

Vermessung von Außenohrübertragungsfunktionen mit reziproker Messmethode

Projektarbeit

Johannes Zaar, Matr.Nr. 0473066

Graz, 10. Dezember 2010

Betreuer: Dr. Franz Zotter



institut für elektronische musik und akustik



Zusammenfassung

Bei der meistverwendeten Technik zur Messung von individuellen Außenohrübertragungsfunktionen (Head-related Transfer Functions, HRTFs) werden rund um die Versuchsperson angeordnete Lautsprecher und In-Ear-Mikrofone verwendet. Im Gegensatz dazu können die HRTFs jedoch auch reziprok gemessen werden, wobei der Messaufbau mit nach außen gewandten Kleinstlautsprechern in den Ohrmuscheln der zu vermessenden Person und ringsum angeordneten Mikrofonen realisiert wird. Die theoretische Basis dieses Austausches von Schallquellen und -senken ist das Helmholtz'sche Prinzip der Reziprozität. Das Verfahren hat den Vorteil einer von der räumlichen Abtastung unabhängigen und dadurch deutlich verkürzten Messdauer und kann so die Fixierung der zu vermessenden Personen erleichtern, sodass sogar eine Vermessung von unterschiedlichen Schulterdrehungen im Bereich des Möglichen liegt.

Unter Verwendung des am IEM entwickelten 64-Kanal-Kugelmikrofonarrays werden in dieser Arbeit ein geeignetes Messsystem und die zugehörige Messsignalverarbeitung und -auswertung erarbeitet. Die dabei auftretenden Probleme werden aufgezeigt und mögliche Lösungen diskutiert und implementiert. Insbesondere auf den Signal-Rausch-Abstand (SNR), der auf Grund der geringen Dimensionen der verwendeten Lautsprecher das zentrale Problem dieser Methode ist, wird näher eingegangen. Um eine sinnvolle Auswertung der Messergebnisse durchführen zu können, wurden die HRTFs mehrerer Versuchspersonen bei einer Referenzmessung am Institut für Schallforschung in Wien vermessen und mit den entsprechenden reziprok gemessenen HRTFs verglichen.

Inhaltsverzeichnis

1	Einführung	5
1.1	Außenohrübertragungsfunktionen - HRTFs	6
1.2	Direkte Messmethode	8
1.3	Reziproke Messmethode	10
2	Hardware	12
2.1	Mikrofon-Array	12
2.2	Messlautsprecher	12
2.2.1	Beschreibung der Wandler	12
2.2.2	Entwurf und Bau	13
3	Übertragungsverhalten der Messlautsprecher	14
3.1	Bündelungsverhalten	14
3.2	Frequenzgänge	16
3.2.1	Ursprüngliches Vermessungskonzept	16
3.2.2	Angewandtes Vermessungskonzept	18
4	Entzerrung der Messlautsprecher	19
4.1	Erstellung der Entzerrungsfiler	19
4.2	Glättung der Entzerrungsfiler	20
4.2.1	Terzbandglättung	20
4.2.2	Phasengang	22
5	Anpassung des Messsignals	23
5.1	Grundsätzliche Überlegungen zum SNR - Motivation	23
5.2	Angepasste Sweepgenerierung	25
5.2.1	Prinzip und Algorithmus	25
5.2.2	Anwendung	27

<i>Vermessung von Außenohrübertragungsfunktionen mit reziproker Messmethode</i>	4
6 Messung	28
7 Signalverarbeitung nach der Messung	30
7.1 Entfaltung	30
7.2 Schneiden der Impulsantworten	30
8 Auswertung	32
8.1 Signal-Rausch-Abstand	32
8.2 Aufteilung in CTFs und DTFs	33
8.3 Beispielhafte Übertragungspfade	35
9 Schlussfolgerungen und Diskussion	37
10 Anhang	38
10.1 Versuchsperson 1	39
10.2 Versuchsperson 2	42
10.3 Versuchsperson 3	45

1 Einführung

Das menschliche Gehör ermöglicht es uns, Schallereignisse sehr genau zu lokalisieren. Eine wesentliche Voraussetzung dafür ist die richtungsabhängige Veränderung von eintreffenden Schallwellen durch Außenohr, Kopf und Torso, sowie die unterschiedlichen räumlichen Positionen der beiden Ohren. Zusammengefasst lässt sich dies als ein Paar richtungsabhängiger Filter betrachten, die Außenohrübertragungsfunktionen oder *head-related transfer functions* (HRTFs, siehe Abb. 1). Diese weisen große interindividuelle Unterschiede auf, da die Form und Größe des Körpers und insbesondere des Kopfes und der Ohren eine entscheidende Rolle spielt. Da außerdem schon sehr geringe Veränderungen von Positionen wahrgenommen werden können, ist die messtechnische Ermittlung von akkuraten HRTFs sehr aufwändig.

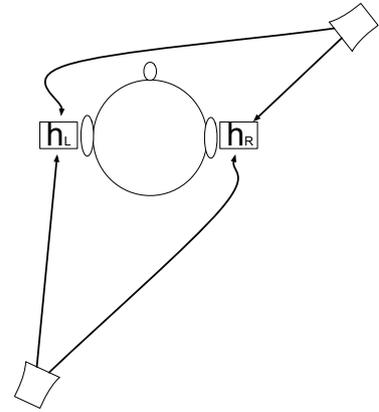


Abbildung 1: Außenohrübertragungsfunktionen

Die räumliche Abtastung für eine vollständige Nachbildung muss auf der Fläche einer Kugel erfolgen, in deren Mitte sich der Kopf der Versuchsperson befindet.



(a) HRTF-Messplatz des ISF Wien mit Lautsprecher-Array



(b) 64-Kanal-Kugelmikrofonarray des IEM

Abbildung 2: Messaufbauten für (a) direkte und (b) reziproke HRTF-Vermessung

Je nach Verwendungszweck sollte der Abstand der Messpunkte zueinander relativ klein sein, idealerweise unendlich klein. Die üblichste Messmethode (im Folgenden als *direkte Methode* bezeichnet, siehe *Abb. 2(a)*) bedient sich einer kreisförmigen vertikalen Anordnung von Messlautsprechern in deren Mitte die Versuchsperson auf einem Drehteller sitzt und in kleinen Schritten um ihre eigene Achse gedreht wird; in jeder Position werden nacheinander Messsignale durch die Messlautsprecher ausgespielt und mit Kleinstmikrofonen in den Ohren der Versuchsperson aufgenommen. Da die Messsignale um Überschneidungen zu vermeiden nacheinander ausgespielt werden müssen - ein gewisses Maß an Verschachtelung ist allerdings möglich, vgl. [MBL06] - ist die Messdauer sehr hoch. Sie ist *abhängig von der räumlichen Abtastung* und beträgt meist mindestens 30 Minuten. Die hier verwendete *reziproke Messmethode* arbeitet mit dem gegenteiligen Prinzip, d.h. einer Abtastung mit einer kugelförmigen Mikrofonanordnung (siehe *Abb. 2(b)*), wobei die Messsignale aus Kleinstlautsprechern in den Ohren der Versuchsperson ausgespielt werden. Dabei ist die Messdauer *unabhängig von der räumlichen Abtastung*, es können alle Übertragungspfade für ein Ohr gleichzeitig aufgenommen werden.

1.1 Außenohrübertragungsfunktionen - HRTFs

Interaurale Laufzeit- und Pegeldifferenz Wichtig für die Lokalisation des menschlichen Gehörs ist zunächst einmal die interaurale Laufzeitdifferenz (interaural time difference, ITD), also der zeitliche Versatz zwischen dem Eintreffen eines Schallereignisses am linken und rechten Trommelfell. Diese ist im Bereich zwischen 80 Hz und 800 Hz durch Phasenunterschiede und darüber - mit einem fließenden Übergang zwischen 800 Hz und 1600 Hz - durch Unterschiede der Gruppenlaufzeiten bestimmt. Der Grund dafür liegt in der Wellenlänge des Schalls, die für Frequenzen unter 800 Hz größer als die doppelte Wegstrecke zwischen den Ohren ist. Unterschreitet die Wellenlänge diesen Wert, so wird die Auswertung von Phasenlaufzeiten ambivalent, was durch die Auswertung von Gruppenlaufzeiten teilweise kompensiert werden kann. Mit steigender Frequenz wird außerdem durch die zunehmende Abschattung die interaurale Pegeldifferenz (interaural level difference, ILD) immer ausgeprägter und daher wichtiger für die Lokalisation.

Die Laufzeit- bzw. Pegelinformation ist allerdings nicht eindeutig, da es mehrere Positionen eines Schallereignisses mit jeweils gleichem Abstand zu beiden Ohren und daher gleicher ITD und ILD gibt. Bei alleiniger Nutzung dieser Information ergeben sich dadurch Verwechslungen mit anderen Positionen; diese sind kegelförmig um die interaurale Achse angeordnet, das Phänomen wird daher auch *cone of confusion* genannt (siehe *Abb. 3 (a)*¹). Diese Mehrdeutigkeit von ITD und ILD wird jedoch in der Realität durch kleine Bewegungen des Kopfes teilweise aufgehoben, das Gehör kann dabei durch Auswertung der Bewegungsrichtung und -geschwindigkeit falsche Positionen ausschließen.

1. <http://193.166.3.2/pub/sci/audio/jwest-thesis/>

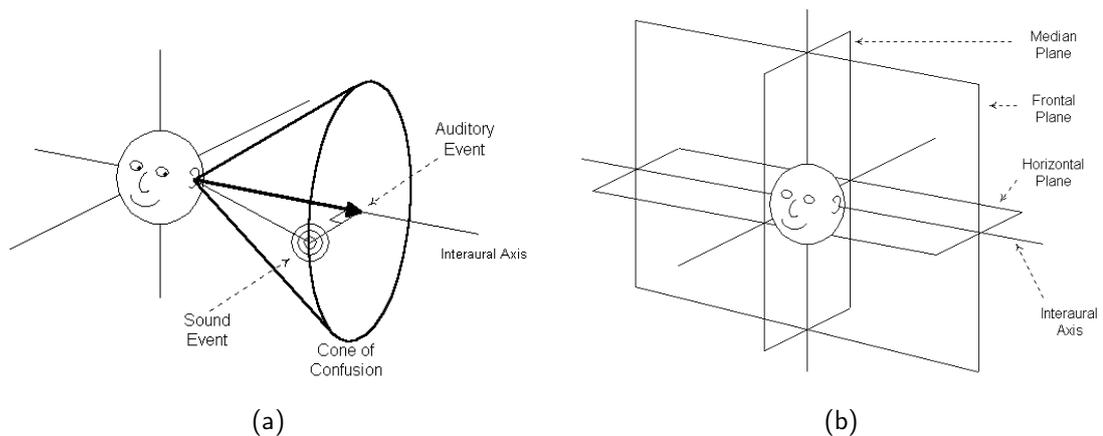


Abbildung 3: (a) Cone of Confusion, (b) Ebenen und Achsen

Frequenzselektive Filterung Die zweite wichtige Lokalisationsinformation für das menschliche Gehör ist die frequenzabhängige Filterung des Schalls durch Torso, Kopf, und Außenohr. Abhängig von der Einfallsrichtung des Schalls werden verschiedene Resonanzen angeregt, jeder Richtung wird dadurch ein anderes frequenzabhängiges Profil aufgeprägt. Durch dieses komplexe akustische Filter wird die Richtungserkennung bereits ohne Kopfbewegungen eindeutig, insbesondere in der Elevation.

Räumliche Auflösung Für eine möglichst genaue Messung der HRTFs gelten relativ hohe Anforderungen an die räumliche Verteilung der Messpositionen. Im Azimut beträgt der gerade noch wahrnehmbare Unterschied (Just Noticeable Difference, JND) zwischen ca. 1° frontal vor der Person und $5^\circ - 10^\circ$ ganz rechts bzw. ganz links von ihr. Die JNDs in der Elevation sind stark signalabhängig und können von 4° (weißes Rauschen) bis zu 17° (durchgehende Sprache einer unbekannt Person) variieren [ZDGG06]. Da die Regionen der höchsten Sensibilität zudem auch individuell leicht unterschiedlich sind, ist eine sehr feine räumliche Abtastung nötig um eine gute Nachbildung zu erreichen.

Formale Darstellung Außenohrübertragungsfunktionen sind abhängig von vier Variablen:

- Frequenz f
- Abstand r
- Azimutwinkel ϕ , Horizontalebene
- Elevationswinkel θ , Medianebene

Dementsprechend lautet die eigentlich korrekte Darstellung $HRTF(\phi, \theta, f, r)$. Der Einfachheit halber verzichtet man aber üblicherweise auf die Darstellung der Distanzabhängigkeit, da die Krümmung des Wellenfeldes ab einer Distanz von ca. 1.3 Metern für die Richtungsabschätzung nicht mehr relevant ist [LAVK06]. Dies hat den messtechnischen

Vorteil, dass man auf einer Kugelfläche mit festem Radius alle Richtungen messen kann. Die mathematische Schreibweise vereinfacht sich so zu

$$HRTF(\phi, \theta, f)$$

Gebräuchlich ist auch eine Unterscheidung zwischen dem richtungsabhängigen (Directional Transfer Function, DTF) und dem richtungsunabhängigen Teil (Common Transfer Function, CTF) einer HRTF. Der Letztere stellt eine Filterkurve dar, die allen HRTFs gemein ist und leicht durch Mittelung der Betragsspektren aller Übertragungspfade berechnet werden kann. Bereinigt man dann die HRTFs um den richtungsunabhängigen Teil, so bleibt nur der lokalisationsrelevante Teil übrig. Die HRTF kann also auch als Multiplikation dieser beiden Übertragungsfunktionen im Frequenzbereich angeschrieben werden:

$$HRTF(\phi, \theta, f) = CTF(f) \cdot DTF(\phi, \theta, f)$$

Anwendung und Standardisierung HRTFs können als Filter für binaurale Wiedergabesysteme verwendet werden um eine virtuelle 3D-Akustik über Kopfhörer zu erzeugen; jedes Signal wird dabei mit den Impulsantworten (Head-related Impulse Responses, HR-IRs) für die gewünschte Richtung gefaltet. Ein großes Einsatzgebiet ist darüber hinaus die Simulation und Nachbildung des menschlichen Gehörs in der Psychoakustik.

Um binaurale Wiedergabesysteme für die breite Masse zu produzieren, können standardisierte HRTFs verwendet werden, für akustische Messungen und Versuche sind diese als Richtwert ebenfalls nützlich. Häufig wird in diesem Feld mit einem Kunstkopf gearbeitet, einem Nachbau eines menschlichen Torsos und Kopfes aus Kunststoff, der mit Ohren und Mikrofonen an der Stelle des Trommelfells ausgestattet ist. Allerdings gibt es im frequenzabhängigen Verhalten der HRTFs starke interindividuelle Unterschiede, weswegen individuelle HRTFs eine wesentlich genauere Lokalisation ermöglichen als standardisierte. Die beste Methode für eine ausreichend genaue Nachbildung von individuellen HRTFs ist daher immer noch die Vermessung der betreffenden Person.

1.2 Direkte Messmethode

Die meist verwendete Technik zur Messung von HRTFs bedient sich sehr kleiner Mikrofone, die genau am Eingang des äußeren Gehörgangs platziert werden (siehe *Abb. 4*). Im Gegensatz zu Messungen mit dem oben erwähnten Kunstkopf wird hier nicht direkt am Trommelfell gemessen und der Gehörgang akustisch verschlossen, die (richtungsunabhängige) Resonanz des Gehörgangs ist daher nicht im Messergebnis enthalten, was z. B. für eine Kopfhörerwiedergabe auch sinnvoll ist. Rings um die Versuchsperson wird ein kugelförmiges Array aus Lautsprechern benötigt, der Kopf muss genau in der Mitte dieser Kugel liegen, sodass der Abstand zwischen Lautsprecher und Mikrofon für jeden Lautsprecher konstant und eine gleichförmige Winkelauflösung gewährleistet ist. Bei einer Dezentrierung des Kopfes kommt es zu Fehlern in der Erfassung der Laufzeiten. In der Praxis verwendet man keine Kugel, sondern zumeist kreisförmige Lautsprecher-Arrays, in denen die Versuchspersonen auf einer Drehscheibe sitzen und in kleinen Schritten um

die vertikale Achse gedreht werden (siehe Abb. 2(a)). Da die Messsignale ohnehin nacheinander abgespielt werden müssen um zu der jeweiligen Position zugeordnet werden zu können wäre eine kugelförmige Anordnung unnötig aufwändig und würde zudem noch zusätzliche Reflexionen hervorrufen.

Als Messsignal wird meist ein exponentieller Sweep verwendet, prinzipiell sind aber auch Messungen mit Maximum Length Sequence, Impulsen, weißem Rauschen oder anderen breitbandigen Signalen mit genügend Energie im relevanten Frequenzband möglich. Wichtig ist hierbei der Schalldruck, der auf die Mikrofonmembran eingepreßt werden kann; je höher dieser ist, desto höher ist der Signal-Rausch-Abstand (SNR) und desto verlässlicher das Messergebnis; Entsprechendes gilt auch für die Dauer eines Sweeps und die Anzahl der Wiederholungen einer Messung (mit anschließender Mittelung).

Um eine gute räumliche Abtastung zu erreichen, müssen extrem viele Übertragungspfade gemessen werden, wodurch die Messdauer leicht sehr groß wird. Während der Messung muss die Versuchsperson stets in der gleichen Haltung verharren, bereits kleine Bewegungen oder Abweichungen von der vorgesehenen Position haben relevante Auswirkungen.



Abbildung 4: In-Ear-Mikrofon

Der Messraum muss möglichst neutral, also reflexionsfrei oder zumindest -arm und nach außen hin akustisch isoliert sein. Um nicht jeden Sweep einzeln nacheinander abspielen zu müssen, wird z.B. am Wiener Institut für Schallforschung und Audiologie mit verschachtelten exponentiellen Sweeps gearbeitet. Die Sweeps werden dabei zu wohldefinierten Zeitpunkten nacheinander aber überlappend abgespielt und durch Fensterung der durch Entfaltung gewonnenen Impulsantworten im Nachhinein wieder getrennt [MBL06]. Dadurch lässt sich eine Beschleunigung der Messprozedur um den Faktor vier erreichen, ohne eine Verschlechterung des SNR in Kauf nehmen zu müssen.

Prinzipiell kann man die Messdauer aber als Produkt aus Messsignallänge und der Anzahl der Übertragungspfade errechnen, wobei je nach Auflösung im Azimut noch die Zeit für das Weiterdrehen berücksichtigt werden muss.

$$T_{ges} = T_{signal} \cdot N_{pos} + T_{rot} \cdot N_{rot} \quad (1)$$

T_{ges}	... Messdauer gesamt
T_{signal}	... Dauer des Messsignals
N_{pos}	... Anzahl der Übertragungspfade/Positionen
T_{rot}	... Dauer des Weiterdrehens
N_{rot}	... Anzahl der Positionen in der Horizontalebene

1.3 Reziproke Messmethode

Prinzip Die reziproke Messmethode bedient sich des Helmholtz'schen Prinzips der Reziprozität von Schallfeldern. Dieses besagt, dass in einem beliebig komplexen linearen und zeitinvarianten Schallfeld der an Punkt A gemessene Schalldruck - erzeugt von einer Schallquelle im Punkt B - der Gleiche ist, der an Punkt B gemessen würde wenn die Quelle im Punkt A läge [ZDGG06].

Es ist also möglich, in einem linearen, zeitinvarianten Schallfeld eine Vertauschung von Quelle und Senke vorzunehmen, ohne dass sich an dem Übertragungspfad zwischen beiden etwas ändert. Basierend auf dieser Erkenntnis kann man sich leicht ein reziprokes Messverfahren für HRTFs vorstellen, bei dem Quellen und Senken vertauscht werden: An Stelle von Mikrofonen werden sehr kleine Lautsprecher und statt eines Lautsprecherarrays ein Mikrofonarray verwendet, die Übertragungspfade werden also von innen nach außen statt von außen nach innen gemessen; das Ergebnis sollte identisch mit dem einer entsprechenden direkten Messung sein.

Vorteile Da bei solch einer Messanordnung nicht die Schallquellen (Lautsprecher), sondern die Schallsenken (Mikrofone) die räumliche Abtastung bestimmen, ist die Messdauer unabhängig von dieser. Im Gegensatz zu der direkten Messmethode ist es so möglich, alle HRTFs eines Ohres gleichzeitig zu messen. Zwar gilt auch hier, dass die Messsignale einzeln nacheinander abgespielt werden müssen, doch deren Anzahl reduziert sich auf zwei (eines für jedes Ohr) während unter Verwendung eines kugelförmigen Mikrofonarrays praktisch jede beliebige räumliche Abtastung realisiert werden kann ohne die Messdauer zu beeinflussen.

$$T_{ges} = 2 \cdot T_{signal} \quad (2)$$

Die Fixierung der Versuchsperson, die bei der direkten Messmethode meist mit aufwendigem Head-Tracking umgesetzt wird, ließe sich durch die geringere Messdauer deutlich erleichtern. Geht man beispielsweise von einem Messsignal mit einer Dauer von 30 Sekunden aus, so liegt die gesamte Messdauer bei gut einer Minute; ein Zeitraum, der das Stillhalten leichter möglich macht als bei der konventionellen Methode (ca. eine Stunde, je nach räumlicher Abtastung) und so auch die Vermessung von HRTFs mit verschiedenen Schulterdrehungen in den Bereich des Möglichen rückt.

Des Weiteren ist auf Grund der geringeren Dimensionen von Messmikrofonen gegenüber Lautsprechern von weniger Reflexionen durch das Array auszugehen.

Nachteile Die Lautsprecher müssen bei der reziproken Messanordnung am Eingang des äußeren Gehörgangs platziert werden, was eine Einschränkung ihrer Größe erfordert. Während Messmikrofone in der nötigen Größenordnung eine relativ gute Qualität erreichen, muss man bei einem Lautsprecher mit einem derart geringen Durchmesser (weniger als 1 cm) mit geringen Schalldrücken rechnen; dies betrifft insbesondere den tiefen Frequenzbereich. Die Gründe dafür sind die geringe Membrangröße und das kleine Luftvolumen hinter der Membran. Durch den geringen Schalldruck ist nur ein kleiner Signal-Rausch-Abstand (SNR) in der Messung erreichbar; dieser Umstand ist wohl das

größte Argument gegen die reziproke Messmethode. Generell ist jedoch bei allen Methoden zur HRTF-Vermessung die Forderung nach gutem SNR und Reflexionsfreiheit insbesondere im tiefen Frequenzbereich schwer zu erfüllen.

Beispiel Da die Methode nicht sehr verbreitet ist, gibt es leider wenig Vergleichsmöglichkeiten. Im Folgenden wird kurz auf eine Umsetzung des Verfahrens von Zotkin et al. an der University of Maryland [ZDGG06] eingegangen.

Hier wurde ein entsprechend kleiner Lautsprecher mit geschlossenem Gehäuse in Silikon-Ohrenstöpsel verbaut, was einerseits den Gehörgang verschließen soll um dessen Resonanz nicht mit zu messen, andererseits soll Messsignal zum Trommelfell hin bedämpft werden. Zum Anschluss des Lautsprechers wurden Drähte mit besonders geringem Durchmesser verwendet um das Messergebnis nicht zu verfälschen. Eine kleine kugelförmige Anordnung bestehend aus Teilen eines Chemiebaukastens und bestückt mit 32 Messmikrofonen diente als Mikrofonarray.



Abbildung 5: links: Mikrofonarray, unten: Microspeaker, oben Mitte: Microspeaker eingebaut, oben rechts: Messmikrofon [ZDGG06]

Als Messsignal wurde ein 96 Samples langer Impuls mit einer maximalen Frequenz von 16 kHz verwendet. Um eine Verbesserung des SNR zu erreichen wurde jeder Messvorgang 48 mal wiederholt und die Ergebnisse dann gemittelt. Um Störgeräusche zu vermeiden wurden die Messungen abends durchgeführt, Computer ausgelagert und die Lüftung abgeschaltet; der Messaufbau wurde zudem in der Mitte des Raumes platziert um die Ankunftszeit der ersten Reflexionen so weit wie möglich zu erhöhen.

Da die Lautsprecher erst ab ca. 1,5 kHz verlässliche Resultate lieferten wurde der untere Frequenzbereich mit einem analytischen Modell (HAT-Modell) erzeugt. Dies ist relativ einfach möglich, da für große Wellenlängen vor Allem die Interaktion zwischen der eintreffenden Schallwelle, Kopf und Torso entscheidend ist; die relativ kleinen Abmessungen der Pinna spielen erst bei höheren Frequenzen eine Rolle. Zwischen den gemessenen und den analytisch erzeugten HRTFs wurde zwischen 1 kHz und 3 kHz überblendet um das vollständige Ergebnis zu erhalten.

Der SNR nach Mittelung von 48 Messungen liegt bei diesem Verfahren bei 37.4 dB, Messungen mit einem Kunstkopf ergaben weitestgehende Übereinstimmungen mit Ergebnissen von direkten Messungen.

2 Hardware

2.1 Mikrofon-Array

Für diese Projektarbeit kann ein kugelförmiges Mikrofonarray, das von Fabian Hohl am IEM entwickelt wurde [Hoh09], eingesetzt werden. Ursprünglich für die Aufnahme und Analyse von richtungsabhängigen Abstrahlcharakteristika von Musikinstrumenten gedacht, eignet es sich auch gut für eine solche Anwendung.

Das Mikrofonarray besteht aus einer kugelförmigen Anordnung von Aluminiumstangen und Verbindungsscheiben, in deren Mitte jeweils die Messmikrofone sitzen. Die Verteilung der Mikrofone auf der Kugeloberfläche ist unregelmäßig, die räumliche Abtastung erfolgt mit 64 Mikrofonen (siehe *Abb. 2(b)*). Der Kugeldurchmesser beträgt ca. 2.7 m, bei Positionierung des Kopfes in der Kugelmitte ist also bei einem Abstand von ca. 1.3 m zwischen einem Ohr und dem nächsten Mikrofon von Fernfeldbedingungen auszugehen. Um die Kammfiltereffekte durch Reflexionen an den Verbindungsscheiben zu minimieren, sind diese mit Absorbieren bedämpft. Die Konstruktion ist an der Decke des Messraumes angebracht, der Raum selbst ist mit Dämmmaterial ausgekleidet und wird als Reflexionsarm angenommen.

2.2 Messlautsprecher

2.2.1 Beschreibung der Wandler

Für den Bau der Messlautsprecher wurden Kleinstlautsprecher vom Typ Pico 8 mm Receiver des Herstellers NXP verwendet (siehe *Abb. 6*). Diese sind mit einem Durchmesser von 8 mm, ihrer kreisrunden Form und einer Höhe von ca. 2 mm von den physischen Dimensionen her gut geeignet zur Verwirklichung des Messvorhabens. Allerdings handelt es sich nicht um professionelle Messlautsprecher, die eigentliche Anwendung liegt im Bereich der Mobiltelefone. Erste Versuche mit Lautsprechern von handelsüblichen Kopfhörerstöpseln scheiterten an deren Dimensionen, da ihr Durchmesser deutlich größer ist als der des Gehörgangs.



Abbildung 6: Pico 8mm Receiver

2.2.2 Entwurf und Bau

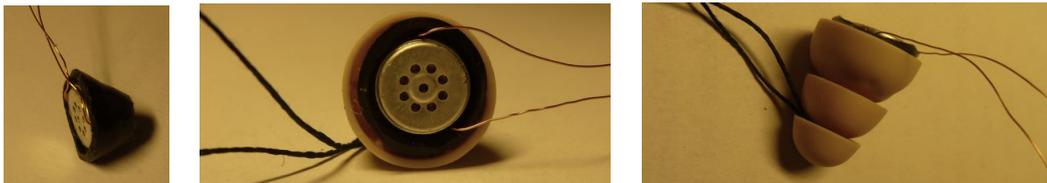
Beim Entwurf der Messlautsprecher sind im Wesentlichen vier Kriterien zu berücksichtigen:

1. Der Lautsprecher muss am Eingang des Gehörgangs sitzen
2. Der Gehörgang muss akustisch vollständig abgeschlossen werden
3. Der Schalldruck nach innen auf das Trommelfell soll möglichst klein sein
4. Der Schalldruck nach außen soll maximal sein

Nachdem der Lautsprecher an der Rückseite offen ist, also kein abgeschlossenes Chassis mit Hohlraum hinter der Membran besitzt, musste sichergestellt werden, dass ein möglichst großes, abgeschlossenes Luftvolumen hinter dem Lautsprecher Platz findet. Wie erwähnt soll der Lautsprecher am Eingang des Gehörgangs sitzen, d.h. das Luftvolumen dahinter befindet sich innerhalb des Gehörgangs, was seine potentielle Größe stark limitiert.

Die letztlich verwendeten Materialien sind ein Paar Spritzwasserschutz-Silikonohrenstöpsel, deren äußere Enden entsprechend zugeschnitten wurden, sowie konisch geformte Gummihülsen, die durch richtigen Zuschnitt auf den Durchmesser der Lautsprecher gebracht wurden. Als Lautsprecherkabel wurden normale, sehr dünne Spulendrähte verwendet. Die Fixierung des Lautsprechers in der Gummihülse und die der Gummihülse auf dem Ohrenstöpsel wurden mit verschiedenen Klebstoffen realisiert. Um den Messlautsprecher problemlos wieder aus dem Ohr entfernen zu können, ist am Ohrenstöpsel ein Bindfaden befestigt, an dem dieser leicht herausgezogen werden kann (siehe *Abb. 7*).

Bei einer Messung zur Ermittlung der unteren Grenzfrequenz wurde festgestellt, dass die Messlautsprecher bei Vollaussteuerung erst ab ca. 400Hz erkennbare Signale liefern (siehe *Abb. 8*).



(a) Wandler in Gummihülse (b) Messlautsprecher - Frontansicht (c) Messlautsprecher - Seitenansicht

Abbildung 7: Messlautsprecher

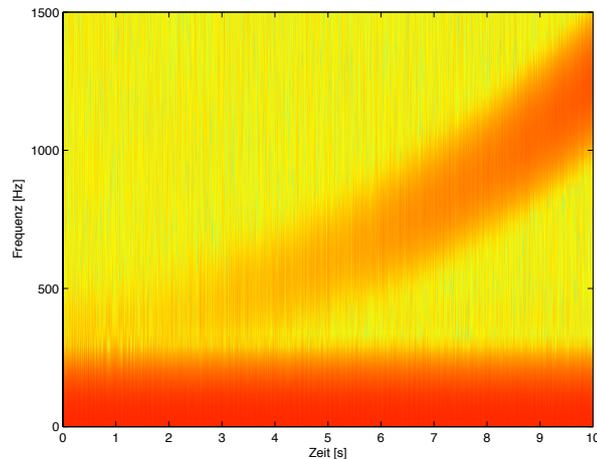


Abbildung 8: Sweepantwort des Messlautsprechers - Spektrogramm (tiefe Frequenzen)

3 Übertragungsverhalten der Messlautsprecher

Das Übertragungsverhalten der Messlautsprecher ist ein wesentlicher Faktor bei der möglichst genauen Ermittlung von HRTFs. Es wird wegen der geringen Dimensionen und der Flexibilität der Gummihülse möglicherweise auch durch die Einbausituation im Ohr bestimmt; daher stellt sich die Frage, ob sich überhaupt ein generell gültiges Übertragungsverhalten ermitteln lässt und wie sich dieses gegebenenfalls messen lässt. Wichtig ist auch die Richtcharakteristik, die Messlautsprecher sollten in der vorderen Halbkugel ein omnidirektionales Abstrahlungsverhalten aufweisen um verlässliche Ergebnisse erzielen zu können.

3.1 Bündelungsverhalten

Zunächst wurde eine Messung zur Überprüfung des Abstrahlungsverhaltens der beiden Messlautsprecher durchgeführt. Dazu wurden diese in einem Stück schalldämmenden Schaumstoffs fixiert und gleichzeitig mit vier Messmikrofonen vermessen. Die Mikrofone wurden in 90°-Schritten rings um den Lautsprecher positioniert und in einem Winkel von 45° relativ zur Lautsprechermembran ausgerichtet.

Als Messsignal wurde ein exponentieller Sweep mit einer Länge von 30 Sekunden verwendet, als Startfrequenz wurde $f_{start} = 300\text{Hz}$ und als Stopfrequenz $f_{stop} = 22.05\text{kHz}$ bei einer Samplingrate von $f_s = 44.1\text{kHz}$ gewählt. Um bei der Division der Spektren der Sweepantworten durch das Anregungsspektrum den Bereich unter 300Hz , in dem ohnehin nur Rauschen zu erwarten ist, nicht unnötig anzuheben, wurden die Sweepantworten vorher mit einem Hochpass 20ter Ordnung und Grenzfrequenz $f_g = 300\text{Hz}$ gefiltert. Hierbei wurden 10 kaskadierte Filter zweiter Ordnung (second order sections) verwendet um trotz der steilen Flanke die Stabilität des Filters zu gewährleisten.



(a) Messaufbau von oben

(b) Messaufbau seitlich

Abbildung 9: Messaufbau zur Messung des Bündelungsverhaltens

Wie an den untenstehenden Grafiken (*Abb. 10, Abb. 11*) deutlich wird, kann das Übertragungsverhalten der Messlautsprecher als näherungsweise gleich für die unterschiedlichen Richtungen betrachtet werden. Obgleich durch den Messaufbau lediglich die Rotationsymmetrie des Abstrahlungsverhaltens nachgewiesen wird, wurde die Bedingung der halbkugelförmigen Omnidirektionalität im Folgenden als ausreichend genau erfüllt angenommen. Die Richtigkeit dieser Annahme wird auch durch die Übereinstimmung mit den frontal gemessenen Frequenzgängen der Messlautsprecher bestätigt (siehe *Abb. 14*), allerdings lediglich bezüglich des Frequenzgangs, nicht des absoluten Schalldruckpegels.

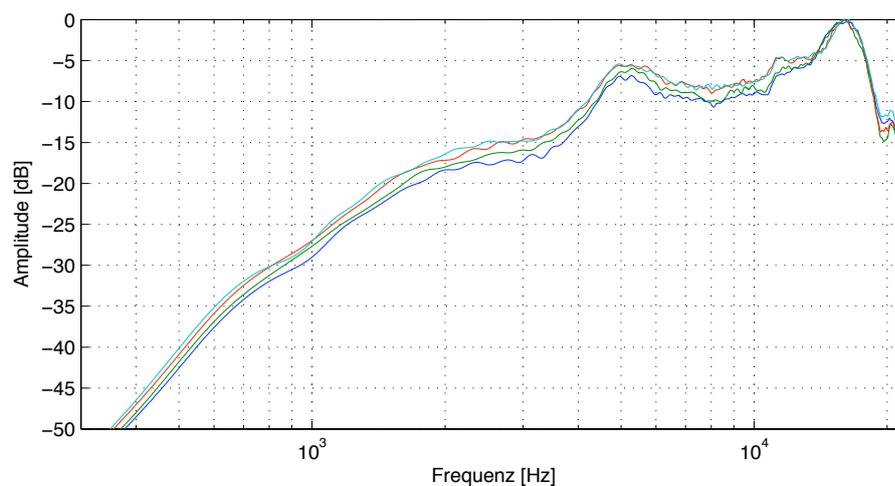


Abbildung 10: Frequenzgänge aus unterschiedlichen Richtungen - linker Messlautsprecher

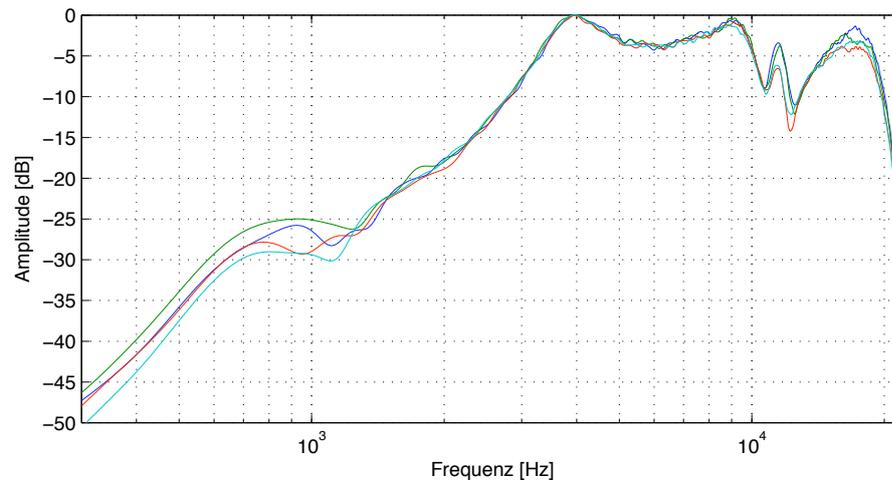


Abbildung 11: Frequenzgänge aus unterschiedlichen Richtungen - rechter Messlautsprecher

3.2 Frequenzgänge

3.2.1 Ursprüngliches Vermessungskonzept

Wegen der oben erwähnten möglichen Abhängigkeit des Übertragungsverhaltens von der Einbausituation wurde ein Vermessungskonzept entwickelt, mit dem einerseits die Unterschiede zwischen unterschiedlichen Einbausituationen eruiert und andererseits wenn möglich die Frequenzgänge der Messlautsprecher im eingebauten Zustand ermittelt werden sollten.

Als Ersatz für den menschlichen Körper diente ein Kunstkopf, es ergab sich allerdings ein Problem bei der Vermessung der im Kunstkopf-Ohr befindlichen Messlautsprecher: Misst man nämlich deren Übertragungsverhalten, so wird diesem unweigerlich die HRTF der entsprechenden Richtung aufgeprägt. Daher sollte zunächst die zugehörige HRTF mittels einer Referenzmessung mit der direkten Methode ermittelt werden. Damit die reziproke Vermessung das selbe Ergebnis liefert, wird der Unterschied der Frequenzgänge zur Entzerrung verwendet.

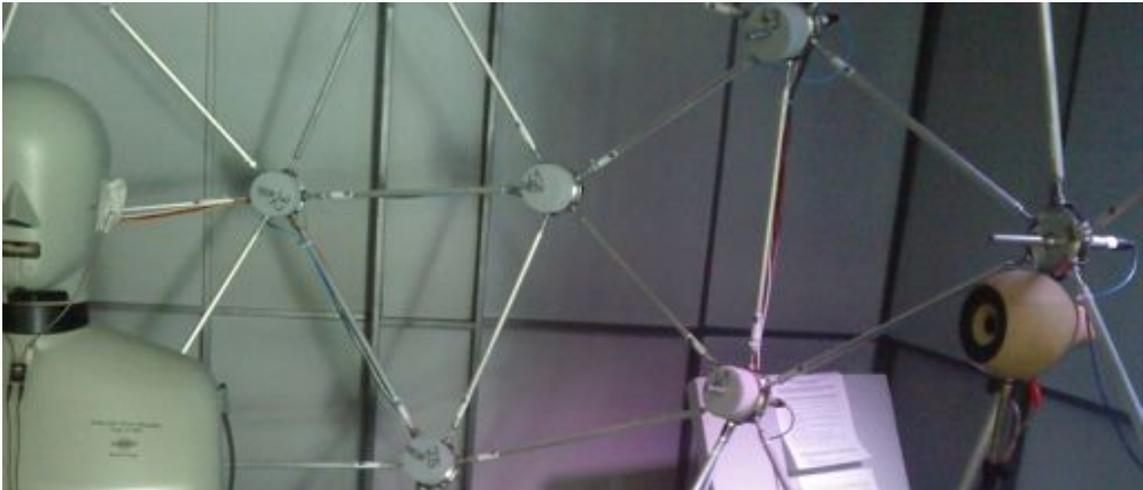


Abbildung 12: Messaufbau

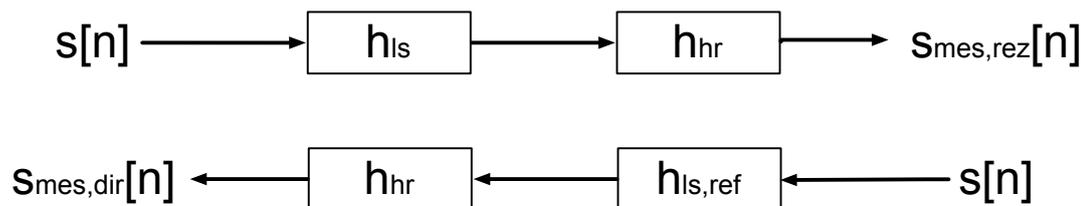


Abbildung 13: Signalflussplan des Vermessungskonzepts

Wie Anhand des Signalflussplans deutlich wird, ergeben sich die gemessenen Sweepantworten der reziproken ($s_{mes,rez}[n]$) und der Referenzmessung ($s_{mes,dir}[n]$) zu

$$s_{mes,rez}[n] = s[n] * h_{ls} * h_{hr} \quad , \quad (3)$$

$$s_{mes,dir}[n] = s[n] * h_{ls,ref} * h_{hr} \quad , \quad (4)$$

bzw. im Frequenzbereich zu

$$S_{mes,rez}(f) = S(f) \cdot H_{ls}(f) \cdot H_{hr}(f) \quad , \quad (5)$$

$$S_{mes,dir}(f) = S(f) \cdot H_{ls,ref}(f) \cdot H_{hr}(f) \quad . \quad (6)$$

$s[n]$ steht dabei für das Messignal, h_{ls} für die Impulsantwort des Messlautsprechers, $h_{ls,ref}$ für die Impulsantwort des Referenzlautsprechers und h_{hr} für die kopfbezogene Impulsantwort.

Der Frequenzgang des Messlautsprechers sollte sich also (bei bekanntem Übertragungsverhalten des Referenzlautsprechers) auf folgende Weise ermitteln lassen:

$$H_{ls}(f) = \frac{S_{mes,rez}(f) \cdot H_{ls,ref}(f)}{S_{mes,dir}(f)} = \frac{S(f) \cdot H_{ls}(f) \cdot H_{hr}(f) \cdot H_{ls,ref}(f)}{S(f) \cdot H_{ls,ref}(f) \cdot H_{hr}(f)} \quad (7)$$

Leider erwies sich dies als zu fehleranfällig, die gemessenen Frequenzgänge unterscheiden sich drastisch von denen der Bündelungsmessung, was möglicherweise auf die leicht unterschiedlichen Positionen von Messmikrofon und Lautsprecher zurückzuführen ist. Eine wichtige Erkenntnis dieser Messung war dennoch, dass sich das Übertragungsverhalten durch zwischenzeitlichen Ein- und Ausbau der Messlautsprecher nicht stark ändert, man also einen generell gültigen Frequenzgang näherungsweise annehmen kann.

3.2.2 Angewandtes Vermessungskonzept

Die tatsächlich verwendeten Frequenzgänge der Messlautsprecher wurden auf Grund der Probleme mit der oben erwähnten Methode analog zu der Bündelungsmessung durchgeführt, allerdings mit nur einem Messmikrofon, welches frontal auf den Messlautsprecher gerichtet war.

Als Messsignal wurde ein exponentieller Sweep mit einer Länge von 30 Sekunden verwendet, als Startfrequenz wurde $f_{start} = 300\text{Hz}$ und als Stopfrequenz $f_{stop} = 22.05\text{kHz}$ bei einer Samplingrate von $f_s = 44.1\text{kHz}$ gewählt. Wie bei der Messung zum Bündelungsverhalten wurde mit einem Hochpassfilter 20ter Ordnung das Rauschen in den Mikrofonsignalen unterhalb von 300 Hz gedämpft. Für jeden Messlautsprecher wurden vier Messungen durchgeführt und die Sweepantworten im Zeitbereich gemittelt.

$$s_{mes,ges}[n] = \frac{\sum_{i=1}^4 s_{mes,i}[n] * h_{hp}}{4} \quad (8)$$

$$H_{ls}(f) = \frac{S_{mes,ges}(f)}{S(f)} \quad (9)$$

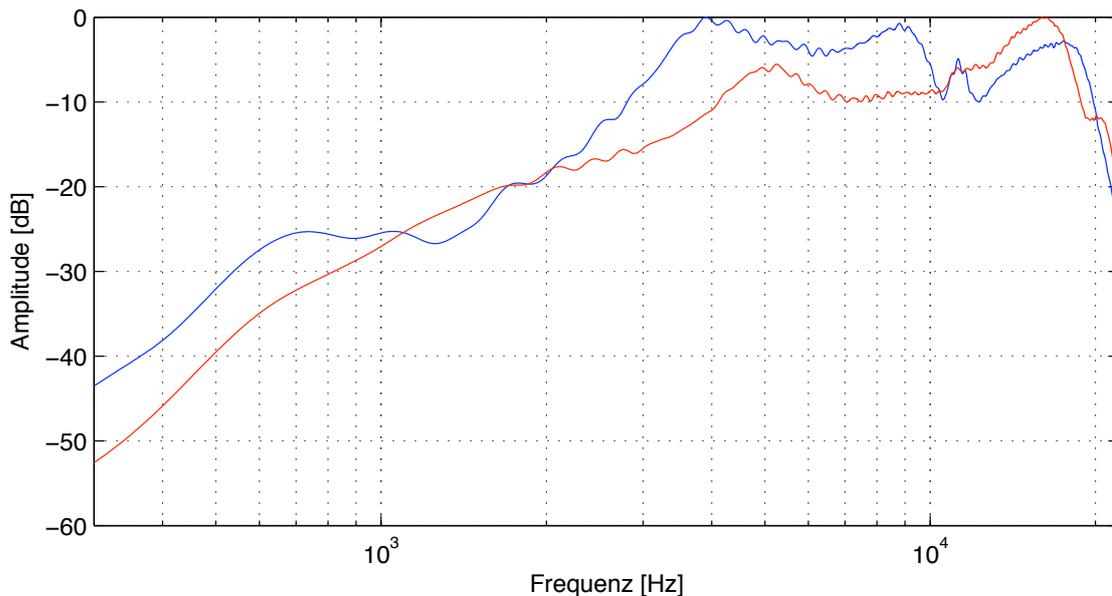


Abbildung 14: Frequenzgänge des linken (rot) und rechten (blau) Messlautsprechers

4 Entzerrung der Messlautsprecher

4.1 Erstellung der Entzerrungfilter

Die Frequenzgänge der Entzerrungfilter ergeben sich aus der Inversion der Messlautsprecherfrequenzgänge. Wie bereits erwähnt, wurde die untere Grenzfrequenz für die reziproke HRTF-Vermessung auf $f_{g,u} = 500\text{Hz}$ festgelegt, eine obere Grenzfrequenz allerdings noch nicht. Da die Amplituden der Messlautsprecher Frequenzgänge ab ca. 19 kHz sehr steil abfallen und die meisten Menschen in diesem Bereich ohnehin nichts mehr hören, wurde die obere Grenzfrequenz auf $f_{g,o} = 19\text{kHz}$ festgesetzt. Somit ist der Bereich, in dem relevante Messergebnisse zu erwarten sind auf 500 Hz bis 19 kHz eingeschränkt. Es ist aus diesem Grund nun nicht mehr nötig, außerhalb dieses Bereichs den Lautsprecher zu entzerren. Der invertierte Lautsprecherfrequenzgang wurde daher bei 500 Hz mit einem Hochpass 20ter Ordnung und bei 19 kHz mit einem Tiefpass 12ter Ordnung gefiltert. Um hierbei möglichst steile Flanken bei hoher Stabilität der Filter zu gewährleisten, wurde wiederum mit kaskadierten Filtern zweiter Ordnung gearbeitet.

$$H_{eq}(f) = \frac{1}{H_{ls}(f)} \cdot H_{hp}(f) \cdot H_{tp}(f) \quad (10)$$

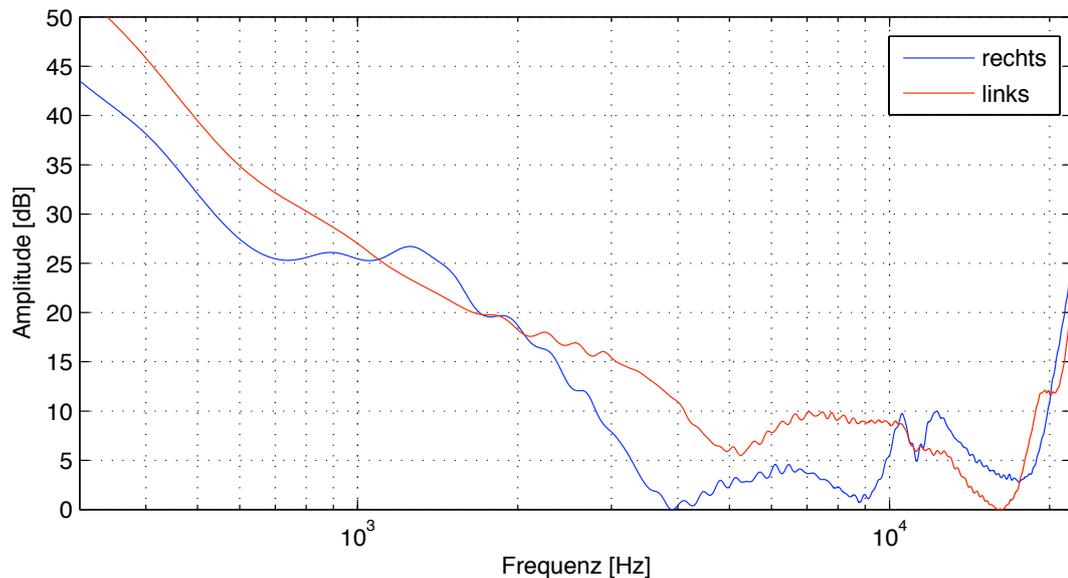


Abbildung 15: Frequenzgänge der Entzerrungfilter

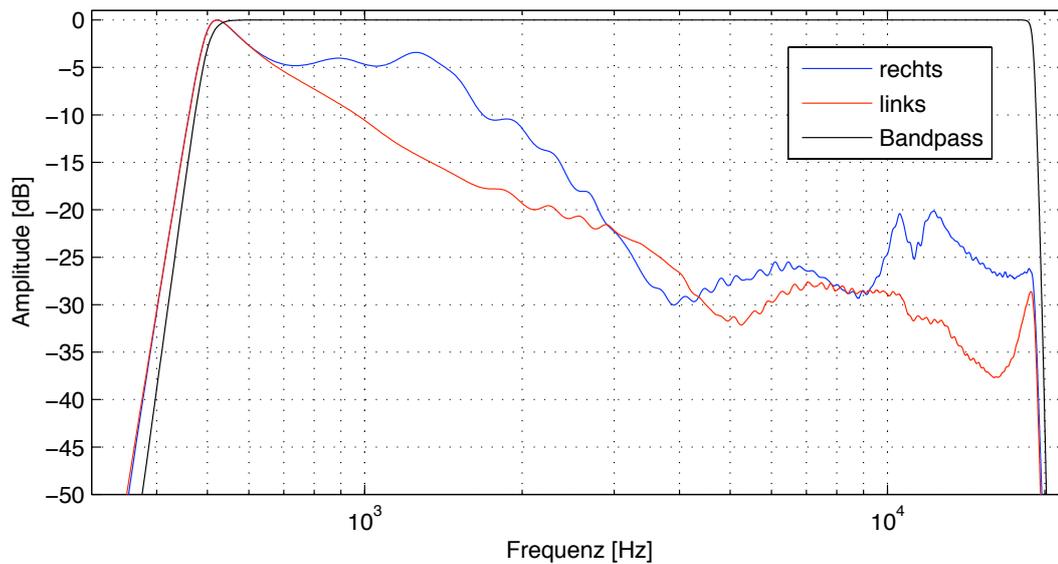


Abbildung 16: Frequenzgänge der Entzerrungsfiler mit Bandpass (normalisiert)

4.2 Glättung der Entzerrungsfiler

Betrachtet man die Frequenzgänge in *Abb. 15* und *Abb. 16*, so ist neben den großen Amplitudenunterschieden in Abhängigkeit von der Frequenz auch zu beobachten, dass die Verläufe stellenweise relativ unruhig sind. Um die Entzerrungsfiler nicht zu komplex zu gestalten und auch in Anbetracht der Tatsache, dass es ohnehin je nach Einbausituation feine Änderungen im Übertragungsverhalten der Messlautsprecher gibt, ist daher eine Glättung der Frequenzgänge sinnvoll.

4.2.1 Terzbandglättung

Nachdem das menschliche Gehör keine äquidistante Frequenzauflösung hat, sondern eher logarithmisch arbeitet, erscheint eine lineare Glättung nicht ideal; die Genauigkeit wäre bei tiefen Frequenzen zu niedrig und bei hohen Frequenzen unnötig hoch. Eine Glättung in Terzbändern hingegen ist dem Gehör deutlich besser angepasst.

Die Oktave einer bestimmten Frequenz ergibt sich aus der Verdopplung derselben. Eine (große) Terz ist genau ein Drittel einer Oktave. Will man also von einer bestimmten Bezugsfrequenz f_{ref} aus die Terz darüber ermitteln, so muss die Bezugsfrequenz mit dem Faktor $2^{\frac{1}{3}}$ multipliziert werden:

$$f_{Oktave} = f_{ref} \cdot 2^1 \quad \Rightarrow \quad f_{Terz\uparrow} = f_{ref} \cdot 2^{\frac{1}{3}} \quad (11)$$

Entsprechendes gilt auch, wenn man von der Bezugsfrequenz aus eine Terz nach unten berechnen möchte, der Exponent wird dabei negativ:

$$f_{Terz\downarrow} = f_{ref} \cdot 2^{-\frac{1}{3}} \quad (12)$$

Die Terzbandglättung wurde nach dem Prinzip eines Moving-Average-Filters realisiert, der Terzbänder zur Mittelung für den jeweiligen Frequenzbin heranzieht. Dadurch vergrößert sich der Bereich, über den gemittelt wird, bei höheren zur Frequenzen. Für tiefe Frequenzen werden nur wenige umliegende Frequenzbins in die Mittelung miteinbezogen, für hohe Frequenzen deutlich mehr. Dies hat den gewünschten Effekt, dass die Genauigkeit von tiefen zu hohen Frequenzen hin abnimmt, die Glättung also mit steigender Frequenz immer stärker greift. Auf einer logarithmischen Frequenzskala betrachtet ist die Genauigkeit im gesamten Bereich gleichbleibend.

Der geglättete Frequenzgang $H_{eq,gl}[k]$ ergibt sich nun folgendermaßen:

$$H_{eq,gl}(k) = \frac{\sqrt{\sum_{k_{min}}^{k_{max}} W_K(k' - k_{min}) \cdot |H_{eq}(k')|^2}}{\sum_{i=1}^K W_K(i)} \quad (13)$$

$$k_{min} = k \cdot 2^{-\frac{1}{6}}$$

$$k_{max} = k \cdot 2^{\frac{1}{6}}$$

$$W_K(k) = \frac{1}{2} \cdot [1 + \cos(\frac{2\pi k}{K})]$$

$$K = k_{max} - k_{min} + 1$$

Für jeden Frequenzbin des diskreten Spektrums werden im gemittelten, quadrierten Betragsfrequenzgang $|H_{eq}[k]|^2$ des jeweiligen Messlautsprechers die Amplituden aller Frequenzbins zwischen einer halben Terz darunter und einer halben Terz darüber aufsummiert, anschließend durch die Anzahl der verwendeten Bins geteilt und die Wurzel gezogen. Um dem betrachteten Frequenzbin am meisten Gewicht zu verleihen wurden alle Bins mit einem terzbreiten Hann-Fenster multipliziert, das um den betrachteten Bin zentriert ist.

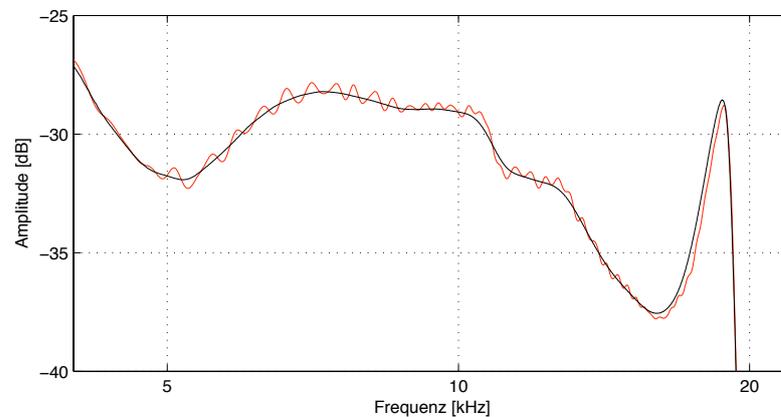


Abbildung 17: Frequenzgang des linken Entzerrungsfilters: unbearbeitet (rot) und geglättet (schwarz)

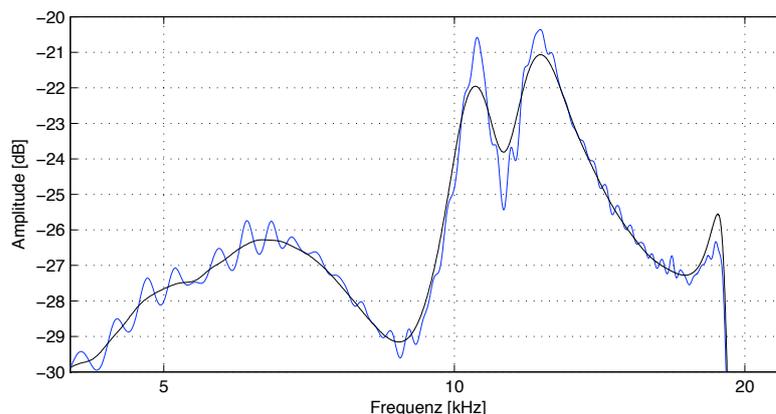


Abbildung 18: Frequenzgang des rechten Entzerrungsfilters: unbearbeitet (blau) und geglättet (schwarz)

4.2.2 Phasengang

Die Terzbandglättung wurde mit den Betragsspektren der Entzerrungsfiler durchgeführt, die Phaseninformation wurde also verworfen. Die geglätteten Entzerrungsfiler sind ebenfalls Betragsspektren, also positiv und reellwertig; transformiert man sie in den Zeitbereich, so erhält man lange, nichtkausale Impulsantworten. Um den Phasengang einer kausalen Impulsantwort für die Entzerrungsfiler zu erhalten, muss aus der Amplitudeninformation eine geeignete Phase berechnet werden.

Unter der Annahme, dass die Messlautsprecher minimalphasige Systeme sind, d.h. sowohl Nullstellen als auch Polstellen der Übertragungsfunktion innerhalb des Einheitskreises liegen, ist eine Erstellung der Phaseninformation über das Cepstrum möglich:

$$h_{eq,gl,ceps}[n] = \begin{cases} 2 \cdot IDFT\{\ln[|H_{eq,gl}[k]|]\} & , \text{ für } 0 < n \leq \frac{K}{2}, \\ 1 \cdot IDFT\{\ln[|H_{eq,gl}[k]|]\} & , \text{ für } n = 0 \text{ und } n = \frac{K}{2}, \\ 0 & , \text{ sonst} \end{cases} \quad (14)$$

$$H_{eq,gl,mp}[k] = e^{DFT\{h_{eq,gl,ceps}[n]\}} \quad (15)$$

$$h_{eq,gl}[n] = IDFT\{H_{eq,gl,mp}[k]\} \quad (16)$$

Die Indizes stehen dabei für Cepstrum (*ceps*) und minimalphasig (*mp*). Auf diese Weise wurden die minimalphasigen und daher kausalen Impulsantworten der Entzerrungsfiler erstellt (siehe Abb. 19(a), Abb. 19(b)). Eine wichtige Eigenschaft von minimalphasigen Systemen ist ihre Invertierbarkeit. Sie spielt bei der Entzerrung der Messlautsprecher eine Rolle. Inverse minimalphasige Systeme behalten die Eigenschaft der Minimalphasigkeit bei.

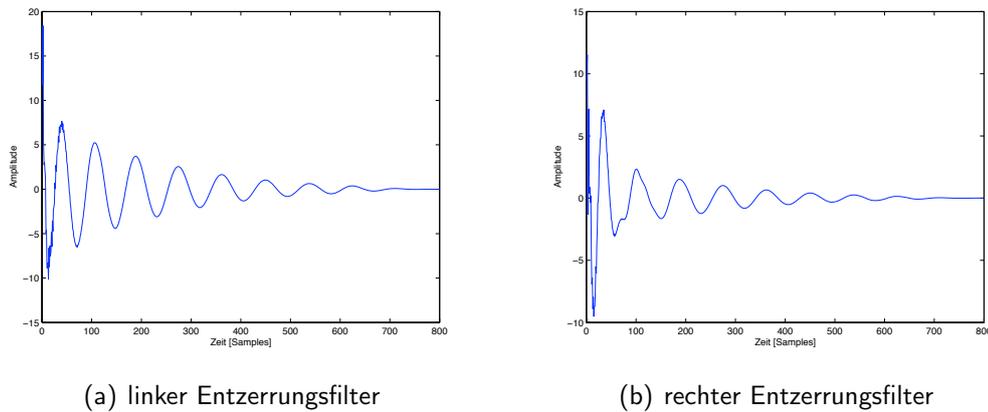


Abbildung 19: Minimalphasige Impulsantworten der Entzerrungsfiler

5 Anpassung des Messsignals

5.1 Grundsätzliche Überlegungen zum SNR - Motivation

Bei der reziproken HRTF-Vermessung ist auf Grund der Lautsprecherdimensionen der Pegel des Messsignals stark eingeschränkt, daher ist auch der Signal-Rausch-Abstand das zentrale Problem dieser Messmethode. Dieser ist definiert als das Verhältnis von Nutzsignal- zu Rauschleistung

$$SNR(f) = E \left\{ \frac{|S(f)|^2}{|N(f)|^2} \right\} = \frac{|S(f)|^2}{E \{|N(f)|^2\}} \quad (17)$$

Um den SNR bei gleichbleibendem Pegel zu erhöhen, kann man bei der Verwendung von Sweeps als Messsignal die Dauer erhöhen - bei einer Verdopplung der Dauer erhöht sich der Anteil des (unkorrelierten) Rauschens um 3 dB, der des (hochkorrelierten) Signals jedoch um 6 dB, was einer Vergrößerung des SNR um 3 dB entspricht. Das Gleiche gilt auch für die Mittelung mehrerer Messergebnisse, ein Verdopplung der Anzahl an Messungen bringt 3 dB SNR-Zuwachs.

Das Gesamtsystem der Messkette stellt sich wie in *Abb. 20* zu sehen ist dar, $n[n]$ steht für das Grundrauschen.

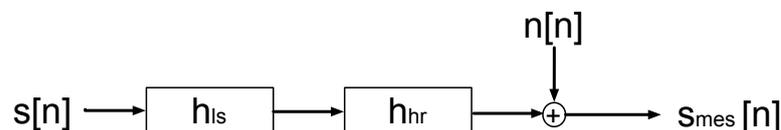


Abbildung 20: Messkette

Das Grundrauschen eines Raumes ist üblicherweise eher tieffrequent, d.h. der SNR nimmt mit abnehmender Frequenz abhängig von den genauen spektralen Verteilungen des Rauschens und des Messsignals ab. Arbeitet man nun mit Lautsprechern, die bei tiefen

Frequenzen zudem sehr schlecht übertragen, so wird der SNR in diesem Bereich noch zusätzlich verschlechtert. Da die Impulsantworten h_{ls} der Messlautsprecher bekannt sind und sich eine Referenz des Grundrauschens $n[n]$ leicht ermitteln lässt, kann beim Entwurf des Messsignals Rücksicht darauf genommen werden. Der ebenfalls vorhandene Einfluss der kopfbezogenen Impulsantworten h_{hr} auf den SNR wird hier nicht kompensiert, da diese noch als unbekannt behandelt werden müssen.

Um die Frequenzabhängigkeit des Hintergrundrauschens und die Übertragungsfunktionen der Messlautsprecher in der Messung zu kompensieren, sollte das Messsignal den Wunsch nach einem konstanten SNR des Betrages K erfüllen:

$$SNR(f) = \frac{|S(f) \cdot H_{ls}(f)|^2}{E\{|N(f)|^2\}} \stackrel{!}{=} K \quad \implies \quad |S(f)|^2 = K \cdot \frac{E\{|N(f)|^2\}}{|H_{ls}(f)|^2} \quad (18)$$

$$|S(f)|^2 = K \cdot E\{|N(f)|^2\} \cdot |H_{eq}(f)|^2 \quad (19)$$

Darin ist $E\{|N(f)|^2\}$ das Leistungsdichtespektrum des Grundrauschens. $H_{eq}(f)$ entzerrt den Frequenzgang des Messlautsprechers $H_{ls}(f)$, damit bei der Messung der akustischen Übertragungspfade alle Frequenzen gleichmäßig angeregt werden. Darüberhinaus sind die Ausdrücke K und $E\{|N(f)|^2\}$ für die Anpassung des Messsignals an das Hintergrundgeräusch verantwortlich. Ein gewisser Fehler in Form von Fluktuationen wird sich auf Grund des Unterschiedes zwischen $N_{ref}(f)$ und $N(f)$ immer ergeben, aber in der statistischen Betrachtung wird der Frequenzverlauf des Rauschens grob korrigiert (siehe *Abb. 21*, *Abb. 22*). Die reelle Konstante K und damit die Höhe des annähernd konstanten SNR ist abhängig von der Signalamplitude, der Länge des Messsignals sowie der Anzahl der gemittelten Messungen.

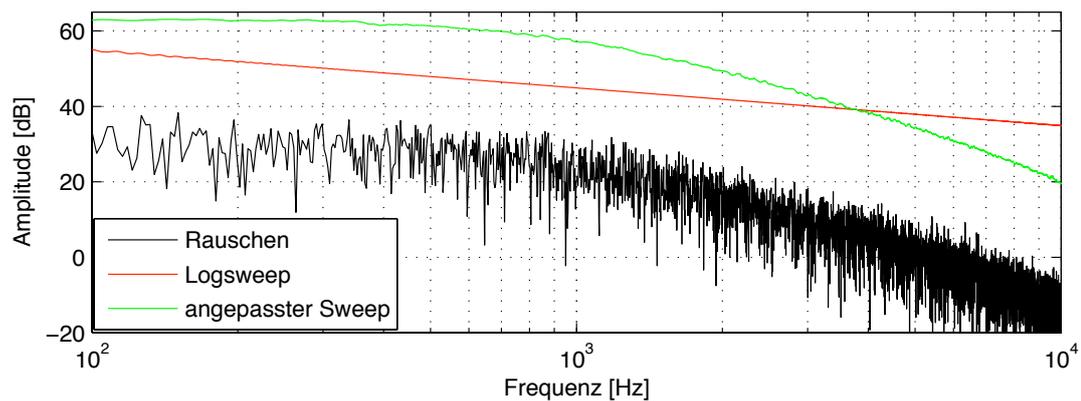


Abbildung 21: Beispielhafte spektrale Verteilungen von Rauschen und Messsignalen

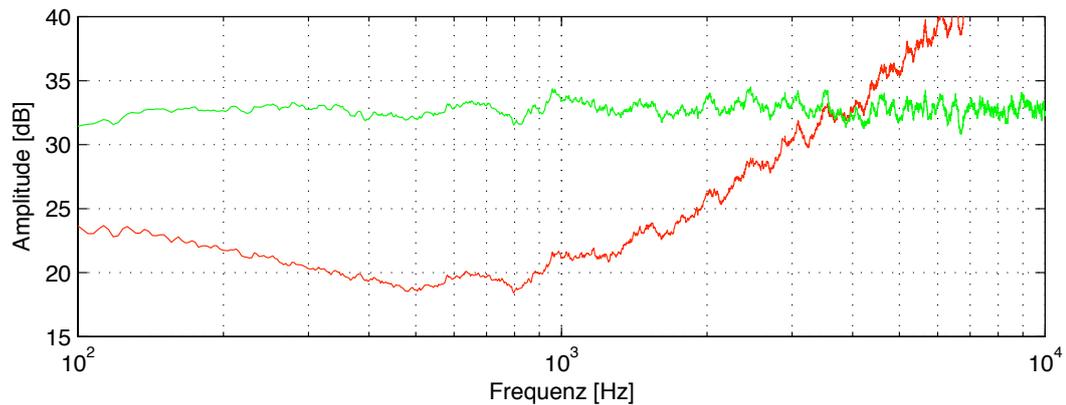


Abbildung 22: Aus Abb. 21 resultierende SNR-Verläufe

5.2 Angepasste Sweepgenerierung

Wie zuvor beschrieben, wird in Messungen meist ein exponentieller Sweep verwendet. Wird dieses Signal direkt mit der Lautsprecherentzerrung (Abb. 16) beaufschlagt, sieht man extreme Pegelschwankungen (Abb. 23). Da die Messlautsprecher bereits an der Aussteuerungsgrenze betrieben werden müssen, ist jedoch eine Steigerung der Vorverstärkung unmöglich. Man würde durch die dargestellte Amplitudenentzerrung des Zeitsignals also einen geringeren Amplitudenwert wählen müssen und dabei den möglichen SNR nicht voll ausschöpfen.

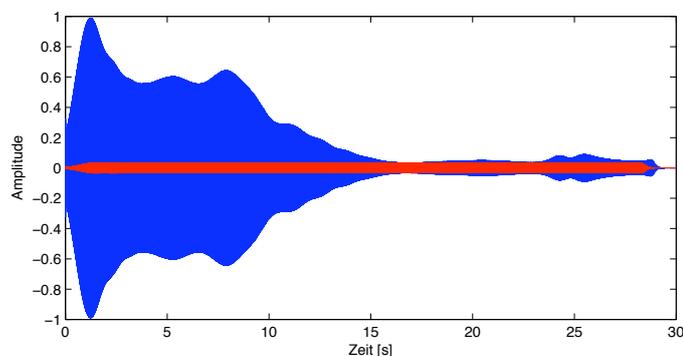


Abbildung 23: Mit Lautsprecher-Entzerrung beaufschlagter Logsweep vor (blau) und nach der Messung (rot)

5.2.1 Prinzip und Algorithmus

Da bei einem Sweep die Energie in einem bestimmten Frequenzband abhängig von seiner Verweildauer in diesem Band ist, lässt sich durch die Vorgabe von Gruppenlaufzeiten, deren Integration über die Frequenz die Phase liefert, ein Sweep mit beliebigem Betragspektrum bei konstanter Amplitude des Zeitsignals erzeugen.

Zur Anpassung der Sweeps auf die Spektren der Entzerrungsfiler kam ein Algorithmus zum Einsatz, der in einem Report von Swen Müller und Paolo Massarani beschrieben wird [MM06]. Dieser basiert auf der einfachen Proportionalitätsbedingung, dass die Gruppenlaufzeit $T_G(f)$ eines bestimmten Frequenzbandes sich zur Gesamtdauer T des Sweeps so verhält, wie seine Energie im entsprechenden Frequenzband zu seiner gesamten spektralen Energie:

$$\frac{T_G(f)}{T} = \frac{|H_{ref}(f)|^2}{\int |H_{ref}(f)|^2 df} . \quad (20)$$

Der Kern des Algorithmus' bestimmt eine vom gewünschten diskreten Betragsspektrum $|H_{ref}[k]|$ abhängige Folge von Gruppenlaufzeiten $T_G[k]$ für jeden Frequenzbin k

$$T_G[k] = T_G[k - 1] + C \cdot |H_{ref}[k]|^2 , \quad (21)$$

wobei die Konstante C nach geeigneter Wahl der Gruppenlaufzeiten für den ersten und den letzten Frequenzbin berechnet wird, N ist dabei die gewünschte Länge des zu erzeugenden Sweeps in Samples und ein Minimalwert für die Länge der inversen diskreten Fouriertransformation.

$$C = \frac{T_G[N/2] - T_G[0]}{\sum_{k'=0}^{N/2} |H_{ref}[k']|^2} = \frac{T}{\sum_{k'=0}^{N/2} |H_{ref}[k']|^2} \quad (22)$$

Hat man die Gruppenlaufzeiten $T_G[k]$ ermittelt, kann daraus durch Integration, bzw. Summation im frequenzdiskreten Bereich, die Phase $\Phi[k]$ des Sweeps berechnet werden

$$\Phi[k] = - \sum_{k'=0}^k T_G[k'] , \quad (23)$$

allerdings muss danach noch sichergestellt werden, dass der Phasenwinkel $\phi[N/2] = 0$ ist (der Frequenzbin $N/2$ repräsentiert $f_s/2$, die halbe Abtastrate), damit kein komplexwertiges Signal entsteht. Das wird durch einen schleichenden Phasenoffset über den gesamten positiven Frequenzbereich realisiert

$$\Phi_0[k] = \Phi[k] - \frac{k}{N/2} \cdot \text{MOD}_{2\pi}(\phi[N/2]) . \quad (24)$$

Das komplexe Sweepspektrum $S[k]$ berechnet sich dann aus dem vorgegebenen Betragsspektrum $|H_{ref}[k]|$ und dem ermittelten Phasengang

$$S[k] = |H_{ref}[k]| \cdot e^{j2\pi\Phi_0[k]} , \quad 0 \leq k \leq \frac{N}{2} , \quad (25)$$

und kann durch Ergänzung der konjugiert symmetrischen Komponenten bei negativen Frequenzen

$$S[k] = S^*[N - k] \quad \frac{N}{2} + 1 \leq k \leq N - 1 \quad (26)$$

auf den gesamten Frequenzbereich erweitert werden. Anschließend wird $S[k]$ durch inverse Fourier-Transformation in den Zeitbereich gebracht

$$s[n] = \frac{1}{N} \sum_{k=0}^{N-1} S[k] \cdot e^{j \frac{2\pi}{N} nk} . \quad (27)$$

Das Ergebnis ist ein Sweep, der im Zeitbereich eine annähernd konstante Amplitude aufweist; die Dauer des Sweeps ist abhängig von der Frequenzauflösung des vorgegebenen Spektrums und kann frei gewählt werden. Sein Betragsfrequenzgang entspricht der Vorgabe $|H_{ref}(f)|$.

5.2.2 Anwendung

Die beiden Sweeps für die HRTF-Vermessung wurden nach dem oben beschriebenen Prinzip mit einer Samplingfrequenz von 44.1 kHz und einer Dauer von 30 Sekunden erstellt, als Referenzspektrums wurden die geglätteten Entzerrungsfiler $H_{eq,gl}(f)$ verwendet. Die Anpassung des Messsignals an das Grundrauschen des Messraumes wurde verworfen, weil der Frequenzgang der Messlautsprecher unterhalb von 500 Hz wenig Signal liefert und das Grundrauschen größtenteils unter 500 Hz liegt (siehe *Abb. 8*).

Die Entzerrungsfiler wurden in Form der minimalphasigen Impulsantworten $h_{eq,gl}$ an den Algorithmus übergeben, um dann durch entsprechende Einstellung der gewünschten Dauer T des Sweeps die DFT-Länge N bestimmen zu können; f_s bezeichnet die Abtastfrequenz.

$$N = T \cdot f_s \quad (28)$$

Um Zeit für Ein- und Ausschwingvorgänge zu lassen wurde am Anfang und am Ende des Zeitsignals jeweils eine halbe Sekunde angefügt.

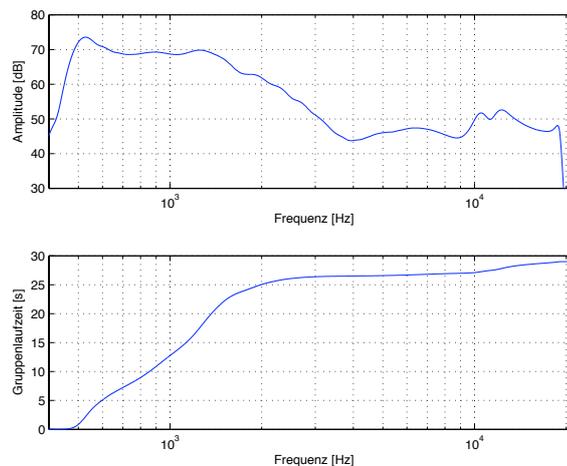


Abbildung 24: Frequenzgang und Gruppenlaufzeit des rechten Sweeps

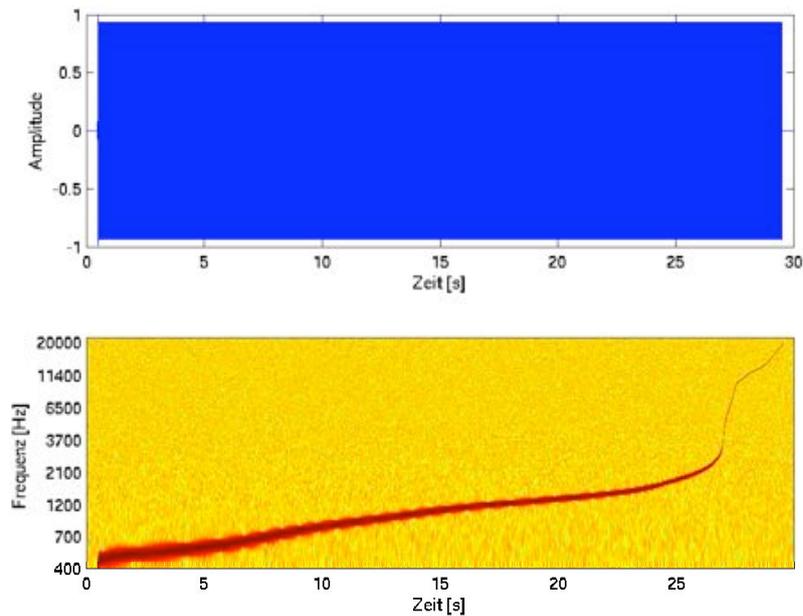


Abbildung 25: Zeitsignal und Spektrogramm des rechten Sweeps

6 Messung

Bei der Durchführung der Messung wurde das Messsignal durch ein analoges Mischpult geführt, sodass sich jede Versuchsperson den Pegel selbst einstellen konnte. Dies war nötig um individuell erträgliche Schalldruckpegel zu erhalten. Der Restschalldruck, der in Richtung Trommelfell abgestrahlt wird, darf nicht vernachlässigt werden. Bei einer Referenzmessung mit dem Innenohr-Mikrofon des Kunstkopfs wurde der höchste eingestellte Pegel ermittelt, dieser beträgt maximal knapp 82 dB SPL (siehe *Abb. 27*).



Abbildung 26: Messlautsprecher im Ohr

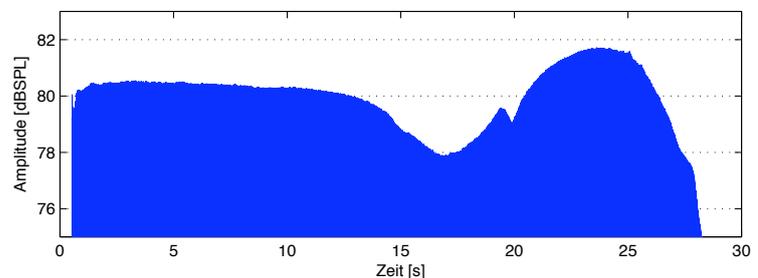


Abbildung 27: Schalldruckpegel am Trommelfell

Die Person wurde dann auf einem Hocker innerhalb des Kugelmikrofon-Arrays so platziert, dass ihr Kopf sich im Mittelpunkt der Kugel mit Blickrichtung nach vorne befand

(siehe Abb. 28). Darüberhinaus wurde die Anweisung gegeben, in Richtung eines Punktes zu sehen, der auf der Schnittgeraden der Horizontal- mit der Medianebene liegt. Während der Messung wurden alle beeinflussbaren Geräuschquellen eliminiert (Computer, Klimaanlage, etc.). Die Messungen wurden Abends durchgeführt um den Störschall von außerhalb des Messraums, wie z.B. schlagende Türen, möglichst gering zu halten.



Abbildung 28: Versuchsproband bei der Vermessung

7 Signalverarbeitung nach der Messung

7.1 Entfaltung

Alle Mikrofon-signale $s_{mes,i}[n]$ wurden zunächst mit einem Hochpassfilter (20. Ordnung, $f_{g,u} = 500\text{Hz}$) und einem Tiefpassfilter (12. Ordnung, $f_{g,o} = 19\text{kHz}$) gefiltert um das mitaufgezeichnete Rauschen in den irrelevanten Frequenzregionen vor der Entfaltung zu minimieren.

Bei der Entfaltung wird gewöhnlicherweise das Spektrum des gemessenen Signals $S_{mes}(f)$ durch das Spektrum des Messsignals $S(f)$ geteilt um dann die Impulsantwort durch die inverse Fouriertransformation zu erhalten. Da hier aber die Messsignale die inversen Betragsfrequenzgänge der Messlautsprecher $H_{eq,gl}(f)$ besitzen, kommen die durch die Messlautsprecher wiedergegebenen Messsignale bereits mit konstanter Amplitude 1 ins akustische System. Aus diesem Grund wird die Entfaltung des Messsignals nur noch auf die Phase angewandt. Zusätzlich wird auch der Phasengang der minimalphasig angenommenen Messlautsprecher berücksichtigt. Mathematisch ergibt sich die gesamte Messkette zu

$$S_{mes}(f) = S(f) \cdot H_{ls}(f) \cdot H_{hr}(f) . \quad (29)$$

In Exponentialdarstellung der komplexen Spektren ist das

$$S_{mes}(f) = |S(f)| \cdot e^{j\phi_S(f)} \cdot |H_{ls}(f)| \cdot e^{j\phi_{H_{ls}}(f)} \cdot H_{hr}(f) . \quad (30)$$

Auf Grund des Zusammenhangs

$$|S(f)| = |H_{eq,gl}(f)| = \frac{1}{|H_{ls}(f)|} \quad (31)$$

bleibt

$$S_{mes}(f) = e^{j\phi_S(f)} \cdot e^{j\phi_{H_{ls}}(f)} \cdot H_{hr}(f) . \quad (32)$$

Das komplexe Frequenzspektrum der Außenohrübertragungsfunktion ergibt sich daher zu

$$H_{hr}(f) = S_{mes}(f) \cdot e^{-j(\phi_S(f) + \phi_{H_{ls}}(f))} . \quad (33)$$

Durch die Anwendung der inversen Fouriertransformation erhält man das Zeitsignal

$$h_{hr}(t) = F^{-1}\{H_{hr}(f)\} . \quad (34)$$

7.2 Schneiden der Impulsantworten

Die Impulsantworten nach der Entfaltung sind in *Abb. 29* zu sehen. Da die Messkoje innerhalb der sich das Kugelmikrofon-Array befindet nicht gänzlich reflexionsfrei ist, klingen die Impulsantworten sehr langsam ab. Die durch Reflexionen erzeugten frequenzabhängigen Verzerrungen, die durch konstruktive oder destruktive Interferenzen entstehen, sollen im Messergebnis nach Möglichkeit nicht enthalten sein. Um diese zu entfernen und außerdem die Impulsantworten auf eine sinnvolle Länge zu bringen, wurden diese auf eine

Länge von 256 Samples reduziert und mit einem asymmetrischen Fenster multipliziert (siehe Abb. 30).

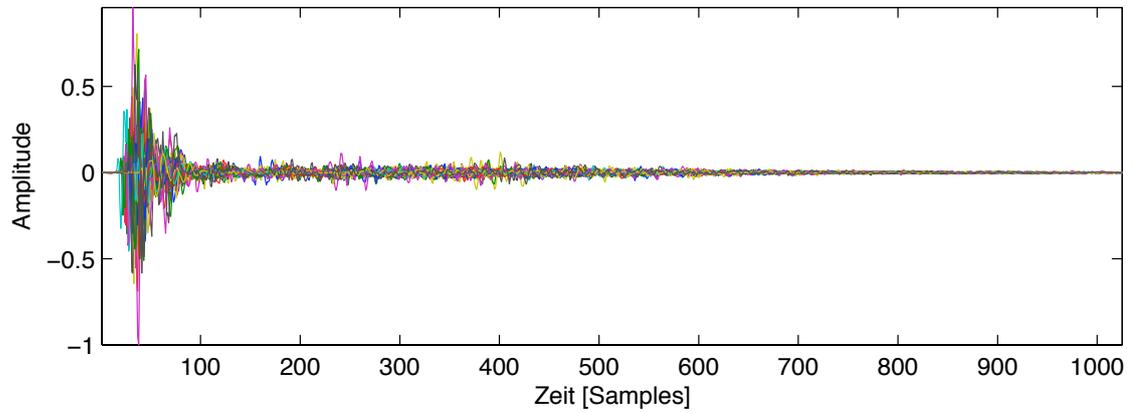


Abbildung 29: unbearbeitete Impulsantworten

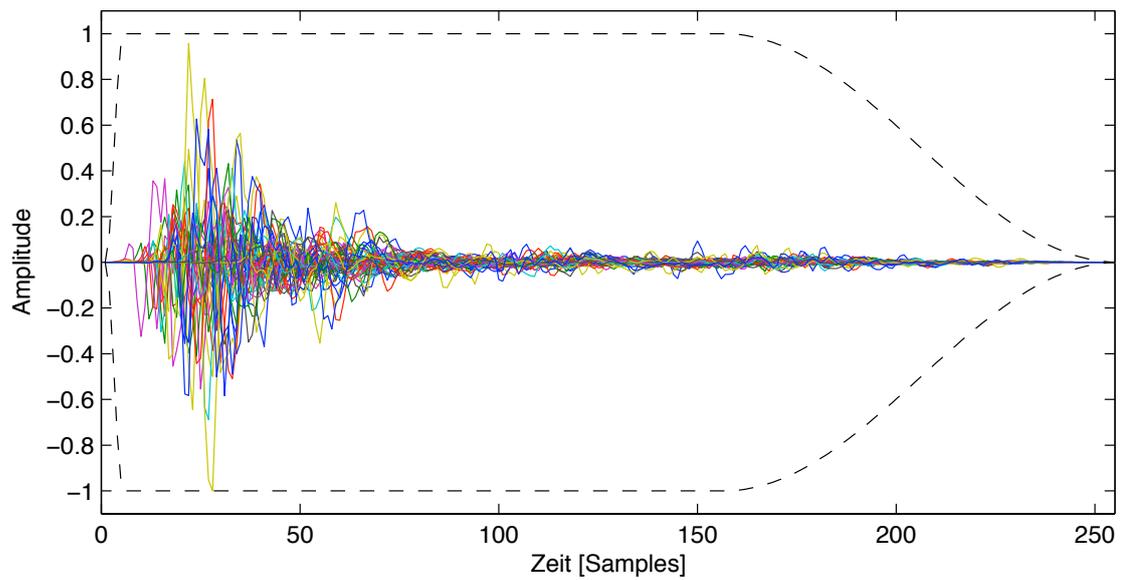


Abbildung 30: geschnittene und gefensterte Impulsantworten mit asymmetrischem Fenster

8 Auswertung

Die im Folgenden dargestellten Grafiken von HRTFs, CTFs und DTFs enthalten der Übersichtlichkeit wegen stets *normierte Frequenzgänge*, da dies den qualitativen Vergleich mit den Daten der Referenzmessung erleichtert. Amplitudenunterschiede zwischen verschiedenen Übertragungspfaden sind daher nicht abzulesen. Es werden zudem ausschließlich *Übertragungsfunktionen des linken Ohres einer Versuchsperson* dargestellt; eine vollständige Aufstellung sämtlicher Übertragungspfade findet sich im Anhang, Kapitel 10. In den Titeln und Unterschriften der Grafiken finden sich *Winkelangaben*, ϕ bezeichnet dabei den Azimutwinkel (von 0° frontal vor der Person bis 359° , von oben betrachtet gegen den Uhrzeigersinn zunehmend), θ den Elevationswinkel (von 0° frontal vor der Person bis -90° nach unten und $+90^\circ$ nach oben).

8.1 Signal-Rausch-Abstand

Bei der Berechnung des Signal-Rausch-Abstands aus den gemessenen Impulsantworten wurde der letztlich als Impulsantwort verwendete Ausschnitt des entfalteten Signals einem lediglich Rauschen enthaltenden Ausschnitt des gleichen Signals gegenübergestellt. Der Ansatz des Algorithmus' über die Gruppenlaufzeit verlangt eine SNR-Berechnung erst nach der Entfaltung. Da das Rauschen auch in der Impulsantwort enthalten ist, wurde der SNR hier auf Grund des Zusammenhangs

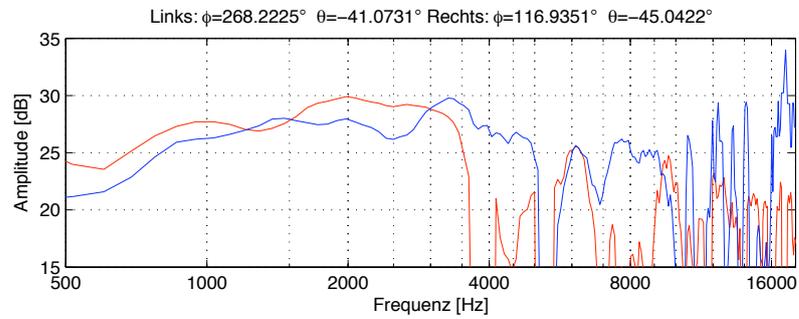
$$\begin{aligned}
 E \left\{ \frac{|S_r(f)|^2}{|N(f)|^2} \right\} &= E \left\{ \frac{|S(f) + N(f)|^2}{|N(f)|^2} \right\} \\
 &= E \left\{ \frac{|S(f)|^2}{|N(f)|^2} \right\} + 2 \cdot E \left\{ \frac{\Re\{S^*(f) \cdot N(f)\}}{|N(f)|^2} \right\} + E \left\{ \frac{|N(f)|^2}{|N(f)|^2} \right\} \\
 &= \frac{|S(f)|^2}{E\{|N(f)|^2\}} + 0 + 1 \\
 &= SNR(f) + 1
 \end{aligned} \tag{35}$$

nach der folgenden Gleichung ermittelt, wobei der Index r für *real* steht:

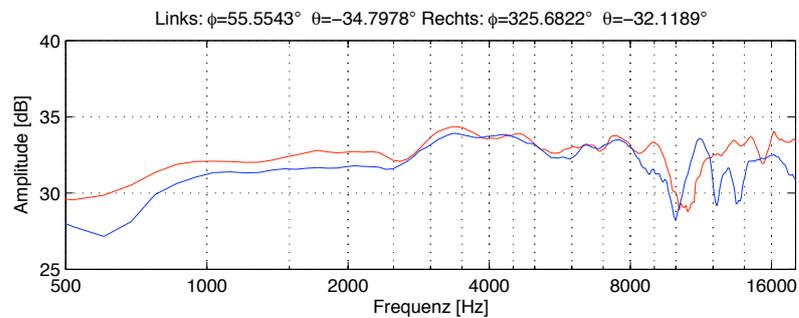
$$SNR(f) = E \left\{ \frac{|S_r(f)|^2}{|N(f)|^2} \right\} - 1 \tag{36}$$

In den unten stehenden Grafiken sind die SNR-Verläufe über die Frequenz von einer Messung zu sehen; der minimale und der maximale SNR bezieht sich jeweils auf einen Kanal, diesen ist deshalb die jeweilige HRTF aufgeprägt und die Verläufe unterscheiden sich. Wie aus den Titeln der Diagramme zu entnehmen ist, handelt es sich bei dem Übertragungspfad mit dem kleinsten SNR um einen Kanal auf der abgewandten Seite, bei dem maximalen SNR um einen Kanal auf der zugewandten Seite. Eine vollständige Aufstellung der SNR-Verläufe aller Versuchspersonen findet sich im Anhang, Abschnitt

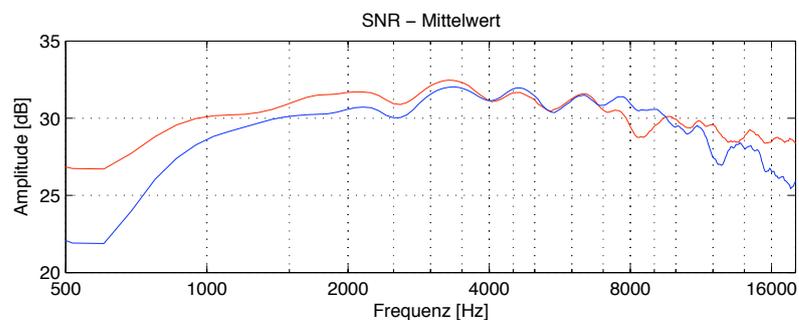
10. Natürlich verhält sich der SNR für jeden Kanal entsprechend dem Frequenzverlauf der jeweiligen HRTF unterschiedlich (siehe *Abb. 31(a)*, *Abb. 31(b)*), im Mittel liegt er aber bei **über 25 dB** und schwankt im betrachteten Frequenzband weniger als 10 dB (siehe *Abb. 31(c)*).



(a) minimaler SNR



(b) maximaler SNR



(c) mittlerer SNR

Abbildung 31: SNR des linken (rot) und rechten (blau) HRTF-Satzes

8.2 Aufteilung in CTFs und DTFs

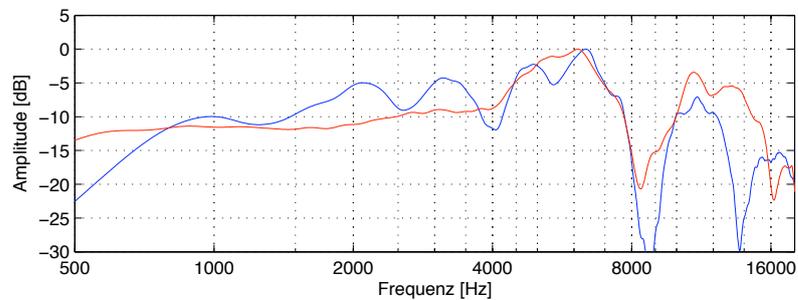
Bei ersten Vergleichen der reziprok gemessenen HRTF-Frequenzgänge mit denen der Referenzmessung am Institut für Schallforschung in Wien (ISF) zeigen sich große Unterschiede (siehe *Abb. 32(a)*). Der Vergleich zwischen den richtungsunabhängigen Anteilen der HRTFs (CTFs) der reziproken Messung und der Referenzmessung zeigt ebenfalls erhebliche Unterschiede (siehe *Abb. 32(b)*). Der naheliegendste Grund dafür ist, dass

die Entzerrung der Messlautsprecher misslungen ist, d.h. ihr Übertragungsverhalten im eingebauten Zustand nicht mit dem in der Entzerrungsmessung ermittelten Übertragungsverhalten übereinstimmt. Auch ein Einfluss von Reflexionen der Messkoje könnte dazu beitragen.

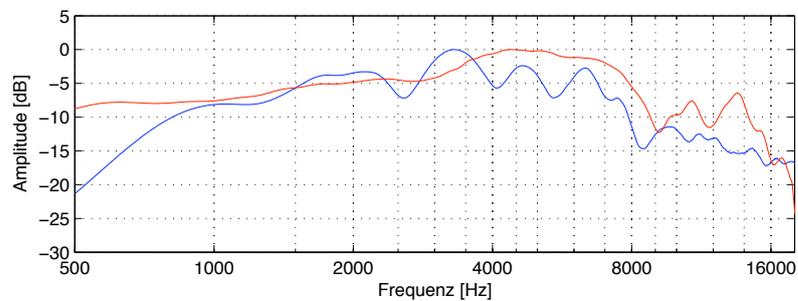
Zur Erleichterung des Vergleichs wurde im Folgenden auf die richtungsabhängigen Anteile der HRTFs (DTFs) zurückgegriffen. Deren Betragsspektren wurden folgendermaßen berechnet

$$DTF(l, f) = \frac{HRTF(l, f)}{\frac{1}{L} \sum_{l'=1}^L HRTF(l', f)} = \frac{HRTF(l, f)}{CTF(f)}, \quad (37)$$

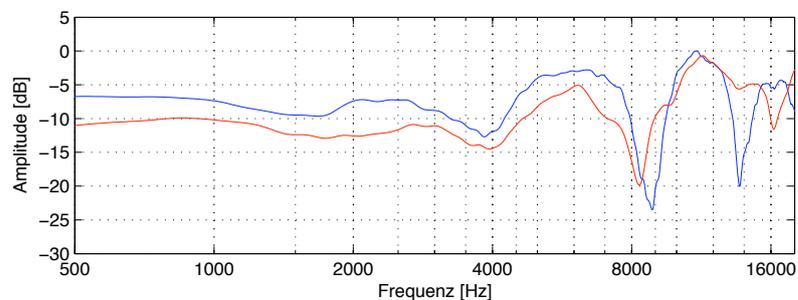
wobei die Variable l für den jeweilige Übertragungspfad und L für die gesamte Anzahl an Übertragungspfaden steht. Wie an *Abb. 32(c)* ersichtlich wird, ist der grafische Vergleich der DTFs deutlich einfacher, da der Grad der Übereinstimmung durch die Entfernung der globalen Filterkurve besser ersichtlich wird.



(a) HRTFs des linken Direktpfades



(b) CTFs linken Ohrs



(c) DTFs des linken Direktpfades

Abbildung 32: Vergleich von HRTFs, CTFs und DTFs der direkten (rot) und reziproken Messung (blau)

8.3 Beispielhafte Übertragungspfade

Da die Verteilung der Mikrofone auf dem Kugelmikrofon unregelmäßig ist, sind die Übertragungspfade etwas schwierig zu kategorisieren. Auch weichen die Positionen dadurch marginal von denen der Referenzmessung ab (maximal 1,5° Winkelabweichung zwischen den Richtungsvektoren der Abtastpunkte). Zum Vergleich werden hier ausschließlich DTFs des linken Ohrs einer Versuchsperson präsentiert.

Soweit die räumliche Abtastung dies erlaubt liegen die folgenden Übertragungspfade ungefähr in der Horizontalebene (Abb. 33) bzw. ungefähr in der Frontalebene (Abb. 34), der Azimutwinkel wird in relativ großen Abständen durchlaufen (siehe Titel der Grafiken).

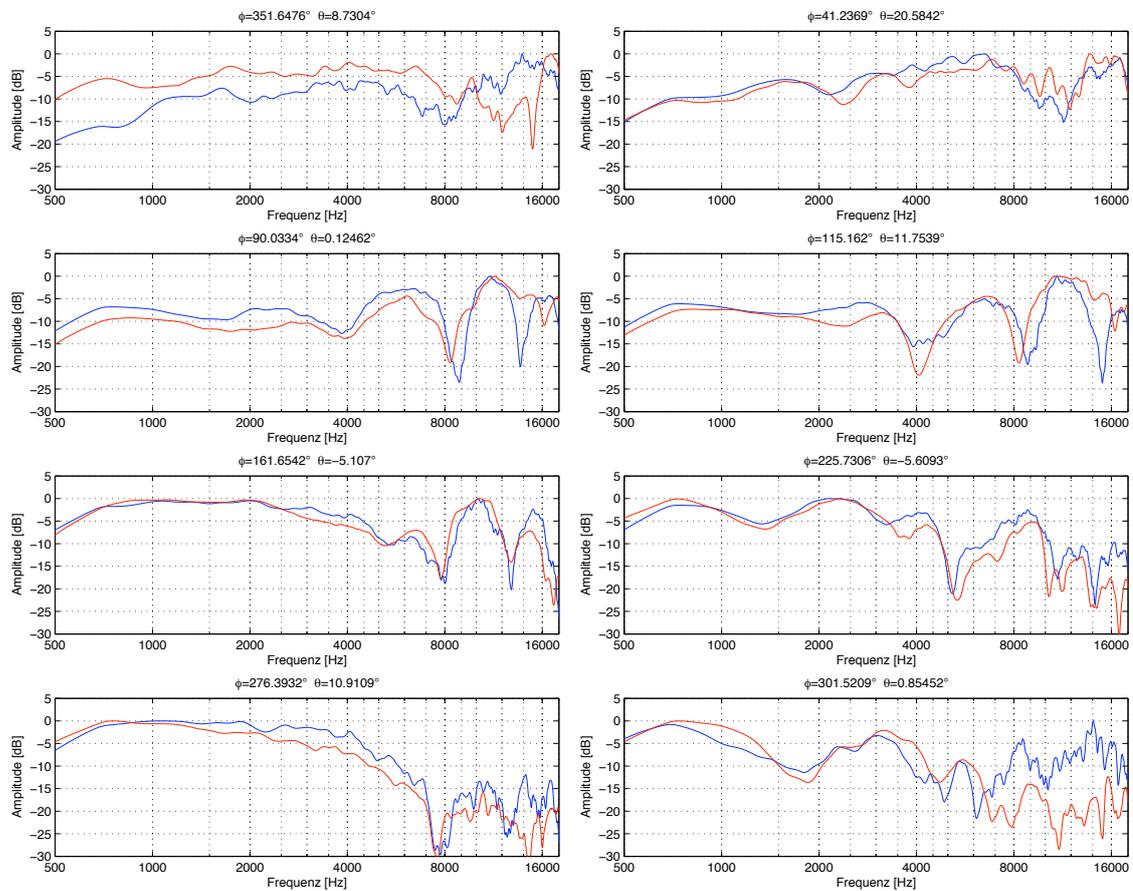


Abbildung 33: Verschiedene Positionen nahe der Horizontalebene der direkten (rot) und reziproken Messung (blau)

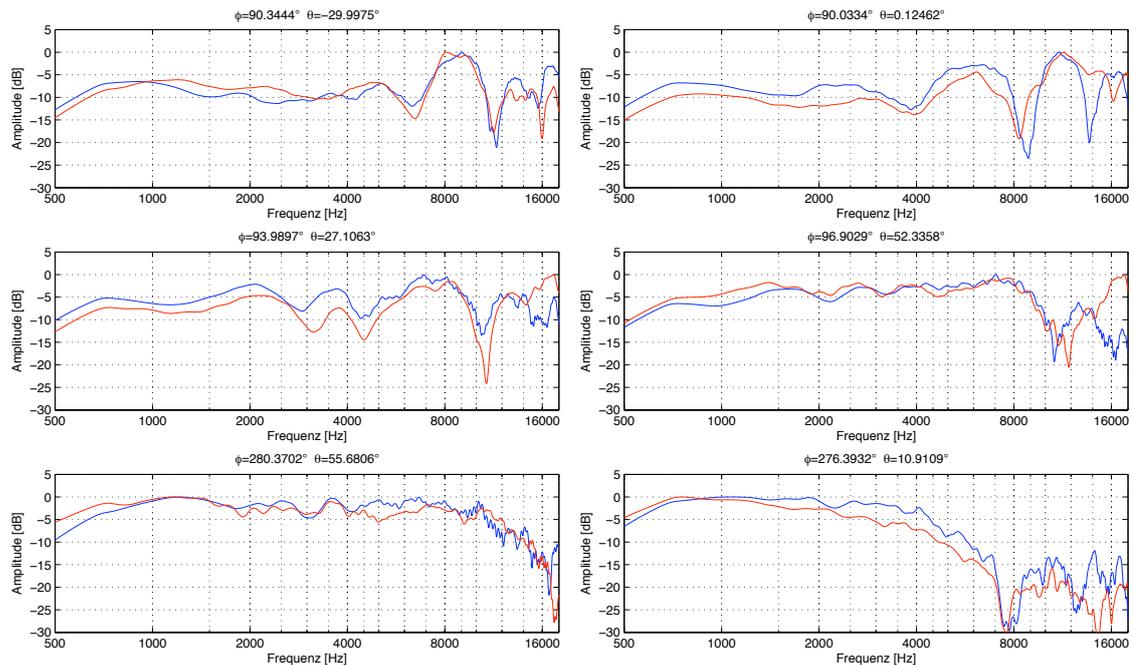


Abbildung 34: Verschiedene Positionen nahe der Frontalebene der direkten (rot) und reziproken Messung (blau)

Wie an den Grafiken deutlich wird, sind die DTFs im Vergleich bis ca. 10 kHz sehr ähnlich, das gilt sogar für die kontralaterale (abgewandte) Seite. Für Frequenzen oberhalb von 10 kHz sind deutliche Abweichungen zu erkennen, teilweise verlaufen die Kurven sogar gegenläufig zueinander, was vermutlich auf einen hochfrequenten Störer bei der Messung (Messrechner) oder eine starke Reflexion im Bereich $\phi = 270^\circ - 359^\circ$ zurückzuführen ist. Hinzu kommen einige Verschiebungen in der Tiefe und Position der Notches in Relation zur Referenzmessung. Dies ist einerseits auf die unterschiedlichen Positionen der Messlautsprecher (reziproke Messung, Abb. 4) und Mikrofone (direkte Messung, Abb. 26) im Ohr der Versuchspersonen zurückzuführen. Der Messlautsprecher ragt weiter aus dem Ohrkanal heraus, dadurch verschieben sich vermutlich die Laufzeiten und Kammfilter. Andererseits sind auch leichte Positionierungsfehler der Versuchspersonen auf Grund fehlender Fixierung wahrscheinlich.

9 Schlussfolgerungen und Diskussion

Zusammenfassung In der vorliegenden Arbeit wurde eine Methode zur Vermessung von individuellen Fernfeld-Außenohrübertragungsfunktionen umgesetzt, welche die Messdauer auf eine Minute reduziert. Die Methode wurde durch die direkte und reziproke Messung von drei Versuchspersonen evaluiert. Es wurde gezeigt, dass die reziprok gemessenen HRTFs aller Versuchspersonen, oder genauer gesagt deren richtungsabhängige Anteile (DTFs), insbesondere im Bereich zwischen 500 Hz und 10 kHz vergleichbare Ergebnisse liefern können.

Ein besonderes Augenmerk wurde dabei auf den Signal-Rausch-Abstand gelegt, der im Mittel bei über 25 dB liegt; in diesem Kontext kam ein Verfahren zur Modellierung von Frequenzsweeps mit konstanter zeitlicher Amplitude zur Anwendung.

Erkenntnisse In der Arbeit wurden Messungen mit Personen gemacht, deren Gehör nicht überlastet werden darf. Beschwerdefreiheit bei der Messung wurde durch einstellbaren Messsignalpegel erreicht. Vor Allem scheint die richtige Wahl der Kleinstlautsprecher von essentieller Bedeutung zu sein sowie deren Entzerrung gleichermaßen. Die Forderung nach einem guten SNR ist darin die zentrale technische Herausforderung, die bei Messungen der Fernfeld-HRTFs gemeistert werden muss.

Kritik und mögliche Verbesserungen Da es sich bei dieser Arbeit um eine prinzipielle Evaluierung der reziproken Messmethode und nicht um ein voll konkurrenzfähiges Messsystem handelt, kann an der Messgenauigkeit noch einiges verbessert werden. Als wichtigster Punkt ist hier eine Verwendung von geeigneten Kleinstlautsprechern zu nennen, idealerweise mit geschlossenem Gehäuse. Diese sollten zur Gänze in den äußeren Gehörgang hineinpassen. Die Messergebnisse zeigen, dass schon kleine Abweichungen in der Position der Messlautsprecher deutliche Fehler verursachen.

Es wurde angenommen, dass durch die kurze Messdauer keine Fixierung benötigt wird. Bewegung der Versuchsperson während der Messung kann nicht ausgeschlossen werden und Fehler hervorrufen. Ziel könnte daher sein, die Messdauer noch weiter zu reduzieren, um die Wahrscheinlichkeit von Bewegungsfehlern zu senken. Auch eine mechanische Fixierung oder optisches Head-Tracking könnten dazu genutzt werden.

Die räumliche Abtastung der verwendeten Mikrofonkugel ist mit 64 Kanälen nicht ausreichend um die Lokalisationsgenauigkeit des Gehörs voll abzubilden, obwohl der Höreindruck mit Interpolation über sphärische Harmonische einen guten räumlichen Eindruck vermittelt. Mit einer höher auflösenden Mikrofonanordnung ist jedoch ohne Vergrößerung der Messdauer eine fast beliebig genaue räumliche Abtastung erreichbar, wenn Reflexionsfreiheit aufrechterhalten werden kann.

Schließlich ist auch eine Ergänzung der gemessenen HRTFs für tiefe Frequenzen mit analytischer Modellierung denkbar (z.B. HAT-Model, [ZDGG06]).

10 Anhang

Im Folgenden sind die Messergebnisse aller drei Versuchspersonen im Vergleich zwischen der direkten Messung am Institut für Schallforschung und der reziproken Messung am Institut für elektronische Musik grafisch dargestellt. Die Achsen sind zugunsten der Übersichtlichkeit nicht beschriftet, vertikal ist stets die Amplitude der normalisierten DTFs in Dezibel und horizontal die Frequenz in Hertz aufgetragen.

Die Position, d.h. Azimut- und Elevationswinkel, ist den Grafiktiteln zu entnehmen. Von 64 mit der Mikrofonkugel gemessenen Pfaden werden nur 52 Pfade dargestellt, da die direkte Messung keine volle Kugel abdeckt und daher für die übrigen Kanäle keine Vergleichsdaten existieren. Nachdem die DTFs zum besseren Vergleich einzeln normalisiert wurden, treten in einigen Grafiken deutliche Pegelunterschiede durch Ausreißer außerhalb des relevanten Frequenzbandes auf. Bei dieser Darstellungsweise soll nur der qualitative Vergleich der DTFs im Vordergrund stehen, vergleiche der Absolutwerte sind nicht zulässig.

10.1 Versuchsperson 1

SNR

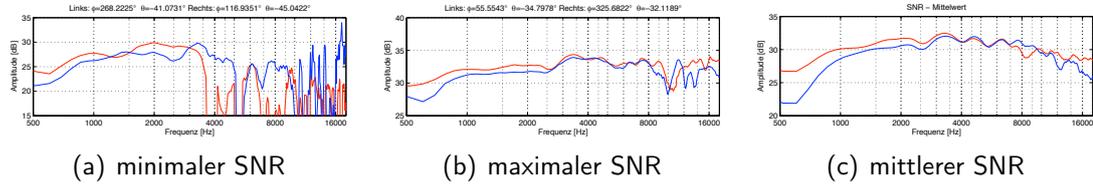


Abbildung 35: SNR des linken (rot) und rechten HRTF-Satzes (blau)

DTF-Satz links

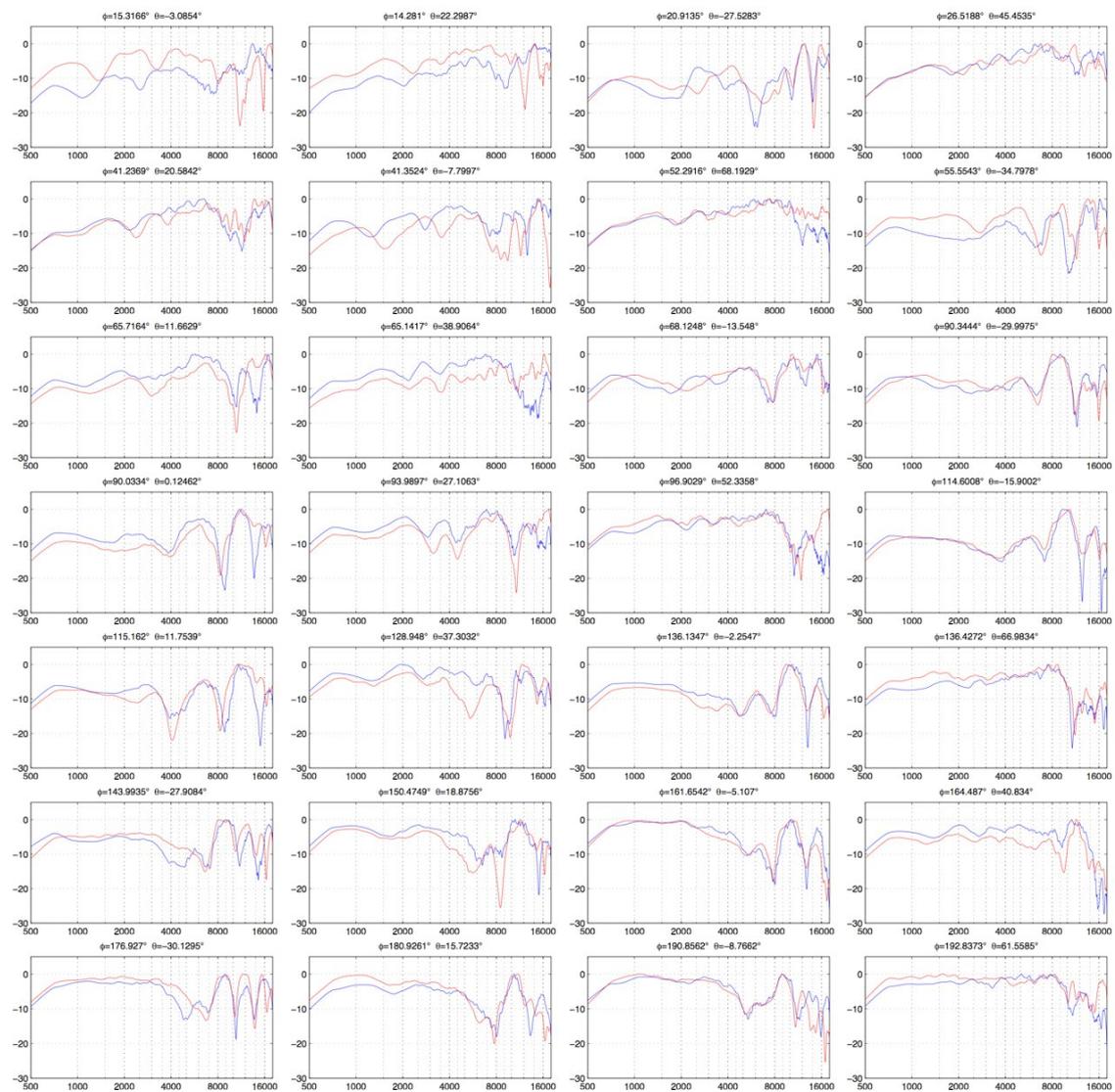


Abbildung 36: DTFs des linken Ohrs der direkten (rot) und reziproken Messung (blau)

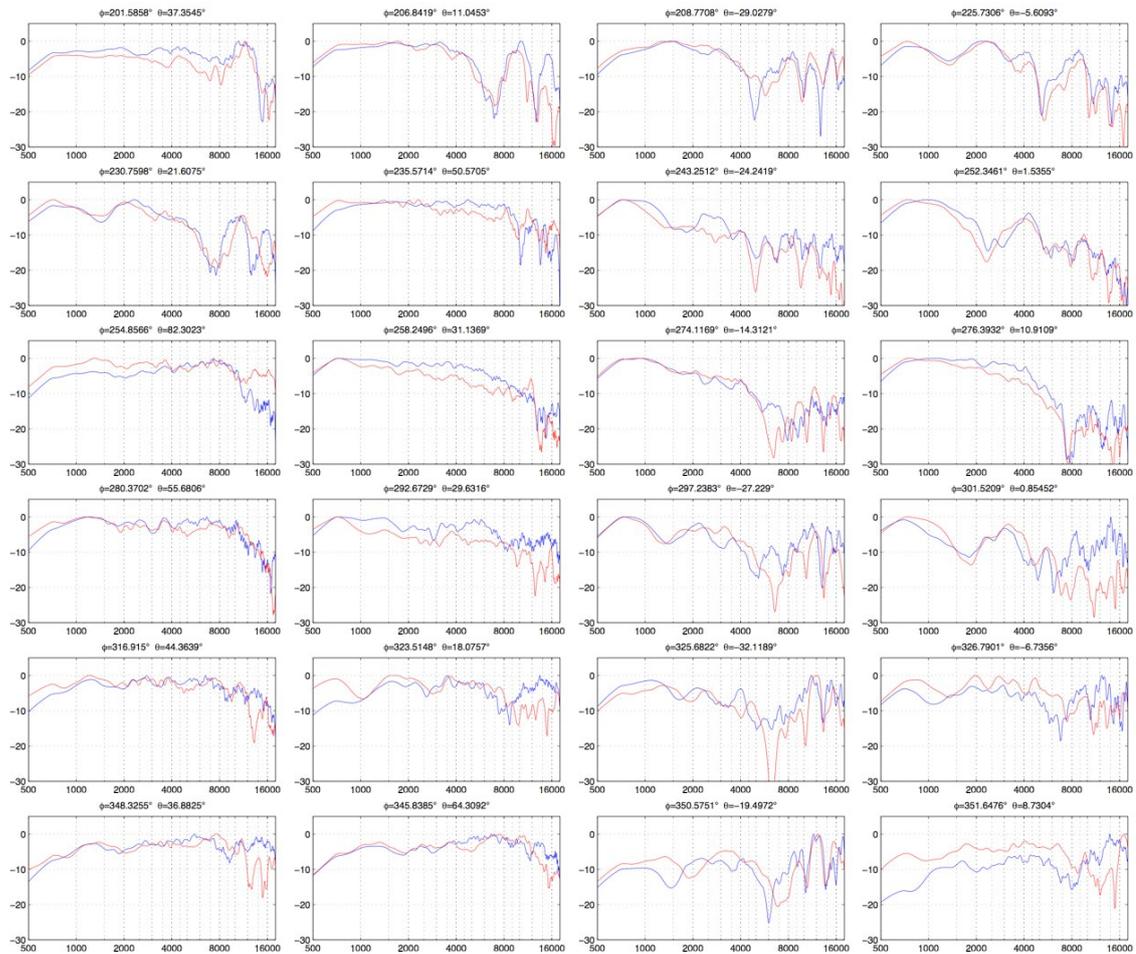


Abbildung 37: DTFs des linken Ohrs der direkten (rot) und reziproken Messung (blau)

DTF-Satz rechts

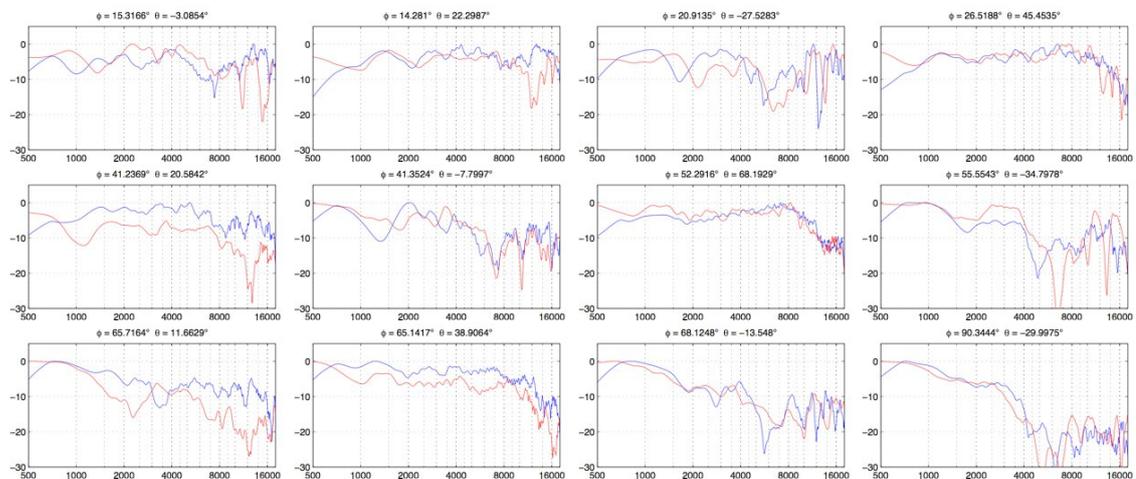


Abbildung 38: DTFs des rechten Ohrs der direkten (rot) und reziproken Messung (blau)



Abbildung 39: DTFs des rechten Ohrs der direkten (rot) und reziproken Messung (blau)

10.2 Versuchsperson 2

SNR

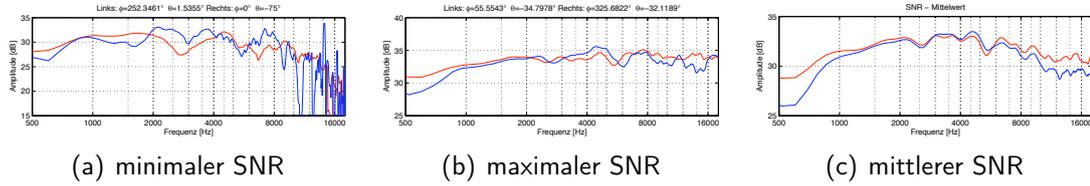


Abbildung 40: SNR des linken (rot) und rechten HRTF-Satzes (blau)

DTF-Satz links

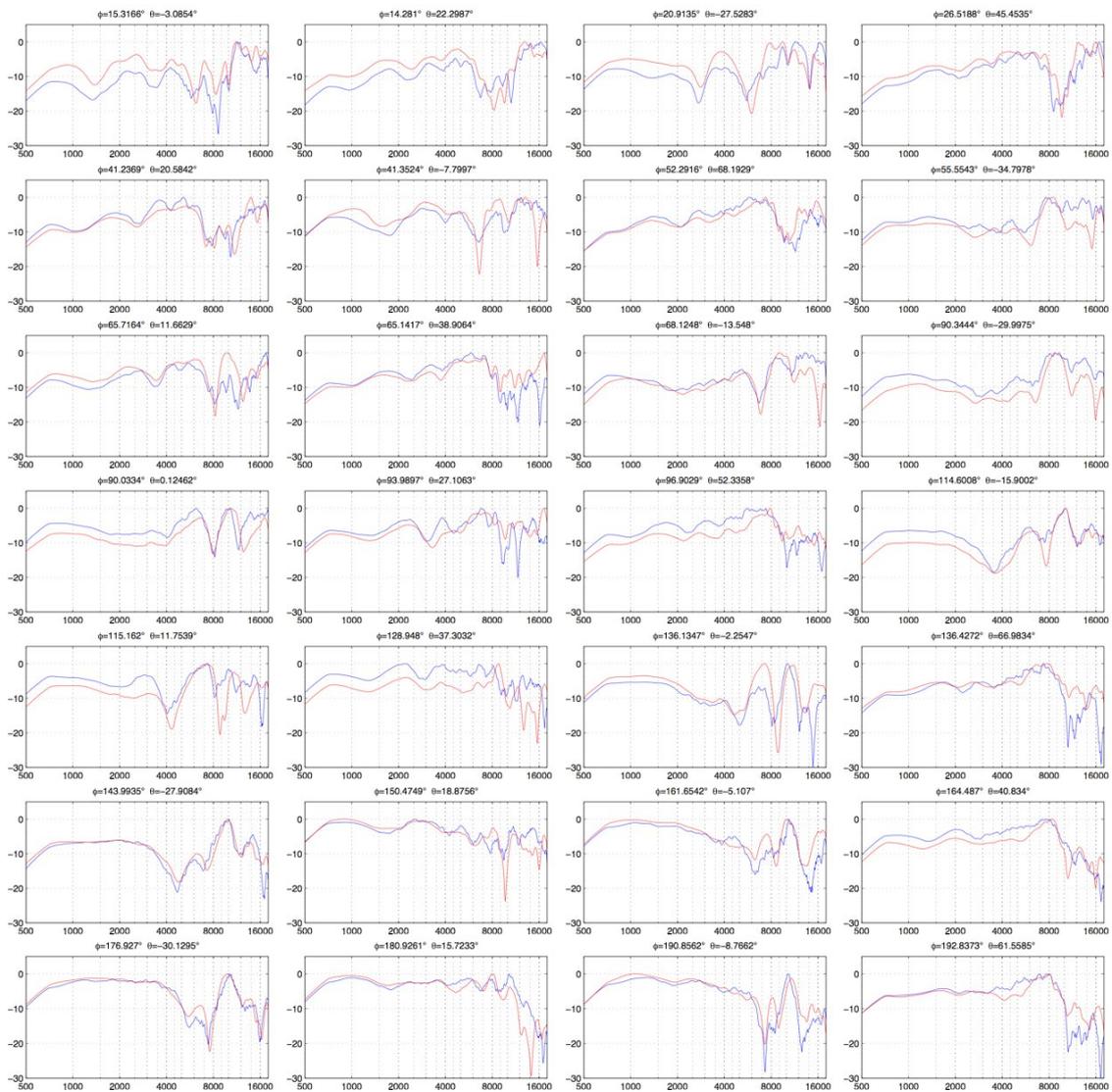


Abbildung 41: DTFs des linken Ohrs der direkten (rot) und reziproken Messung (blau)

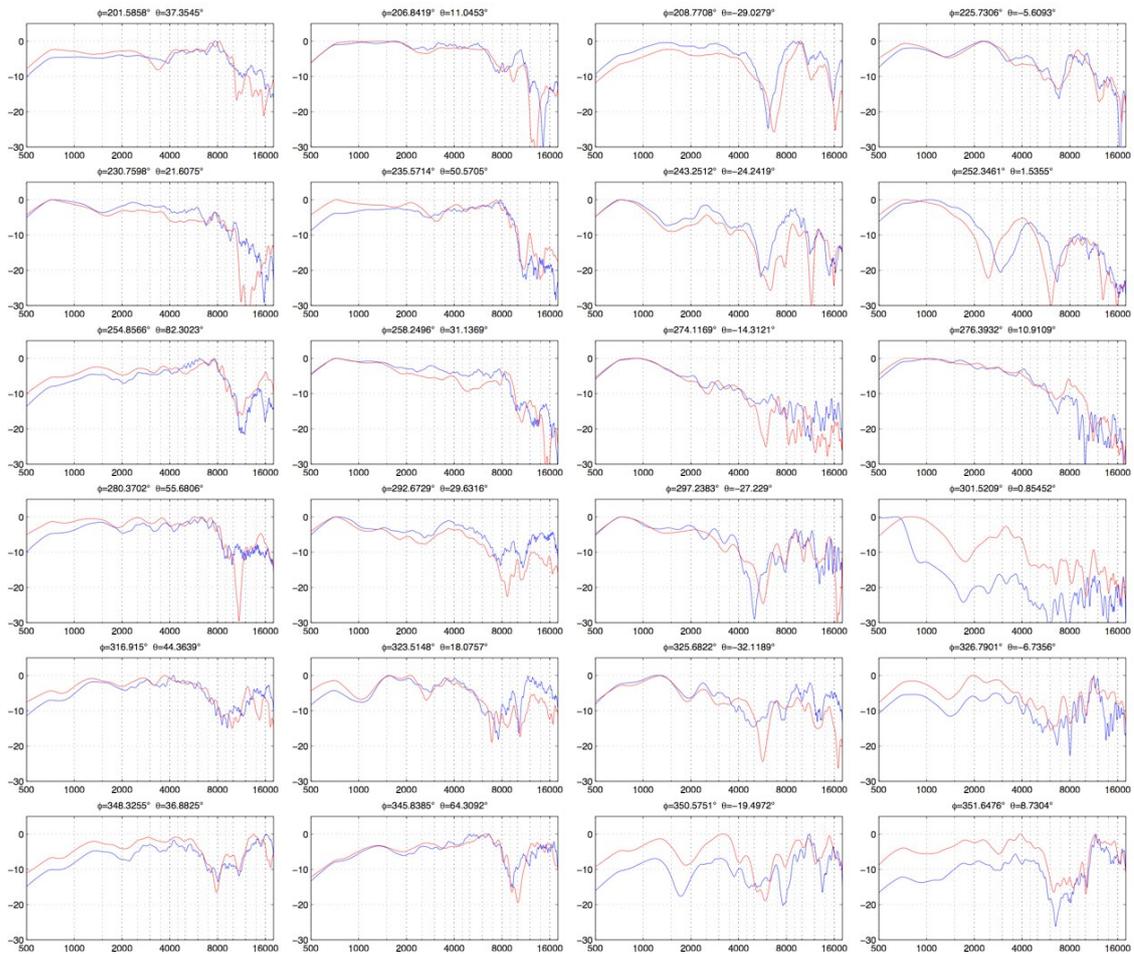


Abbildung 42: DTFs des linken Ohrs der direkten (rot) und reziproken Messung (blau)

DTF-Satz rechts

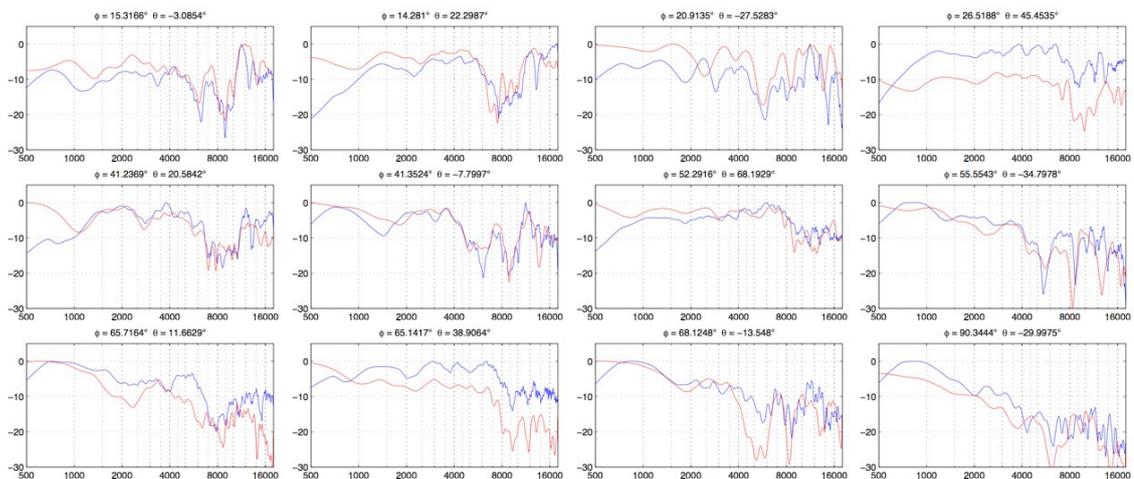


Abbildung 43: DTFs des rechten Ohrs der direkten (rot) und reziproken Messung (blau)

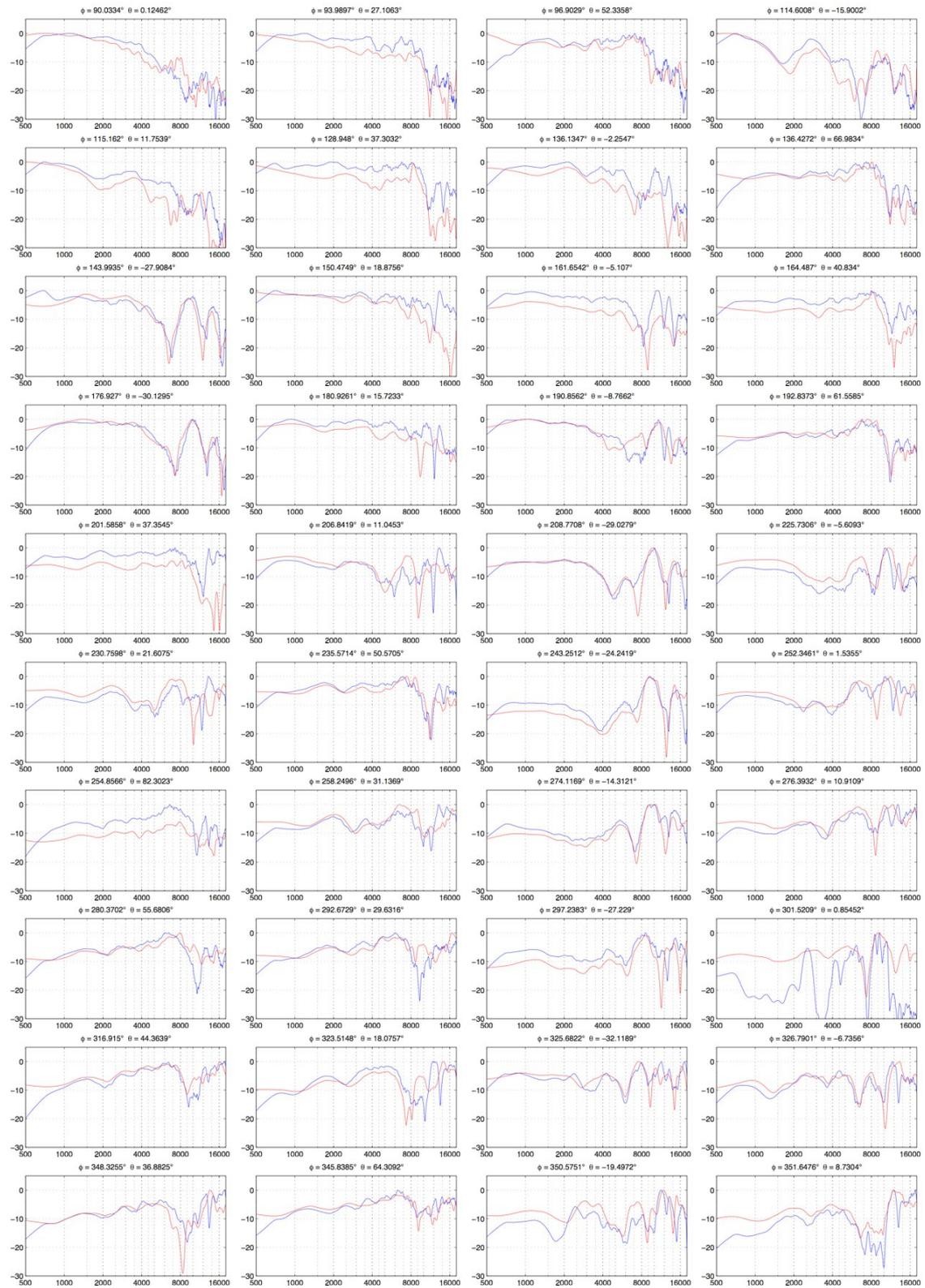


Abbildung 44: DTFs des rechten Ohrs der direkten (rot) und reziproken Messung (blau)

10.3 Versuchsperson 3

SNR

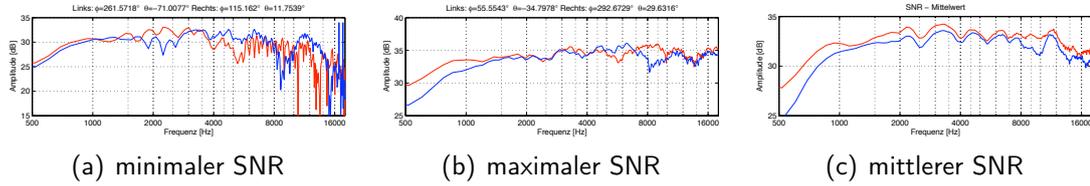


Abbildung 45: SNR des linken (rot) und rechten HRTF-Satzes (blau)

DTF-Satz links

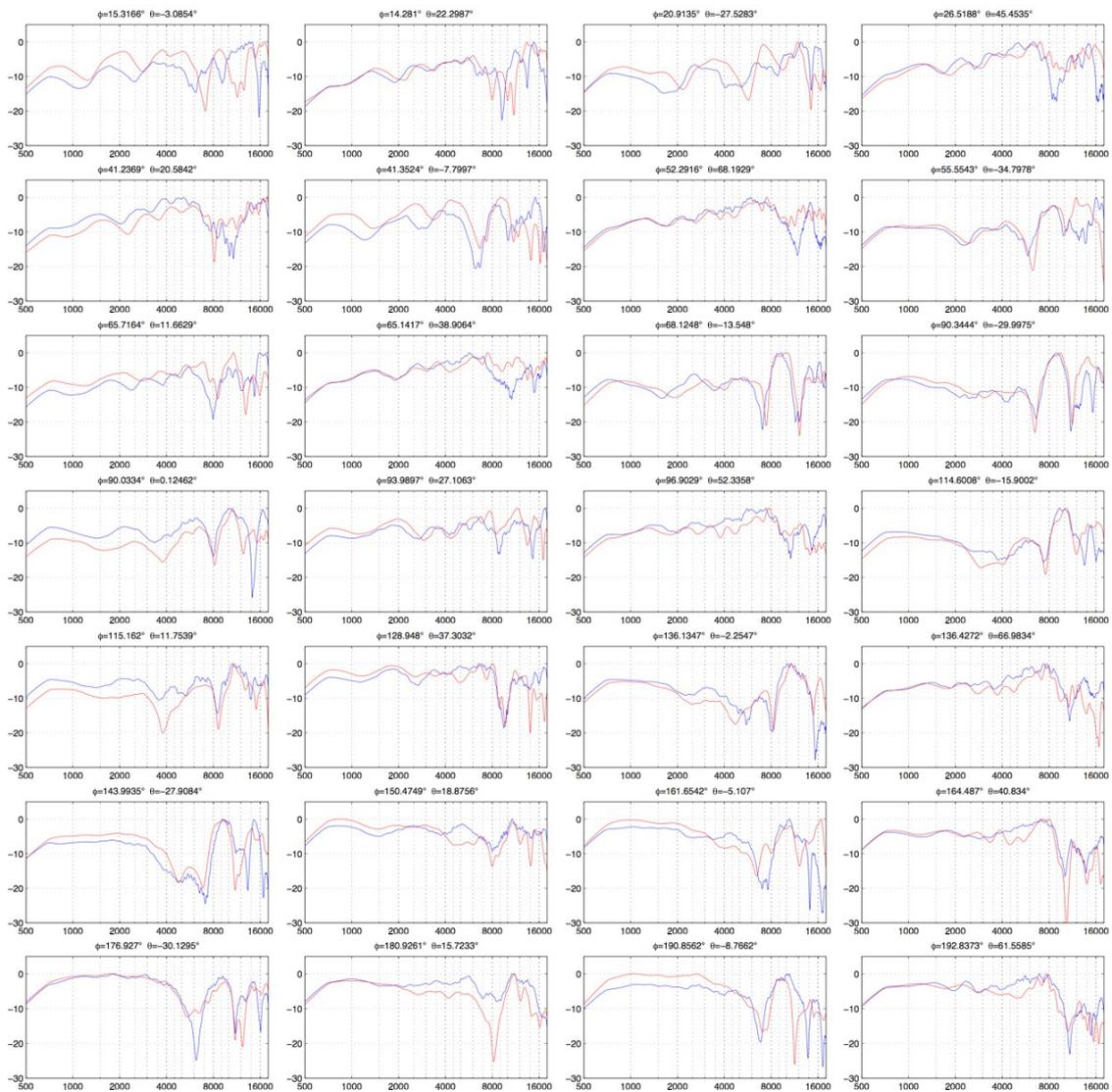


Abbildung 46: DTFs des linken Ohrs der direkten (rot) und reziproken Messung (blau)

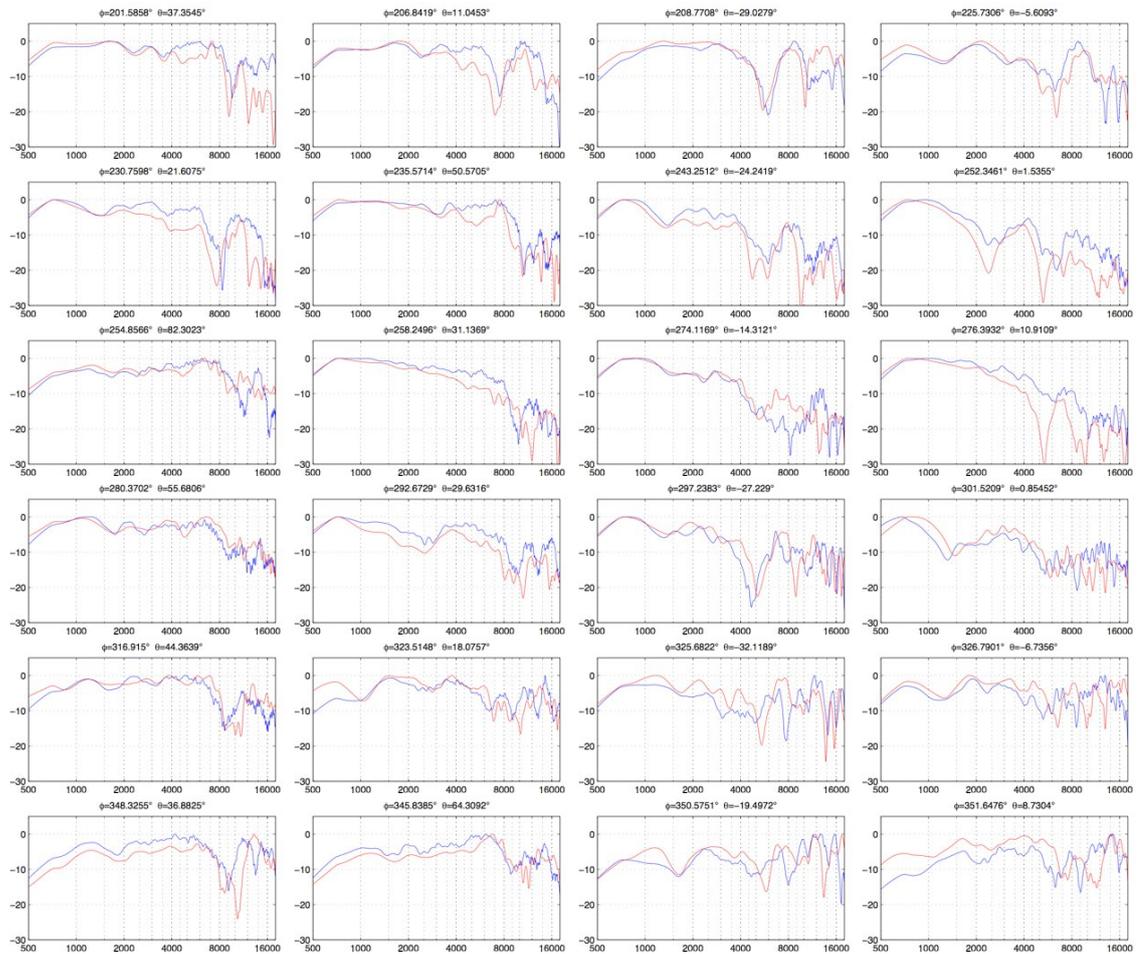


Abbildung 47: DTFs des linken Ohrs der direkten (rot) und reziproken Messung (blau)

DTF-Satz rechts

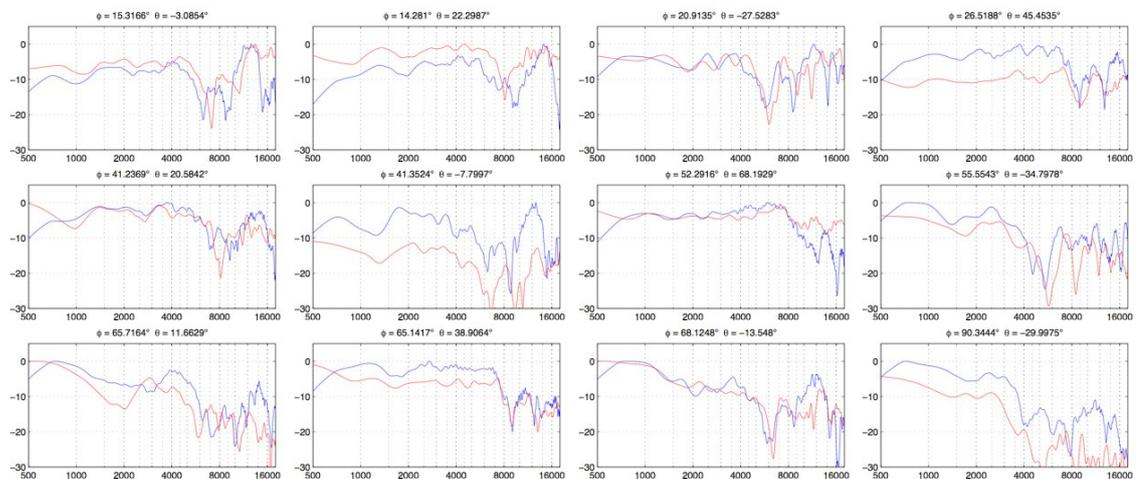


Abbildung 48: DTFs des rechten Ohrs der direkten (rot) und reziproken Messung (blau)

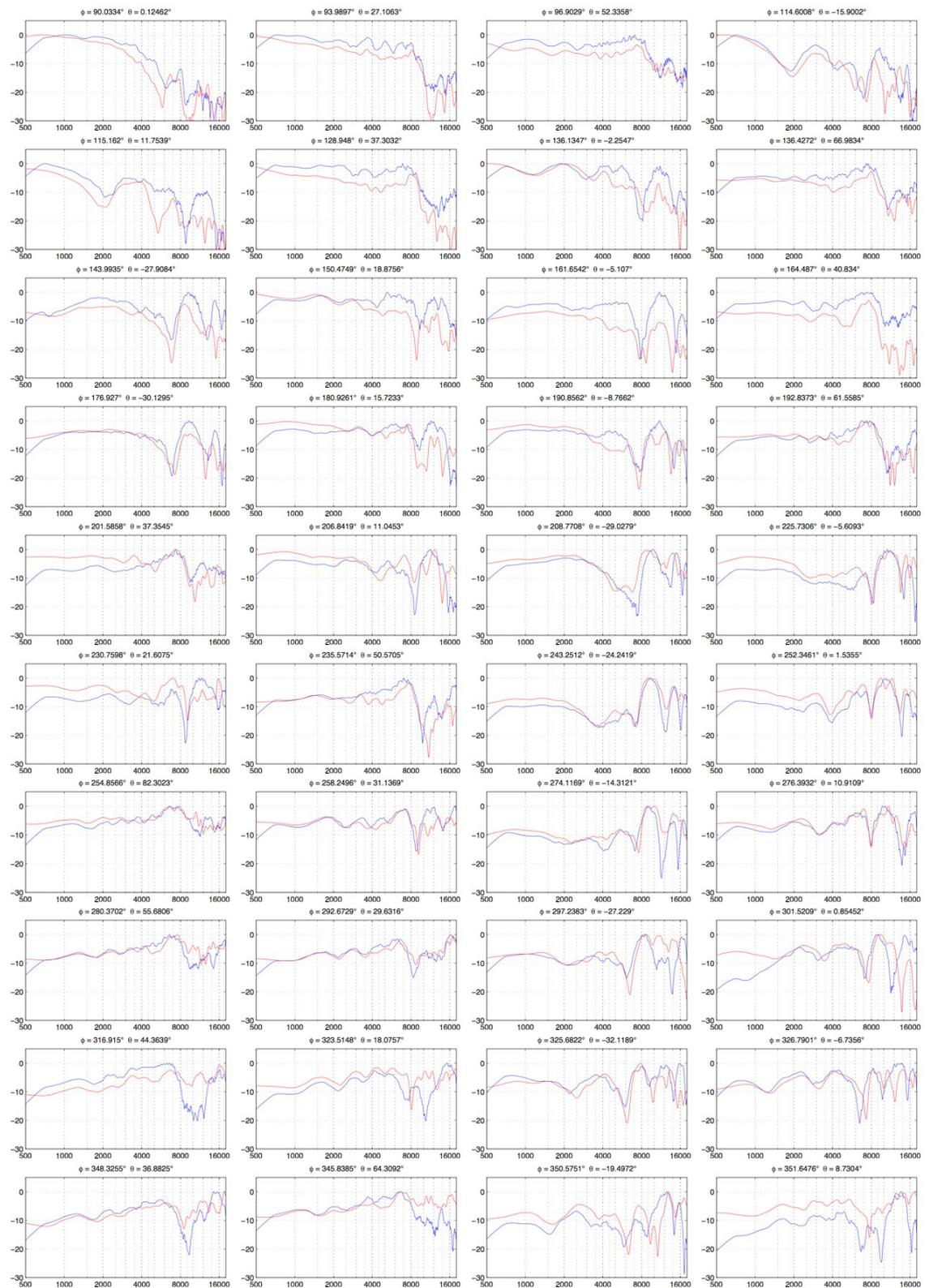


Abbildung 49: DTFs des rechten Ohrs der direkten (rot) und reziproken Messung (blau)

Literatur

- [Hoh09] F. Hohl, "Kugelmikrofonarray zur Abstrahlungsvermessung von Musikinstrumenten," Master's thesis, Institute of Electronic Music and Acoustics, Graz-Austria, 2009.
- [LAVK06] T. Lentz, I. Assenmacher, M. Vorländer, and T. Kuhlen, "Precise Near-to-Head Acoustics with Binaural Synthesis," Institute of Technical Acoustics RWTH Aachen University, Tech. Rep., 2006.
- [MBL06] P. Majdak, P. Balazs, and B. Laback, "Multiple Exponential Sweep Method for Fast Measurement of Head-Related Transfer Functions," Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria, Tech. Rep., 2006.
- [MM06] S. Müller and P. Massarani, "Transfer Function Measurement with Sweeps," Laboratório de Ensaios Acústicos (LAENA) Instituto Nacional de Metrologia, Calibração e Qualidade Industrial (INMETRO), Xerem/Duque de Caxias - RJ, Brazil, Tech. Rep., 2006.
- [ZDGG06] D. Zotkin, R. Duraiswami, E. Grassi, and N. Gumerov, "Fast head-related transfer function measurement via reciprocity," Perceptual Interfaces and Reality Laboratories, Institute for Advanced Computer Studies (UMIACS), University of Maryland at College Park, Tech. Rep., 2006.