



UNIVERSITÄT
FÜR MUSIK UND
DARSTELLENDE KUNST
GRAZ - AUSTRIA

Institut für Rundfunktechnik

Abschlussbericht für

Toningenieur-Projekt (6 SWS)

Realisierung eines virtuellen Kopfhörers

Institut für Elektronische Musik und Akustik (IEM)
der Universität für Musik und darstellende Kunst Graz, Österreich

Institut für Rundfunktechnik GmbH (IRT)
Forschungs- und Entwicklungsinstitut, München, Deutschland

Betreuer:

IEM: **Univ.-Ass. Dipl.-Ing. Dr. Alois Sontacchi**
IRT: **Dipl.-Ing. Klaus Laumann**

Verfasser:

Herbert Petritsch

Matrikelnummer: 0202469

Studienkennzahl: V750

Studium: Elektrotechnik-Toningenieur

E-mail: herbert.petritsch@student.kug.ac.at

Kurzpräsentationstext / Abstract

Ziel dieses Projektes ist die Realisierung eines virtuellen Kopfhörers. Dazu soll ein Messsystem als GUI (Graphical User Interface) in Matlab implementiert werden, das die HRTFs (Head Related Transfer Functions) von allen Schallquellen zu den Mikrofonen eines Kunstkopfes misst und daraus die IHRTFs (Inverse HRTFs) berechnet. Mit letzteren werden aus einem Stereosignal die Quellsignale für die Schallwiedergabe berechnet. Mittels Wellenfeldsynthese werden diese Quellsignale um den Kopf positioniert.

Die Qualität des virtuellen Kopfhörersystems soll für verschiedene Lautsprecher-Array-Geometrien und Parameter evaluiert werden.

Das Projekt wird am IRT durchgeführt und findet in Kooperation mit dem IEM statt.

The aim of this project is the realisation of virtual earphones. A measurement system is implemented as a GUI (Graphical User Interface) in Matlab, which measures the HRTFs (Head Related Transfer Functions) from the acoustic sources to the microphones of a dummy head and calculates the IHRTFs (Inverse HRTFs). After that the acoustic sources for the sound reproduction can be calculated from any stereo signal. Wave field synthesis positions these sources as focussed sources around the head.

The quality of the virtual earphones for different array geometries of secondary sources and for different parameters is evaluated.

The project is carried out at IRT in cooperation with IEM.

Inhaltsverzeichnis

1. Einleitung.....	5
1.1. Motivation.....	5
1.2. Aufbau dieses Berichts	5
2. Aufgabenstellung	7
2.1. Realisierung des virtuellen Kopfhörers.....	7
2.2. Entwicklung eines Messsystems.....	11
2.2.1. Aufgaben des Messsystems.....	11
2.2.2. Bedienung des Messsystems	11
2.3. Messungen	12
2.3.1. Array-Typen.....	12
2.3.2. Abstände	14
2.3.3. Quellenkonstellationen	15
2.3.4. Szenarien	16
2.4. Evaluierung der Messergebnisse	17
3. Audiosignalverarbeitung im Messsystem.....	18
3.1. Erzeugung des exponentiellen Sweeps	18
3.2. Bestimmung aller HRTFs bzw. HRIRs	19
3.2.1. Messungen mit Kunstkopf	19
3.2.2. Latenz.....	20
3.2.3. Bandbreite für die Glättung der Amplitudengänge.....	22
3.3. Berechnung der IHRTFs	24
3.4. Offline-Berechnung der Wiedergabesignale	26
3.5. Bestimmung der Gesamt-Übertragungsfunktionen	26
4. Messungen	29
4.1. Verwendete Hardware bei den Messungen	29
4.2. Messaufbauten im reflexionsarmen Raum.....	30
4.2.1. Linie	30
4.2.2. Türme	32
4.2.3. Oben.....	34
4.3. Parameter-Werte im Messsystem	36
4.3.1. Erzeugung des exponentiellen Sweeps.....	36
4.3.2. Bestimmung aller HRTFs und HRIRs	36
4.3.3. Berechnung der IHRTFs.....	37
4.3.4. Bestimmung der Gesamt-Übertragungsfunktionen.....	38

5. Evaluierung der Messergebnisse	39
5.1. Methodik der Evaluierung	39
5.1.1. Allgemeines	39
5.1.2. Array-Typen.....	44
5.1.3. Abstände	44
5.1.4. Quellenkonstellationen	44
5.1.5. Szenarien	45
5.2. Vergleich und Analyse der Messergebnisse	46
5.2.1. Array-Typen.....	46
5.2.2. Abstände	48
5.2.3. Quellenkonstellationen	49
5.2.4. Szenarien	52
6. Zusammenfassung	55
Anhang: Bedienung des Messsystems.....	57
A.1. Urheber.....	57
A.2. Starten der GUI	57
A.3. Aufbau der GUI.....	60
A.4. Sweep Creation Panel	60
A.5. Measurement Set Panel	61
A.5.1. Name des Messsatzes.....	61
A.5.2. Latenz	62
A.5.3. Anzahl der links angebrachten Lautsprecher	62
A.5.4. Anzahl der rechts angebrachten Lautsprecher	62
A.5.5. Länge der Impulsantwort.....	63
A.5.6. Multiplikationsfaktor.....	63
A.5.7. Bandbreite für die Glättung der Amplitudengänge	63
A.5.8. Daten, die nach jeder Messung geplottet werden sollen	64
A.5.9. Winkel der linken Lautsprecher	64
A.5.10. Winkel der rechten Lautsprecher	65
A.5.11. Messen eines gesamten Messsatzes	65
A.5.12. Plotten aller Impuls- und Frequenzantworten.....	66
A.6. Inverse HRTF Filters Panel	66
A.6.1. Bandpasseinstellungen	66
A.6.2. Berechnung der 3D-Arrays H, G und g	66
A.6.3. Namenserverweiterung	68
A.6.4. Messung der Gesamtübertragungsfunktionen	68
A.6.5. Plot-Optionen	69
A.6.6. Beim Plotten werden geöffnete Figures geschlossen	71
A.6.7. Offline-Berechnung der Wiedergabesignale	71
Referenzen	73
Danksagung	74

1. Einleitung

1.1. Motivation

Binaurale Signale erzeugen bei der Wiedergabe über Kopfhörer einen natürlichen Höreindruck mit genauer Richtungslokalisation. Ein Mangel bei der Kopfhörerwiedergabe ist jedoch der Tragekomfort, der im Wesentlichen zwei Probleme beinhaltet: Erstens ist der Kopfhörer meist über ein Kabel an das Wiedergabegerät angeschlossen, was die Bewegungsfreiheit des Zuhörers einschränkt. Zweitens wird der Druck des Kopfhörers auf das Ohr häufig als störend empfunden.

Das erste Problem kann durch Verwenden eines kabellosen Funk-Kopfhörers behoben werden. Eine andere Möglichkeit, die auch das zweite Problem löst, ist die Wiedergabe des binauralen Signals über einen virtuellen Kopfhörer: Dabei wird der Schall nach Modellen der Wellenfeldsynthese so in der Nähe der Ohren konzentriert, dass für den Zuhörer der Eindruck einer Kopfhörer-Wiedergabe entsteht (siehe Kapitel 2.1.).

1.2. Aufbau dieses Berichts

Kapitel 2 enthält die vollständige Aufgabenstellung:

Kapitel 2.1. behandelt die grundlegende Aufgabenstellung des Toningenieur-Projekts: die Realisierung eines virtuellen Kopfhörers.

Kapitel 2.2. beschreibt Aufgaben, die im Zusammenhang mit der Entwicklung eines Messsystems stehen.

Mit diesem Messsystem und einem Wellenfeldsynthese-Programm sollen dann Messungen, wie in Kapitel 2.3. beschrieben, durchgeführt werden.

Anhand der Ergebnisse, die aus diesen Messungen erhalten werden, soll das virtuelle Kopfhörersystem anschließend evaluiert werden. Wie dabei vorzugehen ist, wird in Kapitel 2.4. besprochen.

Kapitel 3 behandelt Überlegungen und Aspekte der Audiosignalverarbeitung, die der Implementierung des Messsystems zugrunde liegen:

Die Unterteilung in die Kapitel 3.1. bis 3.5. berücksichtigt dabei verschiedene Aufgabenbereiche des Messsystems.

In Kapitel 4 wird die Durchführung der Messungen beschrieben:

Kapitel 4.1. listet zunächst die bei den Messungen verwendete Hardware auf.

In Kapitel 4.2. wird auf die praktische Realisierung der einzelnen Messaufbauten eingegangen.

In Kapitel 4.3. wird die Wahl jener Parameter-Werte im Messsystem diskutiert, die bei allen Messungen verwendet wurden.

Kapitel 5 enthält die Evaluierung der Messergebnisse:

In Kapitel 5.1. wird die Methodik der Evaluierung dargelegt, d.h. es wird besprochen, *wie* evaluiert wurde.

Kapitel 5.2. vergleicht und analysiert die Ergebnisse der Messungen.

In Kapitel 6. werden die wichtigsten Ergebnisse dieses Projekts zusammengefasst.

Im Anhang befindet sich eine Bedienungsanleitung für das Messsystem.

2. Aufgabenstellung

2.1. Realisierung des virtuellen Kopfhörers

Grundsätzliches Ziel dieses Toningenieur-Projektes ist die Realisierung eines virtuellen Kopfhörers. Dieses virtuelle Kopfhörersystem soll für verschiedene Parameter evaluiert werden.

Abbildung 2.1. stellt das gesamte virtuelle Kopfhörersystem im Frequenzbereich für jeweils einen bestimmten Frequenzbin k dar.

Das Kopfhörersignal ist hierbei ein beliebiges (im Optimalfall binaurales) Stereo-Signal (linker Kanal: Y_l , rechter Kanal: Y_r). Dieses wird mit sogenannten Inverse Head Related Transfer Functions (IHRTFs) gefiltert. Die IHRTFs werden hierbei benötigt, um die durch das virtuelle Kopfhörersystem verursachten Head Related Transfer Functions (HRTFs) möglichst zu kompensieren (auf den Begriff der HRTFs wird in diesem Kapitel noch genauer eingegangen). Das Ergebnis sind die Wiedergabesignale.

Mit dem Head-Tracking-System "Polhemus Fastrak" ([1]) werden die Position und der Drehwinkel des Kopfes erfasst. Mit Hilfe der Wellenfeldsynthese werden die Wiedergabesignale als fokussierte Schallquellen um den Kopf eines Zuhörers so positioniert, dass die Abstände und die Winkel der fokussierten Quellen zum jeweiligen Ohr immer konstant bleiben. Als fokussierte Schallquellen werden virtuelle Monopolquellen bezeichnet, die vor einem Lautsprecher-Array positioniert werden [2].

Im Rahmen dieses Projektes sollen verschiedene Lautsprecher-Array-Typen zur Erzeugung der fokussierten Schallquellen durch die Wellenfeldsynthese (WFS) untersucht werden (siehe Kapitel 2.2.). Gemein ist allen Array-Typen, dass diese in eine linke und eine rechte Array-Hälfte mit jeweils 12 Lautsprechern aufgeteilt werden.

Der Wellenfeldsynthese-Block aus Abbildung 2.1 wird in Abbildung 2.2 genauer dargestellt. Im Rahmen der Diplomarbeit von Dipl.-Ing. (FH) Michael Meier wurde ein Wellenfeldsynthese-Programm entwickelt ([3]). Mit diesem werden die Signale jener 12 "linken" bzw. "rechten" realen Lautsprecher (secondary sources) berechnet, die bei Blickrichtung des Zuhörers nach vorne im Azimut-Bereich $[0, 180^\circ]$ bzw. $[0, -180^\circ]$ angebracht sind. Es werden so die Wiedergabesignale $X_1, \dots, X_{K/2}$ bzw. $X_{K/2+1}, \dots, X_K$

mittels Wellenfeldsynthese als fokussierte Quellen $Q_1, \dots, Q_{K/2}$ bzw. $Q_{K/2+1}, \dots, Q_K$ an den vorgesehenen Orten positioniert.

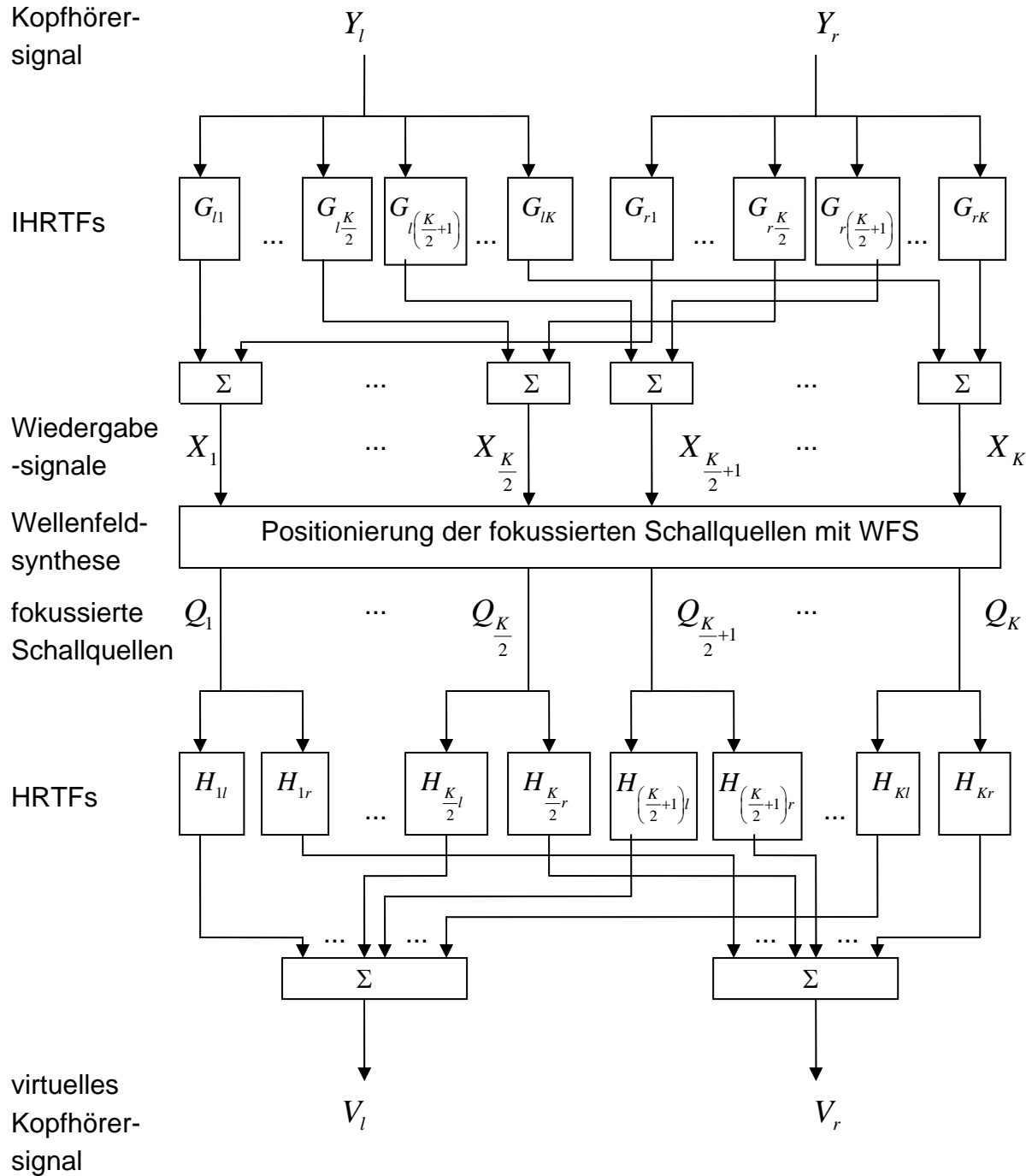


Abbildung 2.1. virtuelles Kopfhörersystem

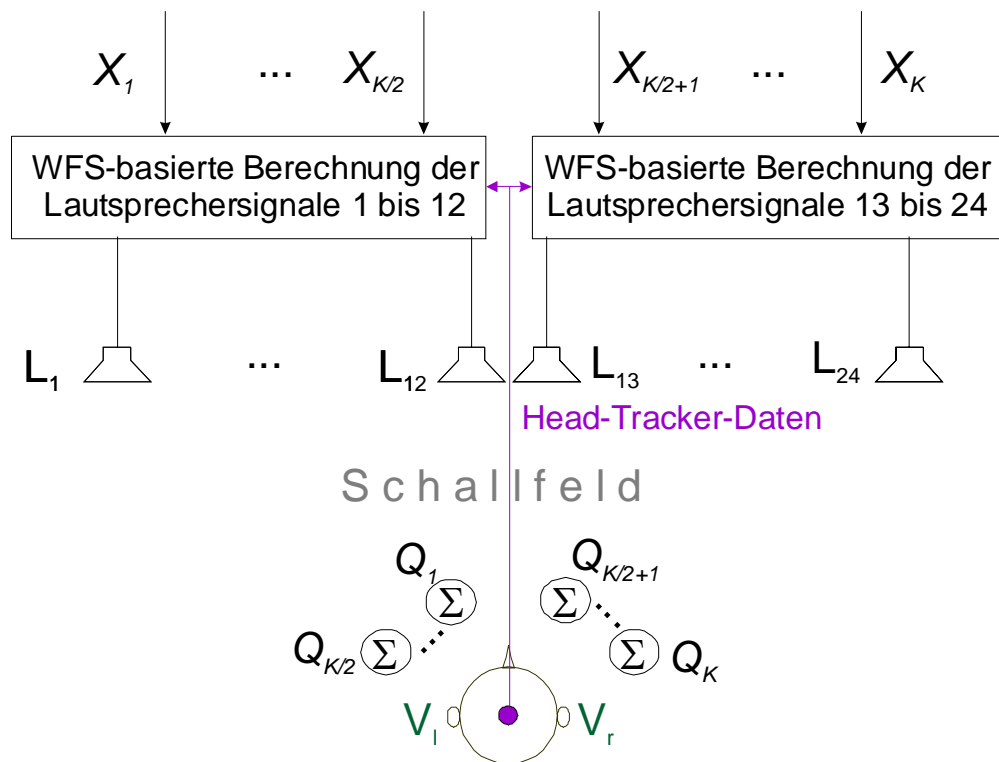


Abbildung 2.2.: Platzierung der Wiedergabesignale an den Orten der fokussierten Schallquellen mit Hilfe der Wellenfeldsynthese

Die fokussierten Schallquellen um den Zuhörer werden in Abbildung 2.2. als Summenblöcke dargestellt. Am Ort der jeweiligen fokussierten Schallquelle werden nämlich alle komplexen Schalldrücke, die sich aufgrund der Lautsprechersignale an diesem Ort ergeben, aufsummiert.

Dass die Schallquellen nicht direkt "am Ohr" fokussiert werden können, hat praktische Gründe (siehe Abbildung 2.3). Die fokussierte Schallquelle (dargestellt als blau gefüllter Kreis) wird erzeugt, indem die Delayzeiten der (in diesem Fall "rechten") Lautsprechersignale, die für die virtuelle Schallquelle hinter dem Array gelten, umgekehrt werden (die virtuelle Schallquelle ist hier als weiß gefüllter Kreis dargestellt). Dabei muss darauf geachtet werden, dass das System kausal bleibt [4]. Weiters muss gewährleistet sein, dass die Schallausbreitung im rot schraffiert dargestellten Bereich ungestört ist: Der Kopf darf hier also kein akustisches Hindernis darstellen.

Außerdem muss gelten, dass sich das linke bzw. (wie in diesem Fall) das rechte Ohr im gültigen Hörbereich (in Abbildung 2.3 hellgrün schraffiert dargestellt) befindet.

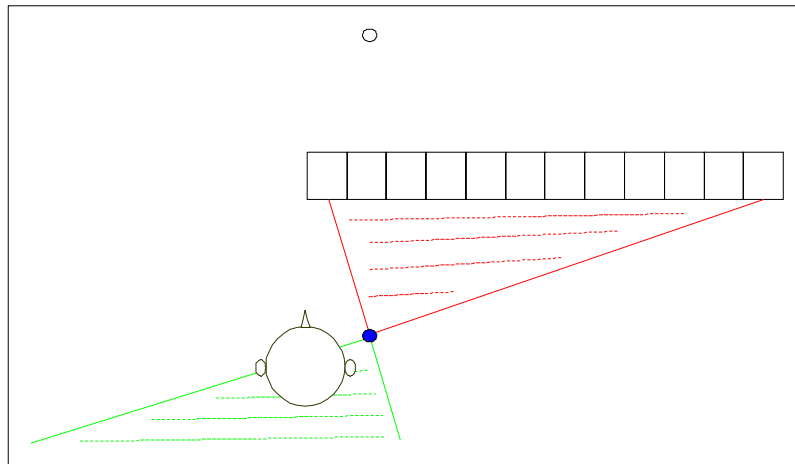


Abbildung 2.3.: fokussierte Schallquelle

Die Übertragungsfunktionen von jeweils einer fokussierten Schallquelle zum Trommelfell des linken bzw. rechten Ohres eines Zuhörers werden in Abbildung 2.1. als Head Related Transfer Functions (HRTFs) bezeichnet.

In die HRTF gehen sowohl der Schallweg von der fokussierten Schallquelle zum jeweiligen Ohr als auch die Einflüsse des Kopfes (die Kopfgeometrie und die momentane Drehung des Kopfes - Azimut und Elevation), der Außenohren und des Gehörgangs auf das am Trommelfell eintreffende Schallsignal ein. Die HRTF ist frequenzabhängig, ihre inverse diskrete Fourier-Transformierte (IDFT) wird als Head Related Impulse Response (HRIR) bezeichnet.

Durch die Überlagerung der HRTF-gefilterten fokussierten Quellsignale ergibt sich an den beiden Trommelfellen des Zuhörers der linke bzw. rechte Kanal V_l bzw. V_r des virtuellen Kopfhörersignals.

Wenn sich der Zuhörer nun im erlaubten Hörbereich der Wellenfeldsynthese befindet, sollte er möglichst die gleiche Hörempfindung haben, wie wenn das ursprüngliche Kopfhörersignal direkt über Kopfhörer wiedergegeben worden wäre.

Mit Hilfe der IHRTFs soll also das Übersprechen von Y_l auf V_r bzw. von Y_r auf V_l so gut wie möglich gedämpft werden, während V_l bzw. V_r möglichst genau Y_l bzw. Y_r entsprechen sollte.

In Anlehnung an [5] wird der Begriff "inverse HRTF" anstatt des Begriffs "Crosstalk Canceler" bzw. "Übersprechkompensationsfilter" verwendet, da letzterer die HRTF-Kompensation nicht erwähnt und damit etwas irreführend ist.

2.2. Entwicklung eines Messsystems

Es soll ein Messsystem in Matlab implementiert werden, mit dessen Hilfe der virtuelle Kopfhörer realisiert und in der Folge evaluiert werden kann.

2.2.1. Aufgaben des Messsystems

Das Messsystem soll die folgenden Bereiche abdecken:

- 1.) Die Erzeugung eines exponentiellen Sweeps.
- 2.) Die Bestimmung aller HRTFs bzw. HRIRs für eine einstellbare Schallquellenanordnung.
- 3.) Die Berechnung der IHRTFs. Dabei sollen die Filterkoeffizienten (IDFTs der IHRTFs) in einem entsprechenden Format abgespeichert werden, sodass mit dem Faltungs-Programm BruteFIR ([6]) für ein beliebiges Kopfhörer-Signal die Wiedergabesignale online berechnet werden können.
- 4.) Zu Testzwecken soll eine Offline-Berechnung der Wiedergabesignale mit dem Messsystem selbst möglich sein.
- 5.) Die Bestimmung der Gesamt-Übertragungsfunktionen, d.h. der Übertragungsfunktionen vom linken bzw. rechten Kanal des Kopfhörersignals zum linken bzw. rechten Kanal des virtuellen Kopfhörersignals.

Aufgrund dieser Aufgaben wird dieses Messsystem als (I)HRTF-Messsystem bezeichnet.

2.2.2. Bedienung des Messsystems

Die Bedienung des (I)HRTF-Messsystems soll über eine GUI (Graphical User Interface) erfolgen.

Eine Bedienungsanleitung zur GUI soll für spätere Anwendungen des Messsystems zur Verfügung stehen (siehe Anhang).

2.3. Messungen

Das (I)HRTF-Messsystem soll gemeinsam mit dem Wellenfeldsynthese-Programm für Messungen im reflexionsarmen Raum (RAR) des IRT verwendet werden.

Die Gesamt-Übertragungsfunktionen sollen für verschiedene Parameter bestimmt werden - welche das sind, wird in den Kapiteln 2.3.1 bis 2.3.4 beschrieben.

2.3.1. Array-Typen

Die verwendeten Array-Typen werden "*Linie*", "*Türme*" und "*oben*" genannt.

Skizzen zu den Messaufbauten mit den verschiedenen Array-Typen sind in den Abbildungen 2.4, 2.5 und 2.6 gegeben. 0° Azimut entspricht einer Blickrichtung des Kunstkopfes nach vorne. Bei -30° Azimut wird der Kunstkopf um 30° im Uhrzeigersinn gedreht.

Beim Array-Typ "*Linie*" (siehe Abbildung 2.4) sind 24 Lautsprecher direkt nebeneinander (Kante an Kante) und auf Ohrhöhe zu positionieren.

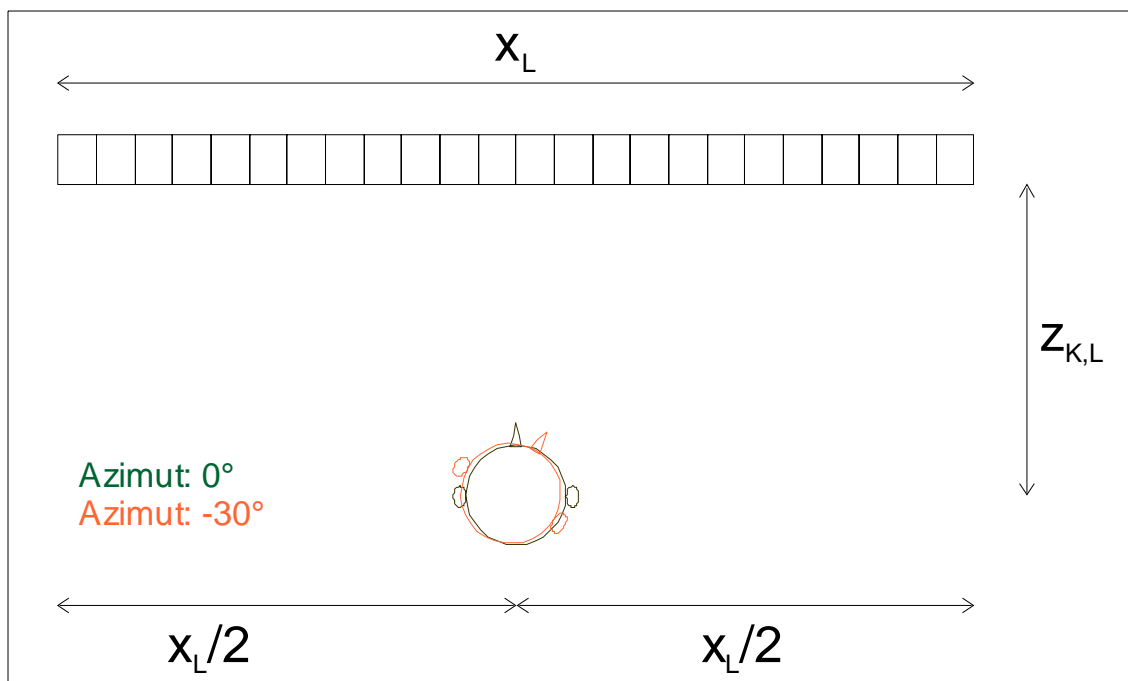


Abbildung 2.4: Messungen mit Array-Typ "*Linie*" (Ansicht von oben)

Der Array-Typ "Türme" (siehe Abbildung 2.5) besteht aus 2 mal 12 jeweils direkt übereinander (Kante an Kante) angebrachten Lautsprechern.

Die Lautsprecher sind in die Richtung der Ohren des Kunstkopfes zu drehen, wenn dieser auf 0° Azimut ausgerichtet ist - dies ist durch die gepunktete, lila Linie in Abbildung 2.5.a angedeutet.

Die Ohrhöhe soll sich genau zwischen dem 6. und 7. Lautsprecher, also in der "vertikalen Mitte" (y-Richtung) jeweils eines Turms befinden (siehe Abbildung 2.5.b).

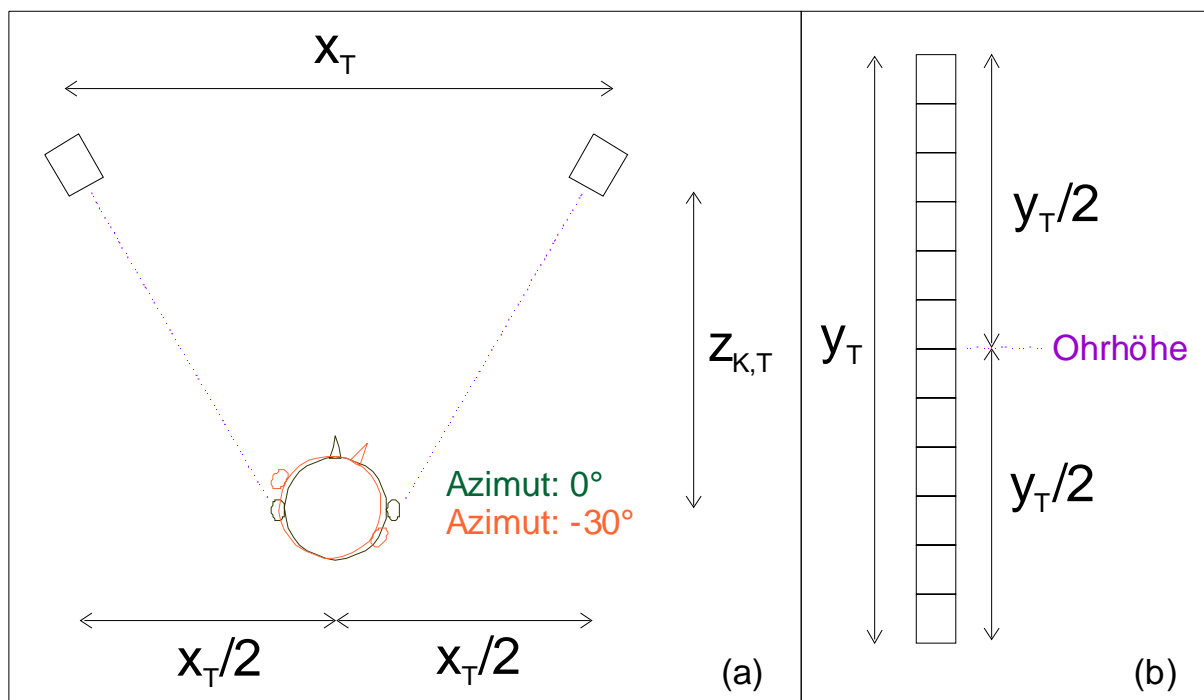


Abbildung 2.5: (a) Messungen mit Array-Typ "Türme" (Ansicht von oben),
(b) Darstellung eines Turms (der Turm wird von vorne angesehen)

Beim Array-Typ "oben" (Abbildung 2.6) sollen 2 mal 12 jeweils direkt hintereinander (Kante an Kante) befestigte Lautsprecher oberhalb des Kunstkopfes angebracht werden.

Wie aus Abbildung 2.6.a zu erkennen ist, befindet sich der Kunstkopf in z-Richtung genau in der "Mitte" des Arrays.

Die Lautsprecher sind in die Richtung der Ohren des Kunstkopfes zu drehen, wenn dieser auf 0° Azimut ausgerichtet ist - das wird durch die gepunktete, lila Linie in Abbildung 2.6.b angedeutet.

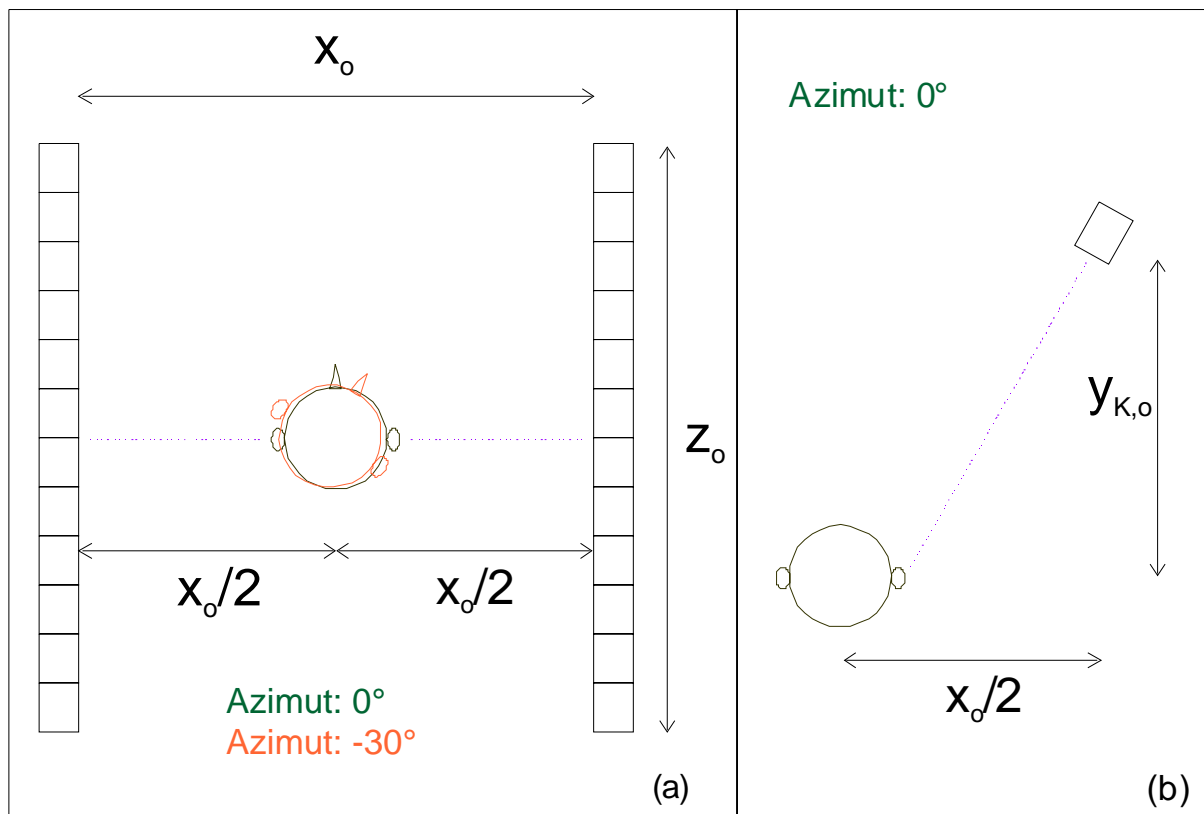


Abbildung 2.6: (a) Messungen mit Array-Typ "oben" (Ansicht von oben),
(b) Darstellung des Kunstkopfes und der rechten "Array-Seite" (Ansicht von hinten)

2.3.2. Abstände

Der Abstand gibt die Distanz in z-Richtung (siehe Abbildungen 2.4 und 2.5) bzw. in y-Richtung (siehe Abbildung 2.6) zwischen dem Eingang des Gehörkanals des Kunstkopfes bei 0° Azimut und der Vorderseite des Lautsprecher-Arrays an. In den Abbildungen 2.4, 2.5 und 2.6 werden diese Abstände mit $z_{K,L}$ bzw. $z_{K,T}$ bzw. $y_{K,o}$ bezeichnet (K... Kunstkopf, L... Linie, T... Türme, o... oben).

Die Messungen sollen bei allen drei Array-Typen für den Abstand 0,6 Meter und beim Array-Typ "oben" außerdem für den Abstand $y_{K,o} = 1,2$ Meter durchgeführt werden.

2.3.3. Quellenkonstellationen

Mit Quellenkonstellation ist die Anordnung der fokussierten Quellen um den Kunstkopf gemeint.

Um die verschiedenen Array-Typen miteinander vergleichen zu können, wird die Quellenkonstellation "4 Quellen (vorne)" bei allen drei Array-Typen realisiert (siehe unten).

Beim Array-Typ "oben" ist es möglich, fokussierte Quellen (auch) *hinter* dem Kopf zu platzieren. Bei den anderen beiden Array-Typen ist dies hingegen nicht möglich: Wenn der Schall ausschließlich von den *vor* dem Kopf aufgestellten, realen Lautsprechern (engl.: secondary sources) und damit von vorne kommt, kann keine Schallausbreitung mit entgegengesetzter Richtung simuliert werden (eine genauere Diskussion dieser Thematik findet sich z.B. in [7]).

Beim Array-Typ "oben" sollen die Messungen auch für die Quellenkonstellation "4 Quellen (außen)" durchgeführt werden (siehe unten).

Beiden Quellenkonstellationen gemein ist, dass bei der Betrachtung der Horizontalebene die fokussierten Quellen vom Kopfmittelpunkt jeweils 0,15 Meter entfernt zu platzieren sind.

Bei den Array-Typen "Linie" und "Türme" sind die Quellen auf Ohrhöhe des Kunstkopfes zu positionieren, beim Array-Typ "oben" um 0,2 Meter darüber - der Gesamtabstand von der fokussierten Quelle zum Mittelpunkt des Kunstkopfes ist beim Array-Typ "oben" also $\sqrt{0,15^2 + 0,2^2} = 0,25$ Meter.

Die Azimute der fokussierten Quellen sollen bei der jeweiligen Quellenkonstellation für Szenario 1 (siehe Kapitel 2.2.4) folgende Werte aufweisen (die Quellen werden im mathematisch positiven Sinn gedreht, d.h. positive Winkel bedeuten eine Drehung gegen den Uhrzeigersinn, negative Winkel eine Drehung im Uhrzeigersinn):

- "4 Quellen (vorne)": $\pm 22,5^\circ$ und $\pm 67,5^\circ$
- "4 Quellen (außen)": $\pm 67,5^\circ$ und $\pm 112,5^\circ$

Wie die Quellen bei den anderen Szenarien positioniert werden, wird basierend auf Szenario 1 in Kapitel 2.3.4. erläutert.

2.3.4. Szenarien

Pro Array-Typ, Quellenkonstellation und Abstand sollen die Messungen für folgende 4 "Szenarien" durchgeführt werden:

"Szenario 1": Der Kunstkopf blickt in 0° -Richtung (Azimut: 0°), es werden die HRTFs für diese Ausrichtung des Kunstkopfes bestimmt und daraus die IHRTFs berechnet (siehe Abbildung 2.7.a).

"Szenario 2": Der Kunstkopf wird um 30° im Uhrzeigersinn gedreht (Azimut: -30°). Es werden die unter Szenario 1 berechneten IHRTFs für 0° verwendet. Die fokussierten Quellen werden bei diesem Szenario *nicht* mitgedreht, d.h. sie bleiben absolut konstant, also an derselben Position wie unter Szenario 1 (siehe Abbildung 2.7.b).

"Szenario 3": Der Kunstkopf bleibt auf -30° Azimut eingestellt. Auch hier werden die unter Szenario 1 berechneten IHRTFs für 0° verwendet. Die relativen Abstände und die Winkel zwischen fokussierten Quellen und Kunstkopf bleiben aber bei diesem Szenario konstant. Die fokussierten Quellen werden also mit dem Kunstkopf mitgedreht (siehe Abbildung 2.7.c).

"Szenario 4": Der Kunstkopf blickt wieder in -30° -Richtung (Azimut). Im Gegensatz zu den Szenarios 2 und 3 werden aber die HRTFs für diese Ausrichtung des Kunstkopfes neu bestimmt und daraus die IHRTFs (nun für -30°) berechnet (siehe Abbildung 2.7.d).

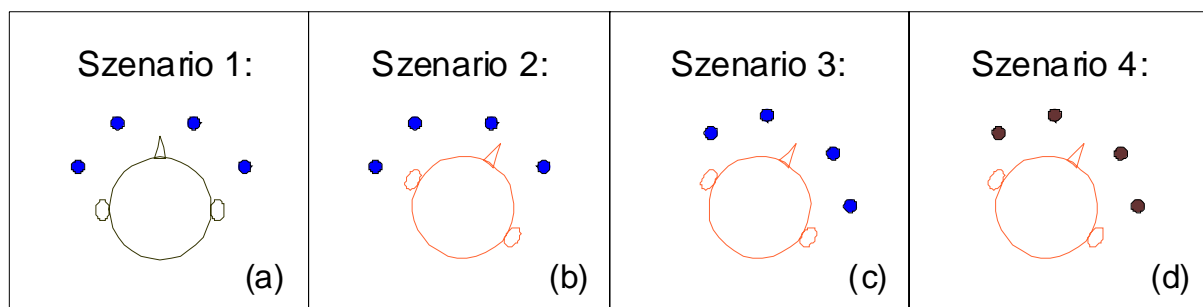


Abbildung 2.7: Szenarien 1 bis 4

2.4. Evaluierung der Messergebnisse

Die Übersprechkompensation beschreibt einen Aspekt der Qualität des virtuellen Kopfhörersystems. Anhand der bestimmten Gesamt-Übertragungsfunktionen soll der Einfluss der Array-Typen, Abstände, Quellenkonstellationen und Szenarien auf die Übersprechkompensation untersucht werden.

Die für die Evaluierung notwendigen Berechnungen sind mit Matlab durchzuführen. Auch die Graphiken sollen mit Matlab erstellt werden. Die erhaltenen Ergebnisse sind auszuwerten und zu analysieren.

3. Audiosignalverarbeitung im Messsystem

3.1. Erzeugung des exponentiellen Sweeps

Verschiedene gängige Methoden zur Messung von Übertragungsfunktionen im akustischen Bereich wurden von S. Müller und P. Massarani untersucht und beschrieben [8]. Das Paper dokumentiert ausführlich, wieso die Messung mit Hilfe von exponentiellen Sweeps für die Praxis am besten geeignet ist.

Die Länge des exponentiellen Sweeps beträgt T Sekunden. Bei einer Samplerate f_s ergibt sich daher eine Länge des Sweeps von $L_s \approx T \cdot f_s$ Samples (die Rundung erfolgt auf ganze Zahlen).

Die Startfrequenz des exponentiellen Sweeps ist f_1 . Da der Sweep exponentiell ist, ist es nicht möglich bzw. sinnvoll, ihn beim Frequenzbin $k=0$, also bei 0 Hertz beginnen zu lassen, da alle Vielfachen von 0 wieder 0 ergeben würden. Davon abgesehen ist ein DC-Anteil nicht hörbar und außerdem in der Audiotechnik generell möglichst gering zu halten, da er keinen praktischen Nutzen hat, aber Audiogeräte beschädigen kann.

Die Endfrequenz des Sweeps liegt bei f_2 .

Durch das harte Ein- bzw. Ausschalten des Sweeps bei der Start- bzw. Endfrequenz entstehen am Anfang und am Ende des Frequenzbereichs des Sweeps so genannte Ripples (siehe Abbildung A.1 im Anhang). Geht man von einem linearen System aus, stören diese jedoch bei der Bestimmung einer HRTF nicht. Aufgrund der Division im Frequenzbereich werden diese Ripples nämlich kompensiert - siehe Formel (3.1).

Hinzu kommt, dass unter Umständen die Lautsprecher / Mikrophone die ersten bzw. die letzten Frequenzen des Sweeps ohnehin noch nicht / nicht mehr wiedergeben / aufnehmen können.

3.2. Bestimmung aller HRTFs bzw. HRIRs

3.2.1. Messungen mit Kunstkopf

Um die HRTFs bestimmen zu können, werden Kunstkopf-Messungen durchgeführt (siehe Abbildung 3.1). Der Kunstkopf weist typische Abmessungen der Kopfgeometrie, der Außenohren und des Gehörgangs auf (die HRTFs sind für verschiedene Menschen etwas unterschiedlich, weswegen beim Messen eine typische Kopfform verwendet wird).

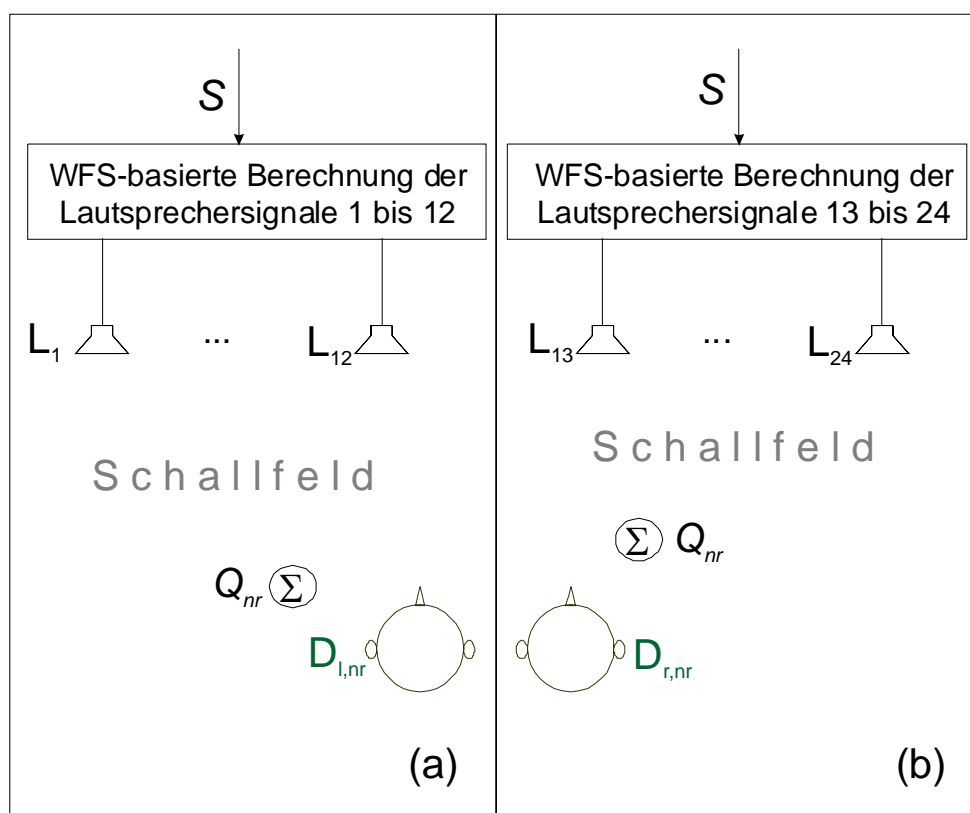


Abbildung 3.1.: Messungen mit Kunstkopf zur Bestimmung der HRTFs mit
(a) linksseitig platzierten fokussierten Quellen und
(b) rechtsseitig platzierten fokussierten Quellen.

Ein Sweep wird in das Wellenfeldsynthese-Programm geschickt, welches die Lautsprechersignale so berechnet, dass der Sweep am Ort der fokussierten Schallquelle abgespielt wird. Für die ersten $K/2$ Kanäle ($nr = 1, 2, \dots, K/2$) werden die 12 linken Lautsprecher verwendet, für die letzten $K/2$ Kanäle ($nr = K/2 + 1, K/2 + 2, \dots, K$) die 12 rechten Lautsprecher (siehe auch Kapitel 2.1. bzw. vgl. Abbildung 2.2.).

Der Kunstkopf wird auf 0° Azimut ausgerichtet. Als Schallempfänger sind Mikrophone in den linken und rechten Gehörgang des Kunstkopfes eingebaut, die die beiden akustischen Signale in analoge elektrische Signale umwandeln. Diese werden anschließend in digitale Signale umgesetzt.

Die HRTF von der fokussierten Schallquelle mit der Kanalnummer nr zum Trommelfell des linken ($b=l$) bzw. rechten ($b=r$) Ohres lässt sich dann folgendermaßen berechnen:

$$H_{(nr)b}[k] = \frac{D_{b,nr}[k] S^*[k]}{S[k] S^*[k]} = \frac{D_{b,nr}[k] \cdot S^*[k]}{\|S[k]\|^2}, \quad k = 0, 1, \dots, L_{lang} - 1 \quad (3.1)$$

$S[k]$ ($k = 0, 1, \dots, L_{lang} - 1$) ist die L_{lang} -Punkte-DFT des Sweeps, d.h. vor der Transformation in den Frequenzbereich wird der Sweep mit Nullen aufgefüllt (Zero Padding). $S^*[k]$ ist die zu $S[k]$ konjugiert komplexe Zahl.

Wie an Abbildung 3.1 zu erkennen ist, wird der Sweep also an der Stelle des fokussierten Quellsignals wiedergegeben. $D_{b,nr}[k]$ ($k = 0, 1, \dots, L_{lang} - 1$) ist die L_{lang} -Punkte-DFT des Signals, das daraufhin am linken ($b=l$) bzw. rechten ($b=r$) Mikrophon des Kunstkopfes ankommt.

Die HRIR wird erhalten, indem die IDFT auf die HRTF aus Gleichung (3.1) angewendet wird. Die HRIR wird anschließend durch Fensterung auf eine Länge von L Samples gebracht. Durch Anwenden der DFT auf diese HRIR erhält die "effektive" HRTF ebenso L Frequenzbins. Da der Radix-2-FFT-Algorithmus auf 2er-Potenzen ausgelegt ist, wird L als 2er-Potenz gewählt.

3.2.2. Latenz

Wie eine HRTF bestimmt wird, ist bereits aus (3.1) bekannt. In der Praxis treten aber Verzögerungen in der Signalkette auf:

Einerseits wird der Sweep nach der Ausgabe in Matlab im Wellenfeldsynthese-Programm digital bearbeitet, DA-umgesetzt und nach entsprechender Verstärkung über einen Lautsprecher ausgegeben. Es tritt also ein Delay zwischen dem Abspielen des Sweeps in Matlab bis zum Eintreffen des (nun analogen) Signals am Verstärker des Lautsprechers auf.

Andererseits tritt eine Verzögerung vom Eintreffen des Signals am Mikrophon des Kunstkopfes bis zum Aufnehmen des (AD-umgesetzten) Signals in Matlab auf.

Die Latenz gibt nun die Kombination dieser beiden Verzögerungen in Samples an. Im Wesentlichen gibt es zwei Möglichkeiten, die Latenz in der Praxis zu berücksichtigen: Eine Möglichkeit ist, das Zeitfenster bei der Messung der HRIR so lange zu wählen, dass die Latenz mit gemessen wird.

Eine andere Möglichkeit ist, die Latenz (einmal) zu messen. In Matlab lässt sich diese Latenz dann bei allen folgenden Messungen über entsprechende Fensterung herausrechnen, wodurch das Zeitfenster bei der Messung der HRIR wesentlich kürzer gewählt werden kann.

Letztere Vorgehensweise ist aus folgenden Gründen sinnvoller: Das kürzere Zeitfenster bewirkt bei der Messung, dass weniger Rauschen aufgenommen wird und dadurch der SNR im Vergleich zu einem längeren Zeitfenster höher ist. Außerdem wird dadurch auch der Aufwand (benötigter Speicherplatz) geringer.

Voraussetzung für dieses Vorgehen ist, dass die Latenz bei einem bestimmten Messaufbau immer konstant bleibt. Bei Verwenden der unter Kapitel 4.1. genannten Hardware und der unter Kapitel 4.3. beschriebenen Parametereinstellungen war dies der Fall.

Der Vollständigkeit halber sei erwähnt, dass folgende Verzögerungen aus praktischen Gründen mit gemessen werden müssen:

1.) Die Verzögerung, die durch den Verstärker verursacht wird (diese sollte sich jedoch nur im Bereich einiger, weniger Samples abspielen): Da passive Lautsprecher verwendet werden, wäre zwar das (für den Lautsprecher) verstärkte Signal theoretisch zugänglich. Jedoch ist ein Audiointerface für gewöhnlich nicht auf derartige Pegel ausgelegt, weshalb zu dessen Schutz das unverstärkte Signal (Line-Pegel) verwendet wird.

2.) Bei Messungen mit durch Wellenfeldsynthese erzeugte fokussierten Quellen muss außerdem jene Verzögerung mit gemessen werden, die durch den größten Abstand zwischen Lautsprecher und fokussierter Quelle und der damit verbundenen Schalllaufzeit entsteht. Dies ist u.a. deshalb notwendig, weil die Schallgeschwindigkeit temperaturabhängig ist und daher die Schalllaufzeit zwischen Lautsprecher und fokussierter Quelle nicht exakt konstant bleibt.

3.2.3. Bandbreite für die Glättung der Amplitudengänge

Die Amplitudengänge der HRTFs können in relativen Frequenzbändern mit der Bandbreite von $B=1/3$ Oktaven geglättet werden (d.h. pro Frequenz werden die Amplitudenwerte einer halben Terz unterhalb dieser Frequenz bis zu einer halben Terz oberhalb dieser Frequenz gemittelt). Der Phasengang wird durch die Glättung nicht beeinflusst.

Eine Glättung in Terz- bzw. 1/3-Oktavbändern ist aus folgenden Gründen sinnvoll:

1.) Starke Amplitudenschwankungen bei den HRTFs haben auch starke Amplitudenschwankungen bei den IHRTFs zur Folge. Nun kann es passieren, dass einzelne Frequenzbänder sowohl im Kopfhörer-Signal als auch in den IHRTFs hohe Amplituden aufweisen. Die Wiedergabesignale werden aus den jeweiligen IHRTF-gefilterten Kanälen des Kopfhörersignals berechnet (siehe Abbildung 2.1.). Werden diese Wiedergabesignale nun abgespielt, so kann dies bei zu hoher Aussteuerung zur Beschädigung eines oder mehrerer Lautsprecher führen. Eine Beschädigung des/der Lautsprecher(s) ließe sich zwar verhindern, indem die Signale leise genug abgespielt werden - allerdings wirkt sich eine Absenkung der Signalpegel ungünstig auf den SNR aus.

Durch Glättung des Amplitudengangs können starke Amplitudenschwankungen und vereinzelte, schmalbandige Notches bzw. Peaks beseitigt werden (vgl. Abbildung 3.2).

2.) Auch das menschliche Gehör glättet (in grober Approximation) etwa in Terzen. Eine Glättung in Terzbändern bildet also bis zu einem gewissen Grad das subjektive Hörempfinden nach.

Die Frequenzbins in Abbildung 3.2 werden durch jene Frequenzen ersetzt, die diesen Frequenzbins entsprechen. Die Skalierung der Frequenzachse ist logarithmisch, da dies dem Gehörempfinden besser entspricht. Die Frequenzachse wird auch bei allen weiteren Amplituden- bzw. Frequenzgängen in diesem Dokument (wie in diesem Absatz beschrieben) dargestellt.

Die Amplituden-geglätteten HRTFs lassen sich für jeweils einen Frequenzbin k zu einer HRTF-Matrix zusammenfassen:

$$\underline{H}[k] = \begin{bmatrix} H_{1l}[k] & \dots & H_{Kl}[k] \\ H_{1r}[k] & \dots & H_{Kr}[k] \end{bmatrix} \quad (3.2)$$

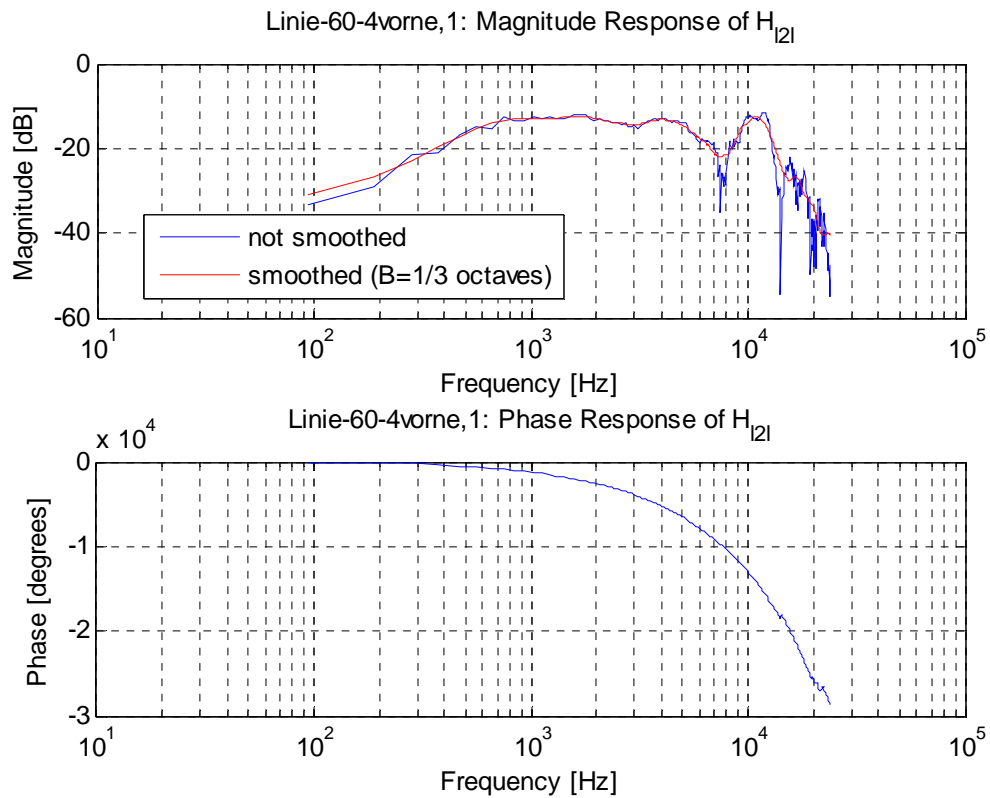


Abbildung 3.2: HRTF ohne bzw. mit Glättung in Terzbändern (hier: Array-Typ "Linie", Abstand 0,6 Meter, 4 vorne fokussierte Quellen, Szenario 1)

3.3. Berechnung der IHRTFs

Wie bereits in Kapitel 2.1. diskutiert wurde, werden die IHRTFs benötigt, um die durch das System verursachten HRTFs zu kompensieren.

Um die IHRTFs zu erhalten, wird zunächst für jeden Frequenzbin k die Pseudoinverse¹ der HRTF-Matrix $\underline{H}[k]$ berechnet. Das Ergebnis ist dann jeweils eine IHRTF-Matrix $\underline{G}[k]$:

$$\underline{G}[k] = \begin{bmatrix} G_{l1}[k] & \dots & G_{lK}[k] \\ G_{r1}[k] & \dots & G_{rK}[k] \end{bmatrix}^T \quad (3.3)$$

In der IHRTF-Matrix sind alle IHRTFs an der Stelle des jeweiligen Frequenzbins enthalten. Wenn man also alle Elemente einer bestimmten Zeile der ersten bzw. zweiten Spalte der Matrizen $\underline{G}[k]$ zu einer Frequenzantwort zusammenfasst, so ist dies dann jeweils eine IHRTF. IHRTFs gibt es folglich genauso viele wie HRTFs. Falls nur 2 Kanäle verwendet werden ($K = 2$) und die HRTF-Matrix somit 2 Spalten hat, ist $\underline{H}[k]$ quadratisch. Die IHRTF-Matrix $\underline{G}[k]$ entspricht in diesem Fall der invertierten HRTF-Matrix $\underline{H}^{-1}[k]$.

Abbildung 3.3. vereinfacht die Darstellung des virtuellen Kopfhörersystems aus Abbildung 1.1., indem die einzelnen Filter durch die Matrizen \underline{G} und \underline{H} ersetzt werden (die Abbildung gilt jeweils für einen Frequenzbin k).

Zur Erzeugung fokussierter Schallquellen mit der Wellenfeldsynthese werden reale Lautsprecher (secondary sources) benötigt. Reale Lautsprecher haben aber eine untere und eine obere Grenzfrequenz, d.h. sie können nicht beliebig tiefe und hohe Frequenzen wiedergeben. Auch A/D-Umsetzer haben eine obere Grenzfrequenz, die in diesem Fall benötigt wird, um Aliasing zu verhindern.

Aufgrund dessen fällt der Amplitudengang bei den HRTFs unterhalb der unteren Grenzfrequenz und oberhalb der oberen Grenzfrequenzen ab.

Dies bewirkt aber bei den inversen HRTF-Filtern umgekehrt ein Ansteigen des jeweiligen Amplitudengangs unterhalb bzw. oberhalb der Grenzfrequenzen. Dieses Ansteigen ist unerwünscht: wenn die Wiedergabesignale (Filterung des Kopfhörersignals mit IHRTFs) abgespielt werden, können Verzerrungen auftreten. Im

¹ Die Berechnung der Pseudoinversen geschieht mit Hilfe der Matlab-Funktion "pinv". Details zur Implementierung dieser Funktion sind der Matlab-Hilfe zu entnehmen.

schlimmsten Fall können die Lautsprecher beschädigt werden. Um diese Probleme zu umgehen, können zwar die Signale entsprechend leise abgespielt werden - dies wirkt sich jedoch wiederum negativ auf den SNR aus.

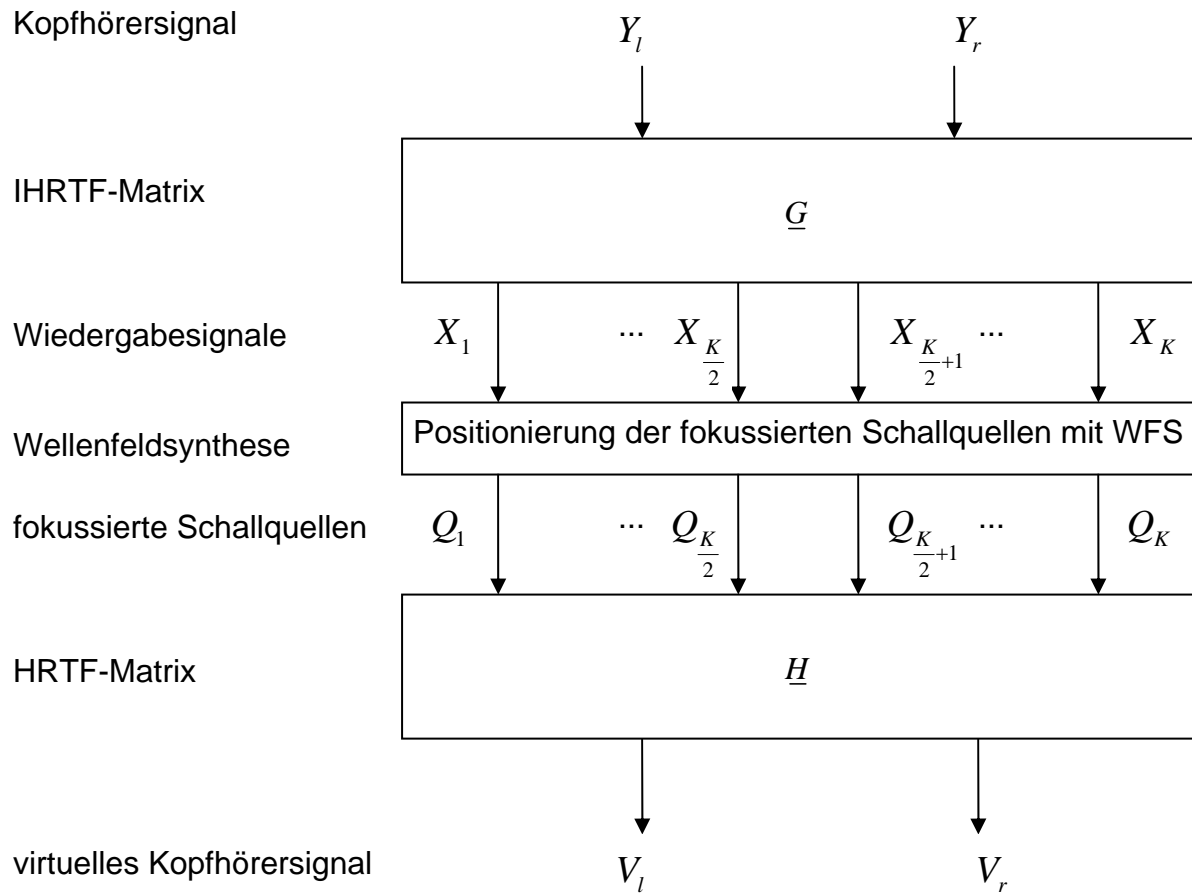


Abbildung 3.3: virtuelles Kopfhörersystem

Aus den genannten Gründen werden die Amplitudengänge der IHRTFs daher Bandpass-gefiltert. Realisiert wird der Bandpass als Butterworth-Filter $2 \cdot M$ -ter Ordnung ($M \in \mathbb{N}$).

Um keine Phasenverzerrungen zu erhalten, wird der Absolutbetrag der Bandpass-Übertragungsfunktionen gebildet und der Butterworth-Filter somit nullphasig gemacht.

Die Filterkoeffizienten, die im Faltungs-Programm BruteFIR ([6]) verwendet werden, sind die IDFTs dieser Bandpass-gefilterten IHRTFs. Diese werden im Folgenden als IHRIRs (Inverse Head Related Impulse Responses) bezeichnet.

3.4. Offline-Berechnung der Wiedergabesignale

Wie aus dem Blockschaltbild in Abbildung 1.1. ersichtlich, lassen sich die Wiedergabesignale im Frequenzbereich folgender Maßen berechnen.

$$X_{nr}[k] = G_{l(nr)}[k] \cdot Y_l[k] + G_{r(nr)}[k] \cdot Y_r[k], \quad k = 0, 1, \dots, L_x, \quad nr = 1, 2, \dots, K \quad (3.4)$$

Alternativ kann (3.4) auch folgender Maßen angeschrieben werden (siehe (3.3) und Abbildung 3.3):

$$\underline{X}[k] = \underline{G}[k] \cdot \underline{Y}[k], \quad k = 0, 1, \dots, L_x, \quad (3.5)$$

In (3.5) sind $\underline{X}[k] = [X_1[k] \dots X_K[k]]^T$ und $\underline{Y}[k] = [Y_l[k] \ Y_r[k]]^T$.

Die beiden Multiplikationen in (3.4) entsprechen linearen Faltungen im Zeitbereich. Die Länge L_x jeweils eines Wiedergabesignals ist daher die Faltungslänge $L_y + L - 1$, wobei L_y der Anzahl der Samples eines Kanals des Kopfhörersignals entspricht und L die Länge einer IHRIR ist (diese ist genauso wie die HRIR 512 Samples lang - siehe Kapitel 4.3.2. und 4.3.3.).

Die K Wiedergabesignale im Zeitbereich werden erhalten, indem pro Kanalnummer nr die IDFT von (3.4) berechnet wird.

3.5. Bestimmung der Gesamt-Übertragungsfunktionen

Die Gesamtübertragungsfunktionen, d.h. die Übertragungsfunktionen vom linken bzw. rechten Kanal des Kopfhörersignals zum linken bzw. rechten Kanal des virtuellen Kopfhörersignals berechnen sich nach (siehe Abbildungen 1.1 und 3.3):

$$H_{eve,ab}[k] = \frac{V_{b,a}[k] Y_a^*[k]}{Y_a[k] Y_a^*[k]} = \frac{V_{b,a}[k] \cdot Y_a^*[k]}{\|Y_a[k]\|^2} \quad \text{für } a = l, r, b = l, r, \quad k = 0, 1, \dots, L_{ges} - 1 \quad (3.6)$$

In (3.6) ist $V_{b,a}[k]$ ($k = 0, 1, \dots, L_{ges} - 1$) die L_{ges} -Punkte-DFT des linken ($b = l$) bzw. rechten ($b = r$) Kanals des virtuellen Kopfhörersignals, wenn nur der linke ($a = l$)

bzw. rechte ($a = r$) Kanal des Kopfhörersignals tatsächlich ein Audiosignal beinhaltet. Der jeweils andere Kanal des Kopfhörersignals beinhaltet ein Nullsignal, ist also stumm geschaltet.

$Y_a^*[k]$ in (3.6) ist die zu $Y_a[k]$ konjugiert komplexe Zahl.

Die Gesamt-Übertragungsfunktionen werden bei der Evaluierung des virtuellen Kopfhörersystems zur Berechnung der Übersprechkompensation verwendet (siehe Kapitel 5.1.). Daher sollte der linke ($a = l$) bzw. rechte ($a = r$) Kanal des Kopfhörersignals möglichst den gesamten Frequenzbereich abdecken, der bei der Evaluierung der Messergebnisse berücksichtigt wird. Ein entsprechend parametrierter exponentieller Sweep erfüllt diese Anforderung, weswegen ein solcher für den linken bzw. rechten Kanal des Kopfhörersignals verwendet wird.

Die Gesamt-Impulsantwort $h_{eve,ab}$ ($a = l, r$, $b = l, r$) wird erhalten, indem die IDFT auf die Gesamt-Übertragungsfunktion aus Gleichung (3.6) angewendet wird. Die Gesamt-Impulsantwort wird anschließend durch Fensterung auf eine Länge von L_{eve} Samples gebracht. Durch Anwenden der DFT auf diese Gesamt-Impulsantwort erhält die "effektive" Gesamt-Übertragungsfunktion L_{eve} Frequenzbins. L_{eve} wird als 2er-Potenz gewählt, da der Radix-2-FFT-Algorithmus auf 2er-Potenzen ausgelegt ist.

Die optische Beurteilung der Amplitudengänge der Gesamtübertragungsfunktionen fällt leichter, wenn diese geglättet werden. Um das zu illustrieren zeigt Abbildung 3.4 die gleichen Gesamt-Übertragungsfunktionen ohne bzw. mit gleitender Terzband-Glättung der Amplitudengänge (die Parameter sind in diesem Beispiel: Array-Typ "Linie", Abstand 0,6 Meter, "4 vorne fokussierte Quellen", Szenario 1). Die Phasengänge werden durch die Glättung nicht beeinflusst.

Wie bereits in Kapitel 3.2.3. diskutiert, bildet eine Glättung in Terzbändern außerdem (grob approximiert) das subjektive Hörempfinden nach. Daher werden die Amplituden-geglätteten Gesamt-Übertragungsfunktionen später zur Berechnung der Übersprechkompensationen verwendet (siehe Kapitel 5.1.).

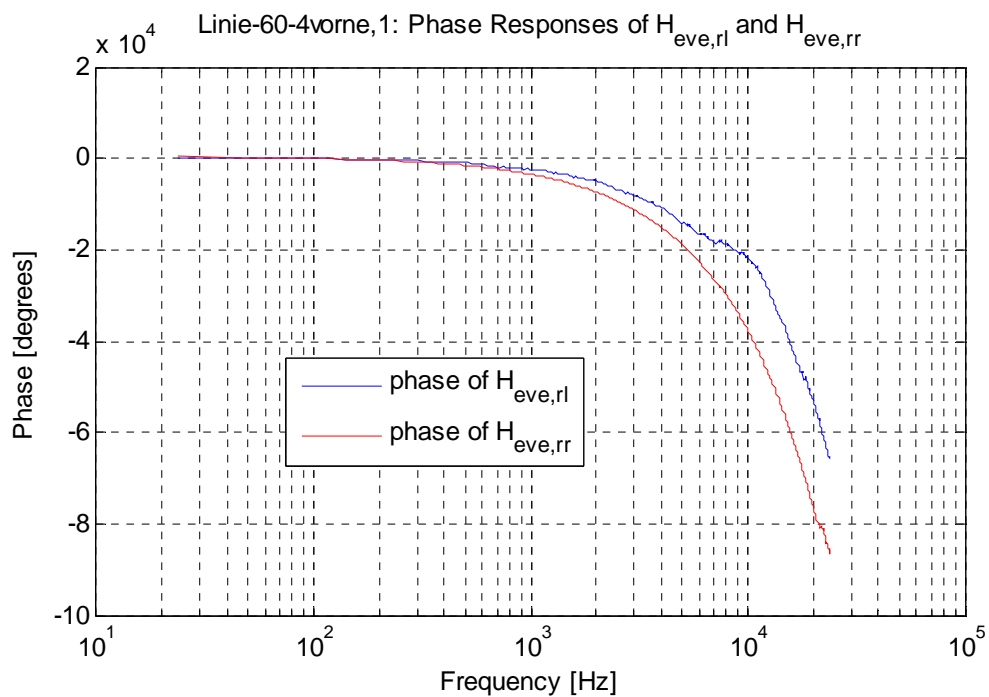
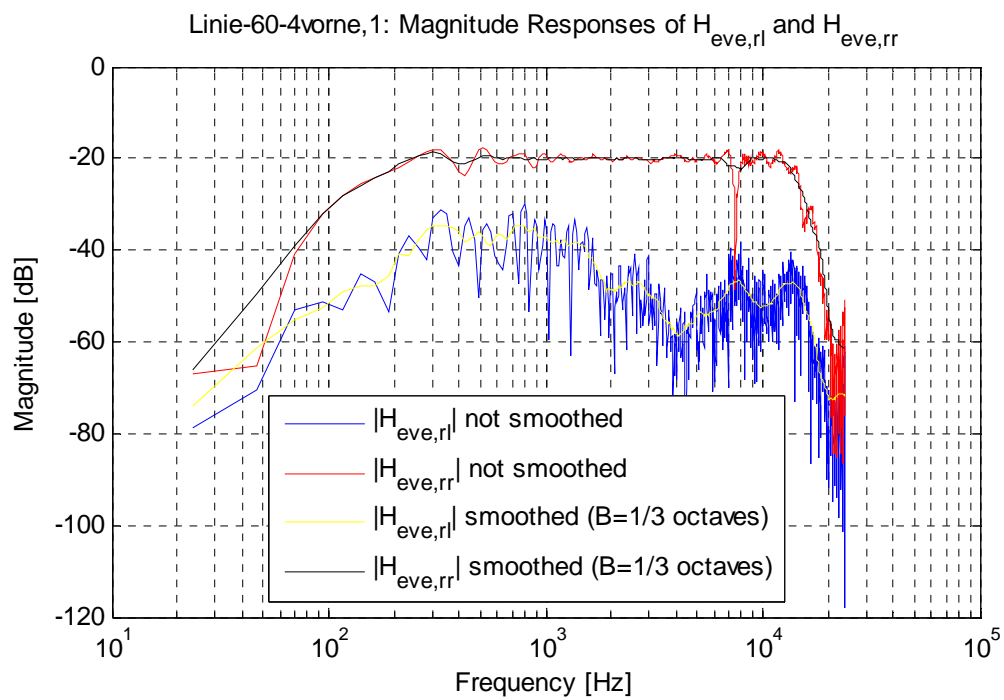


Abbildung 3.4: Gesamt-Übertragungsfunktionen ohne bzw. mit Glättung
(Array-Typ "Linie", Abstand 0,6 Meter, "4 vorne fokussierte Quellen", Szenario 1)

4. Messungen

4.1. Verwendete Hardware bei den Messungen

Zum Durchführen der Messungen im reflexionsarmen Raum wurde folgende Hardware verwendet:

- Laptop: Samsung P55 mit Intel Core 2 CPU (je 2,1 GHz) und 2 GB RAM (Betriebssystem: Windows XP)
- Audio -Interface (für Laptop): RME Fireface 800
- Shuttle XPC mit Intel Core 2 CPU (je 2,66 GHz) und 3 GB RAM (Betriebssystem: Linux-Distribution Ubuntustudio)
- Audio-Interface (für Shuttle XPC): RME HDSP Digiface
- Kunstkopf: Neumann KU 100
- Clock-Generator: Rosendahl nanosync (Synchronisation der Audio-Hardware über Word-Clock)
- DA-Umsetzer (2 mal): creamware A 16 ultra
- Analoge Verstärker (6 mal): Terratec Sine PA 460
- Passive Lautsprecher (24 mal): Elac 301
- Mikrophon-Vorverstärker (2 mal): Stage Line MPA-102 (verwendet als Trittschallfilter mit Grenzfrequenz 100 Hertz)

4.2. Messaufbauten im reflexionsarmen Raum

Bevor Messungen mit den jeweiligen Lautsprecher-Arrays möglich waren, mussten mechanische Vorrichtungen für die Befestigung der Lautsprecher-Arrays im reflexionsarmen Raum (RAR) gebaut werden.

Jene Abstände, die durch die Angabe noch nicht festgelegt wurden, wurden so gewählt, dass sie sich einerseits in einer sinnvollen Größenordnung befanden und andererseits möglichst unkompliziert mechanisch realisierbar waren.

Als Lautsprecher wurden 24 Elac 301 - Boxen (Abmessungen: 9 cm Breite, 12 cm Höhe, 9,5 cm Tiefe) verwendet.

Der Kunstkopf wurde so genau wie möglich entsprechend den Abbildungen 2.4, 2.5 und 2.6 positioniert. Er wurde auf ein oder zwei Holzbretter gestellt, damit er bei den Messungen stabil stehen blieb. Die Reflexionen der glatten Bretter waren aber unerwünscht, da ja Freifeld-Bedingungen simuliert werden sollten - daher wurden die Bretter mit Schaumstoffen überdeckt.

Im Folgenden werden die Messaufbauten der jeweiligen Array-Typen beschrieben (vgl. Abbildungen 2.4, 2.5 und 2.6).

4.2.1. Linie

Die Unterseite des Lautsprecher-Arrays (der Boden der Lautsprecher-Boxen) befand sich 0,9 Meter über dem Gitterboden des reflexionsarmen Raums.

Die Ohrhöhe $y_{K,B,L}$, d.h. der Abstand des Gehöreingangs des Kunstkopfes vom Gitterboden, wurde auf die mittlere Höhe der Lautsprecherboxen, also auf $0,9 + \frac{0,12}{2} = 0,96$ Meter eingestellt.

Die Breite des Lautsprecher-Arrays x_L war sowohl berechnet als auch gemessen ca. $24 \cdot 0,09 = 2,16$ Meter.

In Abbildung 4.1 ist das gesamte Linien-Array zu sehen. Außerdem ist der Kunstkopf abgebildet, der auf 0° Azimut ausgerichtet ist. Der Abstand $z_{K,L}$ ist $z_{K,L} = 0,6$ Meter.



Abbildung 4.1: Array-Typ "Linie", Azimut der Kunstkopfdrehung: 0° , $z_{K,L} = 0,6 \text{ m}$

Die Abbildung 4.2 zeigt das Linien-Array schräg von hinten. Der Kunstkopf ist auf -30° Azimut ausgerichtet. Der Abstand $z_{K,L}$ beträgt auch bei diesem Photo $z_{K,L} = 0,6$ Meter:

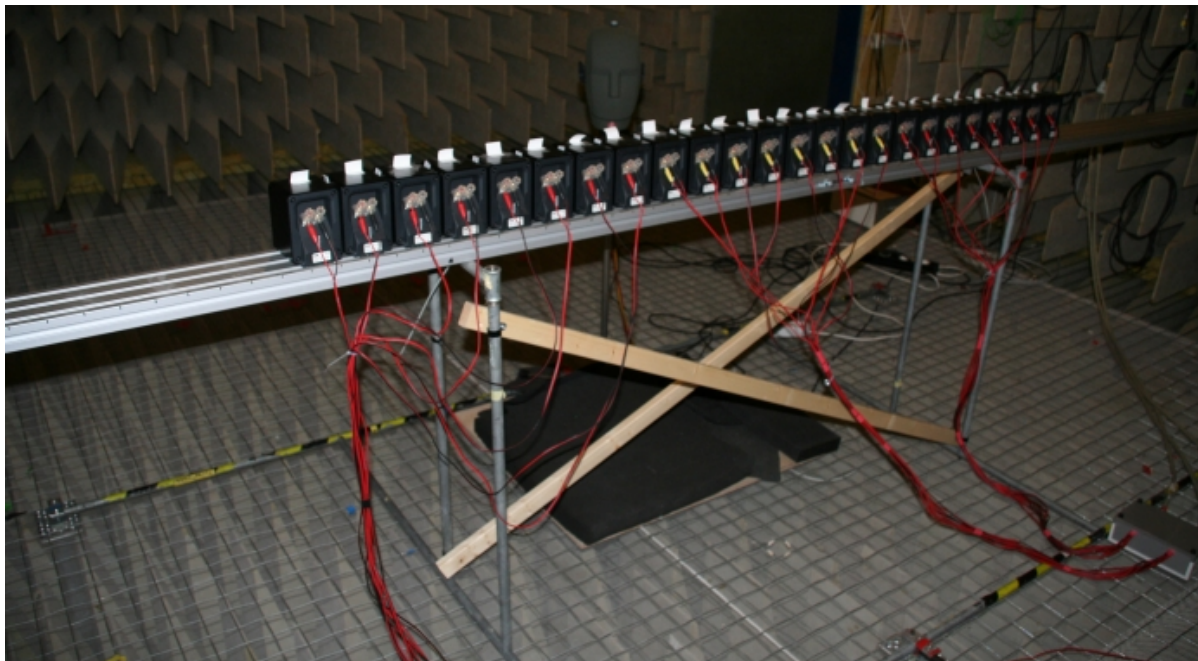


Abbildung 4.2: Array-Typ "Linie", Azimut der Kunstkopfdrehung: -30° , $z_{K,L} = 0,6 \text{ m}$

4.2.2. Türme

Die Grundflächen der beiden Lautsprechertürme befanden sich jeweils 0,325 Meter über dem Gitterboden. Die Lautsprechertürme waren jeweils $y_T = 1,45$ Meter hoch (der gemessene Wert stimmt mit dem berechneten Wert $12 \cdot 0,12 = 1,44$ Meter ca. überein) und reichten daher bis 1,775 Meter über den Gitterboden.

Die Ohrhöhe $y_{K,B,T}$, d.h. der Abstand des Gehöreingangs des Kunstkopfes vom Gitterboden, wurde gemäß der Angabe auf den Abstand zwischen der Höhe eines halben Lautsprecherturms und dem Gitterboden festgelegt. Dies ergab sowohl gemessen als auch berechnet den Wert $0,325 + \frac{1,45}{2} = 1,05$ Meter.

Der Abstand "zwischen" den Türmen x_T (der Abstand in x-Richtung von der halben Breite des einen Turms zur halben Breite des anderen Turms an den Rückseiten der Lautsprecher) betrug $x_T = 1,535$ Meter.

In Abbildung 4.3 ist das Turm-Array von vorne zu sehen: Der Kunstkopf ist auf -30° ausgerichtet, der Abstand $z_{K,T}$ beträgt $z_{K,T} = 0,6$ Meter.

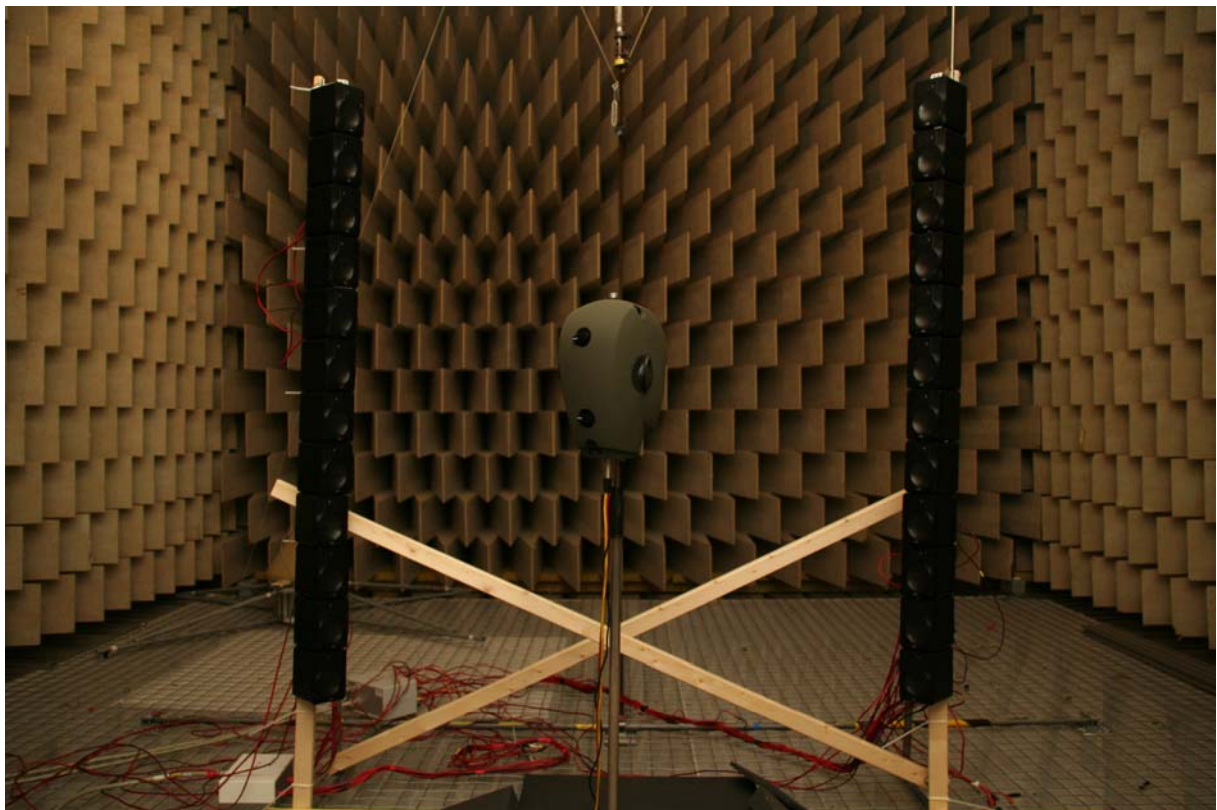


Abbildung 4.3: Array-Typ "Türme", Azimut der Kunstkopfdrehung: -30° , $z_{K,T} = 0,6$ m

Abbildung 4.4 zeigt die Türme von schräg hinten. Der Kunstkopf ist auf 0° Azimut ausgerichtet, $z_{K,T} = 0,6$ Meter.

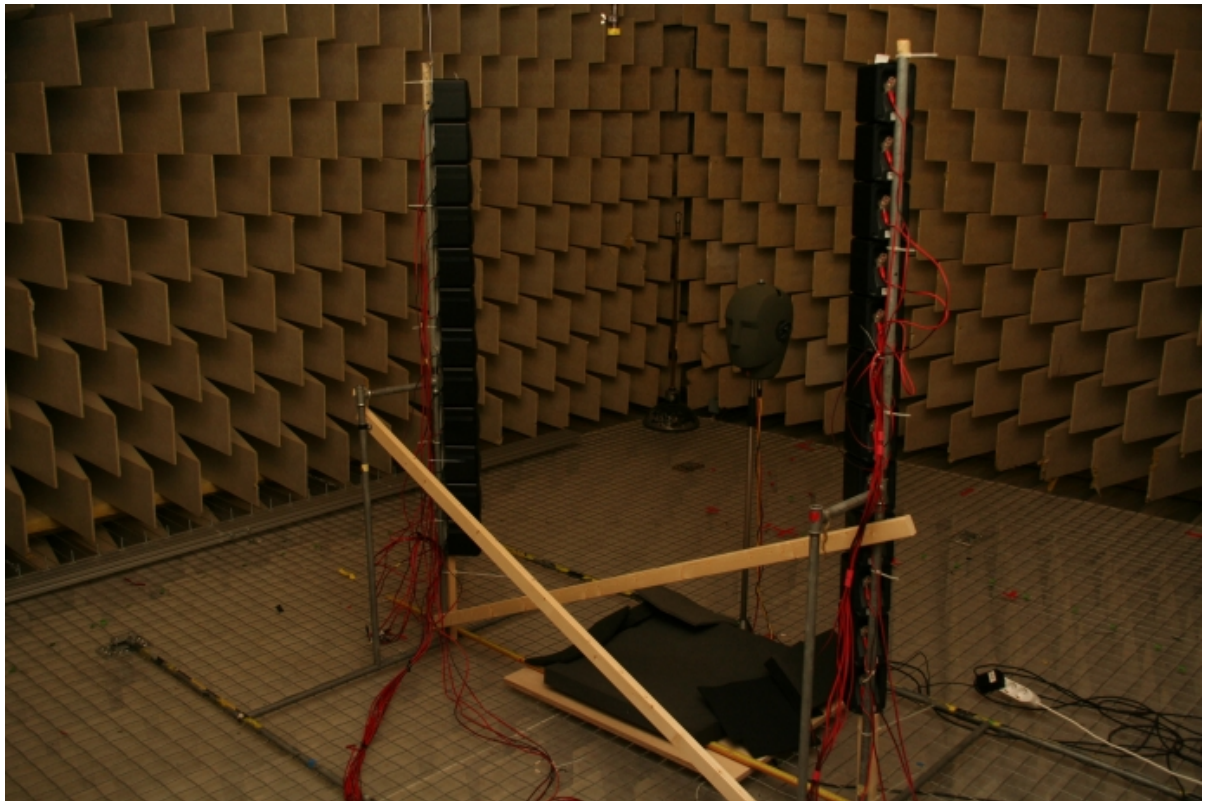


Abbildung 4.4: Array-Typ "Türme", Azimut der Kunstkopfdrehung: 0° , $z_{K,T} = 0,6$ m

4.2.3. Oben

Aufgrund der wesentlich einfacheren praktischen Realisierbarkeit wurden die Lautsprecher nicht an der Decke aufgehängt, sondern an den gleichen Positionen wie beim Array-Typ "Türme" gelassen. Durch das Kippen des Kunstkopfes um 90° wurde jedoch simuliert, dass der Schall von oben kommt.

Gerechtfertigt wird diese Vorgehensweise durch die Tatsache, dass die Schallausbreitung im reflexionsarmen Raum näherungsweise der im Freifeld entspricht. Unter dem begehbaren Gitterboden des reflexionsarmen Raums befinden sich Absorber, weswegen auch die Bodenreflexionen sehr gering sind.

Dass der Kunstkopf nach oben und nicht nach unten schauen sollte, hat folgenden Grund: Die zwar schwachen, aber dennoch vorhandenen Schall-Reflexionen am Gitterboden haben in diesem Fall einen geringeren und weniger störenden Einfluss auf die akustischen Signale, die an den Mikrofonen des Kunstkopfes ankommen.

Der Abstand "zwischen" den jeweils 12 Lautsprechern x_o blieb also wie beim Array-Typ "Türme" $x_o = 1,535$ Meter, z_o wurde nun $z_o = 1,45$ Meter (gleicher Wert wie y_T).

Abbildung 4.5 zeigt die praktische Realisierung dieses Array-Typs. Auf dem Bild ist der Kunstkopf auf 0° Azimut eingestellt, der Abstand $y_{K,o}$ ist $y_{K,o} = 1,2$ Meter.

In Abbildung 4.6 ist die praktische Realisierung dieses Array-Typs für die Ausrichtung des Kunstkopfes auf 30° Azimut und für $y_{K,o} = 1,2$ Meter zu sehen.

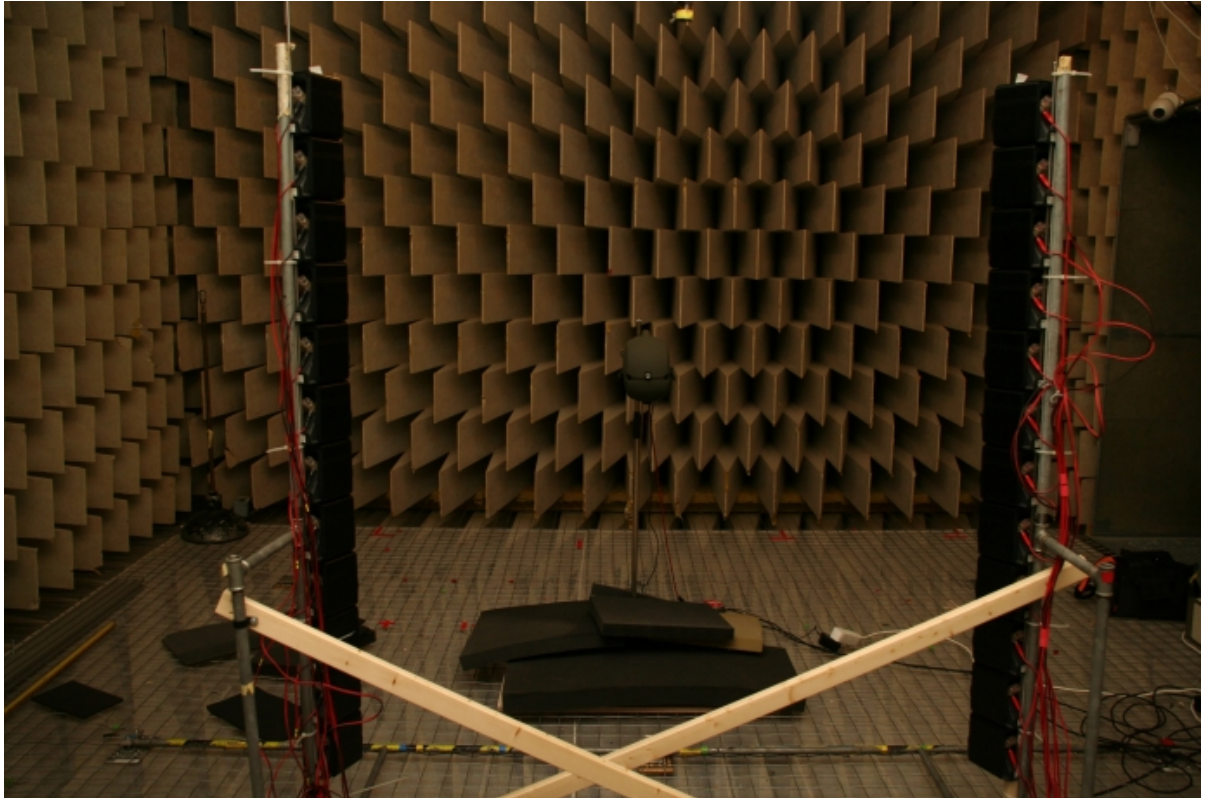


Abbildung 4.5: Array-Typ "oben", Azimut der Kunstkopfdrehung: 0° , $y_{K,o} = 1,2 \text{ m}$

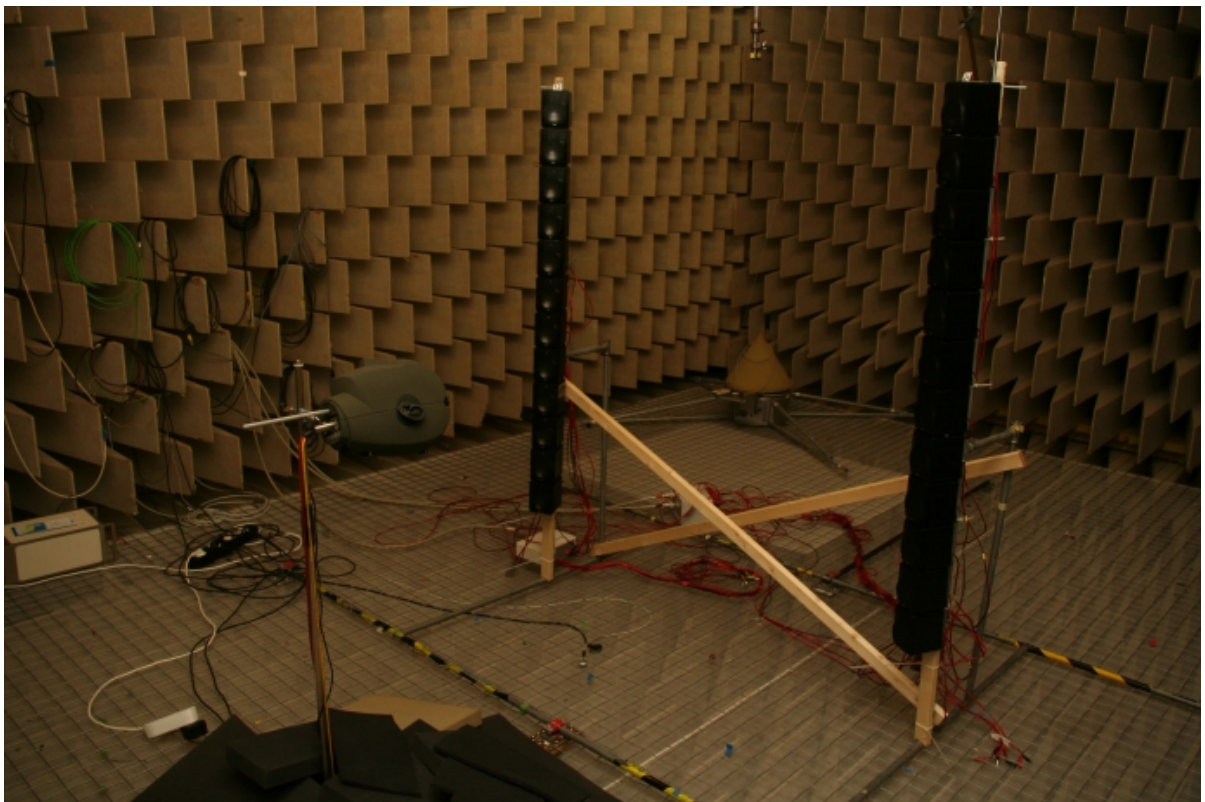


Abbildung 4.6: Array-Typ "oben", Azimut der Kunstkopfdrehung: -30° , $y_{K,o} = 1,2 \text{ m}$

4.3. Parameter-Werte im Messsystem

4.3.1. Erzeugung des exponentiellen Sweeps

Die Samplerate f_s wurde auf 48kHz gesetzt. Diese entsprach der Samplerate des Clock-Generators "Rosendahl nanosync".

Als Startfrequenz für den Sweep wurde $f_1 = 94$ Hz gewählt. Da der Trittschaltfilter Frequenzen unter 100 Hz wegschnitt und die untere Grenzfrequenz des Bandpasses $f_{c1_{bp}}$ bei der Berechnung der inversen HRTF-Filter ohnehin auf 120 Hz eingestellt wurde (siehe Kapitel 4.3.3.), war diese Startfrequenz ausreichend tief.

Die Endfrequenz des Sweeps wurde auf die halbe Samplerate festgelegt, d.h. $f_2 = f_s / 2 = 24$ kHz.

4.3.2. Bestimmung aller HRTFs und HRIRs

Die effektive Länge der HRIRs wurde mit $L = 2^9 = 512$ Samples gewählt. Bei der Wahl des Wertes für L waren folgende Überlegungen ausschlaggebend:

Wie bereits im Kapitel 3.2.2. diskutiert wurde, musste in der HRIR aus praktischen Gründen jene Verzögerung mit gemessen werden, die aus dem größten Abstand zwischen Lautsprecher und fokussierter Quelle und der damit verbundenen Schalllaufzeit resultierte. L musste also lang genug sein, dass die HRIR inklusive der Verzögerung darin enthalten war.

Umgekehrt sollte L auch nicht zu lang gewählt werden, da sonst der Aufwand (benötigter Speicherplatz) höher geworden wäre.

Eine HRIR-Länge von 512 Samples schien hier ein guter Kompromiss zu sein: Das Zeitfenster war lang genug, dass selbst beim Abstand 1,2 Meter (gemeint ist hier der Abstand, wie er in Kapitel 2.3.2. definiert wurde) die HRIR inklusive der Latenz darin enthalten war.

Abbildung 4.7 zeigt als Beispiel eine solche HRIR (hier: Array-Typ "Linie", Abstand: 0,6 Meter, 4 (vorne) fokussierte Quellen, Szenario 1).

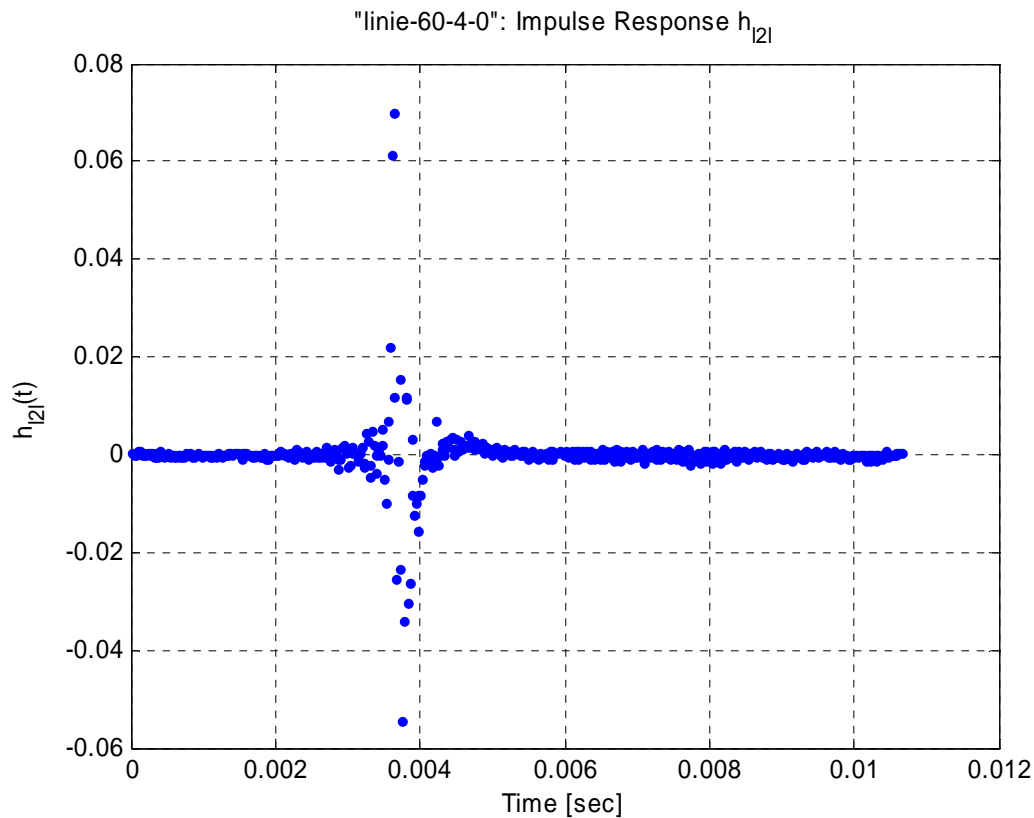


Abbildung 4.7: HRIR (hier: Array-Typ "Linie", Abstand: 0,6 Meter, 4 vorne fokussierte Quellen, Szenario 1)

Die Bandbreite für die Glättung der Amplitudengänge der HRTFs wurde auf eine Drittel-Oktave eingestellt.

4.3.3. Berechnung der IHRTFs

Da die Amplitudenwerte der IHRIRs bereits vor L Samples im Rauschen untergingen, wurde die Länge der IHRIRs genauso wie die der HRIRs mit L Samples gewählt. Ein Beispiel einer solchen IHRIR ist in Abbildung 4.8 dargestellt.

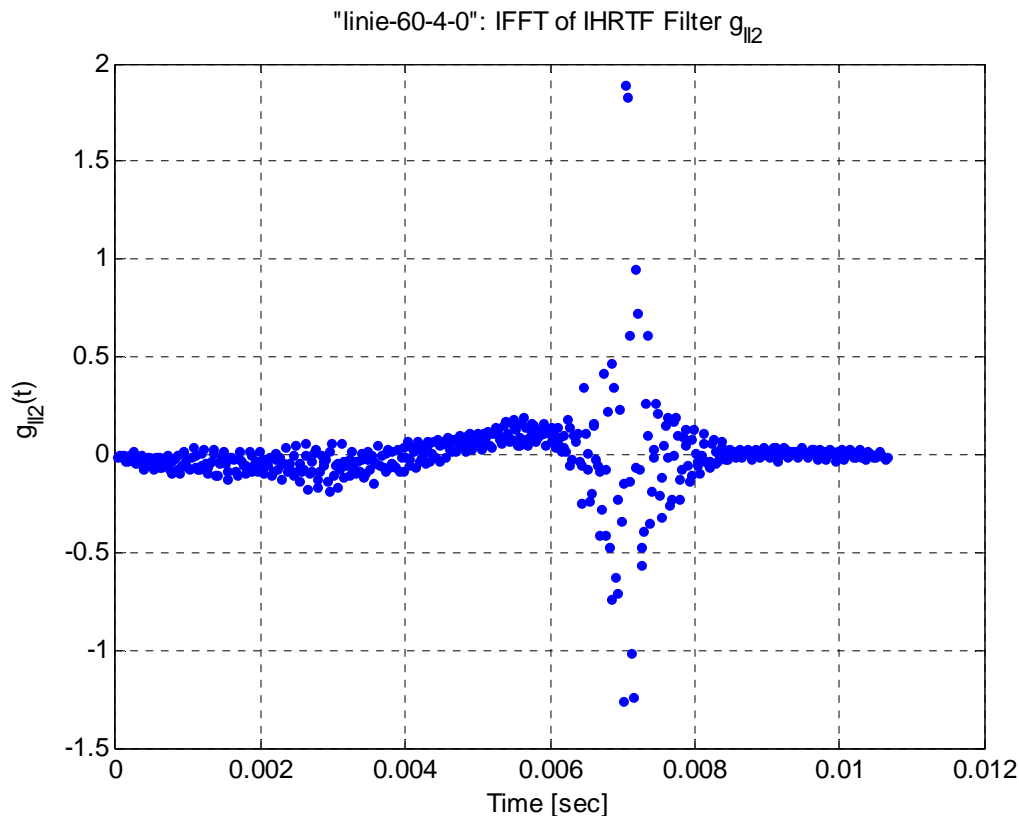


Abbildung 4.8: IHRIR (hier: Array-Typ "Linie", Abstand: 0,6 Meter, 4 vorne fokussierte Quellen, Szenario 1)

Wie bereits in Kapitel 3.3. diskutiert wurde, wurden die inversen HRTF-Filter Bandpass-gefiltert.

Als untere Grenzfrequenz des Bandpasses $fc1_{bp}$ war 120 Hz, als obere Grenzfrequenz $fc2_{bp}$ war 14 kHz eingestellt. M wurde auf 4 gesetzt, die Ordnung des Butterworth-Filters war also 8.

4.3.4. Bestimmung der Gesamt-Übertragungsfunktionen

Die effektive Länge der Gesamt-Impulsantworten bzw. der Gesamt-Übertragungsfunktionen wurde mit $L_{eve} = 2048$ Samples bzw. Frequenzbins gewählt. Der exponentielle Sweep, der zur Ermittlung der Gesamt-Übertragungsfunktionen verwendet wurde, hatte auch wieder eine Startfrequenz von 94 Hz und eine Endfrequenz von 24 kHz.

5. Evaluierung der Messergebnisse

5.1. Methodik der Evaluierung

5.1.1. Allgemeines

Die Evaluierung der Messergebnisse erfolgt anhand der (Gesamt-) Übersprechkompensation, welche im Folgenden als Mittelwert der Übersprechkompensationen der beiden einzelnen Kanäle definiert wird:

$$x_{tc}[k] = \frac{1}{2} \cdot \left(|H_{eve,ll}[k]|_{dB} - |H_{eve,lr}[k]|_{dB} + |H_{eve,rr}[k]|_{dB} - |H_{eve,rl}[k]|_{dB} \right) \quad (5.1)$$

x_{tc} ... cross-talk-cancelation (engl. für Übersprech-Kompensation)

Dabei sind

$$|H_{eve,ab}[k]|_{dB} = 20 \cdot \log(|H_{eve,ab}[k]|), \quad a = l, r, \quad b = l, r, \quad k = 0, 1, \dots, L_{eve} - 1 \quad (5.2)$$

die in Terzbändern geglätteten Amplitudengänge der Gesamt-Übertragungsfunktionen.

Die Übersprechkompensation beschreibt einen Aspekt der Qualität des virtuellen Kopfhörersystems:

Im theoretischen Optimalfall würde der linke bzw. rechte Kanal des Kopfhörersignals nur am linken bzw. rechten Kanal des virtuellen Kopfhörers (d.h. am linken bzw. rechten Trommelfell des Zuhörers) ankommen. Der linke Kanal des Kopfhörersignals würde am rechten Ohr bzw. der rechte Kanal des Kopfhörersignals würde am linken Ohr auf $-\infty$ dB abgedämpft werden. Durch Einsetzen in Gleichung (5.1) würde man also in diesem (theoretischen) Fall eine Übersprechkompensation von ∞ dB für alle Frequenzen erhalten.

Ein derartiges virtuelles Kopfhörersystem ist jedoch praktisch nicht realisierbar. Aus den Überlegungen des vorherigen Absatzes heraus wird aber klar, dass die Qualität des virtuellen Kopfhörersystems umso höher wird, je geringer das Übersprechen bzw. je größer die Übersprechkompensation ist. Für eine tiefer gehende Erörterung dieser Thematik wird auf [5] verwiesen.

Es sei an dieser Stelle auch erwähnt, dass es neben der Übersprechkompensation auch andere Aspekte (Klangverfärbungen, empfundene Stabilität des Systems bei Bewegung im Raum etc.) gibt, die einen Einfluss auf die Qualität des virtuellen Kopfhörersystems haben. Die dafür notwendigen Untersuchungen müssten jedoch im Rahmen einer separaten psychoakustischen Versuchsreihe durchgeführt werden.

Die Übersprechkompensation wird für den Frequenzbereich $f_{uo} = [f_u, f_o]$ berechnet, wobei f_u oberhalb der unteren Grenzfrequenz und f_o unterhalb der oberen Grenzfrequenz des Bandpassfilters (siehe Kapitel 4.3.3.) gewählt wurde. Dies ist deshalb sinnvoll, weil in diesem Frequenzbereich die Übersprechkompensation kaum durch den Bandpassfilter beeinflusst und verfälscht wird.

Wenn die Frequenzbins k_u bzw. k_o mit 9 bzw. 512 gewählt werden, lauten die entsprechenden Frequenzen f_u und f_o

$$f_u = \frac{f_s}{L_{eve}} \cdot k_u = \frac{48000}{2048} \cdot 9 \approx 211 \text{ Hz} \quad (5.3)$$

$$f_o = \frac{f_s}{L_{eve}} \cdot k_o = \frac{48000}{2048} \cdot 512 = 12 \text{ kHz} \quad (5.4)$$

Abbildung 5.1 zeigt die (in Terzbändern geglätteten) Amplitudengänge von $H_{eve,ll}[k]$ und $H_{eve,lr}[k]$, die (ebenfalls in Terzbändern geglätteten) Amplitudengänge von $H_{eve,rl}[k]$ und $H_{eve,rr}[k]$ und die Übersprechkompensation im Frequenzbereich $f_{uo} = [f_u, f_o]$ für die folgende Parameterkombination: Array-Typ "Linie", 4 (vorne) fokussierte Quellen, $z_{K,L} = 0,6$ Meter, Szenario 3.

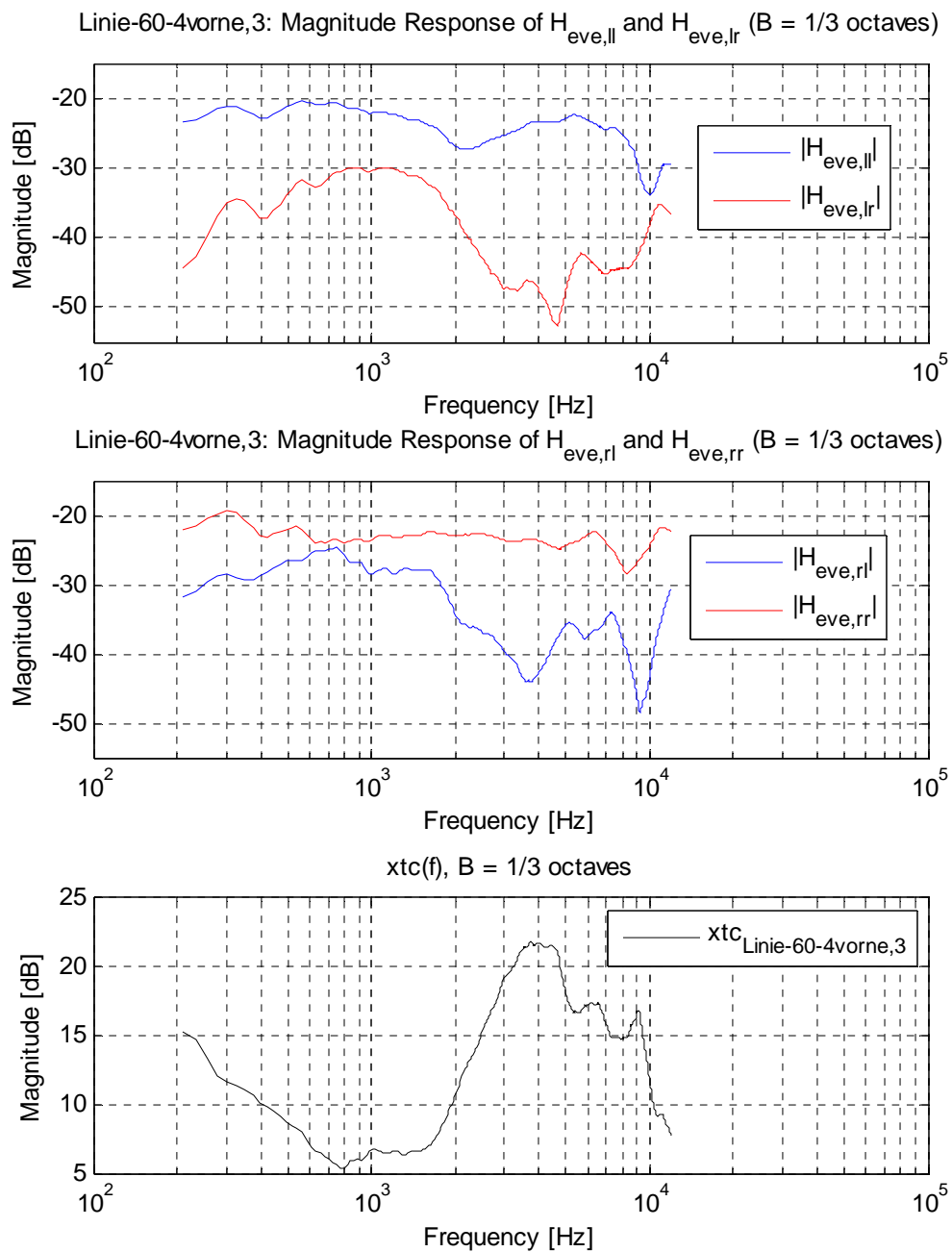


Abbildung 5.1: Amplitudengänge der in Terzbändern geglätteten Gesamt-Übertragungsfunktionen und Übersprechkompensation (hier: Array-Typ "Linie", Abstand 0,6 Meter, 4 (vorne) fokussierte Quellen, Szenario 3)

Im Rahmen dieses Projekts sollen die Übersprechkompensationen verschiedener Parameterkonfigurationen (Array-Typen, Quellenkonstellationen, Abstände, Szenarien) miteinander verglichen werden (vgl. z.B. Abbildung 5.2 in Kapitel 5.2.1.).

Für eine objektive Analyse ist es zweckmäßig, konkrete Zahlen, wie im Folgenden beschrieben, zu vergleichen.

Der betrachtete Frequenzbereich $f_{uo} = [f_u, f_o]$ wird in drei gleich große relative Frequenzbänder (in Tiefen $f_T = [f_u, f_{T,M}]$, Mitten $f_M = [f_{T,M}, f_{M,H}]$ und Höhen $f_H = [f_{M,H}, f_o]$) aufgeteilt.

Die Grenzfrequenzen, die die drei Frequenzbereiche trennen und die dazu gehörigen Frequenzbins berechnen sich zu:

$$f_{T,M} = f_u \cdot \left(\frac{f_o}{f_u} \right)^{\frac{1}{3}} \approx 811 \text{ Hz} \quad (5.5)$$

$$k_{T,M} = \frac{f_{T,M} \cdot L_{eve}}{f_s} \approx 35 \quad (5.6)$$

$$f_{M,H} = f_u \cdot \left(\frac{f_o}{f_u} \right)^{\frac{2}{3}} \approx 3120 \text{ Hz} \quad (5.7)$$

$$k_{M,H} = \frac{f_{M,H} \cdot L_{eve}}{f_s} \approx 133 \quad (5.8)$$

Es werden anschließend die Mittelwerte und Standardabweichungen der Übersprechkompensationen in den entsprechenden Frequenzbereichen berechnet.

$$\overline{xtc}_{f_F} = \frac{1}{k_{F,o} - k_{F,u} + 1} \cdot \sum_{k=k_{F,u}}^{k_{F,o}} xtc[k] \quad (5.9)$$

$$s_{xtc_{f_F}} = \sqrt{\frac{1}{k_{F,o} - k_{F,u}} \cdot \sum_{k=k_{F,u}}^{k_{F,o}} (xtc[k] - \overline{xtc}_{f_F})^2} \quad (5.10)$$

In (5.9) und (5.10) entsprechen dem jeweiligen Frequenzband f_F entweder die Tiefen f_T , die Mitten f_M oder die Höhen f_H . Weiters sind $k_{F,u}$ bzw. $k_{F,o}$ durch den kleinsten bzw. größten Frequenzbin des jeweiligen Frequenzbandes zu ersetzen.

Die aufgrund verschiedener Parameterkonfigurationen erhaltenen Übersprechkompensationen werden in Kapitel 5.2. diskutiert und analysiert, wenn die maximale Differenz der Mittelwerte zumindest in einem der Frequenzbereiche mindestens 6 dB beträgt. Damit soll einerseits die Evaluierung in diesem Dokument in einem übersichtlichen, kompakten Umfang gehalten werden. Andererseits lassen sich durch diese Einschränkung klare und sinnvolle Aussagen treffen.

Tabelle 5.1. listet die verschiedenen Kombinationen von Array-Typen, Abständen und Quellenkonstellationen auf, für die die Übersprechkompensationen ermittelt werden. Die Übersprechkompensationen werden dabei jeweils für alle 4 Szenarien S ($S = 1, 2, 3, 4$) bestimmt.

$xtc[k]$	Array-Typ	Abstand	Quellenkonstellation
$xtc_{Linie-60-4vorne,S}[k]$	Linie	0,6 Meter	4 Quellen (vorne)
$xtc_{Türme-60-4vorne,S}[k]$	Türme	0,6 Meter	4 Quellen (vorne)
$xtc_{oben-60-4vorne,S}[k]$	oben	0,6 Meter	4 Quellen (vorne)
$xtc_{oben-120-4vorne,S}[k]$	oben	1,2 Meter	4 Quellen (vorne)
$xtc_{oben-60-4außen,S}[k]$	oben	0,6 Meter	4 Quellen (außen)
$xtc_{oben-120-4außen,S}[k]$	oben	1,2 Meter	4 Quellen (außen)

Tabelle 5.1: Array-Typ, Abstand und Quellenkonstellation für die jeweilige Übersprechkompensation (an der Stelle des Frequenzbins k)

Die Evaluierung der Einflüsse der verschiedenen Array-Typen, Quellenkonstellationen und Abstände erfolgt für Szenario 1. Einerseits sind die Ergebnisse am besten überschaubar, wenn sie für ein konkretes Szenario erfolgen. Andererseits ist dieses Szenario wohl am typischsten für die Anwendung in der Praxis: Geht man z.B. von einem Arbeitsplatz aus, schaut der Anwender des virtuellen Kopfhörersystems vermutlich meistens nach vorne und vergleichsweise nur sehr selten um 30° nach rechts.

Da jedoch auch die Daten der Szenarien 2, 3 und 4 in einem angemessenen Rahmen diskutiert werden sollen, widmet sich Kapitel 5.2.4. den wichtigsten Ergebnissen, die damit in Zusammenhang stehen. D.h., es werden die Szenarien selbst - einerseits "im Gesamten" (als Überblick) und andererseits für bestimmte Kombinationen von Array-Typen, Quellenkonstellationen und Abständen - miteinander verglichen.

5.1.2. Array-Typen

Ein direkter Vergleich der Übersprechkompensationen der verschiedenen Array-Typen "Linie", "Türme" und "oben" kann nur für jene Kombination aus Abstand und Quellenkonstellation erfolgen, die bei allen 3 Array-Typen verwendet wurde: 0,6 Meter und "4 Quellen (vorne)".

Die Mittelwerte und Standardabweichungen werden berechnet und verglichen, indem $xtc[k]$ in (5.9) und (5.10) durch $xtc_{Linie-60-4vorne,1}[k]$, $xtc_{Türme-60-4vorne,1}[k]$ und $xtc_{oben-60-4vorne,1}[k]$ (siehe Tabelle 5.1) ersetzt wird.

5.1.3. Abstände

Untersucht wird, welchen Einfluss beim Array-Typ "oben" der Abstand des Kunstkopfes vom Array in y-Richtung hat (siehe auch Kapitel 2.2.3.). Dazu werden die Übersprechkompensationen bei den beiden Abständen 0,6 und 1,2 Meter, jeweils separat für die beiden Quellenkonstellationen "4 Quellen (vorne)" und "4 Quellen (außen)", evaluiert.

Es werden die Mittelwerte und Standardabweichungen nach (5.9) und (5.10) berechnet und verglichen, wobei $xtc[k]$ in (5.9) und (5.10) jeweils durch $xtc_{oben-60-4vorne,1}[k]$ und $xtc_{oben-120-4vorne,1}[k]$ bzw. $xtc_{oben-60-4außen,1}[k]$ und $xtc_{oben-120-4außen,1}[k]$ (siehe Tabelle 5.1) ersetzt wird.

5.1.4. Quellenkonstellationen

Es wird der Einfluss der Quellenkonstellation beim Array-Typ "oben" evaluiert. Es werden also die Übersprechkompensationen bei den beiden Quellenkonstellationen "4 Quellen (vorne)" und "4 Quellen (außen)", jeweils separat für die beiden Abstände 0,6 und 1,2 Meter, untersucht.

Es werden die Mittelwerte und Standardabweichungen nach (5.9) und (5.10) berechnet, wobei $xtc[k]$ in (5.9) und (5.10) jeweils durch $xtc_{oben-60-4vorne,1}[k]$ und $xtc_{oben-60-4außen,1}[k]$ bzw. $xtc_{oben-120-4vorne,1}[k]$ und $xtc_{oben-120-4außen,1}[k]$ (siehe Tabelle 5.1) ersetzt wird.

5.1.5. Szenarien

5.1.5.1. Überblick

Die Mittelwerte und Standardabweichungen nach (5.9) und (5.10) werden für alle Kombinationen von Array-Typ, Abstand, Quellenkonstellation (6 Kombinationen - siehe Tabelle 5.1), Szenarien ($S=1,2,3,4$) und Frequenzbereiche (f_T, f_M, f_H) berechnet. Man erhält daraufhin also jeweils $6 \cdot 4 \cdot 3 = 72$ Mittelwerte und Standardabweichungen.

Um einen Überblick zu bekommen, wie gut die Übersprechkompensation bei welchem Szenario funktioniert, werden die Mittelwerte, die sich aufgrund der verschiedenen Szenarien ergeben, miteinander verglichen. Es wird folgende Fragestellung untersucht: In welchem bzw. welchen der 3 Frequenzbereiche treten Unterschiede der mittleren Übersprechkompensationen von mindestens 6 dB auf? Dabei wird gefordert, dass diese Unterschiede bei *allen* 6 Kombinationen von Array-Typ, Abstand und Quellenkonstellation zu finden sind.

5.1.5.2. "Optimale" Kombination der Parameter bei Szenario 3 und f_H

Es wird untersucht, für welche Kombination der Parameter Array-Typ, Quellenkonstellation und Abstand bei Szenario 3, Frequenzbereich $f_H = [f_{M,H}, f_o]$ ($f_{M,H} \approx 3120 \text{ Hz}$, $f_o = 12 \text{ kHz}$) die höchste mittlere Übersprechkompensation nach (5.9) auftritt. Um dies herauszufinden, muss $xtc[k]$ in (5.9) jeweils durch $xtc_{Linie-60-4vorne,3}[k]$, $xtc_{Türme-60-4vorne,3}[k]$, $xtc_{oben-60-4vorne,3}[k]$, $xtc_{oben-120-4vorne,3}[k]$, $xtc_{oben-60-4außen,3}[k]$ und $xtc_{oben-120-4außen,3}[k]$ ersetzt werden.

Dass die Mittelwerte für den Frequenzbereich $f_H = [f_{M,H}, f_o]$ berechnet werden, hat folgenden Grund: Da interaurale Pegeldifferenzen (ILDs) vor allem bei hohen Frequenzen auftreten und damit auch in diesem Frequenzbereich am besten wahrgenommen werden, wird eine große Übersprechkompensation bei hohen Frequenzen als wichtiger erachtet als bei mittleren oder tiefen Frequenzen.

5.2. Vergleich und Analyse der Messergebnisse

5.2.1. Array-Typen

In Abbildung 5.2 sind die Übersprechkompensationen $xtc_{Linie-60-4vorne,1}[k]$, $xtc_{Türme-60-4vorne,1}[k]$ und $xtc_{oben-60-4vorne,1}[k]$ im Frequenzbereich $f_{uo} = [f_u, f_o]$ geplottet.

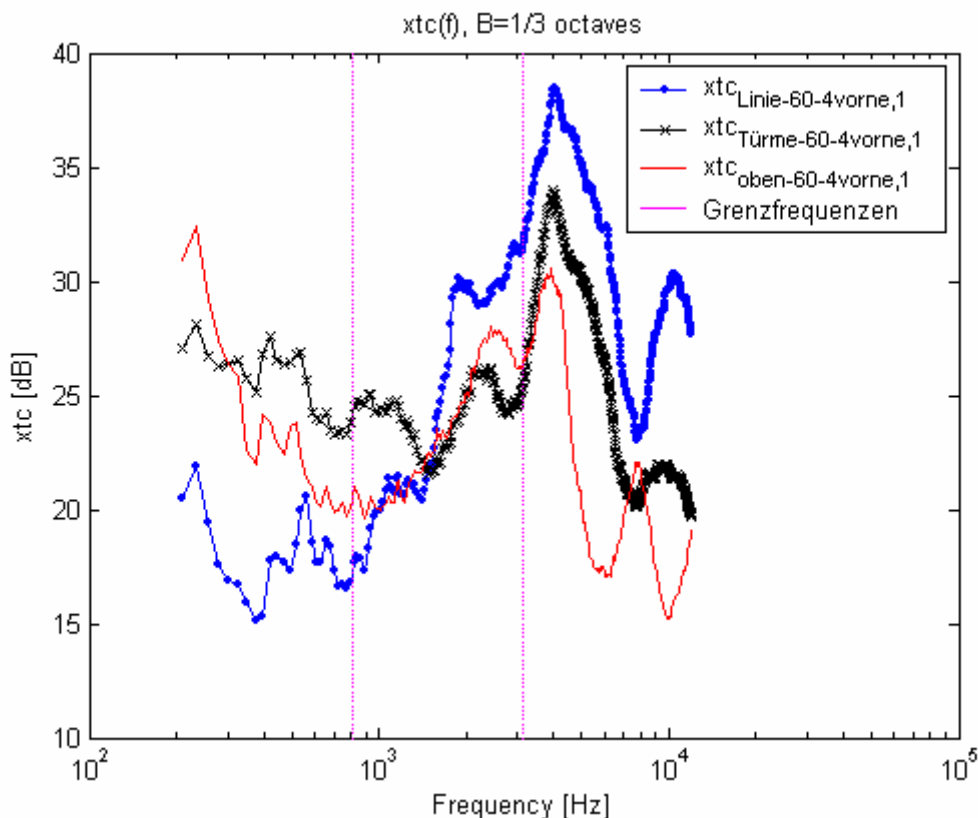


Abbildung 5.2: Übersprechkompensationen bei den drei Array-Typen "Linie", "Türme" und "oben" (Abstand: 0,6 Meter, 4 vorne fokussierte Quellen)

Aus der Abbildung ist Folgendes zu erkennen: Die Übersprechkompensation funktioniert beim Array-Typ „Linie“ für hohe Frequenzen am besten und für tiefe Frequenzen am schlechtesten: Die Dynamik der Übersprechkompensation über den Frequenzbereich $f_{uo} = [f_u, f_o]$ ist also beim Array-Typ "Linie" am größten.

Im mittleren Frequenzbereich ist die Übersprechkompensation bei allen drei Array-Typen in etwa gleich groß.

Außerdem ist bei allen dreien, dem jeweiligen Array entsprechenden Verläufen der Übersprechkompensation, bei ca. 4 kHz ein lokales (Array-Typ "oben") oder sogar

globales (Array-Typ "Linie" und Array-Typ "Türme") Maximum zu erkennen. Das Maximum ist beim Array-Typ „Linie“ am größten (ca. 39 dB). Das Ohr ist in diesem Frequenzbereich sehr empfindlich. Die Frequenzlage dieser Maxima dürfte daher einer guten Qualität des virtuellen Kopfhörers durchaus dienlich sein.

Abbildung 5.3 stellt die nach (5.9) und (5.10) berechneten mittleren Übersprechkompensationen (gelbe Markierung) und die zugehörigen Standardabweichungen (Länge eines blauen, violetten oder roten Striches vom Mittelwert weg gemessen) bei allen Array-Typen und für alle Frequenzbereiche dar.

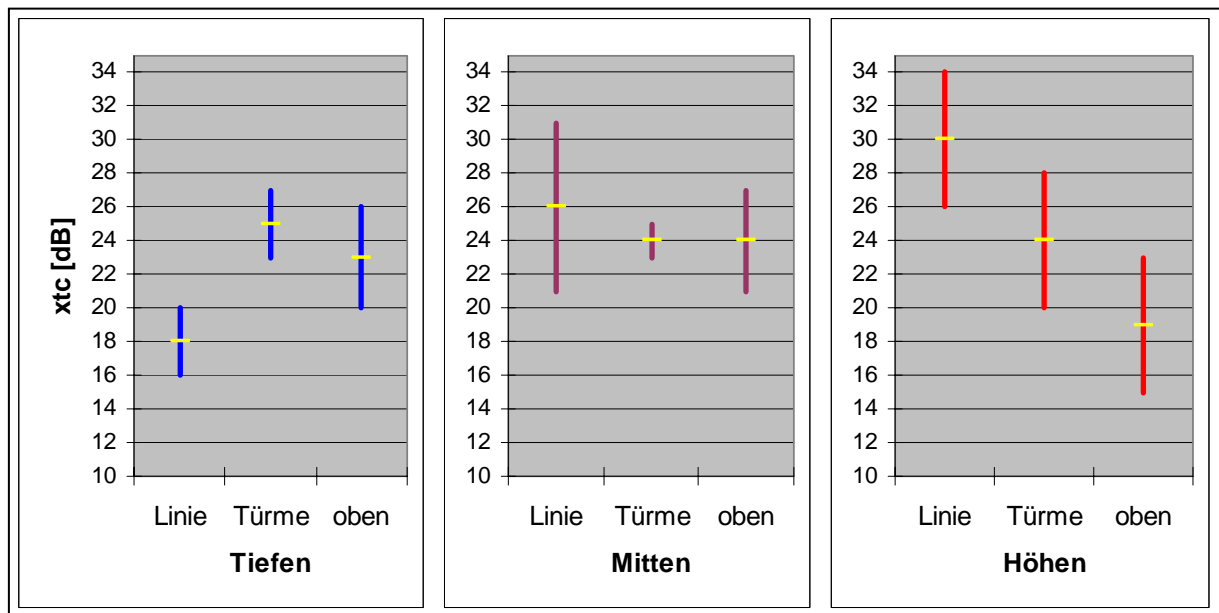


Abbildung 5.3: Mittelwerte und Standardabweichungen von $xtc_{Linie-60-4vorne,1}[k]$,

$xtc_{Türme-60-4vorne,1}[k]$ und $xtc_{oben-60-4vorne,1}[k]$ für alle 3 Frequenzbereiche

Mögliche Erklärungen für die erhaltenen Ergebnisse lauten folgendermaßen:

Beim Array-Typ "Linie" befinden sich alle Lautsprecher und fokussierten Quellen in einer Ebene (auf Ohrhöhe - vgl. Abbildungen 2.2 und 2.4). Dies ist bei den anderen beiden Array-Typen nicht der Fall. Bei hohen Frequenzen wird die Schallabstrahlung der einzelnen realen Lautsprecher gerichtet. D.h., die abgestrahlte Schallenergie in die Richtung, in die der Lautsprecher zeigt, ist größer als jene in die anderen Richtungen (nach oben, unten, zur Seite und nach hinten). Die Fokussierung der Schallquelle funktioniert also vermutlich besser, wenn diese in der Ebene der Lautsprecher liegt.

Interaurale Pegeldifferenzen (ILDs) treten bei hohen Frequenzen stärker auf als bei tiefen. Beim Array-Typ "Linie" haben die ILDs einen höheren Einfluss auf die Übersprechkompensation als bei den anderen beiden Array-Typen, weil die

Lautsprecher alle auf Ohrhöhe positioniert sind. Dadurch bildet sich wiederum eher ein akustischer Schatten auf beiden Seiten des Kunstkopfes aus.

5.2.2. Abstände

Die Übersprechkompensationen $xtc_{oben-60-4vorne,1}[k]$ und $xtc_{oben-120-4vorne,1}[k]$ bzw. $xtc_{oben-60-4außen,1}[k]$ und $xtc_{oben-120-4außen,1}[k]$ zeigen einigermaßen ähnliche Verläufe. In Abbildung 5.4 werden als Beispiel die bei den beiden Abständen 0,6 und 1,2 Meter erhaltenen Übersprechkompensationen miteinander verglichen, wenn 4 vorne fokussierte Quellen verwendet werden.

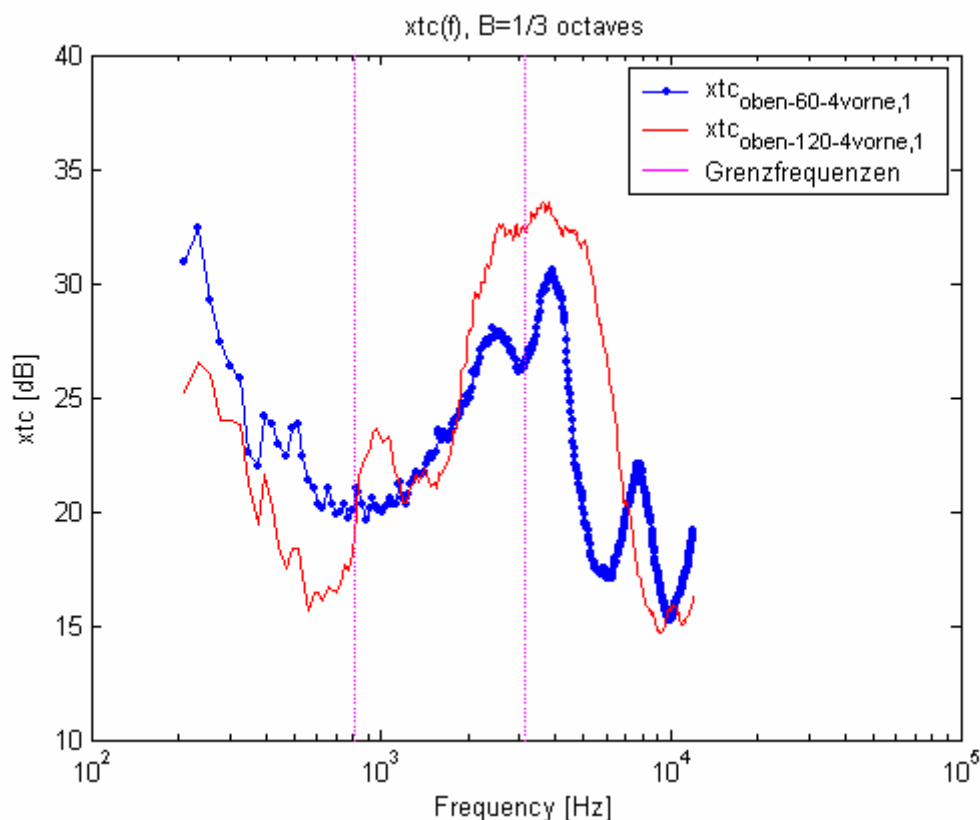


Abbildung 5.4: Vergleich der Übersprechkompensationen bei den beiden Abständen 0,6 und 1,2 Meter (Array-Typ "oben", 4 vorne fokussierte Quellen)

Die Differenzen der nach (5.9) berechneten Mittelwerte liegen bei beiden Quellenkonstellationen in allen 3 Frequenzbereichen jeweils unter 6 dB. Auf eine weitere Analyse des Parameters "Abstände" wird daher in diesem Dokument verzichtet.

5.2.3. Quellenkonstellationen

Die Übersprechkompensationen bei den Quellenkonstellationen "4 Quellen vorne" und "4 Quellen außen" zeigen sowohl für tiefe als auch für mittlere Frequenzen und sowohl für den Abstand 0,6 Meter als auch für den Abstand 1,2 Meter ähnliche Verläufe. Im hohen Frequenzbereich ergeben sich jedoch deutliche Unterschiede.

Die Abbildungen 5.5 und 5.7 zeigen die Übersprechkompensationen bei den beiden Quellenkonstellationen für die Abstände 0,6 und 1,2 Meter (Frequenzbereich:

$$f_{uo} = [f_u, f_o]).$$

Die Abbildungen 5.6 und 5.8 stellen die nach (5.9) und (5.10) berechneten mittleren Übersprechkompensationen und die zugehörigen Standardabweichungen bei den beiden Quellenkonstellationen für die 3 Frequenzbereiche f_T, f_M und f_H (Abstände: 0,6 und 1,2 Meter) dar.

Wie aus diesen 4 Abbildungen (5.5 bis 5.8) zu erkennen ist, ist also das Übersprechen im hohen Frequenzbereich bei 4 außen fokussierten Quellen wesentlich geringer als bei 4 vorne fokussierten Quellen.

Mögliche Begründungen für dieses Ergebnis lauten folgender Maßen:

Da die Lautsprecher seitlich über dem Kopf angebracht sind, ist vermutlich der bei hohen Frequenzen auftretende akustische Schatten am Kunstkopf bei den vier vorderen Quellen (insbesondere bei den beiden Quellen mit $\pm 22,5^\circ$ Azimut) kleiner als bei den vier äußeren Quellen, die seitlich und zudem sowohl vor *als auch* hinter dem Kopf positioniert werden.

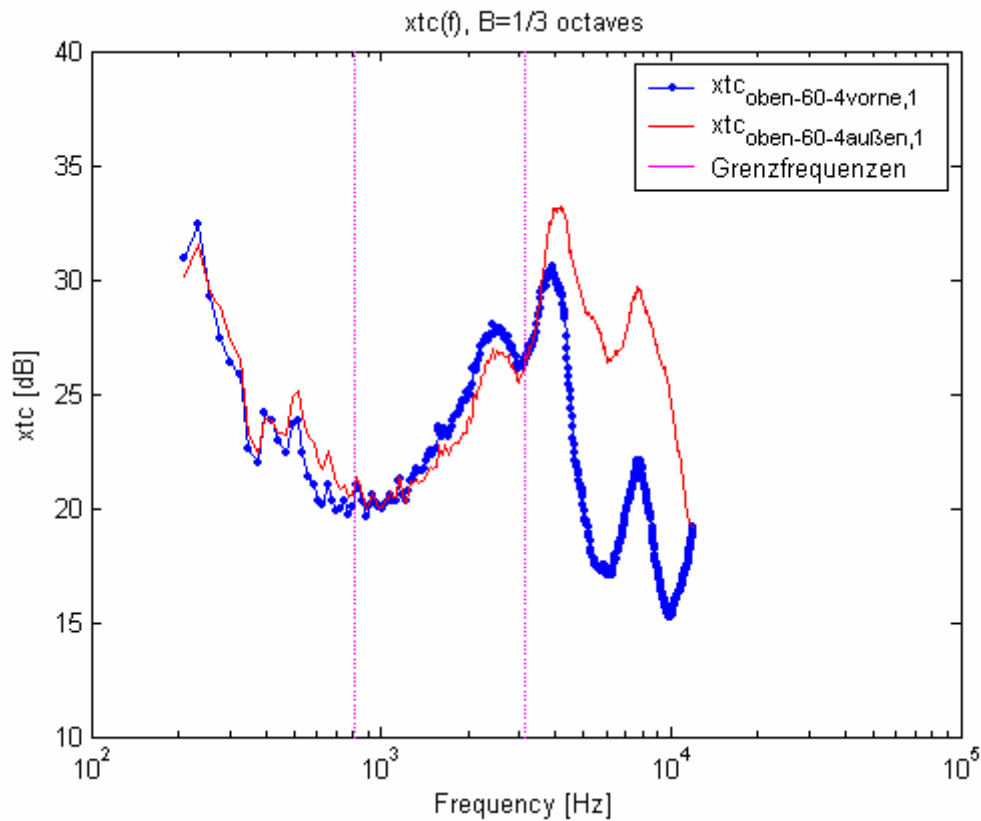


Abbildung 5.5: Übersprechkompensationen bei den Quellenkonstellationen mit 4 vorne und 4 außen fokussierten Quellen (Array-Typ "oben", Abstand: 0,6 Meter)

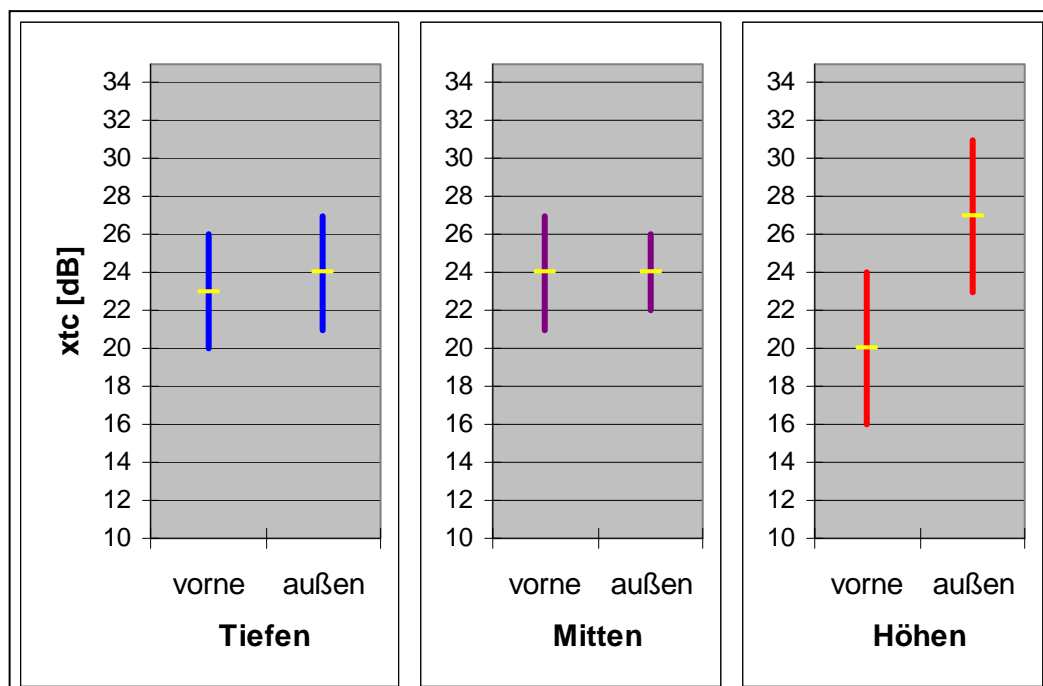


Abbildung 5.6: Mittelwerte und Standardabweichungen von $xtc_{oben-60-4vorne,1} [k]$ und $xtc_{oben-60-4außen,1} [k]$ für alle 3 Frequenzbereiche

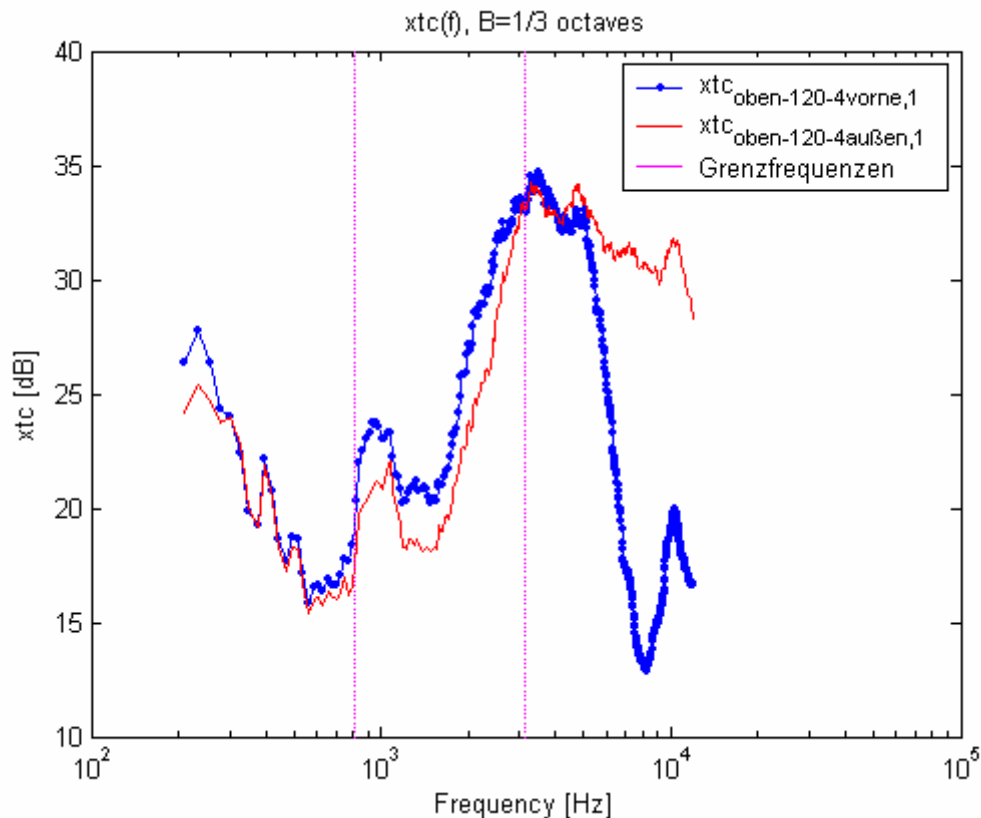


Abbildung 5.7: Übersprechkompensationen bei den Quellenkonstellationen mit 4 vorne und 4 außen fokussierten Quellen (Array-Typ "oben", Abstand: 1,2 Meter)

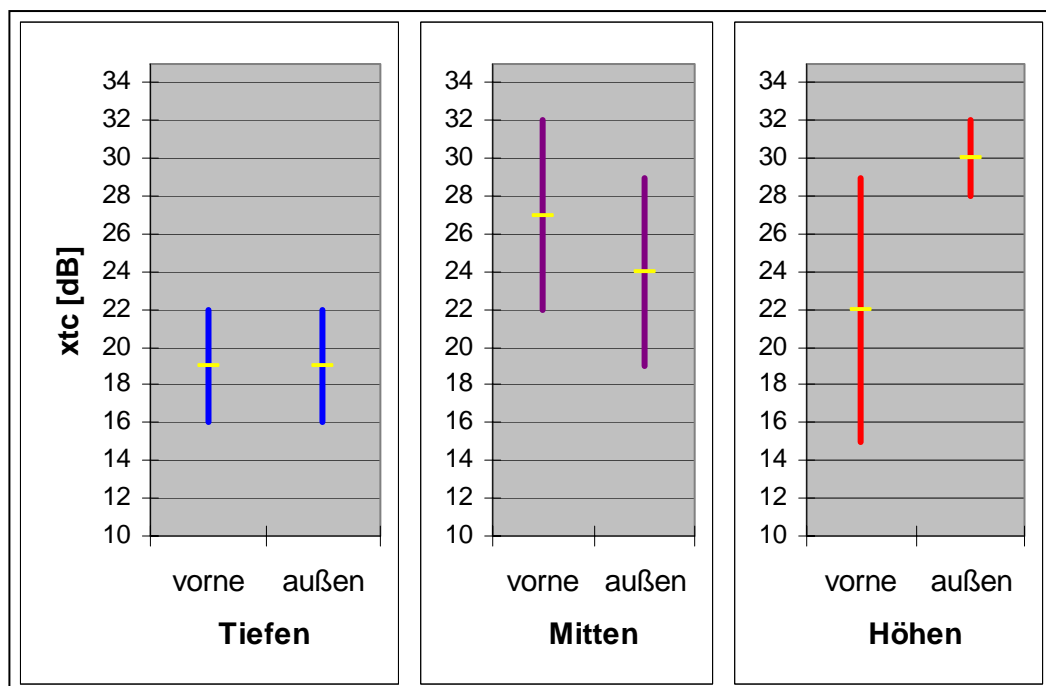


Abbildung 5.8: Mittelwerte und Standardabweichungen von $xtc_{oben-120-4vorne,1} [k]$ und $xtc_{oben-120-4außen,1} [k]$ für alle 3 Frequenzbereiche

5.2.4. Szenarien

5.2.4.1. Überblick

In Abbildung 5.9 werden stellvertretend für alle 6 Kombinationen von Array-Typ, Abstand und Quellenkonstellation die Übersprechkompensationen für den Array-Typ "Türme", für den Abstand 0,6 Meter und für die Quellenkonstellation "4 Quellen vorne" bei allen Szenarien ($S = 1, 2, 3, 4$) im Frequenzbereich $f_{uo} = [f_u, f_o]$ dargestellt.

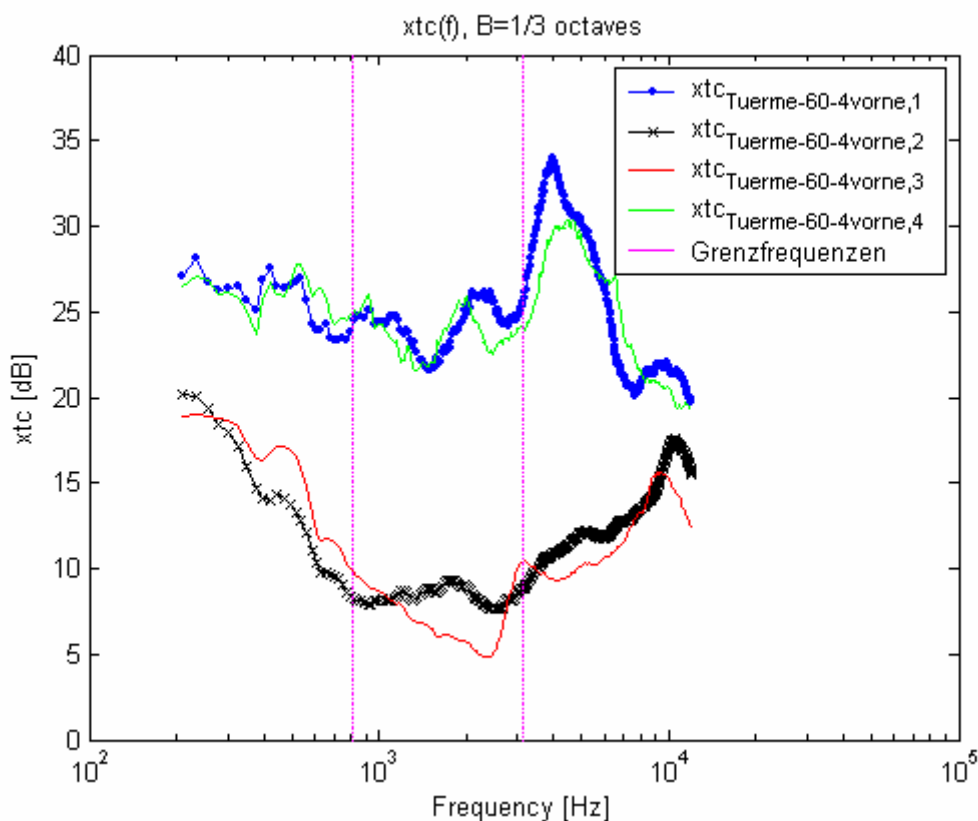


Abbildung 5.9: Übersprechkompensationen bei den Szenarien $S=1,2,3,4$ (Array-Typ "Türme", Abstand 0,6 Meter, 4 vorne fokussierte Quellen)

Folgende Aussagen treffen für *alle* 6 Kombinationen von Array-Typ, Abstand und Quellenkonstellation zu.

Die Szenarien 1 und 4 (die HRTFs wurden für die aktuelle Kunstkopf-Position gemessen und daraus wurden die IHRTFs berechnet) funktionieren in Bezug auf die Übersprechkompensation hervorragend. Bei beiden Szenarien haben die Übersprechkompensationen um 4 kHz herum zumindest lokale Maxima (bei manchen der 6 Kombinationen sind die Maxima sogar global).

Bei den Szenarien 2 und 3 hingegen sind im Allgemeinen um 4 kHz herum keine Maxima vorhanden. Die nach (5.9) berechneten Mittelwerte sind außerdem sowohl bei Szenario 2 als auch bei Szenario 3 für die Frequenzbereiche f_M und f_H um über 6 dB geringer als bei den Szenarien 1 und 4. Damit führt aber das Mitdrehen der fokussierten Quellen (Szenario 3 bzw. Nutzung des Head-Tracking-Systems, siehe Kapitel 2.1.) offensichtlich zu Ergebnissen, die schlechter sind als erhofft.

5.2.4.2. "Optimale" Kombination der Parameter bei Szenario 3 und f_H

Von allen 6 möglichen Kombinationen von Array-Typen, Abständen und Quellenkonstellationen tritt bei $xtc_{Linie-60-4vorne,3}[k]$ für Szenario 3 im hohen Frequenzbereich die höchste mittlere Übersprechkompensation nach (5.9) auf. Diese Kombination wird in Abbildung 5.10 mit jenen beiden Kombinationen verglichen, bei denen der Array-Typ ein anderer ist, aber die Quellenkonstellation und der Abstand gleich sind, d.h. mit $xtc_{Türme-60-4vorne-60,3}[k]$ und mit $xtc_{oben-60-4vorne-60,3}[k]$.

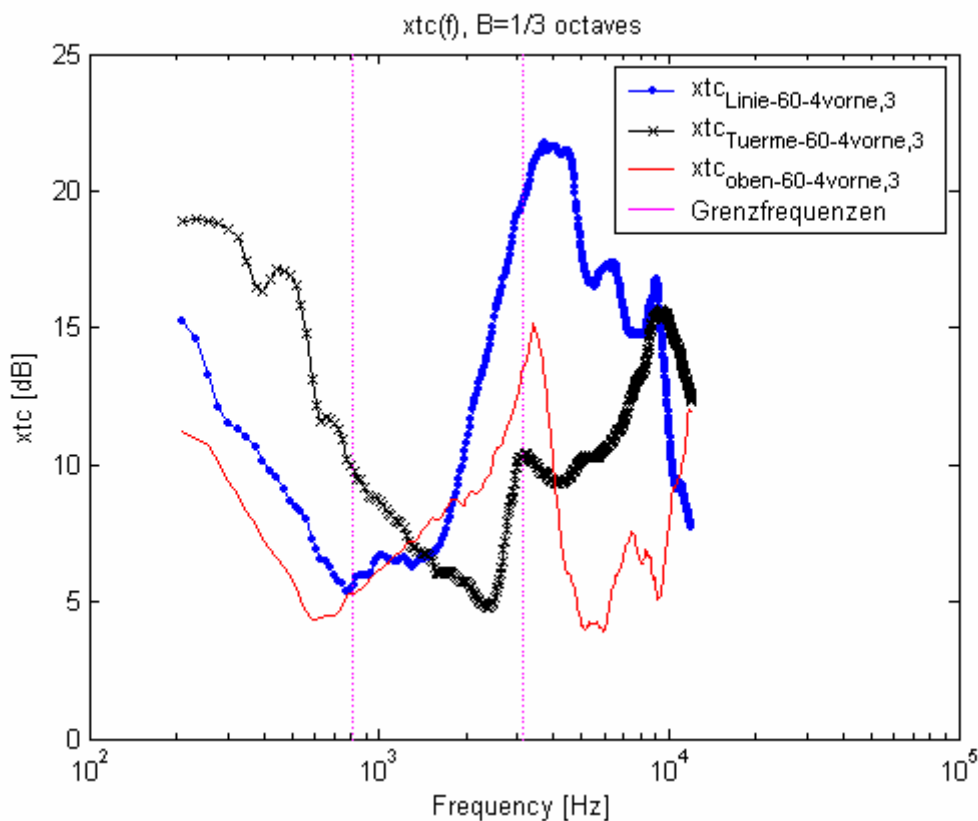


Abbildung 5.10: Übersprechkompensationen bei den drei Array-Typen "Linie", "Türme" und "oben" für Szenario 3 (Abstand: 0,6 Meter, 4 vorne fokussierte Quellen)

In Abbildung 5.11 sind die nach (5.9) und (5.10) berechneten mittleren Übersprechkompensationen und Standardabweichungen bei allen Array-Typen und für die drei Frequenzbereiche dargestellt.

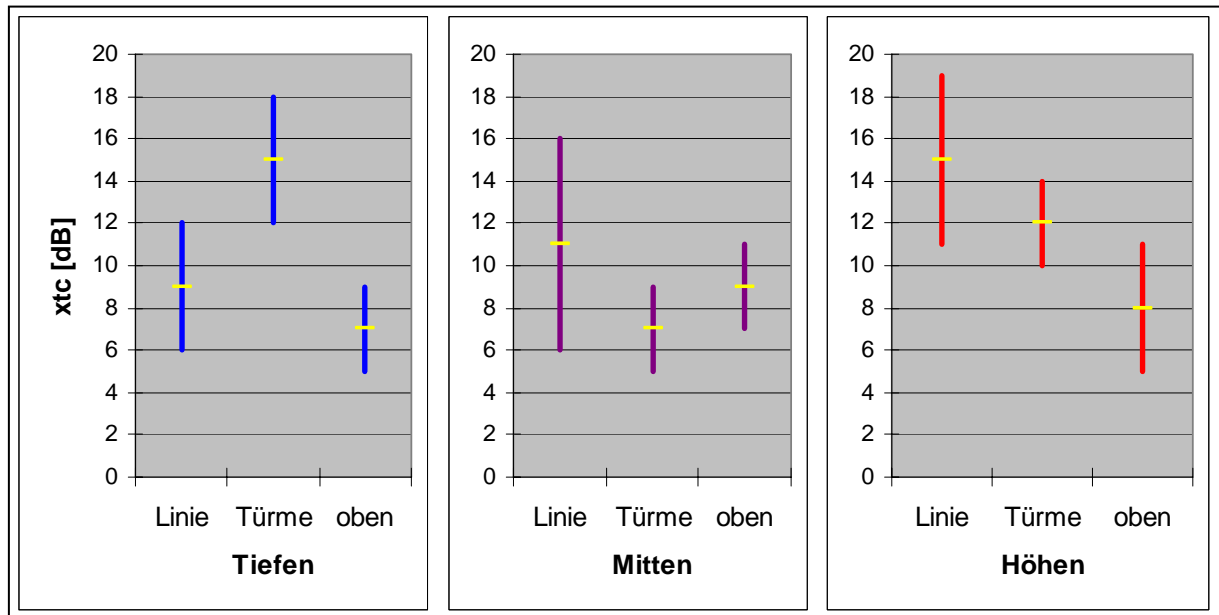


Abbildung 5.11: Mittelwerte und Standardabweichungen von $xtc_{Linie-60-4vorne,3}[k]$,

$xtc_{Türme-60-4vorne,3}[k]$ und $xtc_{oben-60-4vorne,3}[k]$ für alle 3 Frequenzbereiche

Wie aus der Abbildung 5.10 zu erkennen ist, sind die Übersprechkompensationen über weite Teile des oberen Frequenzbereichs (ca. 2 bis 9 kHz) beim Array-Typ "Linie" am größten. Außerdem ist zu erkennen, dass bei diesem Array-Typ die Übersprechkompensation in der Nähe von 4 kHz ein globales Maximum hat, was bei den anderen beiden Array-Typen nicht der Fall ist. Da das Gehör um 4kHz herum besonders empfindlich ist, ist dieser Peak von ca. 22 dB in Bezug auf die Übersprechkompensation ein großer Vorteil.

Die möglichen Erklärungen dafür, dass die Übersprechkompensation für diese Kombination von Array-Typ, Quellenkonstellation und Abstand bei Szenario 3 im hohen Frequenzbereich (bis ca. 9 kHz) am größten ist, entsprechen sinngemäß den Ausführungen bei Szenario 1 (siehe Kapitel 5.2.1.).

6. Zusammenfassung

Das Ziel dieses Projekts bestand darin, einen virtuellen Kopfhörer zu realisieren und die Übersprechkompensation anhand verschiedener Parameter und drei gleich großer relativer Frequenzbänder zu evaluieren. Die Übersprechkompensation beschreibt dabei einen Aspekt der Qualität des virtuellen Kopfhörersystems.

Dazu wurde zunächst ein Messsystem in Matlab implementiert. Dieses ermöglichte die HRTFs für verschiedene Parameterkonfigurationen zu bestimmen und daraus die (Bandpass-gefilterten) inversen HRTF-Filter zu berechnen. Mit letzteren wurden aus einem Stereosignal die Wiedergabesignale (d.h. die Quellsignale für die Schallwiedergabe) berechnet, die mittels Wellenfeldsynthese als fokussierte Schallquellen um den Kopf herum positioniert wurden. Außerdem wurden die Gesamt-Übertragungsfunktionen des virtuellen Kopfhörersystems bestimmt, aus denen in weiterer Folge die frequenzabhängige Übersprechkompensation berechnet werden konnte.

Die Parameter, die untersucht wurden, waren

- der Lautsprecher-Array-Typ:
 - "Linie": das Array befand sich auf Ohrhöhe und vor dem Kunstkopf;
 - (2 Lautsprecher-) "Türme": sie waren links und rechts vor dem Kunstkopf angebracht;
 - "oben": es wurde ein Messaufbau simuliert, bei dem die Lautsprecher seitlich oberhalb des Kunstkopfes angebracht waren
- der Abstand des Kunstkopfes von jener Ebene, in der sich die Vorderseite des Lautsprecher-Arrays befand (verwendete Abstände: 0,6 und 1,2 Meter)
- die Quellenkonstellation, die beschrieb, unter welchen Azimuten die fokussierten Schallquellen um den Kunstkopf positioniert wurden, wenn dieser auf 0° ausgerichtet war (verwendete Quellenkonstellationen: 4 vorne und 4 außen fokussierte Quellen).
- das Szenario, das beschrieb, auf welchen Azimut der Kunstkopf eingestellt war (Szenario 1: 0° , sonst: -30°). Für -30° sagte das Szenario weiters aus, ob die fokussierten Schallquellen mit dem Kunstkopf mitgedreht wurden (Szenario 3) oder nicht (Szenario 2) bzw. ob die HRTFs für diese Kunstkopfdrehung neu gemessen wurden und daraus die inversen HRTF-Filter berechnet wurden (Szenario 4).

Die wichtigsten Ergebnisse, die sich aus der Evaluierung ergaben, sind:

Für die Szenarien 1 und 3 wird im hohen Frequenzbereich bei Verwenden des Array-Typs "Linie" die größte Übersprechkompensation erreicht.

Die Übersprechkompensation ist bei den beiden Abständen 0,6 und 1,2 Meter für Szenario 1, Array-Typ "oben" in den jeweiligen drei Frequenzbereichen fast gleich groß.

Beim Array-Typ "oben" besteht die Möglichkeit, fokussierte Schallquellen vor *und* hinter dem Kopf anzubringen. Im Sinne einer großen Übersprechkompensation bei hohen Frequenzen sollte von dieser Möglichkeit Gebrauch gemacht werden.

In allen drei Frequenzbereichen sind die Übersprechkompensationen am höchsten, wenn die HRTFs für die jeweilige Kunstkopfdrehung bestimmt und daraus die inversen HRTF-Filter berechnet wurden (Szenarien 1 und 4). Es werden sehr gute Übersprechkompensations-Werte bis zu 39 dB erreicht.

Die Übersprechkompensation ist bei den beiden Szenarien 2 und 3 sowohl im mittleren als auch im hohen Frequenzbereich deutlich geringer als bei den Szenarien 1 und 4. Das Mitdrehen der fokussierten Schallquellen (Szenario 3) funktioniert damit weit weniger gut als erhofft. Im Allgemeinen ist auch gegenüber Szenario 2 keine Verbesserung der Übersprechkompensation feststellbar.

Anhang: Bedienung des Messsystems

Wie bereits in Kapitel 2.2.2. beschrieben wurde, wird das (I)HRTF-Messsystem über eine GUI (Graphical User Interface) bedient.

Dieses Benutzerhandbuch bzw. diese Bedienungsanleitung ist in erster Linie als Hilfestellung für die praktische Anwendung des (I)HRTF-Messsystems gedacht.

A.1. Urheber

Alle m-Files des Programms (Unterordner "GUI" des Ordners "(I)HRTF-GUI") außer jene, die im Folgenden speziell erwähnt werden, wurden von Herbert Petritsch im Zeitraum Oktober 2008 bis Jänner 2009 geschrieben. Folgende Dateien gehören zum Programm *pa_wavplay* ([9]):

"pawavplaya.dll", "pawavplayw.dll", "pawavplayx.dll", "pa_wavplayrecord.m", "pa_wavplay.m", "pa_wavrecord.m", "pa_wavplay usage instructions.txt"

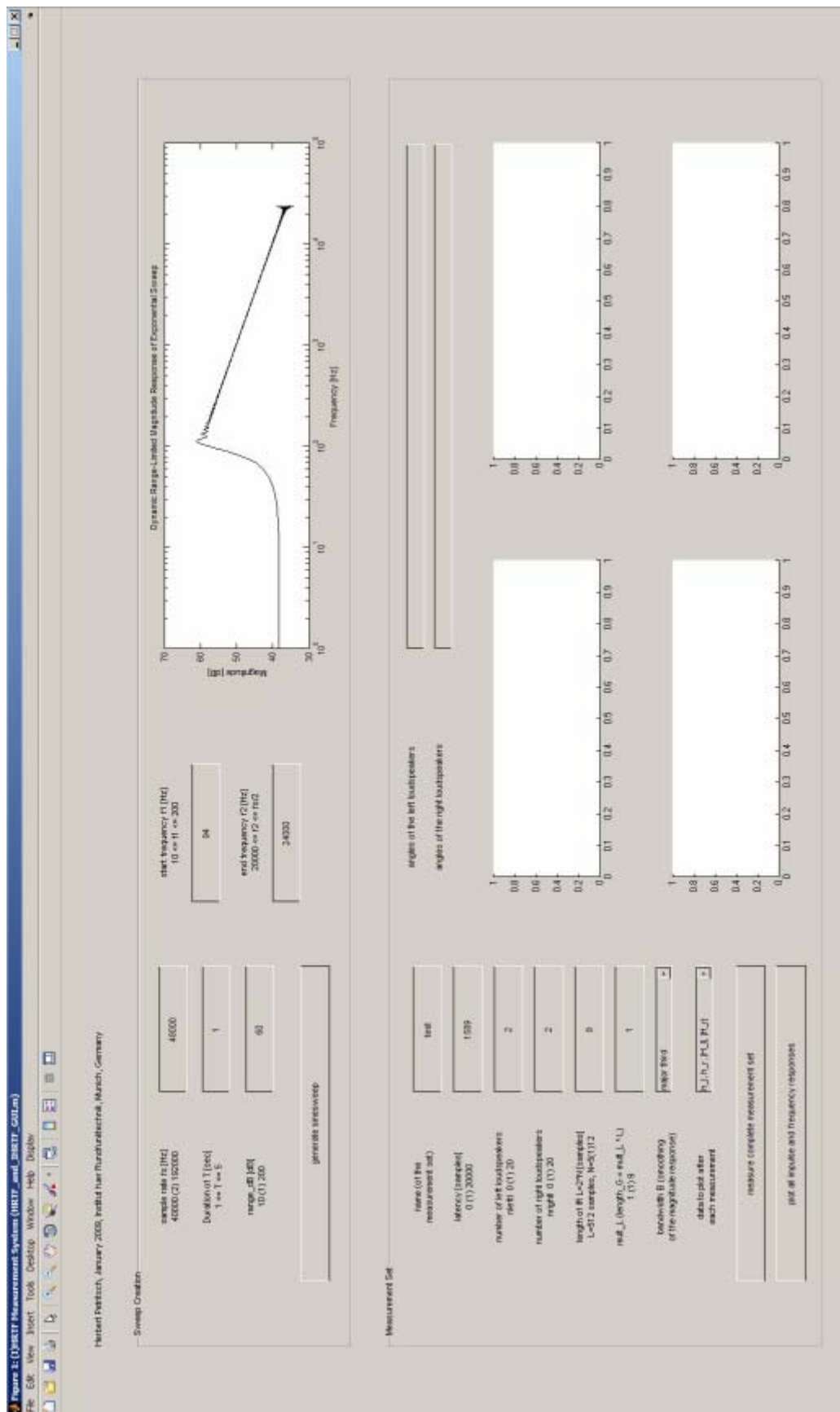
Zum Abspeichern der IHRIRs in einem für das Faltungsprogramm BruteFIR [6] lesbaren Format wurde Source Code von Daniel Menzel (IRT, 2005) als Vorlage genommen und entsprechend adaptiert (betrifft "calc_H_G_and_g.m").

A.2. Starten der GUI

Als Matlab-Verzeichnis ist der Ordner "(I)HRTF-GUI" und davon der Unterordner "GUI" auszuwählen. Im Unterordner GUI befindet sich das komplette Programm. Zum Starten der GUI muss "HRTF_and_IHRTF_GUI" im Matlab Command Window eingegeben werden. Die GUI wird nun angezeigt.

Die Abbildungen A.1 und A.2 zeigen 2 Screenshots der GUI (die Screenshots wurden um 90° im mathematisch positiven Sinn gedreht) direkt nach dem Starten der GUI.

Beim ersten Starten der GUI wird ein Sweep aus den Default-Eingabewerten erzeugt und in der Datei "expsweep.mat" abgespeichert, die sich im Ordner "data" befindet. Der Amplitudengang des Sweeps wird im Sweep Creation Panel (siehe Kapitel A.4) angezeigt.



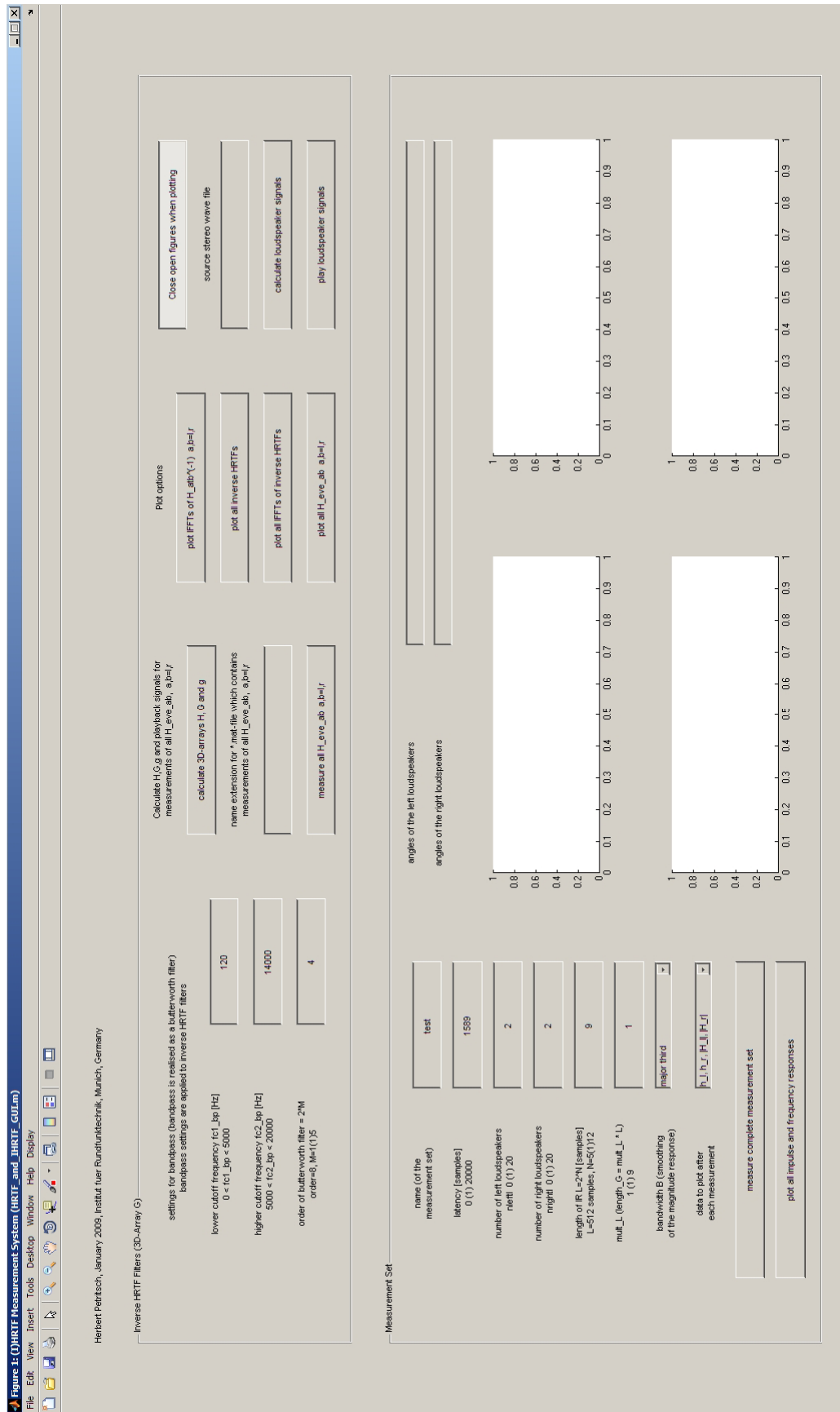


Abbildung A.2: GUI Screenshot 2 (um 90° im mathematisch positiven Sinn gedreht)

A.3. Aufbau der GUI

Die GUI besteht aus 3 Panels: Sweep Creation (siehe Kapitel A.4), Measurement Set (siehe Kapitel A.5) und Inverse HRTF Filters (siehe Kapitel A.6). Zwischen der Anzeige der beiden Panels Sweep Creation und Inverse HRTF Filters (oberes "Drittel" der GUI) können Sie wechseln: Wählen Sie dazu entweder im Display-Menü das entsprechende Panel aus oder drücken Sie Strg+D auf der Tastatur (vgl. Abbildungen A.1 und A.2).

In allen 3 Panels der GUI gibt es mehrere so genannte Edit-Text-Felder, in denen Sie Eingaben machen können. Bei den meisten dieser Felder werden nur skalare Zahlenwerte innerhalb eines bestimmten Wertebereichs akzeptiert.

Bei unzulässigen Eingaben erscheint eine Fehlermeldung und der Wert des Edit-Text-Feldes wird auf den Default-Wert zurückgesetzt. Welche Werte vom Programm akzeptiert werden, ist jeweils links des Edit-Text-Feldes angegeben:

Beispiel a: 40000 (2) 192000. Hierbei ist 40000 der kleinste Wert, der akzeptiert wird und 192000 der größte. Der Wert in der Klammer gibt die Schrittweite an. In diesem Beispiel muss der Wert also eine gerade, natürliche Zahl zwischen 40000 und 192000 sein.

Beispiel b: $0 < \text{fc1_bp} < 5000$: Der Wert für fc1_bp kann eine beliebige Gleitkommazahl größer 0 und kleiner 5000 sein.

Als Default-Werte der Eingabefelder wurden Werte ausgewählt, die bei den Messungen im Rahmen dieses Projekts häufig oder sogar immer benötigt wurden.

Um einen Absturz des Programms möglichst zu verhindern, werden bei zeitaufwändigen Berechnungen und Messungen Waitbars und Message-Boxen angezeigt, die während dieser Zeit User-Eingaben verhindern und außerdem einen Eindruck vom Fortschritt der Berechnungen bzw. Messungen liefern.

A.4. Sweep Creation Panel

Einstellbar sind die Samplerate f_s , die Dauer des Sweeps T , der Dynamikumfang des Sweeps in dB range_dB , die Startfrequenz des Sweeps f_1 und die Endfrequenz des Sweeps f_2 .

Der Parameter range_dB wurde in Anlehnung an [10] eingeführt und dient dazu, Fehler aufgrund von Divisionen durch 0 zu vermeiden.

Wenn Sie alle Werte zu Ihrer Zufriedenheit gewählt haben, drücken Sie den Button "generate sinesweep". Alle eingegebenen (f_s , T , $range_dB$, f_1 und f_2) und berechneten (L_s : Dauer des Sweeps in Samples) Werte sowie der Sweep im Zeit- und im Frequenzbereich (s und S) werden in der Datei "expsweep.mat" abgespeichert, welche sich im Ordner "data" befindet - der zuletzt generierte Sweep wird dabei überschrieben. Der Amplitudengang des neu generierten Sweeps wird geplottet.

A.5. Measurement Set Panel

Das Measurement Set Panel dient dazu, einen kompletten Messsatz von Frequenz- und Impulsantworten zu vermessen.

Im Rahmen dieses Projekts handelt es sich bei den Frequenzantworten um HRTFs (von allen Schallquellen zu beiden Ohren des Kunstkopfes) und deren IFFTs, die als HRIRs bezeichnet werden (Head Related Impulse Responses). Es könnten theoretisch mit der GUI auch andere "Typen" von Frequenz- und Impulsantworten gemessen werden, weshalb die allgemeinen Begriffe in diesem Panel verwendet werden.

Der Vollständigkeit halber sei aber an dieser Stelle erwähnt, dass bei den Berechnungen der IHRTFs im Inverse HRTF Filters Panel davon ausgegangen wird, dass es sich bei den gemessenen Frequenzantworten tatsächlich um HRTFs handelt.

Zur Messung der Frequenz- und Impulsantworten wird jener Sweep verwendet, der in der Datei "expsweep.mat" im Ordner "data" abgespeichert ist.

Sie können folgende Eingaben machen:

A.5.1. Name des Messsatzes

(*name* of the measurement set)

Dieser ist sowohl für die gemessenen Daten des Messsatzes als auch für alle damit zusammenhängenden Berechnungen und Plots sowohl im Measurement Set Panel als auch im Inverse HRTF Filters Panel gültig.

Pro Name eines Messsatzes kann immer nur ein Messsatz gemessen werden. Dies soll verhindern, dass gemessene Daten durch unabsichtliches/unbeabsichtigtes

Klicken des Buttons "measure complete measurement set" gelöscht bzw. überschrieben werden.

A.5.2. Latenz

(*latency* [samples])

Zum Messen der in Kapitel 3.2.2. beschriebenen Latenz geben Sie "latency_signal_path" im Matlab Command Window ein. (Hinweis: Das Signal, das normalerweise über einen Lautsprecher ausgegeben wird, muss natürlich mit einem entsprechenden analogen Input Ihrer Audio-Karte / Ihres Audio Interfaces verbunden werden.)

A.5.3. Anzahl der links angebrachten Lautsprecher

(number of left loudspeakers *nleftl*):

Hier wird die Anzahl jener Lautsprecher angegeben, die im Azimut-Bereich (0° , $+180^\circ$] aus der Blickrichtung des Kunstkopfes aufgestellt sind.

Werden durch Wellenfeldsynthese erzeugte fokussierte Schallquellen verwendet, so ist hier die Anzahl der linksseitig platzierten fokussierten Schallquellen anzugeben (Hinweis: *nleftl* entspricht $K/2$ in Kapitel 2)

A.5.4. Anzahl der rechts angebrachten Lautsprecher

(number of right loudspeakers *nrightl*)

Hier wird die Anzahl jener Lautsprecher angegeben, die im Azimut-Bereich [0° , -180°) aus der Blickrichtung des Kunstkopfes aufgestellt sind.

Werden durch Wellenfeldsynthese erzeugte fokussierte Schallquellen verwendet, so ist hier die Anzahl der rechtsseitig platzierten fokussierten Schallquellen anzugeben. (Hinweis: *nrightl* entspricht $K/2$ in Kapitel 2, $n_{\text{total}} = n_{\text{leftl}} + n_{\text{rightl}}$ entspricht K)

A.5.5. Länge der Impulsantwort

(length of IR $L = 2^N$ [samples])

Eingegeben wird N , wobei $L = 2^N$ ist. L gibt die effektive Länge der Impulsantwort in Samples an.

A.5.6. Multiplikationsfaktor

Der Multiplikationsfaktor $mult_L$ entstand aus folgender Überlegung heraus:

Eine HRTF ist ein FIR-System. Sobald mindestens ein Sample (bei einer Samplenummer ungleich 0) von 0 verschieden ist, hat das inverse System eine unendlich lange Impulsantwort (die z-Transformierte hat in diesem Fall mindestens einen Pol außerhalb des Ursprungs). Das invertierte System ist also ein IIR-System, das aber in der Praxis durch ein FIR-System angenähert werden muss, da eine unendlich lange Impulsantwort nicht in endlicher Zeit gemessen werden kann: Die Impulsantwort wird dort abgeschnitten, wo die Amplitudenwerte im Rauschen "untergehen".

Wird $mult_L$ auf 1 gesetzt, ist $length_G = L$, d.h. die Anzahl der Frequenzbins der inversen HRTF-Filter entspricht jener der HRTFs. Falls jedoch die Amplitudenwerte der IHRIRs erst nach L Samples im Rauschen untergehen, kann $mult_L$ größer 1 gewählt werden und damit das Zeitfenster für die IHRIRs verlängert werden ($length_G = L \cdot mult_L$). Da L eine Zweierpotenz ist und $mult_L \in \mathbb{N}$ ist, ist auch $length_G$ eine Zweierpotenz.

A.5.7. Bandbreite für die Glättung der Amplitudengänge

(bandwidth B)

Durch dieses Pop-up-Menü wird die Bandbreite (große Terz, kleine Terz oder Oktave) für die Glättung der Amplitudengänge der Frequenzantworten des Messsatzes angegeben. Wird "no smoothing" angegeben, findet keine Glättung statt. Der Phasengang wird durch die Glättung nicht beeinflusst.

Für spätere Berechnungen der IHRTFs werden die Amplituden-geglätteten HRTFs verwendet.

A.5.8. Daten, die nach jeder Messung geplottet werden sollen

(data to plot after each measurement)

Hier können Sie angeben, welche Daten nach jeder einzelnen Messung in den 4 Plotfenstern des Measurement Set Panels graphisch dargestellt werden sollen:

h_l , h_r , $|H_l|$, $|H_r|$ (Default-Angabe):

In den oberen beiden Plotfenstern des Measurement Set Panels werden die Impulsantworten vom jeweiligen Lautsprecher bzw. von der jeweiligen, durch Wellenfeldsynthese fokussierten Schallquelle zum linken bzw. rechten Mikrofon geplottet, wobei die im Kapitel 3.2.2. genannten Verzögerungen mit gemessen werden müssen.

In den unteren beiden Plotfenstern des Measurement Set Panels werden die entsprechenden Amplitudengänge der Frequenzantworten geplottet.

H_l , H_r :

In den linken bzw. rechten beiden Plotfenstern werden die Frequenzantwort (oben: Amplitudengang, unten: Phasengang) vom jeweiligen Lautsprecher bzw. von der jeweiligen, durch Wellenfeldsynthese fokussierten Schallquelle zum linken bzw. rechten Ohr geplottet.

none:

Nichts wird geplottet.

A.5.9. Winkel der linken Lautsprecher

(angles of left loudspeakers)

Die hier eingegebenen Werte werden vom Programm nicht weiter bearbeitet, sondern lediglich als Erinnerungshilfe abgespeichert.

Gedacht ist das Edit-Text-Feld für die Eingabe einer 2-Zeilen-Matrix, wobei in der ersten Zeile in der p-ten Spalte der Winkel (in Grad) des p-ten linksseitigen Lautsprechers bzw. der p-ten linksseitigen fokussierten Schallquelle und in der zweiten Zeile in der p-ten Spalte der Abstand (in Meter) des p-ten linksseitigen Lautsprechers bzw. der p-ten linksseitigen fokussierten Schallquelle vom Kopfmittelpunkt stehen soll.

A.5.10. Winkel der rechten Lautsprecher

(angles of right loudspeakers)

Die hier eingegebenen Werte werden vom Programm nicht weiter bearbeitet, sondern lediglich als Erinnerungshilfe abgespeichert.

Gedacht ist das Edit-Text-Feld für die Eingabe einer 2-Zeilen-Matrix, wobei in der ersten Zeile in der q -ten Spalte der Winkel (in Grad) des q -ten rechtsseitigen Lautsprechers bzw. der q -ten rechtsseitigen fokussierten Schallquelle und in der zweiten Zeile in der q -ten Spalte der Abstand (in Meter) des q -ten rechtsseitigen Lautsprechers bzw. der q -ten rechtsseitigen fokussierten Schallquelle vom Kopfmittelpunkt stehen soll.

A.5.11. Messen eines gesamten Messsatzes

(measure complete measurement set)

Nach Drücken des Buttons wird der komplette Messsatz mit jenem Namen, der unter name eingegeben ist, gemessen.

Der Messsatz besteht aus $n_{left} + n_{right}$ Messungen, die jeweils die Impuls- bzw. Frequenzantworten vom jeweiligen Lautsprecher bzw. von der jeweiligen, durch Wellenfeldsynthese fokussierten Schallquelle zum linken und rechten Mikrofon des Kunstkopfes messen.

Die Impuls- und Frequenzantworten sowie die wichtigsten Eingabeparameter und die Samplerate werden in `name_freqANDimpresps.mat` abgespeichert, wobei "name" durch den Namen des Messsatzes zu ersetzen ist.

Jede Messung berücksichtigt jeweils die angegebene Latenz *latency*, den Multiplikationsfaktor *mult_L* und die effektive Länge der Impulsantwort *L*. Falls bei der Bandbreite *B* für die Glättung des Amplitudengangs nicht "no smoothing" angegeben ist, wird außerdem der Amplitudengang geglättet. Danach werden die entsprechenden Daten (Auswahl bei "data to plot after each measurement") in den 4 Plotfenstern des Measurement Set Panels geplottet.

Der Sweep wird bei jeder Messung jeweils über einen eigenen Kanal ausgegeben. Der p -te linksseitig platzierte Lautsprecher hat die Kanalnummer $nr = p$, der q -te rechtsseitig platzierte Lautsprecher hat die Kanalnummer $nr = n_{left} + q$.

A.5.12. Plotten aller Impuls- und Frequenzantworten

(plot all impulse and frequency responses)

Nach Drücken dieses Buttons werden mit Ausnahme der Benutzeroberfläche der GUI alle Matlab-Figures geschlossen. Danach werden alle Impuls- und Frequenzantworten des Messsatzes mit jenem Namen, der unter *name* angegeben ist, jeweils in einem eigenen Matlab-Figure geplottet.

A.6. Inverse HRTF Filters Panel

A.6.1. Bandpasseinstellungen

(bandpass settings)

Es wird nun davon ausgegangen, dass die HRTFs nach (3.1) gebildet wurden. In diesem Fall werden die Bandpasseinstellungen auf die zu berechnenden IHRTFs angewandt (siehe auch Kapitel 3.3.)

A.6.2. Berechnung der 3D-Arrays H, G und g

(calculate 3D-arrays H, G and g)

Durch Klicken dieses Buttons werden einerseits die 3D-Arrays H, G und g berechnet, die alle HRTFs bzw. IHRTFs bzw. IHRIRs beinhalten. Die IHRTFs werden wie in Kapitel 3.3. beschrieben berechnet. Die Amplitudengänge der IHRTFs werden daraufhin Bandpass-gefiltert (siehe Kapitel 3.3. und A.6.1.).

Wie durch den Text, der über dem Button angebracht ist, schon angedeutet wird, werden andererseits auch jene Wiedergabesignale berechnet, mit denen die Gesamt-Übertragungsfunktionen gemessen werden können.

Die wichtigsten Daten werden daraufhin in der Datei "name_H_G_and_g.mat" im Ordner "data" abgespeichert, wobei "name" durch den Namen des Messsatzes zu ersetzen ist, der wiederum im Measurement Set Panel angegeben ist.

Basierend auf dem 3D-Array g werden außerdem alle Dateien geschrieben, die für eine Online-Faltung in BruteFIR benötigt werden. Diese werden im Ordner

"BruteFIR" in einem Unterordner, der mit dem Namen des Messsatzes benannt wird, abgespeichert.

Aufbau des 3D-Arrays H:

Die Zeile $k+1$ enthält alle HRTFs an der Stelle k $\underline{H}_{lpb}[k]$ und $\underline{H}_{rgb}[k]$ von den linken ($p=1,2,\dots,nleftl$) bzw. rechten ($q=1,2,\dots,nrightl$) Schallquellen zum linken ($b=l$) bzw. rechten ($b=r$) Trommelfell. Der Frequenzbin k entspricht der Frequenz $\frac{k \cdot f_s}{length_G}$.

In der Spalte nr sind die HRTFs mit jenem Lautsprecher bzw. mit jener fokussierten Schallquelle als Quelle eingetragen, der bzw. die die Kanalnummer nr hat.

In der Page 1 bzw. 2 sind die HRTFs eingetragen, die zum linken Ohr bzw. zum rechten Ohr gelangen.

Aufbau des 3D-Arrays G:

Die Zeile $k+1$ enthält alle IHRTFs an der Stelle k $G_{alp}[k]$ und $G_{arq}[k]$ vom linken ($a=l$) bzw. rechten ($a=r$) Kanal des Kopfhörersignals zu den linken ($p=1,2,\dots,nleftl$) bzw. rechten ($q=1,2,\dots,nrightl$) Wiedergabesignalen.

In der Spalte 1 bzw. 2 sind jene IHRTFs eingetragen, die den linken bzw. rechten Kanal des Kopfhörersignals filtern (Filterung durch Multiplikation).

Die Page nr enthält jene IHRTFs, die als Ergebnis der Filterung das Wiedergabesignal mit der Kanalnummer nr liefern.

Aufbau des 3D-Arrays g:

Die Zeile $n+1$ enthält alle IHRIRs an der Stelle n $g_{alp}[n]$ und $g_{arq}[n]$ vom linken ($a=l$) bzw. rechten ($a=r$) Kanal eines Kopfhörersignals zu den linken ($p=1,2,\dots,nleftl$) bzw. rechten ($q=1,2,\dots,nrightl$) Wiedergabesignalen.

In der Spalte 1 bzw. 2 sind jene IHRIRs eingetragen, die den linken bzw. rechten Kanal des Kopfhörersignals filtern (Filterung durch Faltung).

Die Page nr enthält jene IHRIRs, die als Ergebnis der Filterung das Wiedergabesignal mit der Kanalnummer nr liefern.

Wie bereits erwähnt werden auch die Wiedergabesignale berechnet, mit denen die Gesamt-Übertragungsfunktionen gemessen werden können.

Dazu wird zunächst ein exponentieller Sweep generiert: Es werden jene Eingaben für T , $range_dB$, f_1 und f_2 verwendet, die aktuell im Sweep Creation Panel eingestellt sind. Beachten Sie, dass Sie in diesem Fall *nicht* auf "generate sinesweep" im Sweep Creation Panel drücken müssen (wenn Sie auf "generate sinesweep" drücken, wird ein neuer Sweep zur Messung der Frequenz- und Impulsantworten generiert).

Aus diesem Sweep wird jeweils ein "linkes bzw. rechtes Kopfhörersignal" gemacht.

Unter Berücksichtigung von *latency* werden die Signale berechnet, die bei der Messung der Gesamt-Übertragungsfunktionen abgespielt werden. Für die Latenz *latency* wird der aktuell im Measurement Set Panel eingestellte Wert verwendet.

A.6.3. Namenserverweiterung

(name extension for *.mat-file, which contains measurements of all H_eve_ab a,b=l,r)

Damit ist also die Namenserverweiterung für die *.mat-Datei gemeint, in der die Messungen der Gesamt-Übertragungsfunktionen abgespeichert werden:

Diese Namenserverweiterung wird benötigt, wenn für die Messungen verschiedener Gesamt-Übertragungsfunktionen die selben inversen HRTF-Filter verwendet werden sollen. Pro Messsatz und Namenserverweiterung kann die Messung der Gesamt-Übertragungsfunktionen immer nur einmal durchgeführt werden. Dies soll ein Überschreiben von Daten durch unbeabsichtigtes Drücken des Buttons "measure all H_eve_ab a,b=l,r" verhindern.

A.6.4. Messung der Gesamtübertragungsfunktionen

(measure all H_eve_ab a,b=l,r)

Es werden die Übertragungsfunktionen vom linken bzw. rechten Kanal des Kopfhörersignals zum linken bzw. rechten virtuellen Kopfhörersignal gemessen und in der Datei "name_H_eve_name_extension_H_eve" abgespeichert, wobei "name" durch den Namen des Messsatzes (einstellbar im Measurement Set Panel) zu ersetzen ist und "name_extension_H_eve" durch die Namenserverweiterung (siehe Kapitel A.6.3.) zu ersetzen ist, die im Inverse HRTF Filters Panel eingestellt ist.

Im Zuge der Messung werden einerseits jene Signale wiedergegeben, die nach Drücken des Buttons "calculate 3D-Arrays H, G and g" (siehe Kapitel A.6.2.) berechnet wurden. Andererseits werden mit den Mikrofonen des Kunstkopfes die jeweiligen Systemantworten aufgenommen, aus denen die virtuellen Kopfhörersignale berechnet werden. Daraufhin können dann die Gesamt-Übertragungsfunktionen berechnet werden (siehe Kapitel 3.5).

A.6.5. Plot-Optionen

Für die zur Auswahl stehenden Plot-Optionen werden die Daten jenes Messsatzes verwendet, dessen Name im Measurement Set Panel zum Zeitpunkt des Drückens auf den jeweiligen Push-Button eingegeben ist.

A.6.5.1. Plotten der IFFTs aller invertierten Übertragungsfunktionen

(plot IFFTs of H_{atb}^{-1} $a,b=l,r$):

H_{atb} bezeichnet die Übertragungsfunktion von der p -ten linken Schallquelle ($a=l$, $p=1,2,\dots,nleftl$) bzw. von der q -ten rechten Schallquelle ($a=r$, $q=1,2,\dots,nrightl$) zum linken ($b=l$) bzw. rechten ($b=r$) Ohr.

Beachten Sie, dass hier die Begriffe "invertiert" (engl.: inverted) statt "invers" (engl.: inverse) und "Übertragungsfunktion" (engl.: transfer function) statt "HRTF" verwendet werden (obwohl es sich bei der Übertragungsfunktion wohl um eine HRTF handelt), damit möglichst wenig Verwechslungsgefahr mit den inversen HRTF-Filtern besteht. Ein Beispiel einer solchen Impulsantwort ist in Abbildung A.3 dargestellt.

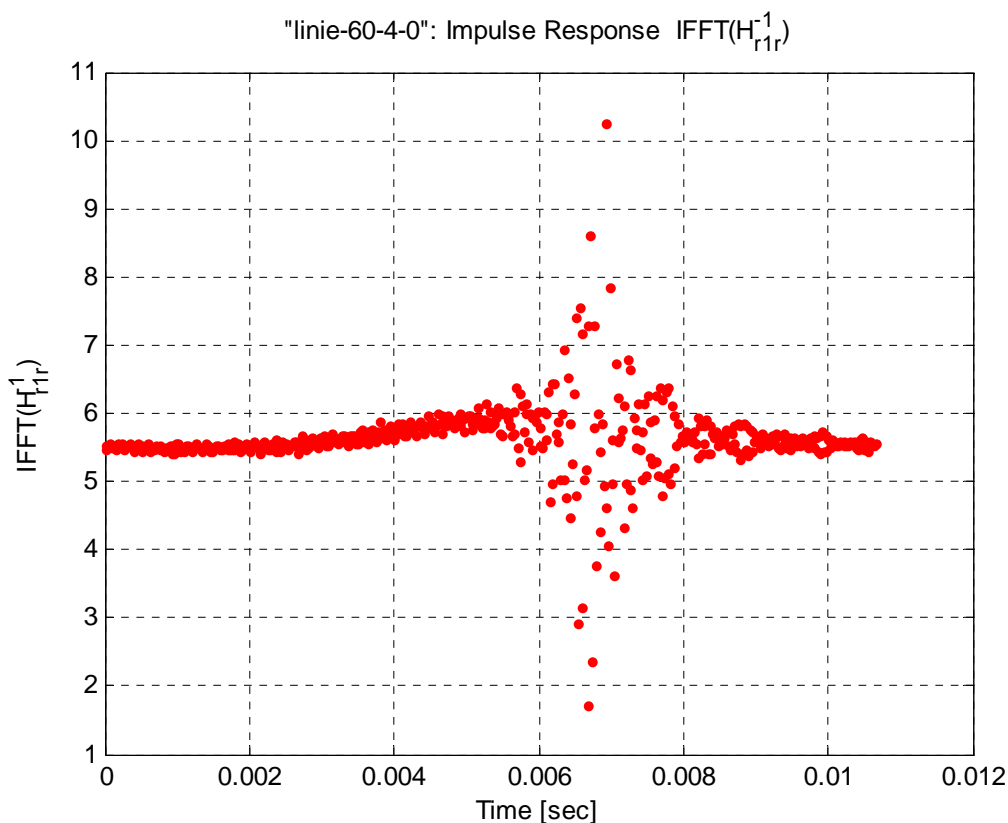


Abbildung A.3: IFFT einer invertierten Übertragungsfunktion

A.6.5.2. Plotten aller inversen HRTFs \underline{G}_{alp} bzw. \underline{G}_{arq}

(plot all inverse HRTFs)

Die inversen HRTFs entsprechen jeweils der Kombination einer Spalte und einer Page aus dem 3D-Array \underline{G} (siehe Kapitel A.6.2.).

In Abbildung A.4 ist ein solches inverses HRTF-Filter dargestellt. Um den Verlauf der Kurve im interessanten Frequenzbereich besser erkennen zu können, wurde beim Amplitudengang in den Bereich von -20 bis +30 dB hineingezoomt.

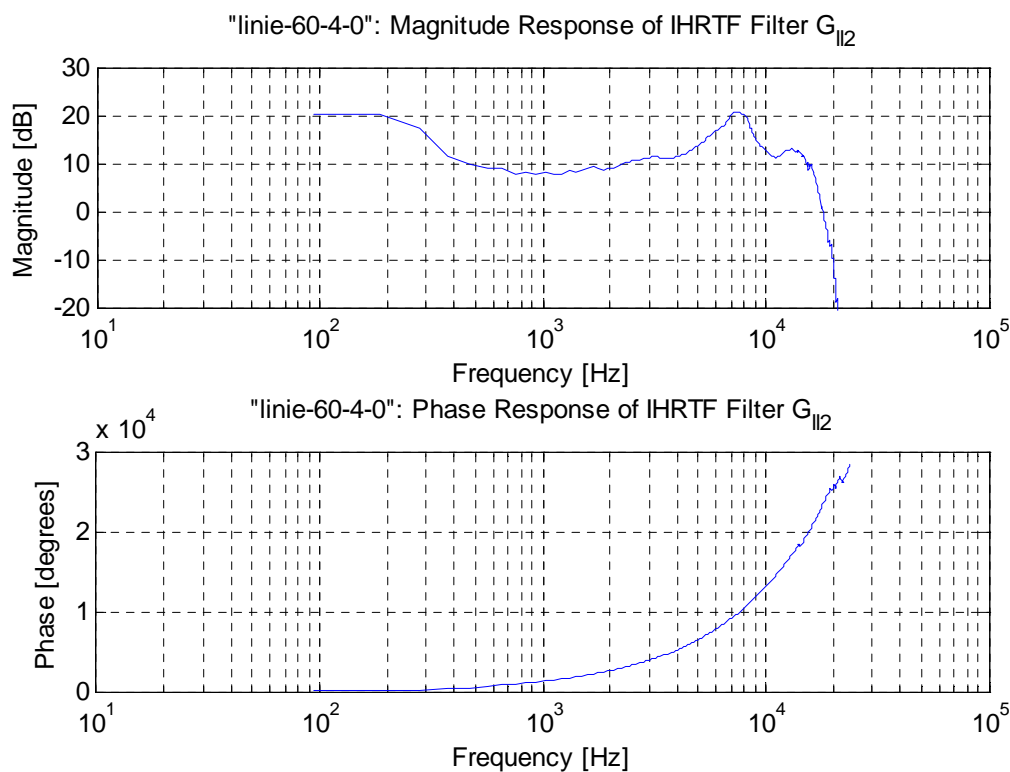


Abbildung A.4: Frequenzantwort eines inversen HRTF-Filters

A.6.5.3. Plotten aller IFFTs der inversen HRTF-Filter \underline{g}_{alp} bzw. \underline{g}_{arq}

(plot all IFFTs of inverse HRTFs)

Sie entsprechen jeweils der Kombination einer Spalte und einer Page aus dem 3D-Array \underline{g} (siehe Kapitel A.6.2.).

Ein Beispiel einer solchen IHRIR ist in Abbildung 4.8 gegeben.

A.6.5.4. Plotten der Gesamt-Übertragungsfunktionen

(plot all H_{eve_ab} a,b=l,r):

Falls bei den Messungen der Gesamt-Übertragungsfunktionen eine Namenserverweiterung (siehe Kapitel A.6.3.) im Inverse HRTF Filters Panel eingegeben wurde, muss diese Namenserverweiterung auch beim Plotten dieser Übertragungsfunktionen im entsprechenden Edit-Text-Feld angegeben sein.

Ob und wenn ja mit welcher Bandbreite B die Amplitudengänge dieser Übertragungsfunktionen geglättet werden sollen, kann (unabhängig von der Bandbreite für die Glättung der Amplitudengänge der HRTFs) im Measurement Set Panel eingestellt werden.

Die Glättung wird direkt nach Drücken des Buttons "plot all H_{eve_ab} a,b=l,r" ausgeführt und dient in diesem Fall nur der Darstellung (die Werte der geglätteten Amplitudengänge überschreiben weder alte Werte noch werden sie separat abgespeichert). Die Abbildung 3.4 in Kapitel 3.5. zeigt die gleichen Übertragungsfunktionen $\underline{H}_{eve,rl}$ und $\underline{H}_{eve,rr}$ ohne bzw. mit Glättung der Amplitudengänge in Terzbändern.

A.6.6. Beim Plotten werden geöffnete Figures geschlossen

(Close open figures when plotting)

Wenn dieser Toggle-Button eingeschaltet ist, werden beim Plotten (d.h., wenn der Push-Button einer Plot-Option gedrückt wird) mit Ausnahme der GUI alle bereits geöffneten Matlab-Figures geschlossen.

Wenn der Toggle-Button deaktiviert ist, bleiben diese Figures geöffnet.

A.6.7. Offline-Berechnung der Wiedergabesignale

Mit dem (I)HRTF-Messsystem können aus einem beliebigen Kopfhörer-Signal (Stereo-Signal) die Wiedergabesignale offline berechnet werden: Zur Berechnung werden die IHRTFs jenes Messsatzes verwendet, dessen Name aktuell im Measurement Set Panel eingegeben ist. Falls die IHRTFs noch nicht berechnet wurden, drücken Sie vorher den Button "calculate 3D-arrays H, G and g" (siehe Kapitel A.6.2.).

Im Edit-Text-Feld unter "source stereo wave file" ist der vollständige Dateiname der gewünschten wave-Datei aus dem Ordner "sound examples" einzutragen (die Endung ".wav" muss im Dateinamen enthalten sein). Diese wave-Datei muss Stereo sein und die gleiche Samplerate wie die IHRTFs haben.

Wenn alle genannten Voraussetzungen erfüllt sind, können die Wiedergabesignale durch Drücken des Buttons "calculate loudspeaker signals" berechnet werden.

Die berechneten Signale werden anschließend im Ordner "sound examples" in der Datei "wav_file_name.mat" abgespeichert, wobei "wav_file" durch den Namen der wave-Datei, die das Kopfhörersignal beinhaltet, und "name" durch den Namen des Messsatzes zu ersetzen ist.

Die Offline-Berechnung der Signale ist v.a. für Testzwecke gedacht - wenn z.B. das (I)HRTF-Messsystem auf einem Windows-PC getestet werden soll (das Online-Faltungs-Programm BruteFIR läuft nur unter Linux). Das Kopfhörer-Signal darf dabei nicht zu lang sein, da sonst der Arbeitsspeicher (RAM) Ihres Rechners nicht ausreichen wird.

Durch Drücken des Buttons "play loudspeaker signals" werden die berechneten Signale abgespielt.

Referenzen

- [1] "Polhemus Fastrak", http://www.polhemus.com/?page=Support_Fastrak (Verwendete Version: 2008)
- [2] M. M. Boone, E. N. G. Verheijen, G. Jansen, "Virtual Reality by Sound Reproduction based on Wave Field Synthesis", Convention Paper 4145, *100th AES Convention*, Copenhagen, Denmark (1996)
- [3] M. Meier, "Entwicklung einer echtzeitfähigen Wellenfeldsynthese-Rendering Software", Diplomarbeit, Fachhochschule Deggendorf (2008)
- [4] J. Ahrens, S. Spors, "Notes on Rendering Focused Directional Virtual Sound Sources in Wave Field Synthesis", *Tagungsband Fortschritte der Akustik - DAGA 2008*, Dresden (2008)
- [5] K. Laumann, G. Theile, H. Fastl, "A virtual headphone based on wave field synthesis", *Acoustics 08 Paris*, pp. 3593-3597 (2008)
- [6] "BruteFIR", <http://www.ludd.luth.se/~torger/brutefir.html> (Letztes Update: 2006)
- [7] E. N. G. Verheijen, "Sound Reproduction by Wave Field Synthesis", Dissertation, Delft University of Technology (1997)
- [8] S. Müller, P. Massarani, "Transfer-Function Measurement with Sweeps", *J. Audio Eng. Soc.*, vol. 49, no. 6, pp. 443-471 (2001)
- [9] Matt Frear, "pa_wavplay", <http://sourceforge.net/projects/pa-wavplay/> (Letztes Update: 2004)
- [10] E. J. Berdahl, J. O. Smith, "Transfer Function Measurement Toolbox", http://ccrma.stanford.edu/realsimple/imp_meas/ (2008)

Danksagung

Ich bedanke mich bei meinen beiden Betreuern Dipl.-Ing. Klaus Laumann vom IRT und Dr. Alois Sontacchi von der Kunstuniversität Graz herzlich für die hilfsbereite Unterstützung und für die Möglichkeit der Durchführung dieses Toningenieur-Projekts im Rahmen eines Praktikums am IRT.

Weiters bedanke ich mich bei Dr. Helmut Wittek von der Firma Schoeps und bei Dr.-Ing. Günther Theile, die dieses Praktikum und in Folge dessen dieses Projekt ermöglicht haben.

Außerdem gilt mein Dank allen Kollegen des IRT, die mir bei diesem Projekt mit Rat und Tat zur Seite gestanden sind. Besonders bedanke ich mich bei Dipl.-Ing. (FH) Michael Meier für die gute Zusammenarbeit bei den Messungen, bei Hans Depner, Werner Flesch und Werner Zelfel für die handwerkliche Unterstützung, bei Dipl.-Ing. Sebastian Goossens, Dipl.-Ing. (FH) Manuel Wiese und Dipl.-Ing. Martin Link für diverse Diskussionen und Hilfestellungen in den Bereichen der Akustik und der Audiosignalverarbeitung.

Mein abschließender Dank gilt meiner Familie und allen Freunden, die mich bei der Organisation meines Aufenthaltes in Deutschland unterstützten.