Abnahme des Bündelungsgrades der Stimme im Bereich von 1 kHz macht sich durch eine Zunahme des Diffusfeldschalldrucks bei diesen Frequenzen bemerkbar.

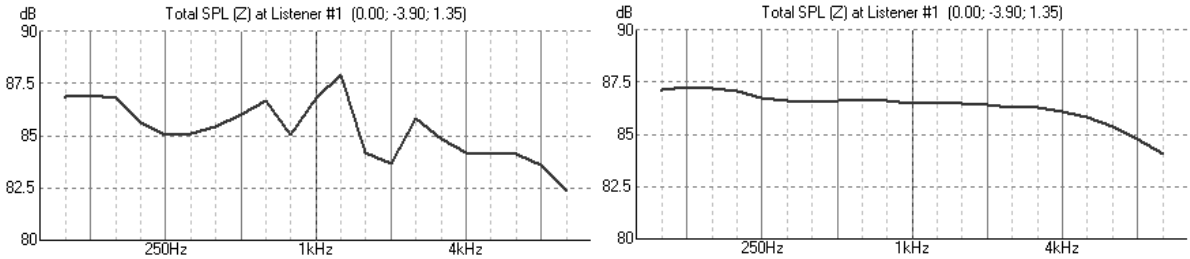

Abbildung 4.11 Simulierter Schalldruckpegel an der Hörerposition für die Stimmenquelle. Links: gerichtet; Rechts: ungerichtet.

Interessant ist auch die Darstellung der vertikalen Reflektionen als Funktion des Eintreffwinkels. In Abbildung 4.12 sieht man, dass zwischen 30° und 40° unterhalb der Horizontalebene und im Bereich von 75° oberhalb der Horizontalebene mehrere starke Reflektionen eintreffen, deren kurze Laufzeiten darauf schließen lassen, dass sie erste

Reflektionen von Decke und Boden sind. Im linken Bild ist die Raumantwort für die gerichteten Quellen dargestellt und im rechten Bild wurde das Schallfeld zum Vergleich mit ungerichteten Quellen berechnet. Bei der Simulation mit gerichteten Quellen treten die frühen Reflektionen für einzelne Instrumente stärker hervor als bei ungerichteten Quellen. Bei 1 kHz ist die erste Reflektion vom Boden beispielsweise für den Bass am intensivsten, während von der Decke der Flügel am stärksten zurückkommt. Im Gegensatz zu ungerichteten Quellen variiert bei gerichteten Quellen das Verhältnis der frühen Reflektionen mit der Frequenz.

Sehr markant ist die Auswirkung der Richtcharakteristik auf die frühen Reflektionen bei der Tuba. Durch ihre starke Bündelung des Schalls nach oben kommt es zu starken Deckenreflexionen, die bei einer ungerichteten Quelle viel weniger intensiv auftreten. Für eine übliche Aufstellung der Tuba bei einer Orchesteraufführung kommt es bei der Simulation einer gerichteten Quelle zu Deckenreflexionen, die bis zu 7 dB lauter als der Direktschall sind. Wegen ihrer kurzen Laufzeiten von 30 bzw. 35 ms rufen sie noch keinen Echoeffekt hervor. Bei einer ungerichteten simulierten Quelle sind die gleichen Reflektionen 8 dB leiser als das Direktsignal. In Hörtests konnte nachgewiesen werden, dass frühe, vertikale Reflektionen Einfluss auf die wahrgenommene Klangfarbe haben (Olive und Toole 1989).

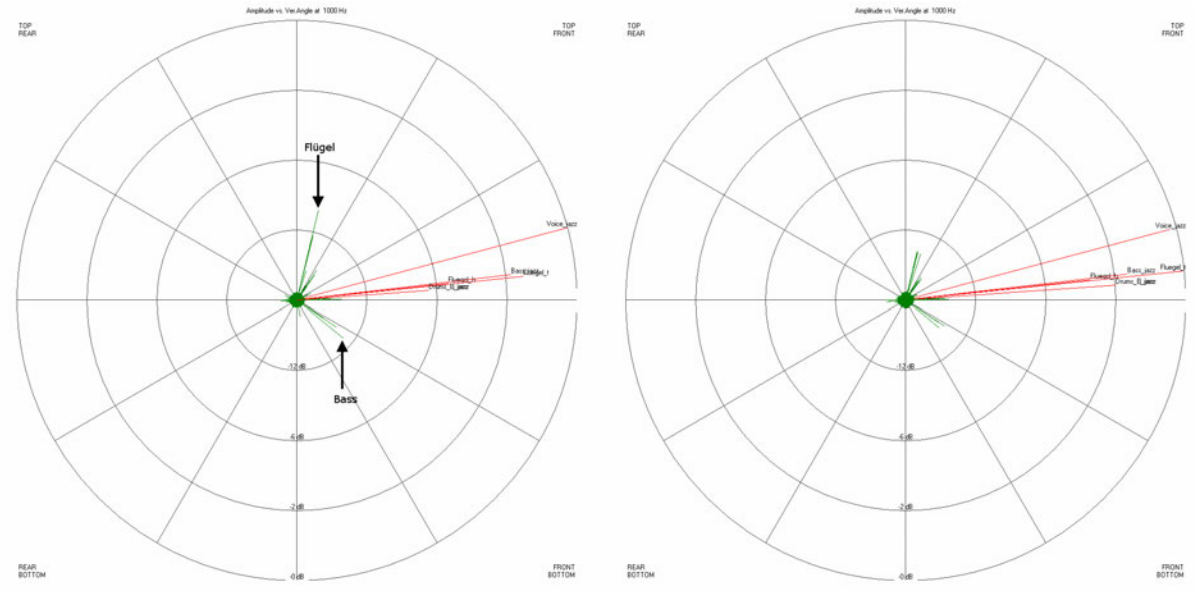

Abbildung 4.12 Darstellung der Richtung und Intensität der vertikalen Reflektionen für „Contredanse“ bei dem Terzband von 1 kHz. Rot: Direktschall; Grün: Reflektionen; Links: Gerichtete Schallquellen; Rechts: gleiche Aufstellung mit ungerichteten Schallquellen.

Das Klarheitsmaß für Musik (im Folgenden: C_{80}) gibt Auskunft über die Durchsichtigkeit des ankommenden Schallereignisses. Dazu trägt nicht nur das Direktsignal sondern auch die Energie der frühen Reflektionen bei. Es berechnet sich durch das Verhältnis der

Schallenergie in den ersten 80 ms nach dem Eintreffen des Direktschalls zur Energie, die nach diesem Zeitabschnitt beim Hörer eintrifft. Durch die stärkere Richtwirkung erhöht sich das frequenzgemittelte C_{80} von 4 dB bei einer ungerichteten Quelle auf 7,5 dB. Bei den mittleren Frequenzen gibt es Abweichungen bis 6 dB und bei den hohen Frequenzen bis 9 dB.

Die wahrgenommene Ausdehnung einer Quelle wird durch den Anteil seitlicher Reflektionen beeinflusst (Barron 1971). Die Richtung einer Quelle wird durch die erste Wellenfront bestimmt. Frühe seitliche Reflektionen fügen eine gewisse Unschärfe in die Lokalisation ein. Zur Abschätzung der wahrgenommenen Ausdehnung einer Schallquelle wurde der akustische Parameter *lateral fraction* (im Folgenden: LF) eingeführt (DIN EN ISO 3382). Er ist ein in Prozent ausgedrücktes Maß für das Verhältnis von seitlich eintreffender Schallenergie zur gesamten Schallenergie für die Zeit von 5 ms bis 80 ms nach dem Eintreffen des Direktschalls. Abbildung 4.13 zeigt die LF für das Schallfeld der Sängerquelle für ein gerichtetes und für ein ungerichtetes Quellmodell. Für die ungerichtete Quelle ist die LF weitestgehend unabhängig von der Frequenz, während sie für die gerichtete Stimme größere Abweichung aufweist. Im Mittel über alle Frequenzen beträgt der Energieanteil des seitlich eintreffenden Schalls für die gerichtete Quelle 9,5 % und für die ungerichtete Quelle 14 %. Man kann vermuten, dass die gerichtete Quelle insgesamt etwas schmaler in ihrer Ausdehnung wahrgenommen wird als die ungerichtete. Der Anstieg der Seitenreflexionen bei 1 kHz und 1,25 kHz ist auf die für die Stimme charakteristische Zunahme der seitlichen Abstrahlung bei diesen Frequenzen zurückzuführen (vgl. Abschnitt 3.6 sowie Abbildung 3.45 und Abbildung 3.48).

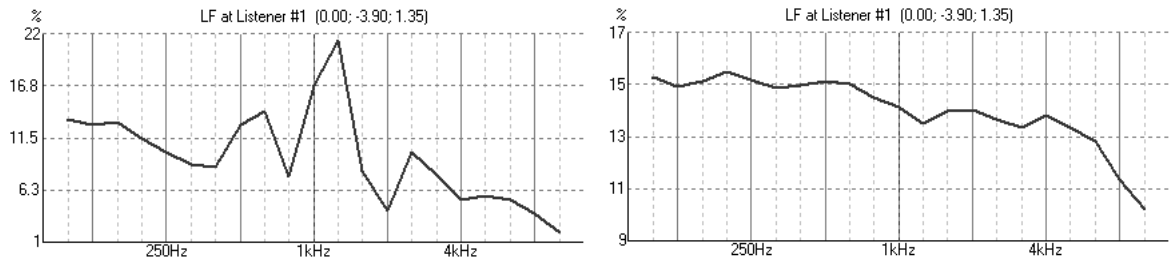

Abbildung 4.13 Anteil der frühen seitlichen Reflektionen (lateral fraction) für die Sängerposition mit korrekter Richtcharakteristik (links) und für ein ungerichtetes Quellmodell. (rechts).

Analysiert man das Schallfeld aller Quellen zusammen, ergeben sich kleinere Einflüsse der Richtcharakteristiken auf die raumakustischen Parameter als bei der Betrachtung der einzelnen Quellen. Das liegt an der Tatsache, dass sich die Eigenschaften der Richtcharakteristiken bei einer Auswertung mit einem „Messsignal“ zum Teil gegenseitig aufheben. In der Realität entsteht die Musik aber aus räumlich verteilten Klangquellen, deren Rhythmus und individuelle Klangfarbe dem Hörer eine Wahrnehmung von differenzierten

Klanggestalten ermöglicht. Für jedes einzelne Instrument ergeben sich deutlich individuelle Unterschiede in den räumlichen Parametern, die wesentlich stärker von der Frequenz abhängen als bei ungerichteten Quellen. Insbesondere machen sich die Richtcharakteristiken im Frequenzgang von Diffus- und Direktfeldschalldruck bemerkbar. Die Übertragung der typischen Frequenzabhängigkeit des Diffus/Direktverhältnisses eines Instrumentes auf die Quellmodelle lässt erwarten, dass daraus eine verzerrungsfreiere Wiedergabe der simulierten Musikinstrumente möglich wird, als es mit ungerichteten Quellen der Fall wäre. Es kann vermutet werden, dass dadurch der Hörerfahrung des Rezipienten über das klangliche Verhalten von Musikinstrumenten in einem Raum eher entsprochen wird.

5 Zusammenfassung

In der vorliegenden Arbeit wurde die Simulation einer Konzertaufführung am Beispiel von zwei Musikstücken durchgeführt. Dabei wurden Quellmodelle generiert, deren Richtcharakteristik eine Auralisation von Musikinstrumenten ermöglicht, die dem räumlichen Klangeindruck eines realen Musikensembles näher kommt als eine Auralisation mit ungerichteten Quellen.

In Abschnitt zwei wurde das nachhallfreie Quellmaterial vorgestellt und eine in Auftrag gegebene Produktion von Musikaufnahmen beschrieben. Informationen über die Besetzung und die Eignung der einzelnen Spuren für eine Auralisation wurden dabei gesammelt.

Im dritten Teil wurden aus Messungen der TU Berlin zur Abstrahlung von Musikinstrumenten Quellmodelle generiert, die in eine akustische Simulationssoftware importiert werden können. Die Messung fand mit einem 32-Mikrofonkugelarray statt. Um die Daten der 32 Messpunkte in ein feinmaschiges Netz aus Kugelkoordinaten zu interpolieren, wurden zwei Methoden vorgestellt und ausgewertet. Die Interpolation in den Koordinaten der Messdaten mit einer anschließenden Abbildung der Werte in das Koordinatensystem für Lautsprechermodelle lieferte dabei die exakteren Ergebnisse. Von 19 der gemessenen Musikinstrumente wurden Quellmodelle erstellt und im EASE-Format auf DVD beigefügt (siehe Anhang A:). Das Modell des Flügels wurde aus den Daten der PTB in das EASE-Format konvertiert und befindet sich ebenfalls auf der DVD. Die Richtcharakteristiken der Blechbläser und der Stimme wurden auf die Hauptabstrahlrichtung normalisiert. Für die Instrumente bei denen keine eindeutige Hauptabstrahlrichtung existiert wurde für die Bezugsrichtung die übliche Aufstellung des Mikrofons bei einer Aufnahme des entsprechenden Instrumentes pauschal abgeschätzt. Zudem wird auf der DVD ein Matlab-Programm zur Verfügung gestellt, mit dem die Richtcharakteristiken auf eine beliebige Bezugsrichtung normiert werden können. Für die zwei Musikstücke, die in Abschnitt zwei vorgestellt wurden, wurden die Mikrofonpositionen in Abschnitt 2 dokumentiert und für eine korrekte Normierung der Richtcharakteristiken in Abschnitt 3 ausgewertet. Am Beispiel der beiden Aufnahmen wurden verschiedene Mikrofonpositionen diskutiert. Es wurde festgestellt, dass Mikrofonpositionen, die das Klangbild eines Instrumentes zu einseitig aufnehmen, für eine Simulation problematisch sein können. Weiterhin wurde der Einfluss der Normierung auf den Frequenzgang der Quellmodelle untersucht, und man kann schlussfolgern, dass eine inkorrekte Bezugsrichtung zu einem schlechteren Ergebnis führen kann als die Auralisation mit einer ungerichteten Abstrahlcharakteristik. Das Resultat der in Teil drei beschriebenen Verfahrensweise zur Simulation von Richtcharakteristiken wurde abschließend mit Modellen der PTB und der TKK verglichen.

Bei der Auralisation der Musikinstrumente in Teil vier konnte durch eine Analyse der gerenderten Impulsantworten festgestellt werden, welchen Einfluss die Richtcharakteristik auf das simulierte Schallfeld hat. Anhand von Frequenzgängen und raumakustischen Parametern wurden Voraussagen über die klangliche Auswirkung der Richtcharakteristiken gemacht. Der Anteil früher Reflektionen wird individuell durch die Richtcharakteristik einer Quelle verändert. Dadurch wird ihre wahrgenommene Ausdehnung und Klangfarbe beeinflusst. Insbesondere konnte gezeigt werden, wie sich der Einfluss der simulierten Quellen auf das frequenzabhängige Diffus- und Direktschallverhältnis an der Hörerposition bemerkbar macht. Es kann davon ausgegangen werden, dass dadurch eine verzerrungsfreiere Reproduktion der Musikaufnahmen möglich wird, die sich in einem natürlicheren Raumeindruck für den Hörer auswirkt. Herauszufinden, ob dadurch der Realismus und die Immersion des Hörers in eine akustische Szene verstärkt werden, bleibt einem Hörversuch vorbehalten.

Literaturverzeichnis

Literatur:

Ahnert, Wolfgang (2009): *EASE 4.3 User's Manual*. Acoustic Design Ahnert. Berlin.

Barron, M. (1971): *The subjective effects of first reflections in concert halls - The need for lateral reflections*. In: *Journal of Sound and Vibration* 15 (4), S. 475–494.

Blauert, Jens (1997): *Spatial hearing.* The psychophysics of human sound localization. rev. Cambridge Mass. u.a: MIT Pr.

Blickensdorff, Johannes (2010): *Raumakustik im Wiener Burgtheater: Ein Vergleich von Messwerten und modellierten Parametern*. Magisterarbeit. Technische Universität, Berlin. Fachgebiet Audiokommunikation.

Bork, Ingolf (2010): *Richtcharakteristik von Musikinstrumenten*. Physikalisch Technische Bundesanstalt. Braunschweig. Online verfügbar unter http://www.ptb.de/de/org/1/16/163/richtchar.htm.

Bronstein, Ilja N.; Semendjajew, Konstantin A. (2008): *Taschenbuch der Mathematik*. 7., überarb. und erg. Frankfurt am Main: Deutsch.

Couvreur, Christophe (1998): *Implementation of a One-Third-Octave Filter Bank in Matlab*. Laboratoire d'Acoustique, Faculté Polytechnique de Mons.

Detzner, Erik; Schultz, Frank; Pollow, Martin; Weinzierl, Stefan (2010): *Zur Schallleistung von modernen und historischen Orchesterinstrumenten II: Holz- und Blechblasinstrumente*. In: Michael Möser (Hg.): Fortschritte der Akustik. 36. Deutsche Jahrestagung für Akustik ; 15.-18. März 2010, Berlin: Dt. Ges. f. Akustik.

Dickreiter, Michael (2008): *Handbuch der Tonstudiotechnik*. 7., völlig neu bearb. und erw. München: Saur.

Fastl, Hugo; Zwicker, Eberhard (2006): *Psychoacoustics.* Facts and models. 3. Aufl. Berlin u.a: Springer.

Goertz, A.; Laws, M. (1998): *Meßtechnische Kriterien zur Beurteilung von PA-Lautsprechern und deren Bedeutung für die klanglichen Qualitäten*. In: 20. Tonmeistertagung, Karlsruhe.

Goldstein, Herbert (1987): *Klassische Mechanik*. 9. Aufl. Wiesbaden: Aula-Verl.

Klemmer, Martin (2008): *Erzeugung von nachhallfreien mehrkanaligen Quellsignalen für die Binauralsynthese*. Berlin: Magisterarbeit im Fachgebiet Audiokommunikation, Technische Universität Berlin.

Krämer, Johannes; Schultz, Frank; Pollow, Martin; Weinzierl, Stefan (2010): *Zur Schallleistung von modernen und historischen Orchesterinstrumenten I: Streichinstrumente*. In: Michael Möser (Hg.): Fortschritte der Akustik. 36. Deutsche Jahrestagung für Akustik ; 15.-18. März 2010, Berlin: Dt. Ges. f. Akustik.

Lentz, T.; Slenczka, K. (2006): *Simulation of Natural Sound Sources*. In: Sabine Langer (Hg.): Fortschritte der Akustik. DAGA '06; 32. Deutsche Jahrestagung für Akustik, 20. - 23. März 2006 in Braunschweig. Berlin: Dt. Ges. für Akustik, S. 685–686.

Lindau, Alexander; Klemmer, Martin; Weinzierl, Stefan (2008): *Zur binauralen Simulation verteilter Schallquellen*. In: Fortschritte der Akustik, DAGA Dresden 2008. 34. Deutsche Jahrestagung für Akustik, 10.-13. März 2008. Berlin: DEGA.

Lindau, Alexander; Kosanke, Linda; Weinzierl, Stefan (2010): *Perceptual evaluation of physical predictors of the mixing time in binaural room impulse responses*. In: AES 128th Convention, May 22-25. London.

Marshall, A. H.; Meyer, J. (1985): *The directivity and auditory impressions of singers*. In: *Acta Acustica united with Acustica* (58), S. 130–140.

Meyer, Jürgen (1995): *Akustik und musikalische Aufführungspraxis.* Leitfaden für Akustiker, Tonmeister, Musiker, Instrumentenbauer und Architekten. 3., erweiterte Aufl. Bergkirchen: PPVMedien (Edition Bochinsky, 24).

Moldrzyk, Christoph; Lentz, Tobias (2004): *Dynamisches binaurales Wiedergabesystem zur wahrnehmungsverifizierten Untersuchung und Erstellung akustischer Räume*. In: Didier Cassereau (Hg.): Proceedings of the Joint Congress CFA/DAGA '04, Strasbourg, March 22 - 25,2004. Paris: Société Française d'Acoustique (SFA).

Möser, Michael (2009): *Technische Akustik*. 8. Aufl. s.l: Springer-Verlag.

Ochmann, M. (1990): *Die Multipolstrahlersynthese - eine effektives Verfahren zur Berechnung der Schallabstrahlung von schwingenden Strukturen beliebiger Oberflächengestalt.* In: *Acustica* (72), S. 233.

Olive, Sean E.; Toole, Floyd E. (1989): *The Detection of Reflections in Typical Rooms*. In: *J. Audio Eng. Soc* 37 (7/8), S. 539–553.

Olson, C. Bruce; Feistel Stefan (2002): *Loudspeaker Device File Formats for EASE 4.0 (Version 4.04e)*. Hg. v. Wolfgang Ahnert. Acoustic Design Ahnert. Berlin.

Otondo, Felipe; Rindel, Jens-Holger (2005): *A new method for the radiation representation of musical instruments in auralizations.* In: *Acta Acustica united with Acustica* (91), S. 902.

Papula, Lothar (2008): *Mathematik für Ingenieure und Naturwissenschaftler.* Band 3 Vektoranalysis, Wahrscheinlichkeitsrechnung, Mathematische Statistik, Fehler- und Ausgleichsrechnung. 5., verbesserte und erweiterte Auflage. Wiesbaden: Vieweg+Teubner Verlag / GWV Fachverlage GmbH.

Pätynen, Jukka (2009): *Directivities of orchestra instruments for auralization*. In: EAA Symposium on Auralization, 15-17 June. Espoo, Finland.

Pätynen, Jukka; Lokki, Tapio (2010): *Directivities of Symphony Orchestra Instruments*. In: *Acta Acustica united with Acustica* 96 (1), S. 138–167.

Pätynen, Jukka; Pulkki, Ville; Lokki, Tapio (2008): *Anechoic Recording System for Symphony Orchestra*. In: *Acta Acustica united with Acustica* 94 (6), S. 856–865.

Pollow, Martin (2009): *Measuring directivities of musical instruments for auralization*. In: Marinus M. Boone (Hg.): Proceedings. NAG/DAGA 2009, International Conference on Acoustics : Rotterdam, 23 - 26 March 2009; [including the 35th German Annual Conference on Acoustics (DAGA)]. Berlin: DEGA.

Pollow, Martin; Behler, Gottfried K. (2009): *Variable Directivity for Platonic Sound Sources Based on Spherical Harmonics Optimization*. In: *Acta Acustica united with Acustica* 95 (6), S. 1082–1092.

Pollow, Martin; Behler, Gottfried K.; Schultz, Frank (2010): *Musical instrument recording for building a directivity database*. In: Möser (Hg.): Fortschritte der Akustik. DAGA, 36. Jahrestagung für Akustik, 15.-18. März. Berlin: Deutsche Gesellschaft für Akustik e.V.

Pollow, Martin; Masiero, Bruno (2009): *Measuring directivities of natural sound sources with a spherical microphone array*. In: Ambisonics Symposium. June 25-27. Graz, Österreich.

Rafaely, Boaz (2005): *Analysis and Design of Spherical Microphone Arrays*. In: *IEEE Transactions on Speech and Audio Processing* 13 (1), S. 135–143.

Renkus-Heinz Inc. (2009): *EASE Version 4.3 User's Guide & Tutorial*. Renkus-Heinz Inc. & Acoustic Design Ahnert. Berlin.

Sandwell, David T. (1987): *Biharmonic spline interpolation of GEOS3 and SEASAT altimeter data*. In: *Geophys. Res. Lett* 14 (2), S. 139.

Tran, Thanh; Le Gia, Quoc Thong (2008): *Interpolation on the sphere: a fast solution technique*. In: *ANZ* 50 (01).

Vorländer, Michael (2008): *Auralization.* Fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality. 1. Aufl. Berlin: Springer (RWTH edition / RWTH Aachen).

Weinzierl, Stefan (Hg.) (2008): *Handbuch der Audiotechnik*. Berlin u.a: Springer (VDI-Buch).

Young, R. W. (1939): *Terminology for Logarithmic Frequency Units*. In: *The Journal of the Acoustical Society of America* 11 (1), S. 134–000. Online verfügbar unter info:bibcode/1939ASAJ…11.134Y.

Zotter, F.; Sontacchi, A.; Noisterning, M.; Höldrich, R. (2007): *Capturing the radiation characteristics of the bonang barung*. In: Proc. 3rd AAAA Congress. 27. - 28. September. Graz.

Normen:

ANSI: *S1.11-2004 - Specification for octave-band and fractional-octave-band analog and digital filters*.

DIN EN 61260, 2003: *Bandfilter für Oktaven und Bruchteile von Oktaven*.

DIN 18041, 2004: *Hörsamkeit in kleinen bis mittelgroßen Räumen*.

DIN 45641, 1990: *Mittelung von Schallpegeln*.

DIN EN ISO 266, 1997: *Normfrequenzen*.

DIN EN ISO 3745, 2003: *Bestimmung der Schallleistungs- und Schallenergiepegel von Geräuschquellen aus Schalldruckmessungen - Verfahren der Genauigkeitsklasse 1 für reflexionsarme Räume und Halbräume*.

DIN EN ISO 3382, 2008: *Messung von Parametern der Raumakustik*.

DIN EN ISO 3741, 2010: *Bestimmung der Schallleistungs- und Schallenergiepegel von Geräuschquellen aus Schalldruckmessungen – Hallraumverfahren der Genauigkeitsklasse 1*.

Software und Musikaufnahmen:

Ahnert, Wolfgang (2010): *EASE*. Berlin: ADA Acoustic Design Ahnert.

Dalenbäck, Bengt-Inge (2011): *CATT-Acoustic*. Gothenburg: CATT.

Drake, Nick (1969): *River Man.* Nick Drake. London: Universal-Island Records Ltd.

Gautier, Thierry (2010): *Contredanse.* Pascale Gautier Quartet. Montpellier.

The MathWorks, Inc (2009): *MATLAB*. Natick, USA: The MathWorks, Inc.

Abbildungsverzeichnis

Abbildung 2.1 Foto der Mikrofonaufstellung für das Streichquartett bei den Aufnahmen zu „River Man“ 8

Abbildung 3.1 Blockdiagramm der Verarbeitungsschritte zur Berechnung der Richtcharakteristiken 11

Abbildung 3.2 Blockdiagramm für die Filterung und Mittelung der Schalldruckpegel. 12

Abbildung 3.3 Frequenzgänge der 21 Terzbandfilter 14

Abbildung 3.4 Richtcharakteristiken des Kontrabasses für sieben exemplarische Terzbänder im Oktavabstand aus dem Blickwinkel vorne/rechts/unten (Azimut=315° und Polarwinkel=110°) 15

Abbildung 3.6 Richtcharakteristiken des Kontrabasses nach Korrektur der Quellposition für sieben exemplarische Terzbänder im Oktavabstand aus dem Blickwinkel vorne/rechts/unten (Azimut=315° und Polarwinkel=110°) 18

Abbildung 3.7 Blockdiagramm für die Filterung und Mittelung der Schalldruckpegel mit Kompensation der Quellposition. 19

Abbildung 3.9 Das Koordinatensystem von EASE 22

Abbildung 3.10 Positionen der 32 gleichmäßig auf einer Kugeloberfläche verteilten Messmikrofone 23

Abbildung 3.11 Interpolation des gemittelten Schalldruckpegels der Stimme in Kugelkoordinaten bei einer Terz von f_m = 4 kHz. 25

Abbildung 3.12 Interpolation des gemittelten Schalldruckpegels der Stimme in Kugelkoordinaten bei einer Terz von f_m = 4 kHz. 26

Abbildung 3.13 Interpolation des gemittelten Schalldruckpegels der Stimme in Kugelkoordinaten bei einer Terz von f_m = 4 kHz. Die Interpolation wurde zur Veranschaulichung der Periodizität in negativer und positiver Richtung um 45° einer Kreisumdrehung erweitert. Das gestrichelte Rechteck entspricht dem herausgeschnittenen Bereich einer vollständigen Kugelfläche. 26

Abbildung 3.14 Interpolation des gemittelten Schalldruckpegels der Stimme in Kugelkoordinaten bei einer Terz von f_m = 4 kHz. Die Messwerte der Polstellen wurden in 2°-Abständen entlang des Azimuts kopiert. 27

Abbildung 3.15 Darstellung der Interpolationsergebnisse für die vier von *griddata* zur Verfügung gestellten Methoden. 28

Abbildung 3.16 Interpolation des gemittelten Schalldruckpegels der Stimme in Kugelkoordinaten bei einer Terz von f_m = 4 kHz 29

Abbildung 3.17 5°-Koordinatennetz vor der Drehung.. 30

Abbildung 3.18 Das 5°-Koordinatennetz von EASE wurde auf das interpolierte 2°-Gitter gelegt.....31

Abbildung 3.19 Interpolation des gemittelten Schalldruckpegels der Stimme bei einer Terz von f_m = 4 kHz im EASE-Lautsprecherformat nach der Methode der Koordinatenabbildung.....32

Abbildung 3.20 Drehung der Koordinaten der Messmikrofone.....33

Abbildung 3.21 Mikrofonarray nach zweimaliger Drehung (links) und Änderung des Blickwinkels zur Korrektur in die ursprüngliche Ansicht (rechts).....34

Abbildung 3.22 Verlauf der Interpolation des gemittelten Schalldruckpegels der Stimme in EASE-Koordinaten bei einer Terz von f_m = 4 kHz 34

Abbildung 3.23 Interpolation des gemittelten Schalldruckpegels der Stimme in EASE-Koordinaten bei einer Terz von f_m = 4 kHz 35

Abbildung 3.24 Verlauf der Interpolation des gemittelten Schalldruckpegels der Stimme in EASE-Koordinaten bei einer Terz von f_m = 4 kHz 36

Abbildung 3.25 Interpolation des gemittelten Schalldruckpegels der Stimme bei einer Terz von f_m = 4 kHz.....37

Abbildung 3.26 Betrag der Abweichung zwischen den Interpolationsverfahren der Koordinatenabbildung und der direkten Interpolation in EASE-Koordinaten für die Stimme bei einem Terzband von 4 kHz 38

Abbildung 3.27 Das Koordinatensystem von CATT 39

Abbildung 3.28 Interpolation des Schalldruckpegels in EASE-Koordinaten für den Flügel bei einer Oktave von f_m = 4 kHz.....40

Abbildung 3.29 Ausrichtung des Koordinatensystems bei der Messung des Flügels 41

Abbildung 3.33 Grafisches Ausmessen des Polarwinkels der zwei Mikrofonpositionen für den Bass. 47

Abbildung 3.38 Normierter Leistungsfrequenzgang des Kontrabasses für zwei Bezugsrichtungen.....62

Abbildung 3.39 Verlauf des Schalldruckpegels für die frontale Richtung bei zwei verschiedenen Normierungen.....63

Abbildung 3.40 Bezeichnung der Achsen, Ebenen und Polstellen in der zweidimensionalen Darstellung der Richtcharakteristik in den Koordinaten des um die Medianachse rotierenden Halbkreisbogens. 64

Abbildung 3.41 a: Horizontales und Vertikales Polardiagramm für drei Vokale. b: Pegeldifferenz zwischen vorderer und seitlicher Abstrahlung für dieselben drei Vokale als Funktion der Frequenz 67

Abbildung 3.42 Terzbandpegel für die Stimme. 68

Abbildung 3.43 Richtcharakteristiken der Stimme für sechs Oktaven 69

Abbildung 3.44 Vertikales Polardiagramm der Stimme für sechs Oktaven.....70

Abbildung 3.45 Horizontales Polardiagramm der Stimme für sechs Oktaven.. 72

Abbildung 3.46 a: Pegelunterschied zwischen frontaler und seitlicher Abstrahlung. b: Pegelunterschied zwischen frontaler und rückwärtiger Abstrahlung. c: Pegelunterschied zwischen frontaler Abstrahlung und Abstrahlung nach oben. d: Richtwirkungsmaß für die Richtung maximaler Abstrahlung.....73

Abbildung 3.47 Normierter Leistungsfrequenzgang. 74

Abbildung 3.48 Richtcharakteristiken der Stimme für drei Terzbänder.....75

Abbildung 4.1 3D-Darstellung des Konzertsaals.....78

Abbildung 4.2 Verlauf der Nachhallzeit für den gewählten Konzertsaal. 79

Abbildung 4.3 Verlauf der Nachhallzeit nach der Optimierung. 79

Abbildung 4.4 Aufstellung der Schallquellen für das Stück „Contredanse“.....80

Abbildung 4.5 Aufnahme des Schlagzeugs für das Stück „Contredanse“. 82

Abbildung 4.6 Aufstellung der Schallquellen für das Stück „River Man“. 84

Abbildung 4.7 Richtwirkungsmaß bei der Abstrahlung auf den Hörer für drei Instrumente....85

Abbildung 4.8 Verhältnis von Direkt- und Diffusschall an der Hörerposition. 85

Abbildung 4.9 Reflektogramm aller Schallquellen von „Contredanse“ in der Form Schalldruckpegel als Funktion der Zeit. (Partikelanzahl: 7371000) 88

Abbildung 4.10 Vergleich der Nachhallzeitberechnung nach Schroeder und Sabine. 90

Abbildung 4.11 Simulierter Schalldruckpegel an der Hörerposition für die Stimmenquelle. Links: gerichtet; Rechts: ungerichtet. 92

Abbildung 4.12 Darstellung der Richtung und Intensität der vertikalen Reflektionen für „Contredanse“ bei dem Terzband von 1 kHz.....93

Abbildung 4.13 Anteil der frühen seitlichen Reflektionen (lateral fraction) für die Sängerposition mit korrekter Richtcharakteristik (links) und für ein ungerichtetes Quellmodell.....94

Tabellenverzeichnis

Tabelle 1 Spurbezeichnungen für das Stück „Contredanse“. 6

Tabelle 2 Positionen der Messmikrofone um die Schallquelle. 24

Tabelle 3 Abweichungen der Interpolationsmethoden von griddata für den Mittelungspegel der Stimme bei f_m = 4 kHz. 28

Tabelle 4 Die Koordinaten der beiden Mikrofone für den Bass 47

Tabelle 5 Die Koordinaten des Mikrofons bei der Aufnahme des Cellos. 48

Tabelle 6 Die Koordinaten des Mikrofons für die Aufnahme der Bratsche. 49

Tabelle 7 Die Koordinaten des Mikrofons für die Aufnahme der Violine.....51

Tabelle 8 Die Koordinaten der Mikrofonposition für die Aufnahme der Gitarre.....52

Tabelle 9 Lage der Maxima 73

Tabelle 10 Abstand der Schallquellen vom Bühnenboden für das Stück „Contredanse“.....82

Tabelle 11 Abstand der Schallquellen vom Bühnenboden für das Stück „River Man“. 84

Tabelle 12 Liste der verwendeten Parameter für AURA Response.....89

Tabelle 13 Zuordnungstabelle von Mikrofonnummer des 32-Mikrofonkugelarray der RWTH Aachen zur Nummer des Audiokanals bei der Messung TU Berlin. 107

Anhang A: Inhalt der DVD

Auf der beiliegenden DVD befinden sich folgende Daten:

1. Diese Arbeit als PDF
2. Matlab-Quellcodes zur Erzeugung von Lautsprechermodellen aus den Daten des 32-Mikrofonkugelflächenarrays
3. Interpolierte Daten aller Instrumente, aus denen mit der Datei „run3_normierung_kenndaten_ascii_export.m“ Quellmodelle mit beliebiger Bezugsrichtung erstellt werden können
4. Lautsprechermodelle im EASE-Format. Es wurde folgende Bezugsrichtung gewählt:

Instrument	**Azimut [°]**	**Polarwinkel [°]**	**EASE Phi [°]**	**EASE Theta [°]**
Violine	0	0	90	90
Viola	0	0	90	90
Violoncello	0	90	0	0
Kontrabass	0	90	0	0
Flöte	0	45	90	45
Oboe	0	45	90	45
Klarinette	0	45	90	45
Bassklarinette	0	45	90	45
Fagott	0	45	90	45
Kontrafagott	0	45	90	45
Englischhorn	0	45	90	45
Horn	225	90	180	135
Trompete	0	90	0	0
Posaune	0	90	0	0
Tuba	0	0	90	90
Harfe	45	90	0	45
Pauke	0	0	90	90
Sopran	0	90	0	0
Gitarre	0	90	0	0
Flügel	0	0	90	90

Anhang B: Patchtabelle

Mikrofon	WAV-Kanal	Richtung
1	1	
2	9	
3	17	
4	18	
5	25	
6	2	
7	10	
8	19	
9	20	
10	26	
11	3	
12	11	
13	21	
14	27	
15	4	Frontal
16	5	
17	12	
18	22	
19	28	
20	6	
21	13	Nordpol
22	29	
23	7	
24	14	
25	23	
26	30	
27	15	
28	24	
29	31	
30	32	
31	8	
32	16	Südpol

Tabelle 13 Zuordnungstabelle von Mikrofonnummer des 32-Mikrofonkugelarray der RWTH Aachen zur Nummer des Audiokanals bei der Messung nach Detzner (2010) und Krämer (2010).

Anhang C: Quellcodes

run1_mittel_und_filterung.m

```
% Mittelung und Filterung der gemessenen Signale der 32-Mikrofonkugelarrays
% der RWTH Aachen für ein Instrument.
%
% 1.) Kompensationsfilter
% 2.) Terzbandfilter, 100 Hz bis 10 kHz
% 3.) Mittelung
% 4.) Speicherung der 21 Terzpegel x 32 Kanäle in einer Matrix
%
% Dependencies:
%   Funktionen: oct3bankext_10k
%   Daten: Mic_Filter_IRs, calibration_micarray_20090512T115222.mat
%
% Roman Schneider, TU Berlin (2011)

close all; clear all; clc;
load Mic_Filter_IRs; % Kompensationsfilter
Path = 'C:\Daten\trompete';     % Input-Pfad
Path_save= 'C:\Instrumente\L_data';     % Output-Pfad
load([Path,'/','calibration_micarray_20090512T115222.mat']); %
Kalibrierungswerte

Fs = 44100;                                     % Samplingfrequenz
L_cal = 94;                                     % Pistonphon-Pegel
[dB_SPL_RMS]
p_cal94 = 10^(L_cal/10);                        % Schalldruckquadrat durch
p0-Quadrat Pa_RMS
p_cal2 = (10.^(absolute_level_channels/10))'; % numerische
Kalibrierungswerte aus dBFS
p_cal2_94 = p_cal94./p_cal2;
NTerz = 21; % Anzahl der Terzbänder

% wav einlesen:
Filenames = dir([Path,'/*.wav']);               % hole Info aller wav aus
best. Ordner
NFiles=size(Filenames);                         % Anzahl Files

% Speicher leeren
x_rms_sum = zeros(32,NTerz);                       % Matrix: Zeile = Kanal;
Spalte = Terzband

for f=1:NFiles(1)                               %do for all files
    disp([num2str(f),'/',num2str(NFiles(1))]);  %Anzeige Status
    file = Filenames(f).name;                   %Filestrings
    [x, Fs] = wavread([Path,'/',file]);         %einlesen eines 32CH File
    for n = 1:32
        % Bewertungsfilter + HP anwenden
        x_bewertet = fftfilt(IR_BEM_BP(:,n),x(:,n));

        % Terzfiltern:
        [x_rms, F] = oct3bankext_10k(x_bewertet);

        % Aufsummierung:
```

```
        for k = 1:NTerz
            x_rms_tmp(n,k) = (x_rms(k) * p_cal2_94(n));
            x_rms_sum(n,k) = x_rms_sum(n,k) + x_rms_tmp(n,k); %
Aufsummierung
        end
    end
end
% Mittelung und Logarithmierung:
L_data = 10*log10(x_rms_sum./NFiles(1));

% Speichern:
save([Path_save,'/','L_trompete'],'L_data');
```

run2_interpolation

```
% Interpolation nach der Mappingmethode
%
% 1.) Kanalnummern patchen
% 2.) Werte kopieren um auf eine vollständige Interpolationsfläche zu
% kommen
% 3.) Interpolation auf ein 2°-Netz in Polarkoordinaten
% 4.) EASE-Lautsprecherkoordinaten erzeugen
% 5.) Nächsten Winkel zwische den beiden Koordinatensystemen suchen
% 5.) Interpolierte Werte den EASE-Koordinaten zuordnen
%
% Dependencies:
%   Funktionen: kugel2kart
%   Daten: chan_numbers
%
% Roman Schneider, TU Berlin (2011)

close all; clear all; clc;
Path = 'C:\\Instrumente\Interp_data'; % Output-Pfad
Path_load = 'C:\Instrumente\L_data'; % Input-Pfad
load([Path_load,'/','L_trompete.mat']); % Input-Data

% Koordinaten des Arrays:
% dihedral angle (between two adjacent pentagons)
dihedral = 2*asin(cos(pi/3) ./ sin(pi/5));

phi = pi/5 * [1 3 5 7 9 1 3 5 7 9 2 4 6 8 0 2 4 6 8 0 ...
    0 ...
    0 2 4 6 8 3 5 7 9 1 ...
    0];

% theta_01 = 180 - 90 - dihedral/2;
theta_01 = pi/2 - 52.62/360*2*pi;
theta_02 = 2 * (pi - pi/2 - dihedral/2);
theta_04 = theta_01 + 2*(pi/2 - dihedral/2);
theta_03 = pi - theta_04;
theta_05 = pi - theta_02;
theta_06 = pi - theta_01;

theta = [theta_01 theta_01 theta_01 theta_01 theta_01 ...
    theta_03 theta_03 theta_03 theta_03 theta_03 ...
    theta_04 theta_04 theta_04 theta_04 theta_04 ...
    theta_06 theta_06 theta_06 theta_06 theta_06 ...
    0 ...
    theta_02 theta_02 theta_02 theta_02 theta_02 ...
```

```
    theta_05 theta_05 theta_05 theta_05 theta_05 ...
    pi];

d = 4.17; % inner diameter array (from tops of mics to tops on other side)
r = d/2;

% Patchen der Channelkoordinaten:
load chan_numbers;
phi_ch=zeros(1,32);
theta_ch=zeros(1,32);
for n=1:32
    phi_ch(chan_numbers(n))=phi(n);
    theta_ch(chan_numbers(n))=theta(n);
end
phi=phi_ch;
theta=theta_ch;

% Akku leeren:
pegel_mapped=zeros(37,72,21); % 37 Phiwinkel, 72 Thetawinkel, 21 Terzen

% Interpolation
for t=1:21 % für alle 21 Terzbänder

    % Fortschrittsanzeige:
    disp([num2str(t),'/21']);

    pegel = L_data(:,t);
    pegel = pegel';

    % Kopieren der Datenpunkte nach links und nach rechts:
    for n=1:32
        phiext_r(n) = phi(n)+(2*pi);
        phiext_l(n) = phi(n)-(2*pi);
    end
    phiextended = [phi phiext_r phiext_l];
    thetaextended = [theta theta theta];
    pegelextended = [pegel pegel pegel];

    % Pole Kopieren:
    % Südpol:
    phi_spol = [2:2:358]./(180/pi);
    theta_spol = ones(1,length(phi_spol)).*pi;
    pegel_spol = ones(1,length(phi_spol)).*pegel(16);

    % Nordpol:
    phi_npol = [2:2:358]./(180/pi);
    theta_npol = zeros(1,length(phi_npol));
    pegel_npol = ones(1,length(phi_npol))*pegel(13);

    phiextended_pol = [phiextended phi_spol phi_npol];
    thetaextended_pol = [thetaextended theta_spol theta_npol];
    pegelextended_pol = [pegelextended pegel_spol pegel_npol];

    % Interpolationsgrid erzeugen:
    phigrid = [0:2:358]./(180/pi); % phi
    thetagrid = [0:2:180]./(180/pi); % theta
    [phimesh,thetamesh] = meshgrid(phigrid, thetagrid);

    % Interpolieren, es ergibt sich eine Matrix mit interpolierten
    % Pegelwerten für Reihe=Theta, Spalte=Phi
```

```
    pegel_i =
griddata(phiextended_pol,thetaextended_pol,pegelextended_pol,phimesh,thetam
esh, 'v4');

    % EASE-Grid erzeugen:
    phi_tmp = [0:5:355]./(180/pi);
    theta_tmp = [0:5:180]./(180/pi);
    [phimesh_tmp,thetamesh_tmp] = meshgrid(phi_tmp, theta_tmp);

    % Grid in die Richtung des Koordinatensystems von EASE drehen:
    [x,y,z] = kugel2kart(phimesh_tmp, thetamesh_tmp, r);
    xtt = z;
    ytt = y;
    ztt = x*(-1);
    xt=xtt;
    yt=ztt*(-1);
    zt=ytt;

    % Mappen der interpolierten Daten auf das Koordinatensystem von EASE:
    xyz_t=[xt(:),yt(:),zt(:)]; % Koordinaten des EASE-Grids
    [x_i, y_i, z_i]=kugel2kart(phimesh,thetamesh,r); % Koordinaten ins
kartesische System überführen
    xyz_i=[x_i(:),y_i(:),z_i(:)]; % Koordinaten des Interpolationsgrids
    npoint=dsearchn(xyz_i,xyz_t); % Finde nächsten Punkt wischen den beiden
Grids
    pegel_i = pegel_i(:);

    % Zuweisung des Pegels zu den EASE-Punkten (pegel_map(h) enthält jetzt
die
    % korrekten Pegelwerte zu den Koordinaten in xyz_t):
    for h=1:length(xyz_t)
        pegel_map(h) = pegel_i(npoint(h));
    end

    % Konvertieren einer eindimensionalen Matrix in eine 37-mal-72-Matrix
    for m=1:37
        for n=1:72
            pegel_map_matrix(m,n)=pegel_map((37*(n-1)+1)+m-1);
        end
    end
    pegel_mapped(:,:,t)=pegel_map_matrix;
end
% Speichern:
save([Path,'/','pegel_mapped_trompete'],'pegel_mapped');
```

run3_normierung_kenndaten_ascii_export

```
% Export eine EASE-Lautsprechermodells aus den interpolierten Daten
%
% 1.) Berechnung des statistischen Richtfaktors
% 2.) Normierung auf Bezugsrichtung
% 3.) ASCII-Export
%
% Roman Schneider, TU Berlin (2011)

close all; clear all; clc;
Path = 'C:\Instrumente\Interp_data'; % Input-Pfad
load([Path,'/','pegel_mapped_trompete.mat']); % Inputdaten
fid = fopen('trompete.xhn', 'w'); % Output-Datei
```

```
% Bezugsrichtung:
phi=1;
theta=1;

% Koordinaten:
phigrid = [0:5:355]./(180/pi); % phi
thetagrid = [0:5:180]./(180/pi); % theta
[phimesh,thetamesh] = meshgrid(phigrid, thetagrid);

% halber Winkelschritt:
deg=(2.5/(180/pi));

% Frequenzbandbezeichung:
FLabel = [ 100 125 160, 200 250 315, 400 500 630, 800 1000 1250, ...
    1600 2000 2500, 3150 4000 5000 6300 8000 10000 ];

% Winkelbezeichnung:
phi_Ease = (0:5:355);

% Schreibe Dateiheader
fprintf(fid, '"FileType","Speaker Types"\r\n');
fprintf(fid, '"Format",4.0\r\n');
fprintf(fid, '"LengthUnit","meters"\r\n');
fprintf(fid, ';\r\n');


for t = 1:21 % für jedes Terzband

    % Berechne Schalldruck einer ungerichteten Quelle mit gleicher
    % Leistung:
    p_2=10.^(pegel_mapped(:,:,t)./10); % Schalldruckquadrat
    a=0; % Akku leeren
    % Aufsummieren der Intensitäten:
    for k = 2:36
        for l = 1:72
            % p^2 * Teilfläche:
            a = a + p_2(k,l)*2*deg*( cos(thetamesh(k,l)-deg) -
cos(thetamesh(k,l)+deg) );
        end
    end
    a=a+p_2(1,1)*2*deg*( cos(thetamesh(1,1)-deg) - cos(thetamesh(1,1)+deg)
);
    a=a+p_2(37,1)*2*deg*( cos(thetamesh(37,1)-deg) -
cos(thetamesh(37,1)+deg) );
    % statistischer Richtfaktor:
    R=(10^(pegel_mapped(1,1,t)/10))/( a/ (4*pi) );

    % Normieren auf Bezugsrichtung:
    for k = 1:37
        for l = 1:72
            att_data(k,l) = (pegel_mapped(k,l,t) -
pegel_mapped(theta,phi,t))*(-1);
        end
    end

    % Schreibe Frequenzheader:
    fprintf(fid, '"SpeakerName","trompete"\r\n'); % <- SPEAKERNAME
    fprintf(fid, '"Frequency",%-5.0f,"Hz"\r\n', FLabel(t));
    fprintf(fid, '"Sensitivity",100,"dB"\r\n');
    fprintf(fid, '"Impedance",8,"ohms"\r\n');
    fprintf(fid, '"Q",%-2.1f\r\n',R);
```

```
    fprintf(fid, '"MaxPower",200,"W"\r\n');
    fprintf(fid, '"DataGood",1\r\n');

    % Schreibe Pegelwerte:
    for n = 1:72
        fprintf(fid, '" %-2.0f°", ', phi_Ease(n));
        for m=1:37
           fprintf(fid, '%-4.2f, ', att_data(m,n));
        end
      fprintf(fid, '\r\n');
    end
    fprintf(fid, ';\r\n');
    fprintf(fid, '"End"\r\n');
    fprintf(fid, ';\r\n');
end
fclose(fid);
```