

Bachelorarbeit

Im Studiengang Audiovisuelle Medien

Virtualisierung einer Regie mit dem Smyth Realiser A16

am Beispiel des Raumes U48 an der Hochschule der Medien Stuttgart

Vorgelegt von Julian Eckstein
an der Hochschule der Medien Stuttgart am 14.08.2025
zur Erlangung des akademischen Grades eines Bachelor of Engineering

Erstprüfer: Prof. Dr. Frank Melchior
Zweitprüfer: Prof. Oliver Curdt

Hiermit versichere ich, Julian Eckstein, ehrenwörtlich, dass ich die vorliegende Bachelorarbeit mit dem Titel: „Virtualisierung einer Regie mit dem Smyth Realiser A16 – am Beispiel des Raumes U48 an der Hochschule der Medien Stuttgart“ selbstständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Ebenso sind alle Stellen, die mit Hilfe eines KI-basierten Schreibwerkzeugs erstellt oder überarbeitet wurden, kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden.

Ich habe die Bedeutung der ehrenwörtlichen Versicherung und die prüfungsrechtlichen Folgen (§ 24 Abs. 2 Bachelor-SPO, § 23 Abs. 2 Master-SPO (Vollzeit)) einer unrichtigen oder unvollständigen ehrenwörtlichen Versicherung zur Kenntnis genommen.

Stuttgart, 14.08.2025

A handwritten signature in blue ink, appearing to read 'J. Eckstein', is written over a horizontal line.

Kurzfassung – Abstract

In dieser Arbeit wird das Thema der binauralen Raumsynthese bzw. Lautsprecher-simulation behandelt. Zunächst erfolgt eine Beschreibung der theoretischen Grundlagen der Binauralsynthese. Anschließend wird ein Lautsprecher-Setup mit dem Smyth Realiser A16 virtualisiert, was beispielhaft am Raum U48 an der Hochschule der Medien in Stuttgart durchgeführt wird. Im nächsten Schritt wird die reale mit der virtuellen Umgebung objektiv verglichen. Zu diesem Zweck wird das Gerät zunächst eingerichtet und anschließend werden Messungen der binauralen Raumimpulsantworten eines Neumann KU100 Kunstkopfes durchgeführt. Zur objektiven Bewertung dieser Messungen werden die binauralen Raumimpulsantworten des Realisers und die der Regie mit dem Programm Room EQ Wizard gemessen und anschließend einander gegenübergestellt. Dabei ergab sich eine weitestgehende Übereinstimmung zwischen den beiden Versionen. Einzelne unerwartete Abweichungen traten auf, wie in etwa ein 2 dB Pegel-Offset zwischen den Ohren in den Messungen der virtuellen Version des Realisers.

This thesis addresses the topic of binaural room synthesis, specifically the simulation of loudspeaker reproduction. The work begins with an overview of the theoretical principles underlying binaural synthesis. Subsequently, a loudspeaker setup is virtualized using the Smyth Realiser A16, demonstrated with reference to Room U48 at the Stuttgart Media University. The real and virtual listening environments are then subjected to an objective comparison. For the virtualization process, the device is first configured, after which binaural room impulse responses are measured using a Neumann KU100 dummy head. To objectively evaluate the measurements, the binaural room impulse responses obtained from the Realiser and those recorded in the actual control room are captured using Room EQ Wizard and subsequently compared. The results reveal a high degree of correspondence between the two versions, with minor but noteworthy deviations, such as a 2 dB interaural level offset observed in the measurements of the virtual environment generated by the Realiser.¹

¹ Diese Übersetzung wurde mit dem ChatGPT angefertigt. Die gesamte Arbeit wurde mit dem KI-Tool DeepL grammatikalisch korrigiert und sprachlich leicht optimiert.

Inhaltsverzeichnis

Kurzfassung – Abstract	I
Abbildungsverzeichnis	IV
Tabellenverzeichnis	V
Abkürzungsverzeichnis	VI
1 Einführung	1
1.1 Motivation	2
2 Räumliches Hören	4
3 Binauralsynthese	6
3.1 Allgemeines Prinzip & Aufbau	6
3.2 Faltung & Entfaltung	8
3.3 Head-Related Transfer Function (HRTF)	9
3.4 Kopfhörerentzerrung	15
3.5 Dynamische Binauralsynthese: Headtracking	16
3.6 Binaurale Raumsynthese	16
4 Praktische Arbeit: Ausgangslage & Geplantes Vorgehen	18
4.1 Raum U48	19
4.2 Neumann KU100	21
4.3 Smyth Realiser A16	21
5 Messung	23
5.1 Funktionsprinzip des Realisers	23
5.2 Einbindung	23
5.3 Konfiguration einer „PRIR“-Messung	24
5.4 Messablauf	29
6 Audiowiedergabe mit dem Realiser	31
6.1 Listening Rooms	31
6.2 Presets & User	32
6.3 Besonderheiten in U48	33

7	Workflow zur Abbildung verschiedener Räume	36
8	Auswertung.....	37
8.1	<i>RoomEQWizard.....</i>	<i>37</i>
8.2	<i>Allgemeines Vorgehen.....</i>	<i>38</i>
8.3	<i>Iteration 1.....</i>	<i>39</i>
8.4	<i>Iteration 2.....</i>	<i>41</i>
9	Fazit & Schlusswort.....	48
 Literatur – und Quellenverzeichnis		50

Abbildungsverzeichnis

Abbildung 1 - Frequenzgänge Außenohr-Übertragungsfunktion, Azimut 90° links, Elevation 0° (Blauert & Braasch, 2008, S.91)	5
Abbildung 2 - Binaurales Rendering-System (in Anlehnung an: Weinzierl & Slavik, 2008, S.672)	7
Abbildung 3 - LTI-System (Li & Peissig, 2020)	8
Abbildung 4 - Außenohrübertragungsfunktion als Impulsantwort & Frequenzgang (Hammerhøi & Møller, 2005, S.239)	9
Abbildung 5 - Multi-Lautsprecher-Setups: Aachen HRTF Database (oben links) (Bomhardt et al., 2016, S.3), Oldenburg HRTF Database (oben rechts) (Denk et al., 2018, S.4) & Einzel-Lautsprecher-Setup: Ircam Listen Datenbank (unten) (IRCAM, o.D.)	14
Abbildung 6 - HRTFs des linken Ohrs gemessen am Trommelfell, offenen und geschlossenen Gehörgang, unten: frequenzweise berechnete Standardabweichung (Hammershøi & Møller, 2005)	15
Abbildung 7 - Kopfhörer-Übertragungsfunktionen: Sony MDR-102 (links), BeyerDynamic DT-990 (rechts), gemessen an 40 Personen (Hammershøi & Møller, 2005, S.231)	15
Abbildung 8 - Preset-Kurve des Trinnox Prozessor	19
Abbildung 9 - Signalfluss in U48	20
Abbildung 10 - Smyth Realiser A16 in der Headstand-Version	22
Abbildung 11 - Realiser Einbindung bei Messungen	23
Abbildung 12 - Realiser Einbindung bei der Audiowiedergabe	24
Abbildung 13 - Messaufbau in U48	28
Abbildung 14 - Neumann KU100 im Messaufbau mit Headtracker (links) & Ohrmikrofone im Detail (rechts)	29
Abbildung 15 - Übersicht der Verhältnisse zwischen User, Presets und Listening Rooms	33
Abbildung 16 - durchschnittlicher Frequenzgang (VAR Glättung): rechtes Ohr, 0°; real (gelb) & virtuell (blau)	40
Abbildung 17 - Peak bei 4kHz (VAR-Glättung), Oben: real (gelb/lila) & virtuell (rot/grün) separiert, Unten links: linkes Ohr, Unten rechts: rechtes Ohr	40
Abbildung 18 - Frequenzgang (VAR Glättung): real, Center, 0°; linkes (rot) & rechtes (grün) Ohr	42
Abbildung 19 – Frequenzgang (VAR Glättung): virtuell, Center, 0°, autoEQ-Kurve; linkes (blau) & rechtes (orange) Ohr	42
Abbildung 20 - Frequenzgang (VAR Glättung): virtuell, Center, 0°, flatEQ-Kurve; linkes (lila) & rechtes (pink) Ohr	43
Abbildung 21 - binaurale Raumimpulsantworten : Center, 0°; real (links) & virtuell (rechts)	44
Abbildung 22 - binaurale Raumimpulsantworten : Center, 90° links; real (grün/rot) & virtuell (blau/orange)	44

Abbildung 23 - Impulsantworten; grün: virtuell, 57° links, Center, rechtes Ohr; blau: virtuell, 90° links, Center, rechtes Ohr; pink: real, 57° links, Center, rechtes Ohr	45
Abbildung 24 - Frequenzgang (VAR Glättung): linkes Ohr, Center, 90° links; real (rot) und virtuell (blau)	45
Abbildung 25 - Frequenzgang (VAR Glättung): rechtes Ohr, Center, 90° links; real (grün) & virtuell (orange)	46
Abbildung 26 - Frequenzgang (VAR Glättung): real, Left wide, 0°; linkes (blau) & rechtes (orange) Ohr	46
Abbildung 27 - Frequenzgang (VAR Glättung): virtuell, Left wide, 0°; linkes (rot) & rechtes (grün) Ohr	46
Abbildung 28 - ETC: Center, 0°, linkes Ohr; virtuell (grün) & real (gelb)	47
Abbildung 29 - ETC: Center, 90° links, rechtes Ohr; virtuell (rot) & real (blau)	47

Tabellenverzeichnis

Tabelle 1 - Kanalbelegung, Positionen & Lautsprecher in U48	21
Tabelle 2 - Lautsprecher-Labels U48 & Realiser	25
Tabelle 3 - Labelbezeichnung "14.1.1" Realiser & U48, Dante-Routing Anpassung	26
Tabelle 4 - Wiedergabe-Routing "14.1.1"	34
Tabelle 5 - Wiedergabe-Routing 9.1.4	35

Abkürzungsverzeichnis

BRIR	Binaural Room Impulse Response
ETC	Energy-Time-Curve
HdM	Hochschule der Medien
HPEQ	Headphone Equalization
HRTF	Head-Related Transfer Function
IF	Interface
ILD	Interaural Time Difference
ITD	Interaural Level Difference
LTI	Linear-Time-Invariant
PRIR	Personal Room Impulse Response
REW	RoomEQWizard
RIR	Room Impulse Response
SNR	Signal-to-Noise Ration
SVS	Smyth Virtual Surround

1 Einführung

Ist 3D-Audio die Zukunft? Bei dieser Frage, gehen die Meinungen innerhalb der Audiobranche vermutlich weit auseinander. Sicher ist jedoch, dass 3D-Audio-Technologien in den letzten Jahren einen regelrechten Boom erlebt haben und ein Gegenstand zahlreicher spannender Entwicklungen sind. Besonders binaurales Audio erfreut sich dabei steigenden öffentlichen Interesses (vgl. Kapitel 1.1).

Die sogenannte Binauralsynthese ist auch das übergeordnete Thema dieser Arbeit. Im speziellen erfolgt eine Kombination verschiedener Themenkomplexe. Im Fokus stehen insbesondere die binaurale Raum- bzw. Lautsprecher-simulation, sowie die Individualisierung von immersivem Audio. Konkret wird mithilfe des Realisers A16 der Firma Smyth Research am Beispiel des Raumes U48 an der Hochschule der Medien in Stuttgart (HdM) eine binaurale Lautsprecher-simulation erstellt und diese objektiv mit der realen Umgebung verglichen. Die Relevanz ergibt sich aus dem eingangs erwähnten zunehmenden Interesse an binauralen Technologien und diesen als Gegenstand eines der Schwerpunkte an der Hochschule. Darüber hinaus ist das Thema an der HdM auch gerade deshalb aktuell, da im Zuge von Sanierungsarbeiten die Studioräume umgebaut wurden und es notwendig ist, die Messungen mit dem Realiser neu durchzuführen und diesen neu einzurichten. Am Ende soll dabei klar werden, wie die Messung von binauralen Raumimpulsantworten mit dem Realiser A16 insbesondere im neuen U48-Setup und allgemein raumübergreifend sinnvoll durchgeführt wird und wie gut die Simulation mit dem realen Raum vergleichbar ist.

Um dies zu beantworten, ist die Arbeit wie folgt strukturiert. Es lassen sich grob drei Abschnitte definieren. Der erste davon behandelt die theoretischen Grundlagen. Von einem kurzen Exkurs über räumliches Hören ausgehend, liegt der Fokus dabei auf den Prinzipien der Binauralsynthese. Im Anschluss folgt der praktische Teil der Arbeit. Dieser beginnt mit dem zweiten großen Abschnitt, der Arbeit mit dem Realiser. Eingangs erfolgt ein kurzer Überblick über das verwendete Equipment und die Umgebung der Messung. Wie eine Messung konkret abläuft und wie der Realiser in das neue Setup integriert wird, wird in diesem Abschnitt erläutert. Anschließend wird die Einrichtung des Realisers zur Audiowiedergabe erklärt. Darauf aufbauend wird ein Workflow zur Messung und Wiedergabe verschiedener Räume erarbeitet, wodurch das Thema der raumübergreifenden Messungen aufgegriffen wird. Im letzten Abschnitt der Arbeit wird schließlich die objektive Vergleichbarkeit der Simulation des Realisers mit dem realen Abhörraum untersucht. Hierzu werden mit der Software „RoomEQWizard“ die virtualisierten und realen binauralen Raumimpulsantworten gemessen und einander gegenübergestellt. Dieser Abschnitt dient der Überprüfung des Messergebnisses.

Zusammengefasst sind die Hauptziele der Arbeit somit die erneute Integration und Konfiguration des Realisers in U48, die Durchführung der Messungen, ein Vergleich der Raumimpulsantworten hinsichtlich ihrer messbaren Qualität, sowie die Erarbeitung eines Workflows zur Messung verschiedener Räume.

1.1 Motivation

Ein Großteil der Motivation für diese Arbeit ergibt sich aus der bereits eingangs erwähnten steigenden Popularität, sowie dem immer größer werdenden Markt für 3D-Audio-Technologien. Daran lässt sich die Aktualität des Themas ablesen. Mit dem expandierenden Markt, geht auch eine zunehmend sinkende Zugangsschwelle für Endnutzer einher. Im Folgenden soll dies anhand einiger Beispiele veranschaulicht werden. Dolby Atmos ist in diesem Zusammenhang ein zentrales Beispiel. Mittlerweile gibt es viele verschiedene Anbieter, die eine Integration von Atmos zur Verfügung stellen. Ganz vorne mit dabei sind Anbieter wie Netflix oder Disney+ für Video- (*Wo kann ich Inhalte in Dolby Atmos anschauen?* o.D.) oder auch Apple Music oder Tidal für Musikstreaming (*Dolby Atmos Music – Musik wie nie zuvor*, o.D.). Neben Atmos ist MPEG-H ein wichtiger Player, wenn es um immersives Audio geht. Im Musik-Streaming-Bereich ist hier Sonys 360 Reality Audio zu nennen, das auf MPEG-H basiert. Daneben ist MPEG-H im brasilianischen und koreanischen Fernsehen als Standard-Codec vertreten und bietet zudem zahlreiche Möglichkeiten zur individuellen Anpassung: die Einstellung der Dialoglautstärke, unterschiedliche Kommentare bei Sport-Livestreams oder die Wahl verschiedener Mixe (Fraunhofer IIS, o. D.). Beide Formate, Atmos und MPEG-H, können sowohl über Lautsprecher als auch binaural über Kopfhörer wiedergegeben werden.

In diesem Kontext lassen sich nun zwei große persönliche Interessenskomplexe ausmachen: Erstens die Personalisierung des Konsums von Audioinhalten im Allgemeinen und von 3D-Inhalten im engeren Sinne und zweitens binaurales Audio. Dabei ist dessen großes Potential zu betonen, denn durch die Möglichkeit 3D-Audio über Kopfhörer wiederzugeben, wird ein recht einfacher Zugang zu dieser Technik für eine breite Masse ermöglicht. Bringt man diese beiden Punkte nun zusammen, ergibt sich ein komplexes Entwicklungsfeld: die Personalisierung der Außenohr-Übertragungsfunktionen (vgl. Kapitel 3.3). Eine gewisse Komplexität liegt in der Natur der Sache (vgl. Kapitel 2), sodass hochqualitative Lösungen vor allem im Konsumentenbereich noch nicht weit verbreitet sind. Vorhandene Lösungen sind beispielsweise Apples „Personalisiertes 3D-Audio“ (*Sound mit personalisiertem 3D-Audio für AirPods und Beats erleben*, 2025) oder Genelecs „Aural ID“ (*Aural ID*, o. D.). Beide verwenden optische Verfahren, um daraus auf die Geometrie des Ohres zu schließen und einen HRTF-Satz (engl. head-related transfer function) zu generieren. Hier lässt sich nun möglicherweise der nächste große Schritt in der 3D-Audio-Welt erkennen, denn offensichtlich besteht

ein marktseitiges Interesse an der Weiterentwicklung in diesem Bereich. Eine erste tatsächliche Messung eines HRTF-Satzes ist im Konsumentenbereich ist mit dem Realiser von Smyth Research möglich. Damit kommen wir zum Thema dieser Arbeit zurück. Der Realiser schließt mit seiner Messroutine erstmals eine Lücke und ermöglicht somit, wenn auch in eingeschränktem Maße, die Messung der eigenen HRTFs. Dies ist also ein weiterer großer Motivationsaspekt für dieser Arbeit, denn die Messroutine des Realisers ähnelt stark den „klassische“ HRTF-Messungen in Reflexionsarmen Räumen und bietet so die Möglichkeit, selbst einmal eine HRTF-Messung durchzuführen.

2 Räumliches Hören

Im Folgenden werden einige Grundlagen des räumlichen Hörens noch einmal aufgegriffen. Die Thematik wird dabei allerdings recht oberflächlich behandelt. Einige Punkte werden außen vor gelassen oder nur sehr grob beschrieben. Der Abschnitt soll ein Gefühl für die Thematik vermitteln und einen Einblick geben, wie unser Gehör Signale verarbeitet. Dies dient vor allem der Einordnung der Binauraltechnik in den übergeordneten Kontext und soll später dabei helfen die notwendigen Verbindungen zur Binauralsynthese, möglichen Herausforderungen und der Analyse der Messungen zu ziehen.

Doch was ist überhaupt räumliches Hören? Eine mögliche Antwort könnte lauten: Räumliches Hören ist unsere alltägliche Wahrnehmung von Schall. Als Menschen nehmen wir unsere Umwelt dreidimensional wahr, wir sehen dreidimensional und hören räumlich. Wenn wir etwas hören, bildet unser Gehirn mit den ihm zur Verfügung stehenden Informationen ein Hörereignis (Blauert, 1974/1996). Räumliches Hören beschreibt nun „(...) die Beziehungen zwischen den Orten sowie den räumlichen Ausdehnungen der Hörereignisse untereinander und zu den korrelierten Merkmalen anderer Ereignisse–vorwiegend Schallereignisse (...)“ (Blauert & Braasch, 2008, S.87). Die wichtigsten Informationen hierfür sind dabei in den beiden Eingangssignalen der Ohren enthalten (Blauert, 1974/1996). Unser Gehirn analysiert diese hinsichtlich vieler verschiedener Aspekte. Doch welche sind dies nun genau?

Abbildung 1 zeigt den Frequenzgang zweier Ohrsignale bei einer Einfallsrichtung von 90° links. Besonders gut zu erkennen sind ein Pegelabfall am rechten Ohr in den hohen Frequenzen, ebenso wie die allgemeine Dämpfung in der Kurve des rechten Ohrs. Diese Abweichungen zwischen den beiden Ohrsignalen werden als interaurale Unterschiede bezeichnet. Dabei handelt es sich vor allem um Laufzeit- und Pegelunterschiede (engl. interaural time difference, ITD & interaural level difference, ILD), welche beispielsweise durch unterschiedliche Abstände der Ohren zur Schallquelle oder durch Abschattung am Kopf entstehen (Blauert, 1974/1996). In Abbildung 1 kommt der Schall nun extrem von links. Fällt der Schall hingegen von vorne ein, so kann es kaum Unterschiede zwischen den Ohrsignalen kommen, da die Laufzeiten zu den Ohren gleich sind und Pegel nicht beeinflusst werden. Somit sind die Ohrsignale relativ identisch. Dies wirft die Frage auf, wie in diesen Fällen eine Ortung erfolgt. Hier kommen monaurale Merkmale, die sogenannten richtungsbestimmenden Bänder, zum Tragen. Grob gesagt, werden anatomisch bedingt je nach Schalleinfallsrichtung bestimmte Frequenzbereiche angehoben und andere abgesenkt. Diese Filterung ist für den jeweiligen Schalleinfallswinkel charakteristisch, sodass das Gehirn anhand dessen ein Hörereignis am entsprechenden Ort bilden kann. Blauert (1974/1996, S.177) bringt die tragenden monauralen und

interauralen Aspekte wie folgt auf den Punkt: „The essential monaural attributes of the ear input signals are time and level differences between the individual spectral components of each individual ear input signal; the essential interaural signal attributes are time and level differences between corresponding spectral components of the two ear input signals.“ Die bisherigen Erläuterungen wurden unter der Annahme einer einzigen zu lokalisierenden Schallquelle vorgenommen. Sind mehrere Schallquellen involviert, also in etwa mehrere Lautsprecher oder aber auch Reflexionen, so sind für die Lokalisation weitere psychoakustische Effekte von Bedeutung. Namentlich sind das vor allem der Präzedenzeffekt („Gesetz der ersten Wellenfront“) und die Summenlokalisierung. Der Präzedenzeffekt besagt, dass die Wellenfront, welche das Ohr zuerst erreicht, richtungsbestimmend für ein Hörereignis ist. Die Summenlokalisierung lässt sich anhand der Stereophonie gut erklären. In einem Stereo-Abhörsetup bildet unsere Wahrnehmung eine Phantomschallquelle in der Mitte zwischen den Lautsprechern (Blauert, 1974/1996). Neben diesen Anhaltspunkten spielen auch Informationen aus anderen Sinnen eine Rolle. Vor allem die Orientierung und Bewegung des Kopfes sind wichtig, um zu entscheiden, ob Schalle von hinten, vorne, oben oder unten kommen (Blauert, 1974/1996).

Das Gehirn wertet diese Information also nach und nach aus und kann so ein Hörereignis an einer bestimmten Position bilden (Frieesecke, 2014). Zusammenfassend lässt sich sagen, dass unsere Hörwahrnehmung ein hochkomplexer Vorgang ist. Da viele Aspekte des räumlichen Hörens durch die Anatomie beeinflusst werden und diese bei jedem Menschen einzigartig ist, ist somit auch das Hören an sich ein hoch individueller Prozess.

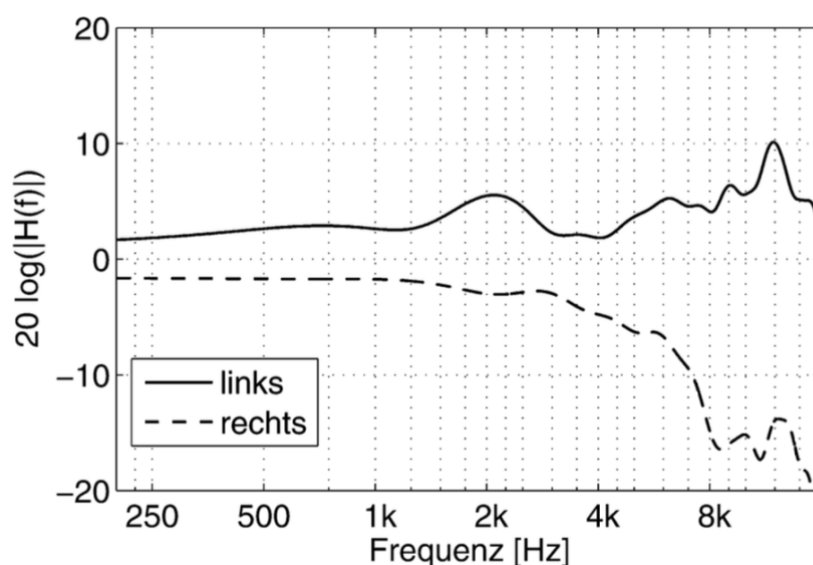


Abbildung 1 - Frequenzgänge Außenohr-Übertragungsfunktion, Azimut 90° links, Elevation 0° (Blauert & Braasch, 2008, S.91)

3 Binauralsynthese

Die Technik hinter binauralem Audio basiert auf den Prinzipien unserer Hörwahrnehmung mit nur zwei Ohrsignale und zielt darauf ab, diese zu reproduzieren (Hammerhøi & Møller, 2005). Zur Erzeugung binauraler Signale gibt es verschiedene Möglichkeiten. Eine einfache und intuitive Möglichkeit ist die direkte Aufnahme der Ohrsignale, entweder am „lebenden Objekt“ oder mithilfe eines Kunstkopfes (Zhang et al., 2017). Es gibt verschiedene Modelle von Kunstköpfen, allen voran wohl am bekanntesten der Neumann KU100 (siehe Abbildung 14, S.29). Sie bestehen im Wesentlichen aus einem Kopf mit zwei Außenohren, in denen sich Mikrofone befinden. In der Regel stellen sie den anatomischen Durchschnitt des menschlichen Kopfes dar (Weinzierl, 2008). Bei dieser binauralen Aufnahmetechnik werden die ohr- und raumspezifische Merkmale also direkt einer Aufnahme aufgeprägt (Hammershøi & Møller, 2005). Das bedeutet jedoch gleichzeitig, dass diese Aufnahmen statisch sind und somit in der Regel wenig kompatibel. Gleichmaßen werden keine dynamischen Cues, also Kopfbewegungen, berücksichtigt und es besteht keine Möglichkeit zur Nachbearbeitung (Weinzierl & Slavik, 2008). An dieser Stelle kommt die Binauralsynthese ins Spiel. Die Idee ist, die Ohrsignale durch digitale Signalverarbeitung zu reproduzieren (Hammershøi & Møller, 2005). Die Übertragung des Schalls von einem Punkt im Raum bis in den Gehörgang lässt sich in Form einer Übertragungsfunktion bzw. einer Impulsantwort beschreiben (Weinzierl & Slavik, 2008). Durch Faltung des Nutzsignals mit diesen Übertragungsfunktionen kann so ein binaurales Signal erzeugt werden, welches „(...) alle durch das Übertragungssystem induzierten auditiven Merkmale (enthält).“ (Weinzierl & Slavik, 2008, S.671). Die Außenohr-Übertragungsfunktion oder HRTF lässt sich messen und enthält sämtliche sogenannten statischen Cues, also ITDs, ILDs und ohrspezifische spektrale Besonderheiten (Hammershøi & Møller, 2005). Dynamische Cues können mittels Headtracking in die Synthese eingebunden werden (Weinzierl & Slavik, 2008). In weiterer Ausführung kann je nach Anwendung ein Raumeindruck direkt mit in die HRTF einfließen. Diese wird dann als binaurale Raumimpulsantwort (engl. binaural room impulse response, BRIR) bezeichnet. Die Wiedergabe von binauralem Audio erfolgt dann in der Regel über Kopfhörer. Die Wiedergabe von binauralisierten Signalen über Lautsprecher ist zwar theoretisch auch möglich, jedoch deutlich schwieriger zu realisieren (Weinzierl & Slavik, 2008).

3.1 Allgemeines Prinzip & Aufbau

Wie oben kurz erwähnt, ist das Prinzip in der Theorie relativ einfach. Ein Nutzsignal wird mit den HRTFs gefaltet, womit ein binaurales Signalpaar entsteht, das in der Regel über Kopfhörer wiedergegeben wird. Binaurales Rendering besteht also im Grunde aus einem Faltungsalgorithmus. In Abbildung 2 ist das gesamte System skizziert. Zunächst müssen die HRTFs gemessen werden. Dies muss an vielen verschiedenen Punkten im Raum geschehen, da HRTFs richtungsgebunden sind. Folglich wird also ein HRTF-Datensatz benötigt (Kapralos, 2003). Ebenso müssen die Rohdaten der HRTF-Messung von den

Einflüssen des Systems befreit werden, was ein Post-Processing erfordert (Weinzierl & Slavik, 2008). Im finalen Rendering müssen für Richtungen, die zwischen den Messpunkten liegen, die HRTFs interpoliert werden (Duraismami et al., 2004). „Richtungsabhängig“ bedeutet im Kontext des Renderings, dass es eine HRTF für eine Schalleinfallrichtung gibt. Bewegt sich der Kopf, so ändert sich diese Richtung und somit müssen die HRTFs dynamisch ausgetauscht werden. Dies kann mittels Headtracking in die Synthese eingebunden werden. Anschließend kann die Faltung erfolgen, wobei die Latenz nicht zu hoch sein darf. Der notwendige Rechenaufwand ist also auch aus Wahrnehmungssicht relevant. Bei der Wiedergabe über Kopfhörer ist schließlich die Begradigung der Frequenzgänge besonders wichtig (Weinzierl & Slavik, 2008). Damit ergeben sich die Punkte, welche in den folgenden Kapiteln behandelt werden. Noch einmal zusammengefasst sind das:

- Messung eines (individuellen) HRTF-Datensatzes & Post-Processing
- Kopfhörer-Entzerrung
- Headtracking & positionsabhängige Auswahl der HRTFs
- Faltung (in Echtzeit und latenzfrei)

Abschließend wird, mit Blick auf die Arbeit mit dem „Smyth Realiser“, zudem noch kurz auf die Thematik der binauralen Raumsimulation eingegangen.

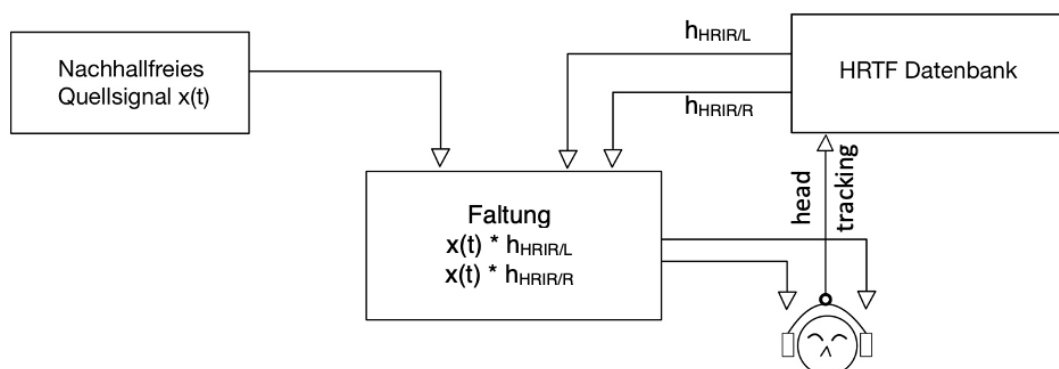


Abbildung 2 - Binaurales Rendering-System (in Anlehnung an: Weinzierl & Slavik, 2008, S.672)

3.2 Faltung & Entfaltung

In der Binauralsynthese ist das Prinzip der Faltung bzw. im weiteren Sinne das der Entfaltung von grundlegender Bedeutung. Dabei werden zwei Funktionen zu einer neuen Funktion kombiniert. Zur Einordnung ist es sinnvoll, das Ganze näherungsweise als lineares, zeitinvariantes System (engl. linear time-invariant, LTI) zu betrachten. Ein LTI-System definiert sich im Zeitbereich über die Gleichung

$$y(t) = x(t) * h(t)$$

mit $y(t)$ als Ausgabe des Systems, $x(t)$ als Nutz- bzw. Eingangssignal und $h(t)$ als Impulsantwort des Systems. Hier ist das Prinzip der Faltung dargestellt: Die Ausgabe ist die Faltung des Nutzsignals mit der Impulsantwort. Im Frequenzbereich steht dem folgende Gleichung gegenüber:

$$Y(f) = X(f) \cdot H(f)$$

Die Ausgabe des Systems ist hier das Produkt aus Eingabe und Übertragungsfunktion. Basierend auf Gleichung 2, lässt sich nun relativ einfach die Berechnung der Übertragungsfunktion formulieren:

$$H(f) = \frac{Y(f)}{X(f)}$$

Die Übertragungsfunktion eines LTI-Systems ist somit das Verhältnis von Ausgangssignal zu Eingangssignal. Man spricht hierbei auch von zirkulärer Dekonvolution (Farina, 2000). Die Impulsantwort erhält man dann durch Fouriertransformation von $H(f)$ in die Zeitdomäne. Alternativ kann man diese auch direkt berechnen:

$$h(t) = y(t) * IFT\left\{\frac{1}{X(f)}\right\} = y(t) * x_{inv}(t)$$

Dabei wird $y(t)$ mit der Umkehrfunktion des Eingangssignals gefaltet. Diese Berechnung bezeichnet man auch als lineare Dekonvolution. In Abbildung 3 sind die Verhältnisse zwischen den Domänen noch einmal graphisch dargestellt. Auf die Binauralsynthese angewendet bedeutet das: binaurale Signale werden durch die Faltung eines Nutzsignals mit einer Übertragungsfunktion, der HRTF, generiert. Diese kann wiederum durch lineare Entfaltung aus dem Eingangs- und Ausgangssignal eines (Mess-)Systems berechnet werden (Li & Peissig, 2020). Wie zu erkennen ist, ist die Übertragungsfunktion des Systems nicht direkt messbar. Es existieren lediglich ein Eingangssignal und ein Ausgangssignal, in welchem die Übertragungsfunktion jedoch enthalten ist und somit berechnet werden kann.

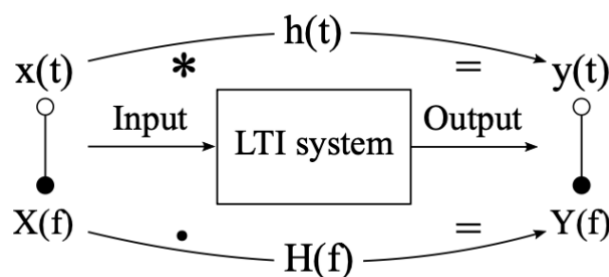


Abbildung 3 - LTI-System (Li & Peissig, 2020)

3.3 Head-Related Transfer Function (HRTF)

Es lässt sich festhalten, dass die Außenohr-Übertragungsfunktion das Grundgerüst der Binauralsynthese ist. Wie ihr Name schon sagt, ist es eine mathematische Funktion, welche die Veränderung des Schalls durch Ohr, Kopf und Torso beschreibt (Møller et al., 1995). Sie definieren sie wie folgt: „A head-related transfer function (HRTF) is a transfer function that, for a certain angle of incidence, describes the sound transmission from a free field to a point in the ear canal of a human subject.“ (Møller et al. 1995, S.300). Zusammengefasst ist eine HRTF also eine richtungsabhängige Übertragungsfunktion von einem Punkt im Freifeld zu einem Punkt im Gehörgang. Diese Übertragungsfunktion kann als Impulsantwort oder Frequenzgang dargestellt werden (vgl. Abbildung 4). Dementsprechend spricht man entweder von einer Außenohrübertragungsfunktion (HRTF, im Frequenzbereich) oder von einer kopfbezogenen Impulsantwort (engl. head-related impulse response, HRIR im Zeitbereich). Mittels Fouriertransformation kann sie zwischen den Domänen hin und wieder zurück überführt werden (Li & Peissig, 2020).

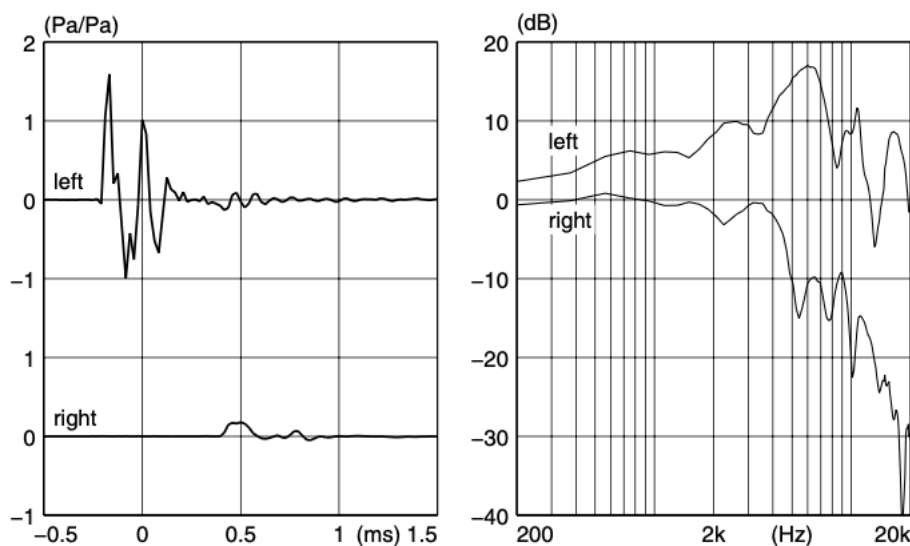


Abbildung 4 - Außenohrübertragungsfunktion als Impulsantwort & Frequenzgang (Hammerhøi & Møller, 2005, S.239)

Zur Generierung eines HRTF Datensatzes gibt es viele verschiedene Ansätze. Im Wesentlichen bestehen diese jedoch stets aus der eigentlichen Messung und der anschließenden Berechnung der HRTFs, die wiederum vom verwendeten Messsignal abhängig ist. Hier wird der Fokus auf der Messung mit Sweeps liegen. Klassischerweise findet die Messung in einem reflexionsarmen Raum statt. In diesem befindet sich ein Mensch oder ein Kunstkopf, welcher mit zwei Ohrmikrofonen ausgestattet ist. An einer bestimmten Position und mit bestimmtem Abstand steht ein Lautsprecher als Schallquelle. Dieser spielt ein Testsignal (= Eingabe) ab, welches das Übertragungssystem Schallquelle – Ohr anregt. Das Testsignal wird von den Ohrmikrofonen aufgenommen und damit die Ausgabe des Systems aufgezeichnet. Schließlich wird das HRTF-Paar für die entsprechende Richtung berechnet (vgl. Kapitel

3.2), welches im Anschluss wiederum einem Post-Processing unterzogen wird. Die wichtigsten Aspekte einer Messung sind also ganz allgemein (Li & Peissig, 2020):

- Methode
- Post-Processing
- Messaufbau

Die Methode beschreibt das genutzte Messsignal, sowie die Art der HRTF-Berechnung nach der eigentlichen Messung. Beispiele für Methoden und Messsignale sind Sweeps, Adaptive Filtering oder Pseudo-Random Sequences. Das Post-Processing ist notwendig, um Beschränkungen des Messsystems zu korrigieren. In etwa werden hier die Frequenzgänge der verwendeten Lautsprecher kompensiert. Der Messaufbau bezieht sich auf das verwendete Setup, also wo gemessen wurde und in welcher Konfiguration. Beispielsweise mit einem Einzel- oder Multi-Lautsprecher-Setup. Ebenso ist hier die Positionierung der Mikrofone von Bedeutung.

3.3.1 Messmethodik

Die Messung von HRTFs mit Sweep-Signalen basiert auf dem in Kapitel 3.2 beschriebenen Prinzip der linearen Entfaltung und stellt eine ebenso klassische wie etablierte Vorgehensweise dar. Deshalb und mit Blick auf den praktischen Teil, ist es sinnvoll, im Folgenden beispielhaft die Messung mit Sweeps zu erläutern, da der Realiser ebenfalls mit diesen arbeitet. Dabei wird erklärt, welche Sweep-Arten es gibt, wie sie generiert werden und wie die Berechnung der HRTFs funktioniert.

Es gibt verschiedene Arten von Sweeps und für diese wiederum unterschiedliche Generierungsmethoden. Für die Messung von HRTFs sind besonders lineare Sweeps und exponentielle Sweeps von Vorteil. Lineare Sweeps haben ein weißes Frequenzspektrum. Ihre Frequenz erhöht sich über die Zeit um einen konstanten Wert. Exponentielle oder auch logarithmische Sweeps besitzen ein pinkes Frequenzspektrum. Das bedeutet, dass ihre Amplitude um 3dB pro Oktave sinkt, wodurch die Energieverteilung über das gesamte Spektrum konstant bleibt. Ihre Frequenz erhöht sich oktavenweise bzw. immer um einen Teil einer Oktave über die Zeit. Das heißt je höher man im Spektrum kommt, desto schneller steigt die Frequenz an. Beide Sweep-Arten können sowohl im Zeit- als auch im Frequenzbereich generiert werden. Im Zeitbereich ist die Generierung relativ einfach. Hierzu wird nach jeder Berechnung eines Samples die Phase einer Sinusfunktion um den entsprechenden Betrag erhöht. Im Falle des linearen Sweeps wird ein fester Wert zur Phase addiert und im Falle des logarithmischen Sweeps wird die Phase mit einem festen Wert multipliziert. Die Generierung im Frequenzbereich ist etwas komplizierter. Zunächst müssen hierzu Magnitude und Gruppenlaufzeit definiert werden.

Anhand dessen können Real- und Imaginärteil des komplexen Spektrums berechnet werden und mittels Fouriertransformation in den Zeitbereich überführt werden (Müller & Massarani, 2001).

Nach der Aufzeichnung der Ohrsignale muss die „rohe“ HRTF zunächst noch aus der Systemausgabe extrahiert werden. Die entsprechende Berechnung ist dabei abhängig vom verwendeten Messsignal. In der Praxis bietet das Prinzip der linearen Dekonvolution in Kombination mit Sweeps große Vorteile. Einerseits ist es sehr einfach die Umkehrfunktion eines Sweeps zu bilden. Diese ist nämlich lediglich der zeitlich invertierte Sweep, wobei bei logarithmischen Sweeps zusätzlich das Amplitudenspektrum verändert wird. Andererseits bietet diese Methode die Möglichkeit, harmonische Verzerrungen, welche durch nicht-lineares Verhalten des Messsystems entstehen, auszuklammern. Denn wie eingangs erwähnt, ist die Betrachtung als LTI-System lediglich eine Annäherung. Die Messung mit Sweeps und die Anwendung dieser linearen Entfaltung (engl. linear deconvolution) kompensiert dies in Teilen (Farina, 2000). Die Methode der Division bzw. die zirkuläre Dekonvolution (vgl. Kapitel 3.2) wird unter anderem von Møller et al. (1995) in ihrer Arbeit „Head-related transfer functions of human subjects“ verwendet. Zwar wurde dabei nicht mit Sweeps gearbeitet, jedoch soll auch diese Berechnungsmethode hier eine Erwähnung finden. Møller et al. berechnen dabei nicht das direkte Verhältnis von Ausgangs- zu Eingangssignal, sondern nutzen stattdessen ein Referenzsignal, welches an der Position des Kopfes gemessen wird. Auf diese Weise werden die Einflüsse des Messsystems bereits bei der Berechnung der HRTF ausgeklammert. Dies stellt bereits eine Überschneidung mit dem Post-Processing dar, welches im folgenden Abschnitt genauer ausgeführt wird.

3.3.2 Post-Processing

Eine HRTF ist direkt nach der Messung nur bedingt verwendbar. Um einen nutzbaren Datensatz oder eine Datenbank zu generieren, müssen die Rohdaten noch weiterverarbeitet werden. Grund dafür sind vor allem Beschränkungen durch das Messsystem. Von Interesse sind dabei insbesondere (Li & Peissig, 2020):

- Entfernung der systembedingten Einflüsse auf den Frequenzgang
- Entfernung von Reflexionen
- Neumodellierung des tieffrequenten Bereichs.

Wie bereits im vorherigen Absatz erwähnt, ist es möglich, die Einflüsse des Messsystems auf den Frequenzgang bereits bei der Berechnung der HRTFs zu neutralisieren. Dies ist nötig, da die einzelnen Komponenten – Verstärker, AD/DA-Wandler, Lautsprecher und Mikrofone – ihre individuellen Frequenzgänge der Messung aufprägen. Somit liegt also nativ kein „neutraler“ Datensatz vor. Neben der eigentlichen HRTF-Messung muss daher auch das hierfür benötigte Referenzsignal, die akustische

Übertragungsfunktion zwischen Systemeingang und -ausgang, gemessen werden (Hammershøi & Møller, 2005). Dazu kann entweder eine Messung in Abwesenheit der Testperson durchgeführt werden, bei der die Mikrofone an der Stelle des Kopfmittelpunkts positioniert werden (Møller et al., 1995), oder die Übertragungsfunktionen der Messkomponenten werden einzeln gemessen (Li & Peissig, 2020). Die Berechnung ist in beiden Fällen wie oben beschrieben. Neben Møller et al. wurde diese Methode auch zur Erstellung der Cipic-Datenbank genutzt (Algazi et al., 2002). Dabei spricht man von einer Freifeldentzerrung der HRTFs. Neben der Freifeldentzerrung gibt es die Diffusfeldentzerrung. Auch hier werden die Einflüsse des Messsystems entfernt, allerdings wird die Referenzfunktion dabei über alle Richtungen gemittelt. Das Ergebnis ist ein HRTF-Satz, bei dem der Einfluss des Diffusschalls auf die Hörwahrnehmung berücksichtigt wird (Kapralos, 2003). Auf den Punkt gebracht, ist eine freifeldentzerrte HRTF also eine „rohe“ HRTF, welche nur die Veränderung des Schalls durch die Anatomie beschreibt („wie sie unter Freifeld-Bedingungen auftritt“). Eine diffusfeldentzerrte HRTF integriert zusätzlich eine Gewichtung des Frequenzgangs hinsichtlich des diffusen Schallfeldes. Letztere entspricht somit eher der natürlichen Wahrnehmung („wie sie unter Diffusfeld-Bedingungen auftritt“).

Ein weiterer wichtiger Punkt des Post-Processings ist die Fensterung. Dabei wird die HRTF im Zeitbereich in ihrer Länge eingeschränkt. Das Ziel dabei ist, frühe Reflexionen, beispielsweise vom Messequipment, zu eliminieren (Algazi et al. 2002). Ebenso werden die On-Set-Delays, also in etwa die Laufzeiten vom Lautsprecher zu den Mikrofonen ausgeglichen (Møller et al., 1995). Zudem wird damit die Tap-Länge der HRTF verringert (Armstrong et al., 2018). Dies kann unter Umständen zu einer Beschneidung des LF-Bereichs führen, welcher jedoch später ohnehin nachmodelliert wird (Li & Peissig, 2020).

Die tiefen Frequenzen müssen aus verschiedenen weiteren Gründen neu modelliert werden. Ausschlaggebend ist vor allem die Größe der verwendeten Lautsprecher und die damit einhergehende unzureichende Basswiedergabe. Zudem hat jeder reflexionsarme Raum ein Dämpfungslimit, also eine Cut-Off-Frequenz, im tieffrequenten Bereich welche von der Länge der beim Bau verwendeten Absorber abhängt. Grundsätzlich gibt es auch dabei wieder viele verschiedene Herangehensweisen (Li & Peissig, 2020), hier wird jedoch ebenfalls die Vorgehensweise von Armstrong et al. (2018) beispielhaft beschrieben. Da die Anatomie von Kopf, Ohren und Torso kaum Einfluss auf die tiefen Frequenzen nimmt, kann der LF-Bereich mit linearem Frequenz- und Phasengang nachgebildet werden. Zunächst wird eine Crossover-Frequenz zwischen dem LF-Modell und der HRTF gewählt. Anschließend wird ein Impuls generiert, dessen Amplitude und Delay der durchschnittlichen Gruppenlaufzeit und Amplitude der HRTF um die Crossover-Frequenz entsprechen. Schließlich wird der Impuls zeitlich verschoben, bis

die Phasen an der Crossover-Frequenz übereinander liegen, und mit High-Order Filtern der Fade zwischen Modell und HRTF realisiert. (Armstrong et al., 2018).

3.3.3 Messaufbau & Mikrofonposition

Bei den Messaufbauten unterscheidet man klassischerweise zwischen Multi- und Einzel-Lautsprecher-Setups. Daneben besteht die Möglichkeit der reziproken Messung mithilfe eines Mikrofonarrays und In-Ear-Lautsprechern. Multi-Lautsprecher-Setups gibt es entweder als omnidirektionales Setup oder als Lautsprecherring. Bei Ersterem wird eine volle Kugel an Lautsprechern um den Messpunkt herum gebildet, während ein Lautsprecherring entweder auf horizontaler oder vertikaler Achse aufgebaut wird. Um die HRTFs aus verschiedenen Richtungen messen zu können, muss entweder der Ring um den Messpunkt drehbar aufgehängt sein oder das Messsubjekt muss sich auf einer drehbaren Platte befinden. Auch bei omnidirektionalen Setups ist es in der Regel notwendig das Subjekt zu rotieren, um hochauflösende HRTF-Sätze messen zu können. Einzel-Lautsprecher-Setups finden vor allem bei Kunstkopf-HRTF-Messungen Anwendung, können jedoch auch bei echten Menschen verwendet werden. Gerade bei einer Kunstkopfmessung ist es einfach möglich, den Lautsprecher statisch zu belassen und den Kopf drehbar zu verbauen. Am Menschen würde man den Kopf bewegen und die Position mit Headtracking aufzeichnen. Eine weitere Möglichkeit besteht darin, den Lautsprecher auf einer vertikalen Schiene entlangfahren zu lassen und zusätzlich das Messsubjekt zu rotieren. Grundsätzlich hängt die Wahl des Setups oft von der Anwendung ab. Für Messungen an Kunstköpfen bietet es sich beispielsweise an ein Einzel-Setup zu wählen, da der Aufwand gering ist und die Dauer der Messung meist keine Rolle spielt. Multi-Lautsprecher-Setups verkürzen die Messdauer, was sie für Messungen am Menschen sinnvoll macht. Abgesehen von diesen „direkten“ Messmethoden, hat die reziproke Messung den Vorteil, dass alle Richtungen gleichzeitig gemessen werden können wodurch der Vorgang ebenfalls sehr zeiteffektiv ist. Gleichzeitig wird die Möglichkeit von Reflexionen am Messequipment minimiert. Nachteilig bei dieser Methode sind jedoch die Basswiedergabe und die Lautstärke der kleinen Lautsprecher (Li & Peissig, 2020). In Abbildung 5 werden einige Setups beispielhaft aufgezeigt.

Bei der Verwendung von Ohrmikrofonen spielt die Positionierung eine wichtige Rolle. Etabliert hat sich dabei die Position am verschlossenen Eingang des Gehörkanals. Wie von Hammershøi und Møller (2005) beschrieben, lässt sich die Schallübertragung von einer Quelle zum Ohr in einen richtungsabhängigen und einen richtungsunabhängigen Teil zerlegen. Dabei ist die Übertragung vom Eingang des Gehörkanals zum Trommelfell der richtungsunabhängige Teil, womit in einer HRTF, welche am geschlossenen Gehörgang gemessen wird, bereits alle räumlichen Informationen enthalten sind. Dies hat enorme praktische Vorteile: es muss kein Mikrofon in den Gehörgang eingeführt werden und

es können größere Mikrofone verwendet werden. Darüber hinaus konnte gezeigt werden, dass die Messung am geschlossenen Gehörkanal eine geringere Varianz zwischen verschiedenen Testpersonen aufweist. Diese Positionierung sorgt also auch für eine allgemein bessere Verwendbarkeit und Kompatibilität der HRTFs. In Abbildung 6 ist dies dargestellt. Hier wurden die HRTFs des linken Ohres von 12 Personen gemessen, jeweils direkt am Trommelfell, am offenen Gehörgang und am geschlossenen Gehörgang. Es ist zu erkennen, dass die HRTFs in den ersten beiden Fällen deutlich mehr Unterschiede aufweisen als die HRTFs, welche am geschlossenen Gehörgang gemessen wurden. Dies betrifft vor allem den Bereich der hohen Frequenzen. Die unteren drei Grafiken zeigen die Standardabweichung der HRTFs und bestätigen dieses Bild. Ihr Wert ist gerade in den hohen Frequenzen am Trommelfell und am offenen Gehörgang sehr hoch.

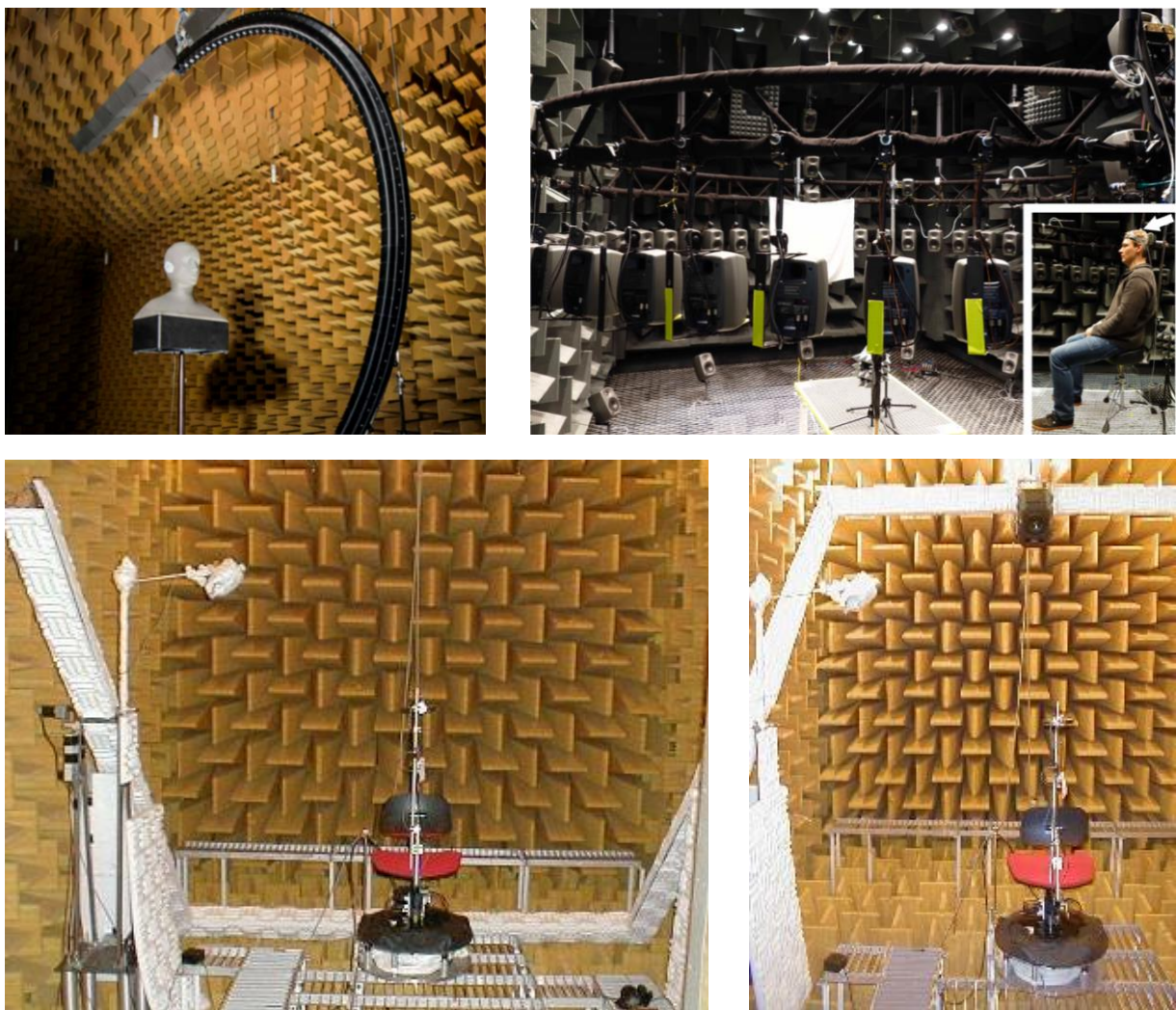


Abbildung 5 - Multi-Lautsprecher-Setups: Aachen HRTF Database (oben links) (Bomhardt et al., 2016, S.3), Oldenburg HRTF Database (oben rechts) (Denk et al., 2018, S.4) & Einzel-Lautsprecher-Setup: Ircam Listen Datenbank (unten) (IRCAM, o.D.)

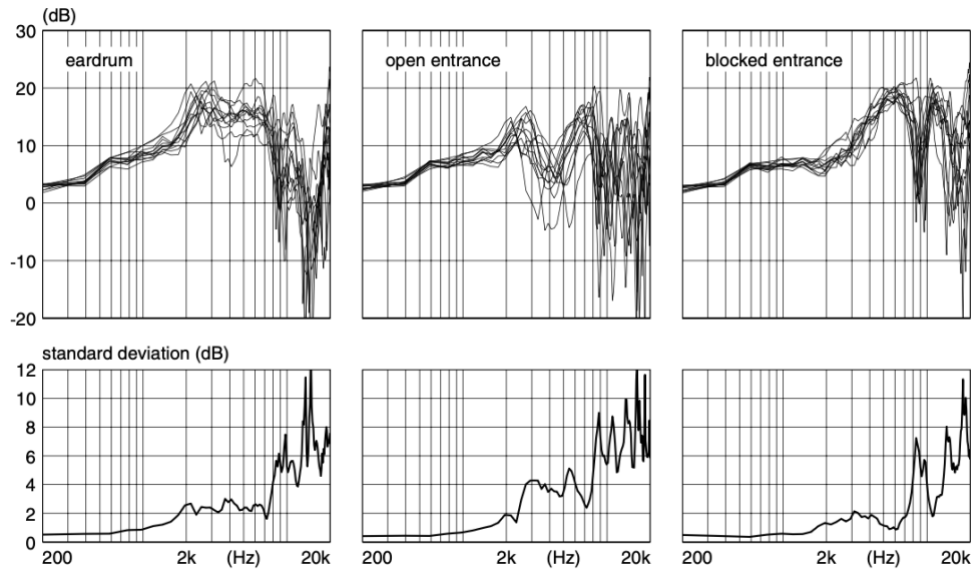


Abbildung 6 - HRTFs des linken Ohrs gemessen am Trommelfell, offenen und geschlossenen Gehörgang, unten: frequenzweise berechnete Standardabweichung (Hammershøi & Møller, 2005)

3.4 Kopfhörerentzerrung

Binaurales Audio wird in der Regel über Kopfhörer wiedergegeben. Dabei ist klar, dass Kopfhörer meist keinen flachen Frequenzgang haben, sondern stets einen individuellen bzw. vom Hersteller implementierten. Außerdem konnte gezeigt werden, dass die Übertragungsfunktion eines Kopfhörers von Mensch zu Mensch teils stark variiert (siehe Abbildung 7) und sich durch wiederholtes auf- und absetzen zusätzlich verändert. Daher ist eine Entzerrung erforderlich. Dabei wird die Übertragungsfunktion des Kopfhörers gemessen und eine passende inverse Funktion erstellt. Mittels Dekonvolution werden beide Funktionen miteinander verrechnet, wodurch der Frequenzgang begradigt wird. Durch individuelle Messungen der Übertragungsfunktion des Kopfhörers kann die Varianz zwischen mehreren seiner Nutzern berücksichtigt werden. Unterschiede durch Auf- und Absetzen können durch eine Wiederholung der Messung und eine Mittelung der Übertragungsfunktion ausgeglichen werden. (Hammershøi & Møller, 2005)

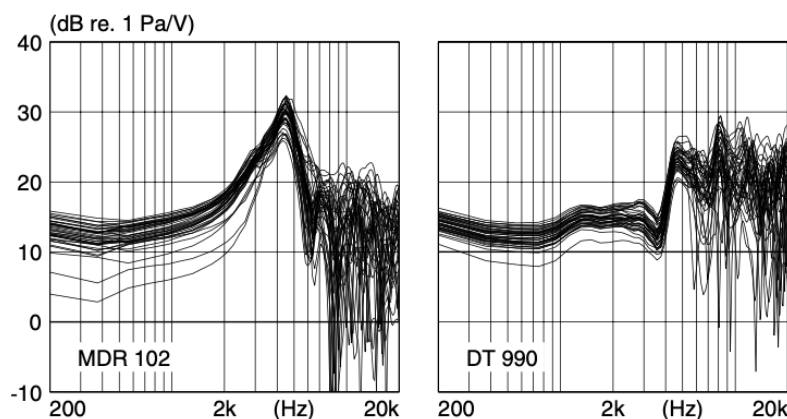


Abbildung 7 - Kopfhörer-Übertragungsfunktionen: Sony MDR-102 (links), BeyerDynamic DT-990 (rechts), gemessen an 40 Personen (Hammershøi & Møller, 2005, S.231)

3.5 Dynamische Binauralsynthese: Headtracking

Wie Hammershøi und Møller (2005) schreiben, kann durch Binauralsynthese an sich schon eine sehr gute Nachbildung des räumlichen Hörens erreicht werden. Einschlägig verbessert wird diese jedoch durch die Einbindung der Kopfbewegung. Die dynamische Binauralsynthese umfasst im Großen und Ganzen die bisher erläuterten Prinzipien und Konzepte, erweitert diese jedoch um das sogenannte Headtracking. Dabei wird die Orientierung des Kopfes erfasst und an den Renderer übermittelt, welcher dann die HRTF entsprechend der sich ändernden Schalleinfallrichtung auswählt (Hammershøi & Møller, 2005).

Kopfbewegungen im Kontext des räumlichen Hörens können verschiedene Hintergründe haben. Sie können unbewusst, hin zu einem Hörereignis erfolgen. Dadurch verringert sich die Lokalisationsunschärfe, allerdings muss bereits eine relativ genaue Ortung des Hörereignisses erfolgt sein. Dem gegenüber stehen bewusste Peilbewegungen, die darauf abzielen, ein Hörereignis genauer zu lokalisieren. Generell wird die Lokalisation durch Kopfbewegungen verbessert. Dabei ändern sich die Ohrsignale und ein Hörer ist durch diese charakteristische Veränderung in der Lage, die Ortung zu verbessern. In die Analyse fließen zudem die Informationen über die Kopfposition, -orientierung und -bewegung mit ein. Ohne die Möglichkeit den Kopf zu bewegen, kommt es zu einigen unerwünschten Effekten. Dazu zählen Front-Back-Confusion, bei der ein Hörer nicht sagen kann, ob ein Hörereignis vor oder hinter dem Kopf ist, sowie die Im-Kopf-Lokalisation (Blauert, 1974/1996).

Um eine wirklich „naturgetreue“ Binauralsynthese umzusetzen, ist Headtracking also unbedingt notwendig. Headtracker arbeiten mit verschiedenen Prinzipien. So gibt es Systeme, die mit Ultraschall arbeiten, Systeme auf Basis von Trägheitssensoren – also Gyroskop, Kompass und Beschleunigungssensoren – oder Systeme, welche die Bewegung mit einer oder mehreren Kameras erfassen. Je nach Implementierung sind dabei oft volle 6-Degree-of-Freedom möglich. (Hess, 2012)

3.6 Binaurale Raumsynthese

Mit Blick auf den praktischen Teil dieser Arbeit darf das Thema der binauralen Raumsimulation nicht unerwähnt bleiben. Dabei werden mithilfe der Binauralsynthese die akustischen Eigenschaften eines Raumes für Kopfhörer reproduziert, wozu die entsprechenden Rauminformationen in die Synthese integriert werden müssen. Dies erfolgt durch sogenannte binaurale Raumimpulsantworten (engl. binaural room impulse response, BRIR). Eine BRIR ist im Grunde eine HRTF, welche zusätzlich die Reflexionen des Raumes enthält. Diese können entweder berechnet oder gemessen werden. Eine Messung findet dementsprechend in dem zu simulierenden Raum statt (Kapralos, 2003).

Das Post-Processing ist an die neuen Anforderungen angepasst. So werden lediglich die Frequenzgänge der Mikrofone und der Kopfhörer kompensiert, die Amplituden werden normalisiert und es wird eine Fensterung vorgenommen (Lindau, 2014). Die Frequenzgänge der Lautsprecher sollen beispielsweise, gerade im Falle eines Studios oder Kinosaals, enthalten bleiben, sodass die gesamte Signalkette mit abgebildet wird. Ein besonderes Augenmerk liegt hier auf der Fensterung. Man könnte annehmen, dass die BRIR in ihrer Länge der Nachhallzeit des Raumes entsprechen muss. Es hat sich jedoch gezeigt, dass lediglich die frühen Reflexionen für den richtigen Raumeindruck respektive die korrekte Ortung von Hörereignissen wichtig sind (Zotkin et al., 2004). Somit kann die Länge der BRIR entsprechend reduziert werden. Dies ist entscheidend für das finale Rendering, da es hier ausreicht, die gekürzte BRIR, welche nur den Direktschall und die frühen Reflexionen enthält, mit dem Nutzsignal zu falten. Das Gleiche gilt bei der Auswahl der passenden BRIR, falls Headtracking verwendet wird (Lindau, 2014). Für den späten Teil der Reflexionen, also den diffusen Anteil, ist es ausreichend, diesen einmal zu berechnen und dem Signal hinzuzufügen (Zotkin et al., 2004). Angewandt auf die Simulation einer 3D-Regie bedeutet das nun also Folgendes: Jeder Lautsprecher kann als virtuelle Schallquelle wiedergegeben werden. Die am Renderer ankommenden Signale werden dann entsprechend ihrer Positionierung mit dem zugehörigen Filter, also der BRIR, gefaltet und über Kopfhörer abgespielt.

Bei dieser Form der Binauralsynthese gibt es diverse weiterführende Effekte, die die Wahrnehmung beeinflussen. Vor allem der sogenannte „Room Divergence Effect“ ist dabei von Bedeutung. Dieser beschreibt die Veränderung der Schallwahrnehmung, wenn der simulierte Raum und der Wiedergaberaum nicht übereinstimmen. So wurde festgestellt, dass der Grad der Externalisierung abnimmt, wenn dies der Fall ist. Auch das Vorhandensein von visuellen Cues, also zu Sehen wo sich beispielsweise eine Schallquelle befindet, spielt eine Rolle. Die Sichtbarkeit der Lautsprecher in etwa erhöht den Grad der Externalisierung um bis zu 16%. (Werner et al., 2016).

4 Praktische Arbeit: Ausgangslage & Geplantes Vorgehen

Im Zuge der Sanierungsarbeiten an der Hochschule der Medien (HdM) werden auch die Tonstudios im Untergeschoss umgebaut. Neben den Räumlichkeiten in U54 wurde dabei auch der Raum U48 saniert. Zusätzlich zu einem 3D-Lautsprechermonitoring besteht hier die Möglichkeit der Simulation dieser Lautsprecher mit dem Smyth Realiser A16. Aufgrund des Umbauten müssen die hierfür erforderlichen Messungen neu durchgeführt werden.

Die Bearbeitung des praktischen Teiles verläuft wie folgt. Zunächst findet eine Einarbeitungsphase in Raum U48 statt, in der der Workflow mit dem Realiser und der neuen Regie erarbeitet werden soll. Um ein Gefühl für die Thematik zu erlangen, werden Testmessungen durchgeführt, welche am Neumann-Kunstkopf KU100 und im Eigenversuch erfolgen. Zudem gilt es, eventuelle Probleme bei der Einbindung des Realisers in das neue Regie-Setup zu lösen, etwa das eines passenden Routings. Das Ziel dieser Phase ist es, herauszufinden, wie ein möglichst sauberes Ergebnis gelingen kann. Schließlich erfolgt die Hauptmessung am Kunstkopf und damit die Erstellung des finalen Produkts dieser Arbeit. Die Konfiguration des Realisers wird mit seiner Einrichtung zur binauralen Wiedergabe von Audiomaterial abgeschlossen. Priorität hat dabei die möglichst einfache Anwendbarkeit für die Studierenden. Die Evaluierung der Messung ist durch einen Vergleich der BRIRs mit den tatsächlichen Raumimpulsantworten (RIR) geplant. Diese werden hierzu mit dem Programm RoomEQWizard (REW) gemessen und anschließend einander gegenübergestellt. Dabei sollen eventuelle Ungereimtheiten und Abweichungen der Virtualisierung sichtbar gemacht werden.

In den folgenden Kapitel findet sich zunächst ein kurzer Überblick über den Raum U48, sowie den Realiser und den Neumann KU100, bevor sich der anschließende Teil vorerst allein mit der Messthematik beschäftigt. Dabei wird geklärt, welche Einstellungen getätigt werden müssen, wie die Messungen ablaufen und wie im konkreten Fall vorgegangen wurde. Anschließend werden die nötigen Einstellungen und die konkrete Konfiguration für die Wiedergabe beschrieben. Zusätzlich soll nachfolgend der Frage nachgegangen werden, inwieweit es sinnvoll möglich ist, mehrere Räume, insbesondere die der HdM, mit dem Realiser zu virtualisieren. Zuletzt erfolgt die Erläuterung der Evaluierung mit RoomEQWizard. Dabei werden das Vorgehen und die Einrichtung der Messungen erklärt und anschließend die Ergebnisse ausgewertet. Ganz allgemein werden Besonderheiten, welche sich durch die Arbeit in U48 ergeben, betont beschrieben.

4.1 Raum U48

Der Raum U48 ist die Filmtone- und 3D-Regie der HdM. Insgesamt sind hier 20 Genelec-Lautsprecher verbaut. Neun davon befinden sich auf Höhe der Ohren, fünf sind jeweils nach oben versetzt bzw. an der Decke angebracht. Ebenfalls ist ein LFE-Subwoofer vorhanden. Es sind also verschiedene Abhörkombinationen wie 7.1.4 oder 9.1.4 möglich. Die Positionen, die Kanalbelegung und die verwendeten Lautsprecher sind in Tabelle 1 noch einmal dargestellt. Insgesamt ließe sich das Setup somit als „14.1.5“ beschreiben. Das Routing und der Signalfluss erfolgen über ein Dante-Netzwerk. Die Quellsignale laufen zunächst über eine Matrix/Kreuzschiene, welche mit der Dadman-Software gesteuert wird. Hier können verschiedene Quellen ausgewählt werden. Von dort aus werden die Signale durch den Trinnov-Prozessor geleitet, welcher für die Lautsprecherentzerrung und das Bassmanagement zuständig ist. Seine Preset-Kurve ist in Abbildung 8 dargestellt. Dieser leitet die Signale weiter auf die Kopfhörerverstärker und das Interface (IF), welches ebenfalls mit Dadman gesteuert werden kann. Die Signale können von drei verschiedenen Quellen kommen. Einem DAW-Rechner, einem RMU-Rechner und einem frei belegbaren Slot im Dante-Routing für externe Geräte wie beispielsweise private Rechner. Wie in Abbildung 9 dargestellt, erfolgt die Steuerung der DAWs getrennt von der Steuerung des übrigen Signal-Processings. Der RMU-Rechner steuert also nur den Trinnov-Prozessor, sowie den DA-Wandler und die Kreuzschiene. Hier ist ebenfalls externes Atmos-Rendering möglich.

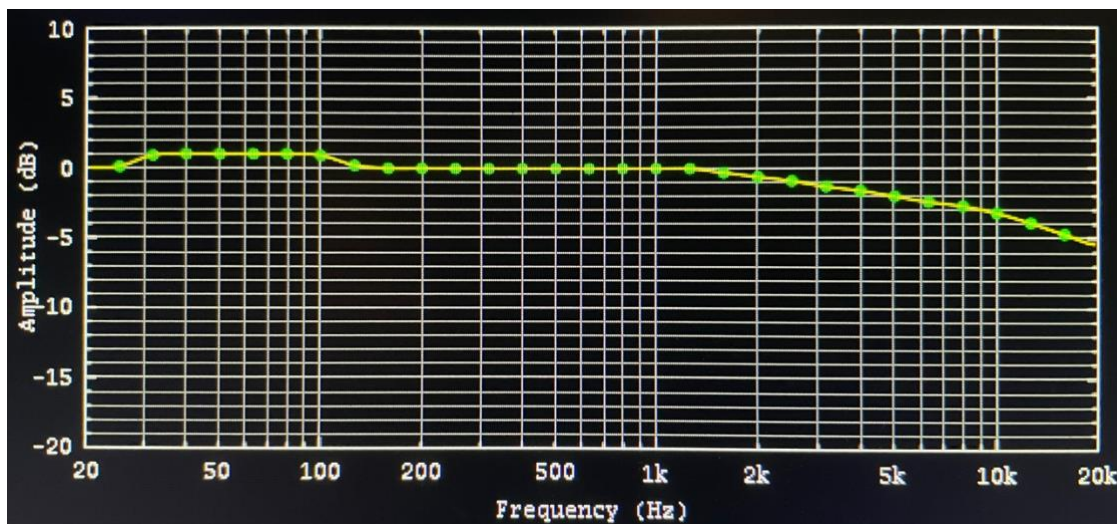


Abbildung 8 - Preset-Kurve des Trinnov Prozessor

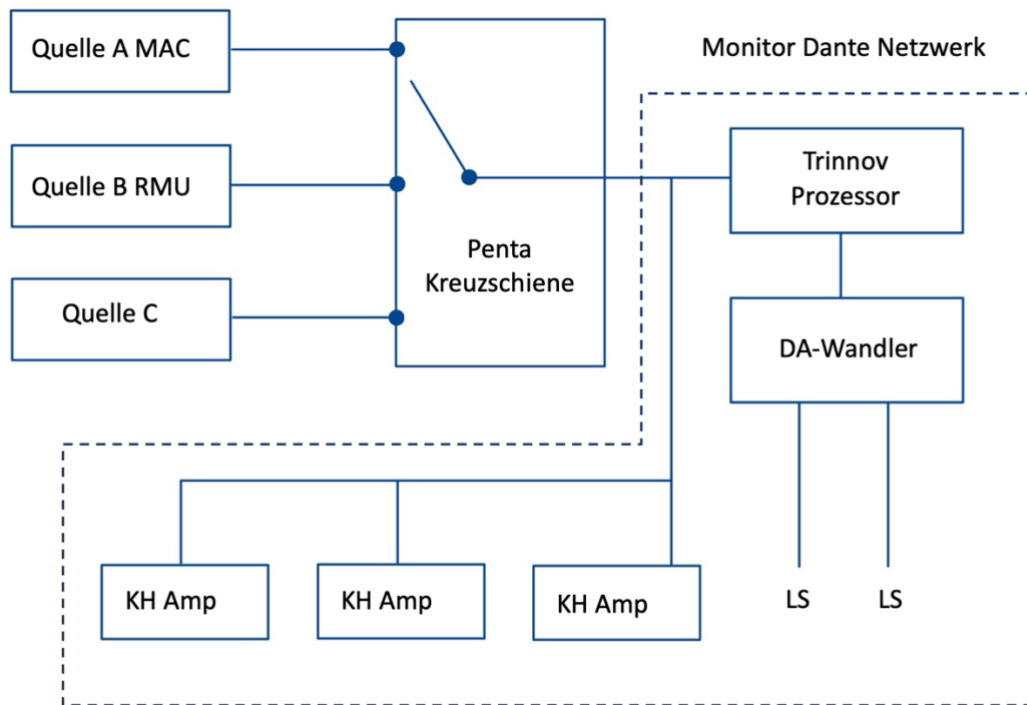


Abbildung 9 - Signalfluss in U48

Kanal	Bezeichnung	Idealisierte Position (Az.,Ele.)	Layer	Lautsprecher
1	L	30,0	Middle	Genelec 8341 A
2	R	-30,0	Middle	Genelec 8341 A
3	C	0,0	Middle	Genelec 8341 A
4	LFE	xx	Middle	Genelec 7370
5	Ls	114,0	Middle	Genelec 8341 A
6	Rs	-114,0	Middle	Genelec 8341 A
7	Lr	150	Middle	Genelec 8341 A
8	Rr	-150	Middle	Genelec 8341 A
9	Ltf Tfl	45,45	Top	Genelec 8331 A
10	Rtf Tfr	-45,45	Top	Genelec 8331 A
11	Ltr Tbl	135,45	Top	Genelec 8331 A
12	Rtr Tbr	-135,45	Top	Genelec 8331 A
13	Lw	65,0	Middle	Genelec 8341 A
14	Rw	-65,0	Middle	Genelec 8341 A
15	Lh	34,30	Upper	Genelec 8040 B
16	Rh	-34,30	Upper	Genelec 8040 B
17	Ch	0,30	Upper	Genelec 8040 B
18	Lsh	135,30	Upper	Genelec 8040 B

19	Rsh	-135,30	Upper	Genelec 8040 B
20	TopCenter	0,90	Top	Genelec 8331 A

Tabelle 1 - Kanalbelegung, Positionen & Lautsprecher in U48

4.2 Neumann KU100

Der KU100 ist die dritte Generation der Kunstkopf-Mikrofone von Neumann. In ihm sind zwei Kondensator-Mikrofonkapseln mit Kugelcharakteristik verbaut. Zudem lässt sich eine 10dB-Dämpfung einschalten, sowie ein Hochpassfilter bei 40 Hz oder 150 Hz. Der Kunstkopf ist diffusfeldentzerzt und somit sind mit ihm gemachte Aufnahmen gut lautsprecherkompatibel (Georg Neumann GmbH, 2020).

4.3 Smyth Realiser A16

Der Realiser A16 von der Firma Smyth Research ist ein Gerät zur binauralen Reproduktion von mehrkanaligem Audio. Der Fokus liegt dabei auf der Wiedergabe von Audiomaterial über ein binaurale Lautsprechersimulation, also ein virtuelles Lautsprecher-Setup. Das Gerät ist somit in erster Linie ein binauraler Renderer. Es implementiert jedoch ebenso die Möglichkeit eigene HRTFs bzw. BRIRs zu messen, sprich selbst ein Lautsprecher-Setup zu virtualisieren. Der zugrundeliegende Rendering-Algorithmus wurde von Smyth Research entwickelt und heißt „Smyth Virtual Surround“ (SVS).

Der Realiser ist entweder als 2U Rack-Version oder, wie an der HdM, als „Headstand“-Version (Abbildung 10) erhältlich. Mitgeliefert werden zudem eine Fernbedienung, zwei Messmikrofone mit verschiedenen Ohreinsätzen und ein Erdungskabel, sowie ein Headtracker mit „Set-Top“ als Referenzeinheit.

In der „Headstand“-Version sind standardmäßig 16 analoge In- und Outputs verbaut, wobei die Rack-Version auch mit AES3- oder Dante-Schnittstelle erworben werden kann. Stets verbaut sind mehrere HDMI-Inputs, ein SPDIF-Eingang und ein USB-2.0-Input. Für die Messung von „Personal Room Impulse Responses“ (PRIR²), wie Smyth Research es nennt, sind zudem Eingänge für die Messmikrofone und das Erdungskabel verbaut. Zwei Kopfhörerbuchsen sind für die binaurale Wiedergabe von Multichannel-Audio vorhanden.

Neben dem binauralen Rendering von eingespeistem 16-kanaligem Audio, kann der Realiser Bitstreams zweier NGA-Formate („Next Generation Audio“) dekodieren: Dolby Atmos und DTS:X. In Atmos sind im „Dual User Mode“ ebenfalls bis zu 16 Kanäle renderbar, im „Single User Mode“ überdies bis zu 24 Kanäle. Bei DTS:X werden bis zu 12 Kanäle unterstützt.

² Die Begriffe PRIR und BRIR werden im Folgenden meist synonym verwendet.

Mehrkanaliges Audio wird also analog oder digital in den Realiser eingespeist, wo es vom integrierten SVS-Renderer binaural aufbereitet und über Kopfhörer wiedergegeben wird. Der Realiser kann jedoch auch als reiner Decoder für Dolby Atmos oder DTS:X verwendet werden, wobei das Audio über Lautsprecher wiedergegeben wird.

Für die Messung eigener BRIRs implementiert der Realiser einige Programmroutinen, unter anderem zur Kalibrierung der Lautsprecher, zur Messung der BRIRs und zur Kopfhörerentzerrung (HPEQ). Diese lassen sich mit wenigen Knopfdrücken bedienen.

Für die Wiedergabe können Presets für mehrere Nutzer konfiguriert werden, die Einstellungen für verschiedene Abhörkonfigurationen in den jeweiligen Formaten (Atmos, DTS:X, PCM) zulassen (vgl. Kapitel 6.1).

Die Bedienung sämtlicher Funktionen und Einstellungen erfolgt über den integrierten Bildschirm mit der mitgelieferten Fernbedienung (Smyth, 2021). Das Hauptmenü gliedert sich in folgende Punkte:

- Presets (für User A & B): Einstellung der Abhörpresets für die Userprofile (Kapitel 6.2)
- Audio Source: Einstellung des Quelleingangs
- Listening Rooms: Konfiguration der Listening Rooms (vgl. Kapitel 6.1)
- Apps: Programmroutinen für die Messung von „PRIRs“ (vgl. Kapitel 5.3.3 & 5.4)
- Files: PRIR- und HPEQ-Dateisystem
- Settings: Einstellungen insb. des Sound Rooms, Teile der Messeinstellungen, User-Namen und Kopfhörer-Namen (vgl. Kapitel 5.3.1 & 5.3.3)



Abbildung 10 - Smyth Realiser A16 in der Headstand-Version

5 Messung

5.1 Funktionsprinzip des Realisers

Für die nachfolgenden Kapitel ist es von Vorteil zunächst das Funktionsprinzip des Realisers kurz zu erläutern. Bei einer Messung generiert der Realiser ein File mit binauralen Raumimpulsantworten oder, wie oben bereits beschrieben, in der „Realiser-Sprache“ „Personal Room Impulse Responses“. Um diese zu erhalten, werden Sweeps nacheinander über die angeschlossenen Lautsprecher abgespielt. Dabei ist jedem Ausgangskanal ein Label zugewiesen, welches mit dem angeschlossenen realen Lautsprecher korrespondiert. So können die einzelnen BRIRs einem Lautsprecher zugeordnet werden. Pro Lautsprecher wird dann für verschiedene Kopforientierungen eine BRIR gemessen und daraus ein Datensatz erstellt, welcher letzten Endes den virtuellen Lautsprecher definiert. Bei der Wiedergabe ist diesem dann ein bestimmter Eingangskanal zugeordnet, an dem wiederum das Signal anliegt, das im realen Abhörraum am korrespondierenden realen Lautsprecher anliegen würde. Dieses Signal wird dann mit den BRIRs gefaltet und auf diese Weise eine binaurale Reproduktion der Lautsprecheranordnung umgesetzt. Mithilfe von Headtracking wird schließlich die Kopfposition erfasst und die BRIRs der verschiedenen Messpunkte entsprechend interpoliert.

5.2 Einbindung

Zur Durchführung von Messungen laufen die Signale vom Realiser kommend über ein Ferrofisch-Interface in das Dante-Netzwerk und auf die Lautsprecher (vgl. Kapitel 4.1). In Dante wird der Realiser bzw. das IF auf den freien Slot geroutet. Dazu muss am RMU-Rechner auf der Penta Kreuzschiene noch die entsprechende Quelle („C“) gewählt werden. In die Gegenrichtung, beispielsweise von der DAW kommend, laufen die Signale direkt über Dante auf das Ferrofisch-IF und von dort auf den Realiser. Eine Grafik dieser Signalkette ist unten beigefügt. Das konkrete Routing unterliegt einigen Besonderheiten und wird in den Kapiteln 6.3.2 für Messungen und in Kapitel 7.3 für die Audiowiedergabe genauer ausgeführt.

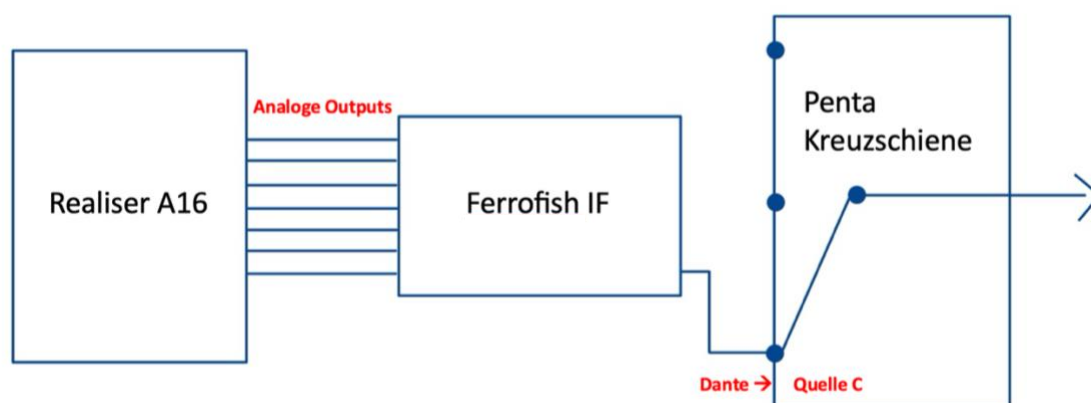


Abbildung 11 - Realiser Einbindung bei Messungen

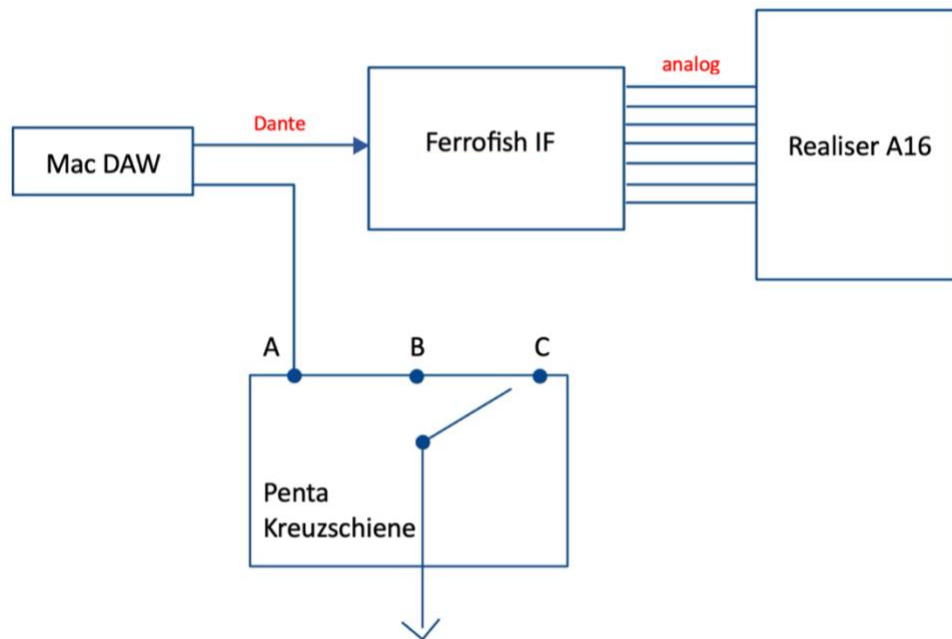


Abbildung 12 - Realiser Einbindung bei der Audiowiedergabe

5.3 Konfiguration einer „PRIR“-Messung

5.3.1 Sound Room

Der Sound Room ist eine der wichtigsten Messeinstellungen. Zwar werden hier im Wesentlichen lediglich eine Reihe von Metadaten voreingestellt, diese werden jedoch später in die Messdateien eingebunden und sind sehr wichtig, um die virtuellen Lautsprecher korrekt unterscheiden zu können. Zunächst wird dem Sound Room hierfür ein Name (z.B. „U48“) und eine Beschreibung (z.B. „9.1.4“) vergeben. Anschließend werden Angaben zu bis zu 16 Lautsprecher gemacht, beispielsweise zu Position und Abstand. Ebenfalls wird festgelegt, welcher Lautsprecher an welchem Ausgangskanal des Realisers anliegt. Hierzu werden den 16 Kanälen Lautsprecher-Labels zugewiesen. Diese Informationen dienen später dazu, die gemessenen BRIRs einem virtuellen Lautsprecher zuzuordnen. Theoretisch muss der „Sound Room“ nur dann umkonfiguriert werden, wenn sich das Setup ändert, also wenn ein neues verbaut wird oder ein anderer Raum gemessen werden soll. Hierzu lassen sich im Gerät jedoch zwei unterschiedliche „Sound Rooms“ definieren. Im Falle von U48 ist es leider nicht möglich sämtliche Lautsprecher mit der Konfiguration eines „Sound Rooms“ zu erfassen. Daher wurden beide „Sound Rooms“ mit je unterschiedlichen Einstellungen genutzt. In Sound Room 1 wurde mit den neun Lautsprecher auf Ohrhöhe, den fünf Höhen- und dem Top-Lautsprecher an der Decke ein Custom-Setup („14.1.1“) konfiguriert und in Sound Room 2 ein 9.1.4-Setup. Somit sind alle Lautsprecher mindestens einmal vorhanden. Hinsichtlich der oben erwähnten Labels, ist deren Wahl für die Wiedergabe von Audiomaterial mit dem Realiser von großer Bedeutung. Dieser unterstützt nämlich nur bestimmte Lautsprecher-Layouts mit vorgegebenen Labels. Die Labels im Sound Room müssen mit diesen

übereinstimmen, um die Einstellungen für die Wiedergabe unkompliziert und fehlerfrei zu halten. Dies wird im Kapitel 6.1 genauer ausgeführt. Aus diesem Grund kann es sein, dass nicht alle Labels des Sound Rooms mit denen in U48 übereinstimmen. Eine Übersicht der Realiser- und der U48-Labels ist der folgenden Tabelle zu entnehmen (ohne Beachtung der korrekten Kanalbelegung! hierzu siehe Kapitel 5.3.2). Beispielsweise nutzt der Realiser für 9.1.4-Layouts die Labels Lss und Rss statt Ls und Rs. Anderer Labels heißen grundsätzlich anders. So wird in etwa „left rear“ zu „left back“ oder „right surround height“ zu „right height rear“. Zwar gibt es das Label „Rhs“, dieses bedeutet im Realiser jedoch „right height side“ und die eingespeicherte Standardposition weicht stark von der Position eines Surround-Lautsprechers ab. Allgemein mussten einige dieser Positionen an U48 angepasst werden.

U48	Realiser
L	L
R	R
C	C
LFE	SW
Ls	Lss
Rs	Rss
Lr	Lb
Rr	Rb
Ltf / Tfl	Ltf
Rtf / Tfr	Rtf
Ltr / Tbl	Ltr
Rtr / Tbr	Rtr
Lw	Lw
Rw	Rw
Lh	Lh
Rh	Rh
Ch	Ch
Lsh	Lhr
Rsh	Rhr
Top	T

Tabelle 2 - Lautsprecher-Labels U48 & Realiser

5.3.2 Kanalbelegung und Anpassung des Routings für Messungen

Die Kanalbelegung zwischen dem Realiser und U48 unterliegt einigen Beschränkungen. Vor allem für den Sound Room 1 muss deshalb für eine Messung das Dante-Routing angepasst werden. So liegen beispielsweise die Höhen- und der Top-Lautsprecher in U48 auf den Kanälen 15-20, während der Realiser jedoch nur über 16 Kanäle verfügt. Die Belegung dieses Sound Rooms und die erforderlichen Anpassungen im Dante-Routing sind der Tabelle 3 zu entnehmen. Für den Sound Room 2 gestaltet sich das Routing etwas einfacher. Da hier nur 14 Kanäle belegt sind und sämtliche verwendeten Lautsprecher auch in U48 auf den Kanälen 1-16 liegen, wurde die Kanalbelegung des Realisers einfach an U48 angeglichen. Somit ist das Dante-Routing hier 1 zu 1. Die entsprechende Kanalbelegung kann der Tabelle 1 entnommen werden.

Kanal	Realiser	U48	Ferrofisch-Output (Sender)	Penta-Input (Empfänger)
1	L	L	1zu1	
2	R	R		
3	C	C		
4	SW	LFE		
5	Lss	Ls		
6	Rss	Rs		
7	Lb	Lb		
8	Rb	Rb		
9	Lh	Ltf	9	15
10	Rh	Rtf	10	16
11	Ch	Ltr	11	17
12	Lhr	Rtr	12	18
13	Rhr	Lw	13	19
14	Lw	Rw	14	13
15	Rw	Lh	15	14
16	T	Rh	16	20
17		Ch		
18		Lsh		
19		Rsh		
20		Top		

Tabelle 3 - Labelbezeichnung "14.1.1" Realiser & U48, Dante-Routing Anpassung

5.3.3 Messeinstellungen

Vor der Messung müssen verschiedene Einstellungen überprüft werden. Diese sind in mehreren Untermenüs zu finden. So unter anderem in den Systemsettings unter dem Punkt „Measurement Settings“. Hier sind vor allem das Mikrofongain, die Lautstärke der Sweeps und die Einstellung des De-Noising zu beachten. Ebenso kann hier die Lautstärke der Ansagen des Geräts während einer Messung verändert werden (vgl. folgendes Kapitel). Weitere Einstellungen erfolgen in den Programmroutinen unter dem Menüpunkt „Apps“. Zunächst wird hier festgelegt, wie viele Winkel pro Lautsprecher gemessen werden. Allerdings ist die Messung auf Messpunkte in der Horizontalebene beschränkt. Die Elevation, also die Medianebene, wird derzeit nicht unterstützt (Smyth, 2021). Die gewünschten Lautsprecher können hier zudem an- oder abgewählt werden, sowie die Art des Sweeps und seine Länge verändert werden. Diese liegt entweder bei 4 oder bei 12 Sekunden, jeweils mit oder ohne Überlappung. Daneben werden ein „Subject“, also die Person, an der die Messung durchgeführt wird, der zu messende Raum (= „Sound Room“) und die später verwendeten Kopfhörer eingestellt. Diese können innerhalb des „Settings“ Menü vorher benannt werden. Zuletzt muss zwischen den beiden Messmethoden „Synchronous“ und „Asynchronous“ gewählt werden. Bei Ersterer wird der Sweep vom Realiser generiert und bei Letzterer von einem externen Gerät. Diese Methode ist vor allem dann sinnvoll, wenn der Realiser nicht direkt an die Lautsprecher angeschlossen werden kann. Für diese Arbeit wurden die Messeinstellungen mit der Intention gewählt, ein möglichst akkurates Ergebnis zu erzielen. Insbesondere wurde hierzu das Gain der Mikrofone auf 6dB reduziert, um die SNR („Signal-to-Noise-Ratio“) zu verbessern und das De-Noising eingeschaltet. Darüber hinaus wurde die Zahl der Messwinkel weitmöglichst erhöht (11 mal 15° Schritte in beide Richtungen), um die „Auflösung“ des BRIR-Files zu vergrößern und ein 12 Sekunden Sweep mit Overlapping gewählt um die Messdauer im Rahmen zu halten und dennoch Genauigkeit zu gewährleisten. Als Messmethode wurde die „Synchronous“-Methode verwendet.

5.3.4 Messaufbau

Zum Aufbau des Messequipments, werden die Ohrmikrofone mit passenden Ohreinsätzen versehen und am Realiser angeschlossen, ebenso wie (bei einer Messung am Menschen) das Erdungskabel, welches noch mit einem Band am Handgelenk befestigt wird. Zudem muss der Headtracker verbunden und auf einem Haltebügel montiert werden, welcher dann dem Messsubjekt aufgesetzt wird. Er unterstützt den Nutzer später dabei, den Kopf im korrekten Winkel zu drehen. Bei den im Rahmen dieser Arbeit durchgeführten Messungen, wurde ferner der Kunstkopf auf einem Stativ montiert und am Sweetspot positioniert. Das Stativ wurde so eingestellt, dass die beiden künstlichen Ohren auf Höhe der Lautsprecher liegen. Die Höhe wurde im Anschluss markiert, um sicherzustellen, dass sie durch die Drehung des Kopfes bzw. des Stativs während der Messung nicht verändert wird. Schließlich wurden

noch die Ohrmikrofone ein- und der Headtracker aufgesetzt. Innerhalb der Lautsprecheranordnung wurden sämtliche Gegenstände, soweit möglich, entfernt, vornehmlich also die Stühle an den Rand gestellt. Um Reflexionen entgegenzuwirken, wurden die Bildschirme nach hinten geklappt und mehrere Absorber auf dem Tisch platziert.

5.3.5 Erläuterungen zur Verwendung des Kunstkopfes

In diesem Kapitel soll die Frage geklärt werden, wieso ein Kunstkopf verwendet wird, ist doch die Möglichkeit einer mehr oder weniger vollständigen Individualisierung von binauralem Audio ein großes Alleinstellungsmerkmal des Realisers. Da aber an der HdM viele verschiedene Studenten den Realiser verwenden können sollen, ist es notwendig, eine größtmögliche Kompatibilität unter den verschiedenen Nutzern zu erreichen. Der Kunstkopf ist für diese Anwendung ideal geeignet, da er eine Durchschnittsanatomie implementiert und die mit ihm durchgeführten Messungen somit für jeden nutzbar sind.



Abbildung 13 - Messaufbau in U48



Abbildung 14 - Neumann KU100 im Messaufbau mit Headtracker (links) & Ohrmikrofone im Detail (rechts)

5.4 Messablauf

Eine Messung beginnt mit der Kalibrierung der Lautsprecher. Hierzu wird die „App“ „Calibrate Speakers (CAL)“ aufgerufen. Der Realiser spielt je nach Konfiguration des „Sound Rooms“ auf jedem Lautsprecher mehrere Sweeps ab, um die Lautstärke einzustellen. Diese wird über die Ohrmikrofone detektiert. Ist der Pegel mindestens eines Mikrofons im idealen Bereich, wird der nächste Kanal/ Lautsprecher bespielt. Der LFE folgt als Letztes. Gegebenenfalls muss die Kalibrierung mehrmals durchgeführt werden, beispielsweise wenn das Signal clippt oder permanent zu leise ist. Dann müssen entweder die Abhör-Lautstärke und/ oder das Gain der Mikrofone angepasst werden. Durch Verlassen der Routine wird das Speaker-Gain gespeichert. Das heißt, dass bei gleichen Einstellungen theoretisch nicht vor jeder Messung neu kalibriert werden muss. An dieser Stelle besteht auch noch einmal die Möglichkeit, die Kanalbelegung und das Routing zu überprüfen.

Anschließend erfolgt die eigentliche Messung, die in der App „PRIR Measurement (SPK)“ gestartet wird. Der Realiser sagt zunächst die gewünschte Messposition an (z.B. 45° left) und gibt anschließend mehrere Sinustöne wieder, deren Lautstärke mit der Annäherung an diese Position abnimmt. Diese wird über den Headtracker erfasst. Ist kein Ton mehr zu hören, wird die Messung mit der Enter-Taste gestartet. Darauf spielen immer vier Lautsprecher zeitversetzt und mit einem gewissen Overlay einen Sweep ab. Wurden alle Lautsprecher angespielt, wird die nächste Position angesagt und der Kopf bzw. in diesem Fall der Kunstkopf entsprechend gedreht. Bei Letzterem ist darauf zu achten, dass dieser vor

dem Start des nächsten Sweeps vollends stillsteht. Nachdem alle Messwinkel erfasst wurden, wird die Messung mit der Enter-Taste gespeichert und dabei eine entsprechende Datei erstellt. Hierzu erzeugt der Realiser für jeden realen Lautsprecher (im Sound Room angegeben) ein virtuelles Pendant und speichert dieses in einem „PRIR“-File (vgl. Kapitel 5.1). Die Datei liegt zunächst im Pufferspeicher und muss noch auf den festen Speicher übertragen werden.

Direkt im Anschluss ist es sinnvoll, die Kopfhörerentzerrung (HPEQ) durchzuführen, damit sich die Position der Mikrofone durch Ab- und Anlegen nicht verändert (Smyth, 2021). Das Kapitel 3.4 erklärt bereits die Notwendigkeit dieser für die korrekte binaurale Wiedergabe von Audiomaterial. Um die HPEQ mit dem Realiser durchzuführen, werden die Ohrmikrofone angelegt bzw. nach der BRIR-Messung in den Ohren belassen und die gewünschten Kopfhörer darübergezogen. Im App-Menü wird „Headphone EQ (HP)“ ausgewählt und eine der drei möglichen Messmethoden gewählt. Neben der hier verwendeten „AutoEQ“-Methode, gibt es die „FlatEQ“- und die „ManualEQ“-Methode. Wie der Name sagt, wird bei der „FlatEQ“-Methode lediglich eine flache EQ-Kurve generiert, während bei einem „ManualEQ“ die Flat- und Auto-EQs noch einmal manuell überarbeitet werden können. Bei der „AutoEQ“-Methode wird neben einer zusätzlichen flachen EQ-Kurve, eine Frequenzkurve erstellt, die den Frequenzgang des Kopfhörers ausgleicht (vgl. Kapitel 3.4). Dies läuft folgendermaßen ab. Nachdem die Messung gestartet wurde, spielt der Realiser zwei kurze Sweeps um die Gains zu überprüfen. Anschließend folgt pro Seite ein längerer Sweep. Danach wird eine HPEQ-Datei erstellt und ebenfalls erst in den Pufferspeicher gelegt. Diese muss auch hier noch in den internen Speicher kopiert werden.

Konkret wurden in U48 zwei Messungen durchgeführt. Die erste mit dem Sound Room 1, also dem custom „14.1.1“-Setup und die zweite mit dem Sound Room 2 und dem 9.1.4-Setup. Somit existiert für jeden realen Lautsprecher mindestens ein virtuelles Pendant. Zudem wurde die Kopfhörerentzerrung mit drei verschiedenen Kopfhörern am KU100 gemessen. Zunächst mit dem AKG K702, welcher später jedoch durch den AKG K712 ersetzt wurde (vgl. Kapitel 8.4), sowie dem Beyer Dynamics 770D in der 80-Ohm-Variante. Dies alles wurde an einem Stück durchgeführt, um eine veränderte Positionierung der Ohrmikrofone zu vermeiden und somit in sich konsistente Messungen zu erhalten.

6 Audiowiedergabe mit dem Realiser

Die vollständige Konfiguration des Realiser erfordert nach der Messung einige weitere Einstellungen, um die Audiowiedergabe zu ermöglichen. Diese werden im Folgenden erläutert. Zunächst gilt es sich einen Überblick über die verschiedenen Begriffe zu verschaffen. Darüber hinaus sollten ein Kopfhörer, sowie der Headtracker und das Set-Top angeschlossen sein.

6.1 Listening Rooms

Insgesamt können 32 „Listening Rooms“ konfiguriert werden. Diese definieren im Endeffekt das gewünschte Abhör-Setup, welches *unabhängig* vom zuvor eingestellten Sound Room ist. Das heißt, wenn beispielsweise ein 9.1.4-Setup verbaut ist, im Sound Room als solches konfiguriert und anschließend vermessen wurde, können sämtliche möglichen kleineren Abhörkombinationen genutzt werden. Zudem können Listening Rooms aus virtuellen Lautsprechern verschiedener „PRIR“-Files bestehen. Somit ist es also möglich, mit den beiden Messungen aus U48 jeden der Lautsprecher virtuell abzubilden und diese auf beliebige Weise miteinander zu kombinieren. Im Listening Room Menü wird hierzu zunächst das gewünschte Speaker-Layout aus einer Liste gewählt. Pro Wiedergabeformat stehen hier bestimmte vorgefertigte zur Auswahl oder im Falle von PCM auch vier Custom-Layouts. Weiter unten erscheint dann eine Liste der Kanäle und der damit assoziierten Lautsprecher. Im nächsten Schritt müssen die Lautsprecher und Kanäle des Listening Rooms nun mit den entsprechenden virtuellen Lautsprechern im „PRIR“-File verknüpft werden. Smyth Research nennt diesen Vorgang „Speaker Matching“. Im Endeffekt wird dabei jedem Eingangskanal des Realisers ein virtueller Lautsprecher zugewiesen. Hierzu bestehen zwei Möglichkeiten. Die erste ist es, dies den Realiser automatisch machen zu lassen. Nach einem Klick auf einen der Kanäle können das gewünschte „PRIR“-File und anschließend die Option „Select all matching speakers“ ausgewählt werden. Der Realiser gleicht dann die Labels des „PRIR“-Files mit den Labels des Listening Rooms ab und verknüpft diese passend. Die zweite Möglichkeit besteht darin, einzelne Lautsprecher zu verknüpfen. Hierzu wird der entsprechende Kanal im Listening Room gewählt und nach der Auswahl des „PRIR“-Files die Option „select one matching speaker“. Nun kann ein einzelner virtueller Lautsprecher gewählt werden.

Wie erkennbar ist, ist das automatische Speaker-Matching um einiges weniger aufwändig. Da dieses jedoch nur über die Speaker-Labels funktioniert, ist es wichtig bereits bei der Konfiguration des Sound Rooms auf die Übereinstimmung der Labels mit denen der gewünschten Listening Rooms zu achten. Stimmen diese nicht überein, müssen sie manuell angepasst werden, da der Realiser den betroffenen virtuellen Lautsprecher sonst nicht rendert. Welche Labels die verschiedenen Speaker-Layouts der Listening Rooms verwenden, ist in den Tabellen in Appendix A des Manuals nachzulesen. Eine Ausnahme bilden die Custom-Layouts der PCM Listening Rooms. Hier kann das automatische Speaker-

Matching stets ohne Weiteres angewendet werden. Möchte man hingegen einen Listening Room aus verschiedenen „PRIR“-Files aufbauen, so ist das manuelle Speaker-Matching die Methode der Wahl. Darüber hinaus ergibt sich ein weiteres Problem im Zusammenhang mit den Labels. Soll Audio über HDMI eingegeben werden, so müssen die verwendeten Speaker-Labels mit den Wiedergabeformaten konform sein. Andernfalls kann es vorkommen, dass der virtuelle Lautsprecher nicht mit dem entsprechenden Audio-Stream assoziiert wird und somit nichts gerendert werden kann (Smyth, 2021). Auch hier können die passenden Labels den Tabellen aus Appendix A entnommen werden.

6.2 Presets & User

Der Realiser kennt unter dem Begriff „User“ zwei verschiedene Arten: Die erste sind die aktiven Benutzer, also die jeweiligen Kopfhörerausgänge (User A & User B), welche lediglich zur Unterscheidung dieser beiden dienen. Die zweite sind die frei definierbaren Nutzernamen. Insgesamt können acht dieser Nutzer-Slots vergeben werden. Dies sind auch die „User“, welche bei der PRIR- und der HPEQ-Messung gewählt werden können und in dem „PRIR“-File gespeichert werden. Sie dienen somit der Zuordnung einer Messung oder eines HPEQ-Files zu einem Messsubjekt. Ferner kann für die beiden Kopfhörerausgänge (User A & B) der Nutzernamen ausgewählt werden. Somit kann pro User A/B einer von acht User-Namen gewählt werden, für den dann eines der 16 möglichen Presets eingestellt werden kann. Darin kann für die drei unterstützten Formate (Atmos, DTS:X & PCM) je der gewünschte „Listening Room“ eingestellt werden. Ebenfalls wird pro Preset eine Kopfhörerentzerrung gewählt. Zusammengefasst, kann also in 16 Presets ein HPEQ ausgewählt werden, pro Wiedergabeformat 32 Listening Rooms und diese wiederum für acht User. Für jeden Kopfhörerausgang kann zudem der aktuelle User ausgewählt werden. Diese Verhältnisse sind in Abbildung 15 noch einmal übersichtlich zusammengestellt.

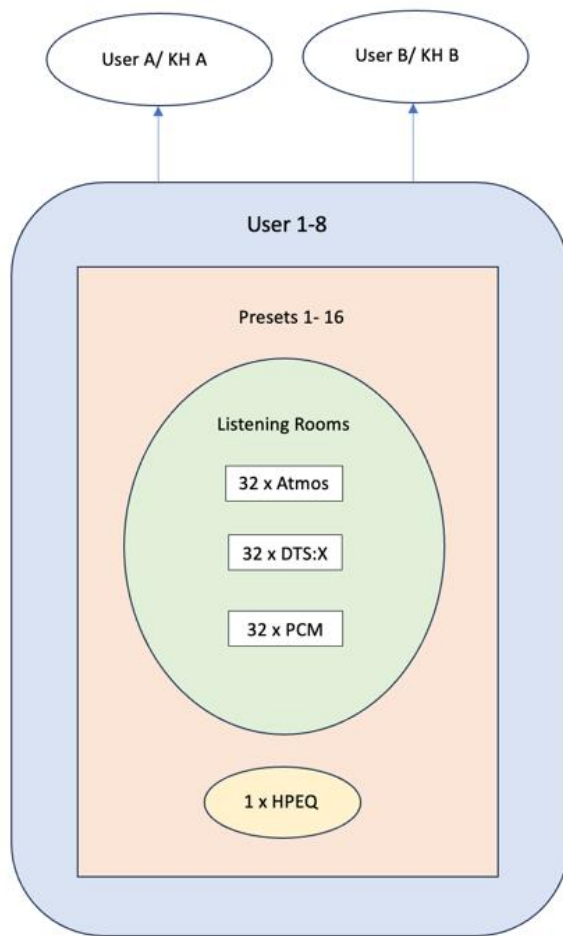


Abbildung 15 - Übersicht der Verhältnisse zwischen User, Presets und Listening Rooms

6.3 Besonderheiten in U48

In U48 wurde die Konfiguration der Nutzer und Listening Rooms etwas angepasst. Ein Grund dafür ist, eine möglichst einfache Bedienung zu gewährleisten. Zudem ist offensichtlich, dass nicht für jeden Studierenden ein User-Slot definiert werden kann. Der angepasste Aufbau ist also der Folgende: Für den KU100 wurden zwei User-Profile angelegt. Neben der Bezeichnung des Users (KU100), sind auch der Name eines Kopfhörers und des Raumes enthalten. Beispielsweise also „U48 KU100 AKG“. Die Idee ist, dass in beiden User-Slots immer die gleiche Kopfhörer-Entzerrung eingestellt ist, sodass die Studierenden dies nicht selbst ändern müssen. Innerhalb dieser Profile gibt es die Folgenden Presets: Das erste Preset enthält als Factory Listening Room einen in der BBC gemessenen Raum. Dieser wurde in der Konfiguration 9.1.4 angelegt. Es folgt ein Listening Room mit dem 9.1.4-Layout der HdM und ein Custom Listening Room mit dem „14.1“-Layout. Das Ansprechen der verschiedenen Sub-Layouts erfolgt über das Dante-Routing. Dabei werden immer die jeweils benötigten Kanäle verbunden. Auf diese Weise müssen die Benutzer des Realisers lediglich das Dante-Routing an das von Ihnen gewünschte Setup anpassen, bei der User-Wahl auf den passenden Kopfhörer achten und schließlich das entsprechende Preset auswählen.

Ähnlich wie bei der Messung (siehe Kapitel 5.3.2), muss auch bei der Wiedergabe das Dante-Routing angepasst werden. Dies ist auch hier den verschiedenen Kanalbelegungen zwischen U48 und dem Realiser geschuldet, zusätzlich aber auch den Listening Rooms. Dabei gestaltet sich das Routing des Custom-Layouts relativ einfach: Entsprechend Tabelle 3 in Kapitel 6.3.2, ist es ausreichend die In- und Outputs zu vertauschen. Der einzige Unterschied besteht darin, dass das Signal direkt von der Penta-Quelle auf das Ferrofisch-IF geroutet wird. Das angepasste Routing ist in Tabelle 4 am Beispiel des DAW-Macs als Quelle dargestellt. Das Routing des 9.1.4-Setups ist hingegen vom Listening Room abhängig. Durch die Konfiguration mittels automatischem Speaker-Matching, hat dieser eine vom Realiser festgelegt Kanalbelegung. Hieran muss das Routing angepasst werden. Die Kanalbelegung des Listening Rooms, sowie das angepasste Routing sind in der Tabelle 5 dargestellt.

DAW-Output (Sender)	Ferrofisch-Input (Empfänger)	Listening Room „14.1.1“
1	1	L
2	2	R
3	3	C
4	4	SW (=LFE)
5	5	Lss
6	6	Rss
7	7	Lb
8	8	Rb
15	9	Lh
16	10	Rh
17	11	Ch
18	12	Lhr
19	13	Rhr
13	14	Lw
14	15	Rw
20	16	T

Tabelle 4 - Wiedergabe-Routing "14.1.1"

Kanal	U48	9.1.4 Listening Room	DAW-Output (Sender)	Ferrofish-Input (Empfänger)
1	L	L	1zu1	
2	R	R		
3	C	C		
4	LFE	SW (=LFE)		
5	Ls	Lss		
6	Rs	Rss		
7	Lb	Lb		
8	Rb	Rb		
9	Ltf	Lw	13	9
10	Rtf	Rw	14	10
11	Ltr	Ltf	9	11
12	Rtr	Rtf	10	12
13	Lw		/	13
14	Rw		/	14
15	Lh	Ltr	11	15
16	Rh	Rtr	12	16

Tabelle 5 - Wiedergabe-Routing 9.1.4

7 Workflow zur Abbildung verschiedener Räume

Grundsätzlich ist es mit dem Realiser möglich verschiedene Räume zu virtualisieren. Aufgrund der zahlreichen Studios an der HdM, ist dies gerade hier eine interessante Nutzungsmöglichkeit. Im Folgenden sollen die entsprechenden Parameter diskutiert werden.

Zunächst müssen einige grundlegende Dinge erledigt werden. Der Realiser muss im neuen Raum aufgebaut und das Ferrofisch-IF an die Lautsprecher angeschlossen werden (i.d.R. über Dante). Alle weiteren Schritte hängen vor allem davon ab, welches Lautsprecher-Setup verbaut ist und ob bzw. wie der Sound Room umgestellt werden muss. Grundsätzlich muss der Sound Room jedoch nicht unbedingt verändert werden. Da mit den 16 Ausgangskanälen und den beiden Setups 9.1.4 und „14.1.1“ bereits alle üblichen Lautsprecherpositionen abgedeckt und die verfügbaren Kanäle ausgelastet sind, ist eine komplette Umkonfiguration nicht sinnvoll. Es sollte ausreichen, die Lautsprecherpositionen im Sound Room bei Bedarf anzupassen und den Namen des Sound Rooms und seine Beschreibung zu ändern, sodass die Dateien später unterschieden werden können. Sollten in einem Raum weniger Lautsprecher als in den oben genannten Layouts vorhanden sein, können die übrigen vor der Messung abgewählt werden. Während der Messung werden dann nur die gewünschten Lautsprecher angespielt. Davon ausgehend, ist der nächste Schritt die Anpassung des Dante-Routings. Hierbei müssen die Kanalbelegungen des Realisers mit denen des Studios abgeglichen und entsprechend neu geroutet werden. Anschließend ist eine Messung des Raumes möglich, die wie oben beschrieben durchgeführt werden kann.

Die Konfiguration der Wiedergabe erfolgt wie in Kapitel 6 beschrieben. Einzig die Kanalbelegung muss, wie auch bei der Messung, zwischen Studio und dem gewählten Listening Room abgeglichen werden. Gegeben falls muss auch hier das Dante-Routing angepasst werden.

Im Folgenden sollen einige Möglichkeiten zur Archivierung der Dateien erläutert werden. Grundsätzlich besteht die einzige Möglichkeit, die Dateien aus dem Realiser zu kopieren darin, diese auf eine SD-Karte zu laden. Eine Möglichkeit wäre also, eine SD-Karte als Archiv zu nutzen. Dies ist die direkteste und einfachste Lösung. Möchte man die Dateien anderweitig extern, beispielsweise in einer Online-Datenbank, archivieren, so ist stets der vorherige Transfer der Files auf eine SD-Karte notwendig. Hierzu gibt es eine eigene Lösung von Smyth Research: die „Realiser Exchange“-Plattform. Über diese Online-Plattform können eigene PRIR-Files hochgeladen und anderen zur Verfügung gestellt werden. Ebenso ist es möglich, die PRIR-Files anderer Personen und Räume herunterzuladen und zu nutzen.

8 Auswertung

Die Ergebnisse der Messung sollen im Folgenden überprüft werden. Unter Zuhilfenahme des Programms RoomEQWizard werden dafür die tatsächlichen Raumimpulsantworten mit den simulierten Raumimpulsantworten verglichen. Dieses Vorgehen hat als Ziel eine objektive, messtechnische Beurteilung der Ergebnisse. Mit dem REW wird das Ergebnis also nicht hinsichtlich seiner wahrgenommenen Qualität, sondern hinsichtlich seiner messbaren Unterschiede und Gemeinsamkeiten bewertet.

Das Vorgehen ist wie folgt geplant. Der Neumann-Kunstkopf wird als Messmikrofon verwendet und auf dem Sweet Spot positioniert. Mit verschiedenen Kopforientierungen werden dann in zwei Schritten Messungen durchgeführt. In einem ersten Schritt wird die tatsächliche Akustik gemessen, indem Messsignale direkt über die Lautsprecher wiedergegeben werden. Anschließend wird in einem zweiten Schritt die Akustik erneut gemessen. Diesmal wird jedoch dem Kunstkopf ein Kopfhörer aufgesetzt und die Messsignale durch den Realiser geschickt. Auf diese Weise wird die von diesem simulierte Raumakustik vermessen. Anschließend werden die Ergebnisse miteinander verglichen und ausgewertet.

8.1 RoomEQWizard

RoomEQWizard ist eine kostenlose Software zur Messung der Raumakustik. Sie wurde von John Mulcahy entwickelt. Und diente ursprünglich der automatisierten Generierung von EQ-Filtern (Menolotto, 2022). Inzwischen ist sie ein weit verbreitetes Tool um akustische Probleme in Studios oder Heimkinos durch Messung zu erkennen und deren Verbesserung zu erleichtern. Hierzu implementiert das Programm einige nützliche Funktionen. Der Kern der Software sind dabei die Messung und Berechnung verschiedener akustischer Parameter eines Raums, wie Frequenzgang und Impulsantwort. Daneben bietet der REW einige nützliche Tools. Es können Wasserfalldiagramme erstellt oder bereits passende inverse Filterkurven berechnet werden. Ebenso sind ein Signal-Generator und ein Realtime-Analyser implementiert.

8.2 Allgemeines Vorgehen

Exemplarisch wurden die realen und virtuellen Lautsprecher des 9.1.4-Setups vermessen. Hierzu wurde der Kunstkopf am Sweet Spot aufgestellt, mithilfe eines Interfaces an den Rechner angeschlossen und dieses im REW als Input-Quelle gewählt. Die Ausgabe des Messsignals von REW erfolgte je nach Iteration via Dante oder direkt über das IF (vgl. Iteration 1 & Iteration 2). Da REW (in der freien Version) nur einen Eingang unterstützt, wurden die BRIRs dann einzeln für die beiden Ohren des Kunstkopfes gemessen. Daneben kann REW das Signal immer nur auf einen Kanal und damit auf einen Lautsprecher ausgeben, was zudem messtechnisch sinnvoll ist. Daher wurden die Lautsprecher nacheinander angespielt. Auf diese Weise wurde über folgende Lautsprecher gemessen:

L C Rs Lw Lb Rtf Rtr

Das Signal wurde bei der Messung jeweils an den gewünschten Lautsprecher geschickt. Überdies wurden die BRIRs für die Kopforientierungen 0°, 90° links und 150° links gemessen, sowie für 56° links, um die Interpolation des Realisers zu erfassen. Zusammengefasst gibt es also pro Orientierung für jeden Lautsprecher ein reales und ein virtuelles BRIR-Paar.

Zunächst wurden die Levels überprüft. Hierzu lässt sich im REW ein kurzes Testsignal abspielen. Im Anschluss gibt der REW eine Bewertung ab, ob das Signal zu schwach, in Ordnung oder zu laut war. Dabei wurde darauf geachtet, dass die Signale vom Realiser und von den Lautsprechern eine ähnliche Lautstärke haben. Anschließend wurden die BRIRs für jedes Ohr zunächst real und anschließend virtuell gemessen. Dazu wurde dem Kunstkopf ein am Realiser angeschlossener Kopfhörer (mit der passenden HPEQ) aufgesetzt und die Ausgabe über die Lautsprecher gemutet. Um zwischen den beiden Ohren zu wechseln, wurden die Ausgangskabel des Kunstkopfes entsprechend am IF getauscht. Während der Messung spielt der REW einen Sweep ab. Dessen Länge wurde auf der Standardeinstellungen von 5 Sekunden belassen. Anschließend liegt innerhalb weniger Augenblicke das Ergebnis der Messung vor.

8.3 Iteration 1

In einem ersten Messdurchlauf wurden die Messsignale über Dante ausgegeben. Hierzu wurde der Rechner über Dante auf die Penta-Kreuzschiene Quelle C gepatcht und diese in der Dadman-Software ausgewählt. Um das Messsignal an den Realiser zu senden, wurde ebenfalls in Dante der Rechner auf das Ferrofisch-IF gepatcht (vgl. Tabelle 5). REW erkennt bereits bei der Auswahl des Output-Geräts, wie viele Ausgänge dieses hat und definiert eine dementsprechende Anzahl an Ausgangskanäle. Die Kanalbelegung in REW kann dann entsprechend angepasst werden und wurde der in U48 angeglichen. Auf diese Weise bleibt der Signal-Output zwischen REW und U48 linear und der Eingriffspunkt für das korrekte Routing zum Realiser besteht wie bisher in Dante. Zudem muss auf diese Weise während einer Messung nicht umgeroutet werden. Die Auswahl des jeweiligen Ausgabekanals geschieht in REW.

Bei einer ersten Betrachtung der Messungen fiel, wie in Abbildung 16 dargestellt, ein Peak bei ca. 4000 Hz auf. Dieser zeigte sich bei verschiedenen Messungen unterschiedlich stark ausgeprägt sowohl am rechten, als auch am linken Ohr. Ein Vergleich der durchschnittlichen Frequenzgänge der realen und virtuellen Umgebung ergab eine Abweichung von ca. 10 dBFS pro Ohr (vgl. Abbildung 17). Eine mögliche Ursache könnte ein Fehler im Kopfhörer oder dessen Entzerrung sein.

Darüber hinaus konnten im Nachhinein zwei Fehler in der Messdurchführung ausgemacht werden. So zeigte sich erstens ein leichter Offset zwischen den Pegeln der realen und virtuellen BRIRs. Rückblickend konnte dadurch festgestellt werden, dass die Levels während der Messung nicht genau genug aneinander angeglichen wurden. Zweitens wurde lediglich in relativen Zeiten ohne Referenz gemessen, was zur Folge hat, dass die Laufzeiten der Impulsantworten nicht miteinander verglichen werden können. Auf diese Weise werden alle Impulse an ihrem Peak ausgerichtet und liegen somit übereinander. Daher können in etwa keine interauralen Laufzeitunterschiede betrachtet werden. Um diese Fehler auszugleichen und zur Erörterung der Ursache des Peaks bei 4 kHz, wurde eine weitere Iteration der Messungen durchgeführt.

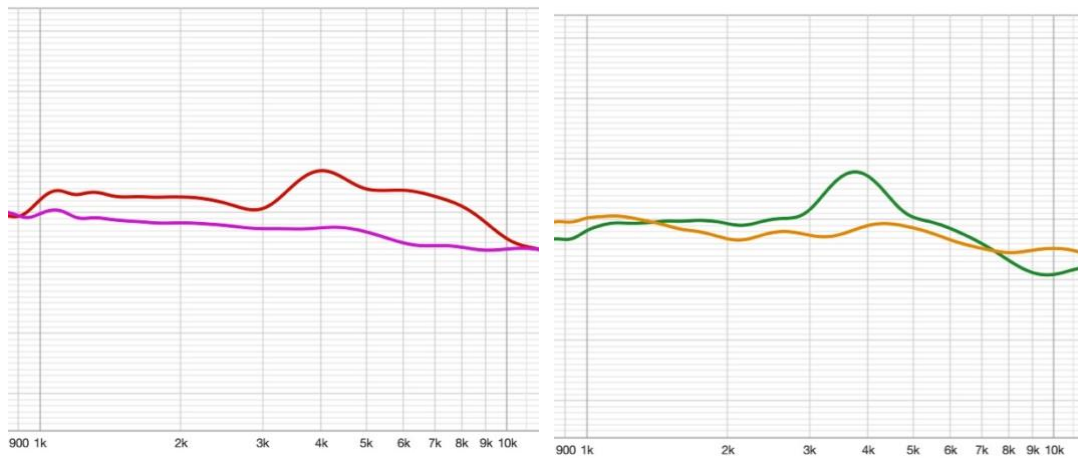
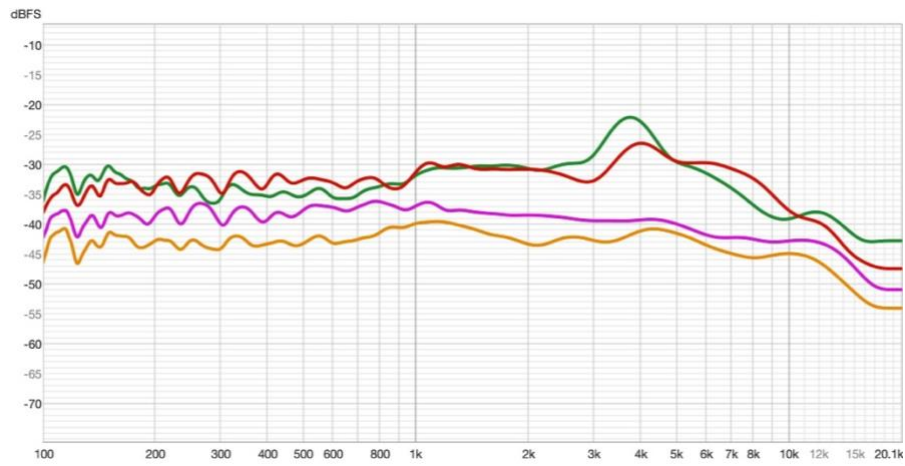


Abbildung 16 - Peak bei 4kHz (VAR-Glättung), Oben: real (gelb/lila) & virtuell (rot/grün) separiert, Unten links: linkes Ohr, Unten rechts: rechtes Ohr

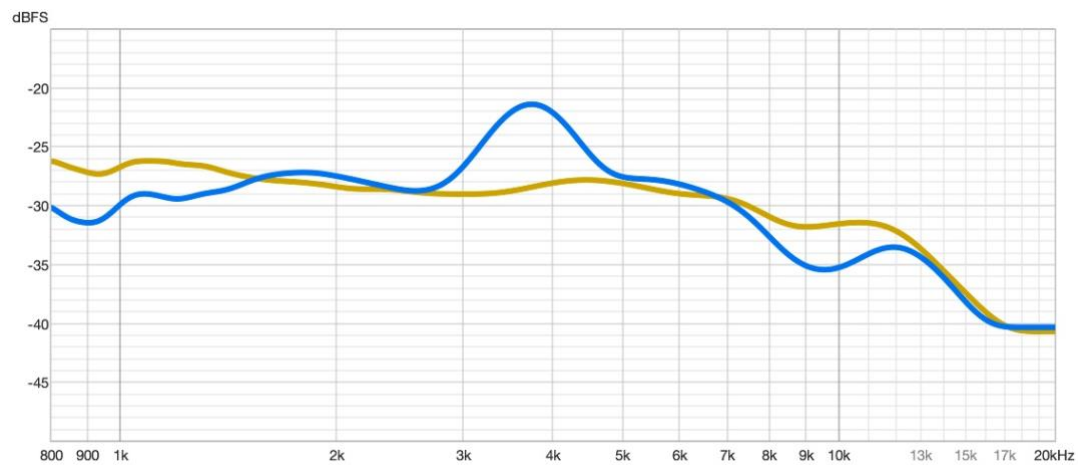


Abbildung 17 - durchschnittlicher Frequenzgang (VAR Glättung): rechtes Ohr, 0°; real (gelb) & virtuell (blau)

8.4 Iteration 2

Bei der erneuten Durchführung der Messungen musste das Setup leicht verändert werden. Um die absoluten Zeiten der Impulse zu messen, wurde ein Loopback als zeitliche Referenz eingerichtet. Dafür wurden ein Ausgang und ein Eingang des IFs miteinander verbunden und vor der Messung die entsprechende Option in REW gewählt. Zusätzlich muss dann in der Software der für den Loopback genutzte Ein- und Ausgangskanal angegeben werden. Dieser Aufbau ermöglicht es REW die Laufzeiten der Raumimpulsantworten mit der Referenz-Laufzeit vom Ausgang zum Eingang des IFs abzugleichen. Unterschiedliche Laufzeiten, zum Beispiel zwischen den Ohren, werden auf diese Weise dargestellt. Diese Konfiguration hat jedoch den Nachteil, dass das gesamte Messsetup umgebaut werden muss, denn im REW ist es nur möglich ein einzelnes Ausgangsgerät zu wählen. Da das Loopback-Signal über das Interface ausgegeben werden muss, muss daher nun auch das Messsignal darüber ausgegeben werden. Eine Konfiguration wie anfangs über Dante ist somit nicht mehr möglich. Daher wurde der zweite Ausgang des IFs mit den analogen Eingängen einer in U48 verbauten Stagebox verbunden. In der Dadman-Software können diese Eingänge auf die Penta-Kreuzschiene gepatcht werden, wodurch sie in das Dante-Netzwerk gelangen. Im Dante-Routing liegt das Signal dann an der Kreuzschiene an. Diese ist also der Sender, und wird auf die Empfänger-Eingänge der Quelle C der Kreuzschiene geroutet. Somit ist die Kreuzschiene gleichzeitig Sender und Empfänger. Um die Signale an den Realiser zu senden, wurde die Kreuzschiene zudem auf das Ferrofish-IF entsprechend der Tabelle 5 geroutet. Um die einzelnen Lautsprecher anzusprechen, wird in der Dadman-Software, der analoge Input auf den jeweils gewünschten Kanal der Kreuzschiene gepatcht.

Die Levels des Realisers und der Lautsprecher wurden vor den Messungen einmalig aneinander angeglichen und für die restliche Messung so belassen. Der restliche Ablauf ist entsprechend der obigen Beschreibung gleichgeblieben.

Um die Frage des Peaks bei 4 kHz zu klären, wurden einige Testmessungen mit einem anderen Kopfhörer und entsprechender Entzerrung durchgeführt. Im Vergleich zu den Messungen aus dem ersten Durchlauf ist hier nun kein Peak mehr bei 4 kHz vorhanden. Offensichtlich hatte der zuvor verwendete Kopfhörer also einen Defekt, weshalb alle weiteren Messungen mit dem neuen Kopfhörer durchgeführt wurden. Dabei handelt es sich um den AKG K712.

Die erste Betrachtung der neuen Ergebnisse zeigte zunächst eine recht signifikante Pegelabweichung zwischen den Ohren in den virtuellen Messungen. Dies wurde vor allem am Center-Lautsprecher mit 0° Kopforientierung sichtbar (vgl. Abbildung 18 & Abbildung 19). Daher wurden einige Tests gemacht, um die Sache zu überprüfen und mögliche Messfehler oder Defekte zu finden bzw. auszuschließen. Da

die Messungen der realen Raumimpulsantworten diese Abweichungen nicht aufweisen, konnten der Kunstkopf, das Mess-Interface sowie die gesamte Signalübertragung als Fehlerquelle ausgeschlossen werden. Daher musste der Fehler im Ferrofisch-IF oder Realiser liegen. Ein erstes Anzeichen für die Fehlerquelle lieferte ein Vergleich der Pegel, die am Kunstkopf ankamen. Dabei zeigte sich, dass der Pegel am linken Ohr konstant um 2 dB leiser war als der Pegel am rechten Ohr. Daraufhin wurde die Kopfhörerentzerrung als mögliche Fehlerquelle geprüft. Hierzu wurde statt der „AutoEQ“-Kurve die „flatEQ“-Kurve eingestellt, wodurch schließlich kein Offset mehr erkennbar war (vgl. Abbildung 19 & Abbildung 20). Um sicherzugehen, dass hier nicht die Messdurchführung oder ein Defekt am Kopfhörer die Ursache waren, wurde dies auch mit einem anderen Kopfhörer getestet, was das gleiche Ergebnis zur Folge hatte. Es scheint also, dass der Realiser auf die „AutoEQ“-Frequenzkurven des linken Ohres einen konstanten Offset von -2 dB rechnet, wofür es jedoch bisher keine Erklärung gibt.



Abbildung 18 - Frequenzgang (VAR Glättung): real, Center, 0°; linkes (rot) & rechtes (grün) Ohr

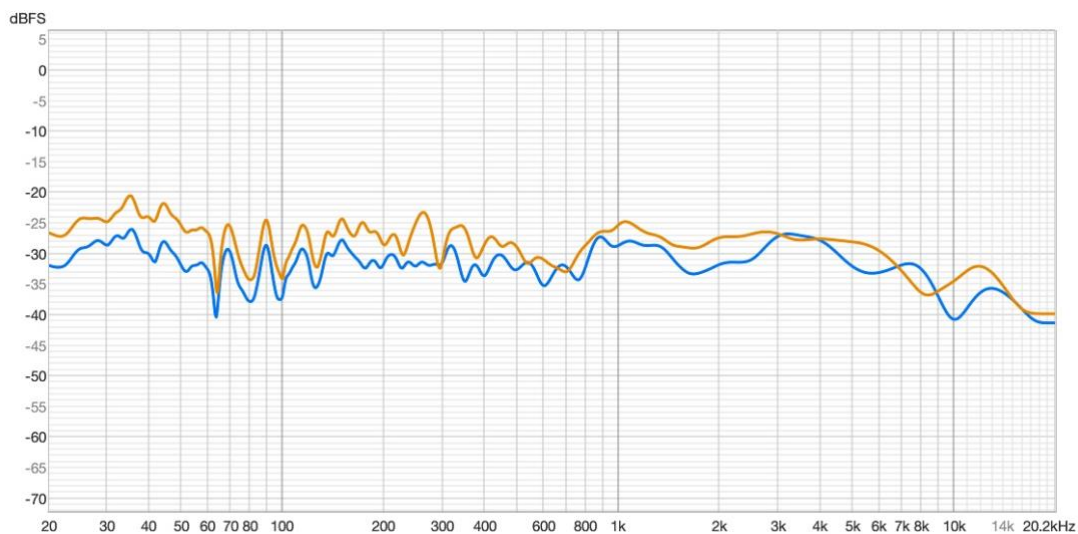


Abbildung 19 – Frequenzgang (VAR Glättung): virtuell, Center, 0°, autoEQ-Kurve; linkes (blau) & rechtes (orange) Ohr

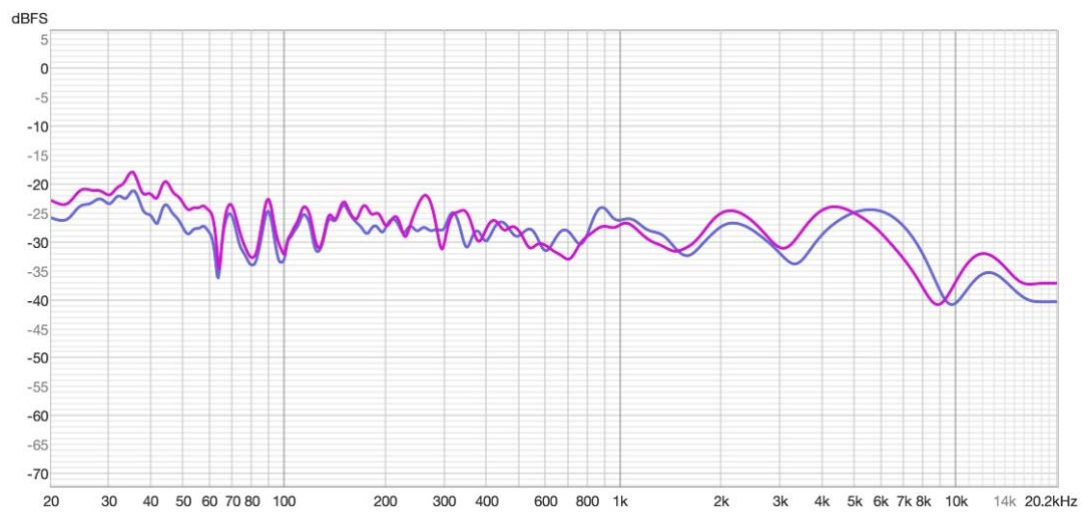


Abbildung 20 - Frequenzgang (VAR Glättung): virtuell, Center, 0°, flatEQ-Kurve; linkes (lila) & rechtes (pink) Ohr

Betrachtet man nun die Messungen im Großen und Ganzen, zeigt sich jedoch, dass der Realiser durchaus mit der realen Umgebung vergleichbare Ergebnisse liefert. Die Impulsantworten zeigen keine unerwarteten Offsets (vgl. Abbildung 21) und ähneln sich stark. Die dennoch vorhandenen Unterschiede sind auf das Processing des Realisers zurückzuführen. Auch ein Vergleich der interauralen Laufzeiten (vgl. Abbildung 22) zeigt nur marginale Differenzen zwischen real (0,79 ms) und virtuell (0,74 ms) auf. Deutlich auffällig sind jedoch die Unterschiede in den Latenzen zwischen beiden Versionen. Das Signal kommt in diesem Fall beim Realiser ca. 3,8 Millisekunden früher an, als bei den Lautsprechern. Dies kann sowohl auf die unterschiedlichen Signalwege zurückgeführt werden, als auch auf eine eventuelle interne Latenzkompensation des Realisers. In diesem Kontext ist die Betrachtung der Latenzen der Messung bei 57° links auffällig, da die Differenz zwischen virtuell und real ca. 4 Millisekunden größer ist. Da diese Messung nicht im selben Durchgang wie die übrigen durchgeführt wurden, wurde dies mit anderen aus diesem Durchgang abgeglichen. Dies bestätigte die veränderte Latenz. Die virtuellen Impulsantworten aus diesem Durchgang liegen alle bei ca. 31 ms, während die des ersten Durchgangs bei ca. 35 ms liegen. Die realen liegen stets bei ca. 39 ms (vgl. Abbildung 23).

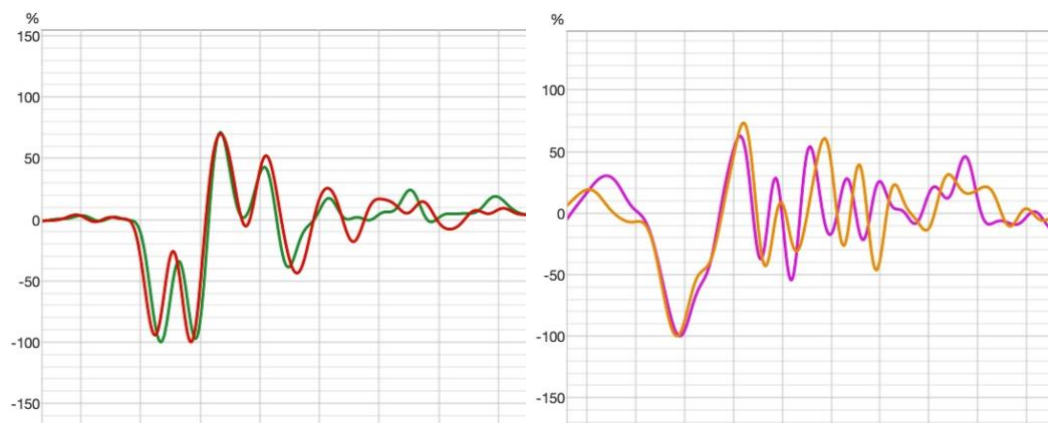


Abbildung 21 - binaurale Raumimpulsantworten : Center, 0°; real (links) & virtuell (rechts)

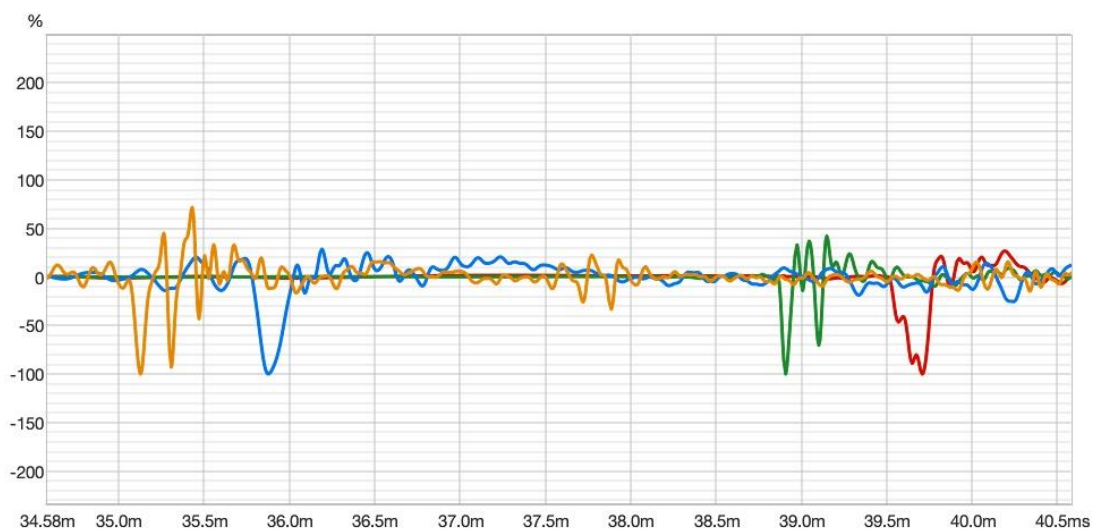


Abbildung 22 - binaurale Raumimpulsantworten : Center, 90° links; real (grün/rot) & virtuell (blau/orange)



Abbildung 23 - Impulsantworten; grün: virtuell, 57° links, Center, rechtes Ohr; blau: virtuell, 90° links, Center, rechtes Ohr; pink: real, 57° links, Center, rechtes Ohr

Bei der Betrachtung der Frequenzgänge muss der oben beschriebene Pegel-Offset durch die Kopfhörerentzerrung berücksichtigt werden. Grundsätzlich kann man jedoch sagen, dass der objektive Verlauf der Frequenzkurven trotz dieses Offsets gut übereinstimmt. Dies fällt vor allem auf, wenn man den Frequenzgang der einzelnen Ohren miteinander vergleicht (Abbildung 24 & Abbildung 25). Gerade hier wird auch der Offset am linken Ohr noch einmal deutlich sichtbar. Betrachtet man nur den Verlauf, so ist der des virtuellen Frequenzgangs dennoch relativ ähnlich zum realen. Rechnet man den Offset heraus, sollten hier die Kurven ähnlich deckend sein, wie am rechten Ohr. In Abbildung 26 und Abbildung 27 können die realen und virtuellen ILDs abgelesen werden. Während in der realen Version gut erkennbar ist, ab welcher Frequenz der Kopf Einfluss auf die Pegel nimmt, ist dies aufgrund des Offsets in der virtuellen Version schwieriger. Ebenso verschiebt sich diese Grenzfrequenz hier von ca. 300 Hz hinzu 600 Hz.

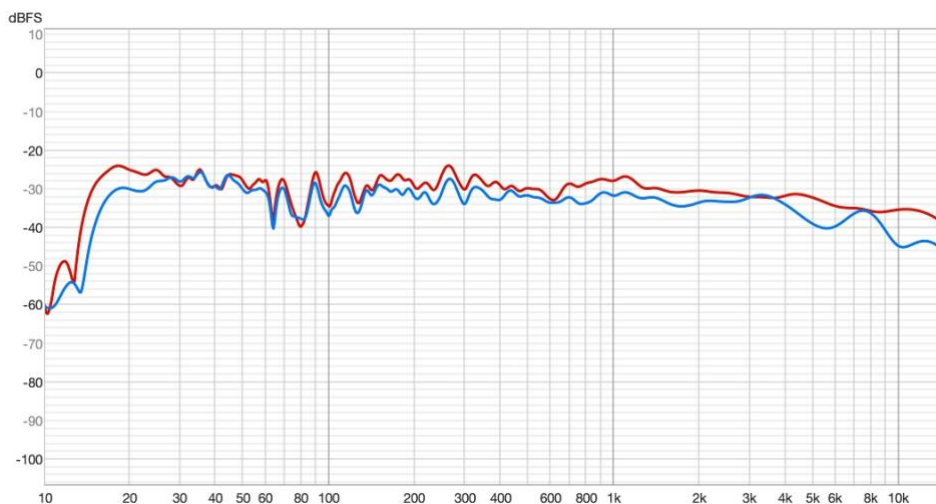


Abbildung 24 - Frequenzgang (VAR Glättung): linkes Ohr, Center, 90° links; real (rot) und virtuell (blau)



Abbildung 25 - Frequenzgang (VAR Glättung): rechtes Ohr, Center, 90° links; real (grün) & virtuell (orange)

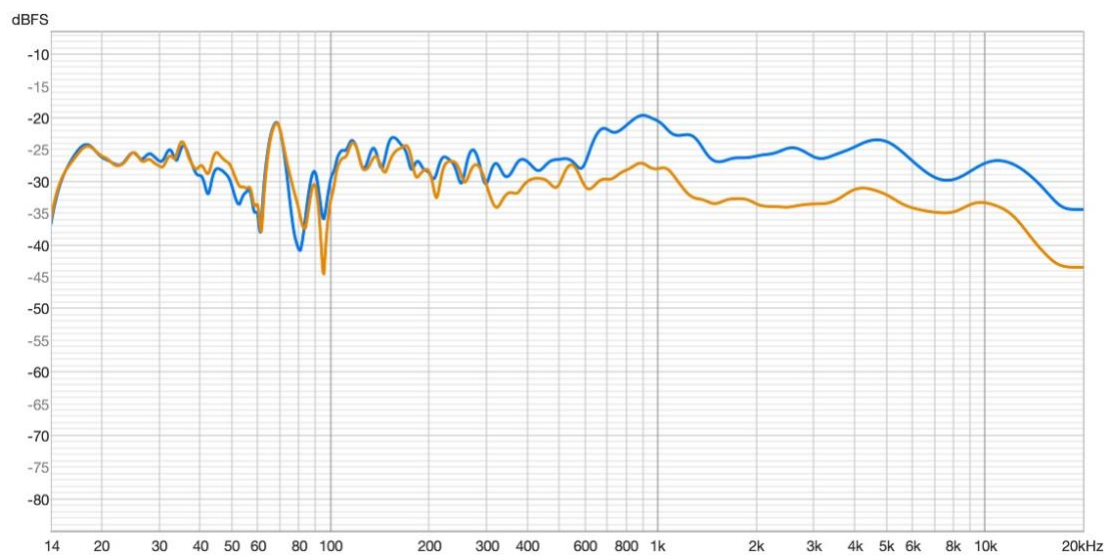


Abbildung 26 - Frequenzgang (VAR Glättung): real, Left wide, 0°; linkes (blau) & rechtes (orange) Ohr



Abbildung 27 - Frequenzgang (VAR Glättung): virtuell, Left wide, 0°; linkes (rot) & rechtes (grün) Ohr

Die Betrachtung der ETC (energy-time curve) der Messungen (Abbildung 28 & Abbildung 29) zeigt ebenfalls ein interessantes Bild. Hier bietet die virtuelle Version sogar eine Art Vorteil, denn im Durchschnitt ist der Noise-Floor der virtuellen Messungen deutlich niedriger (zwischen 5 und 10 dBFS) als der, der realen. Eine mögliche Ursache hierfür könnte das in Kapitel 5.3.3 beschriebene De-Noising sein.

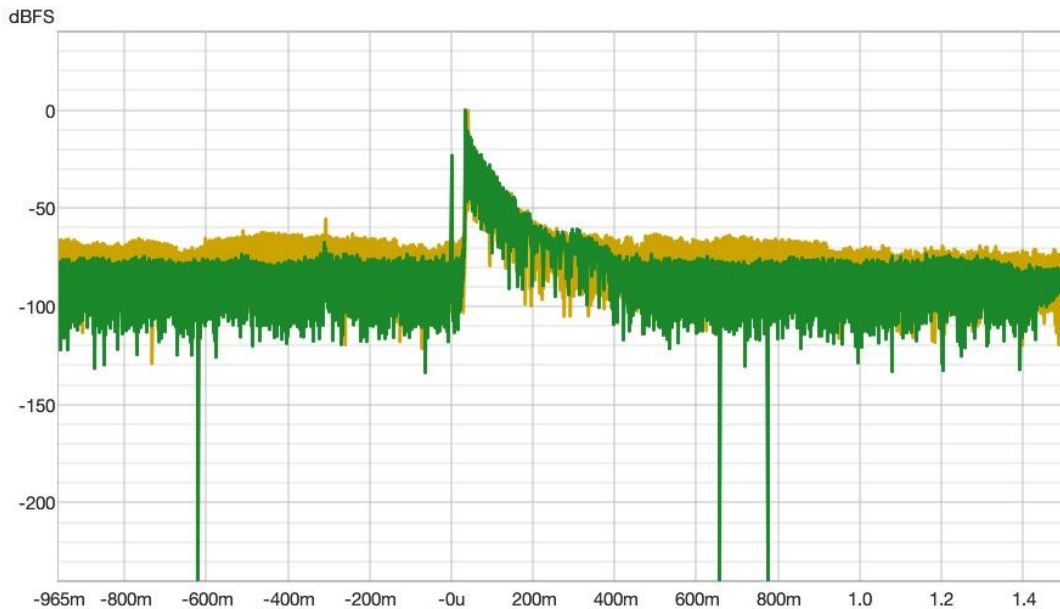


Abbildung 28 - ETC: Center, 0°, linkes Ohr; virtuell (grün) & real (gelb)

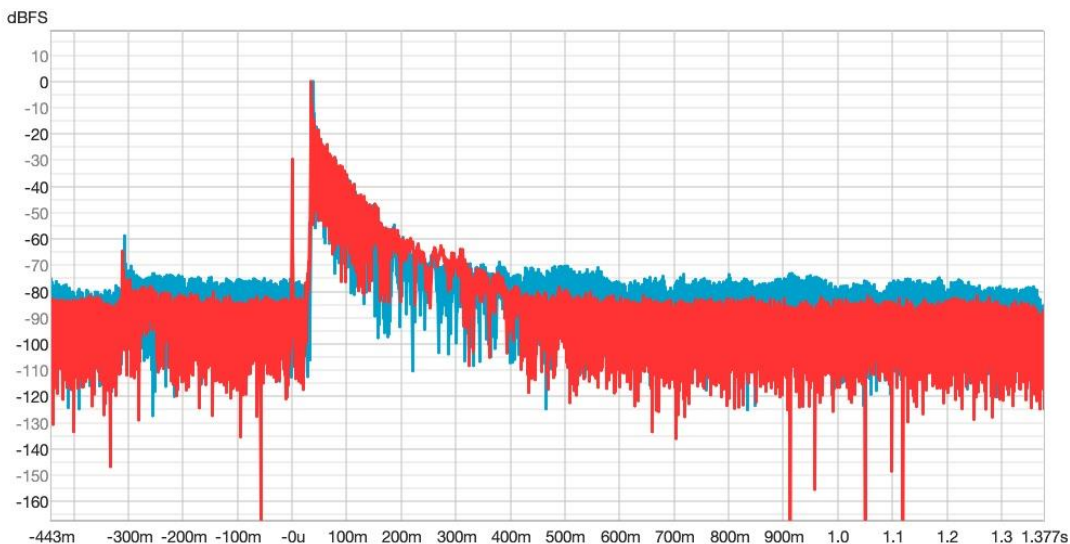


Abbildung 29 - ETC: Center, 90° links, rechtes Ohr; virtuell (rot) & real (blau)

Die gesamten Messdaten und weitere Bilder können im elektronischen Anhang dieser Arbeit eingesehen werden.

9 Fazit & Schlusswort

Rückblickend lässt sich Folgendes festhalten: Die Erstellung einer binauralen Lautsprecher-simulation konnte umgesetzt und gelöst werden. Dabei wurden für den Raum U48 an der Hochschule der Medien in Stuttgart alle Lautsprecher im Rahmen zweier Messungen mit dem Smyth Realiser A16 erfasst und können nun für das binaurale Monitoring dort genutzt werden. Die Konfiguration und Einbindung für Wiedergabe und Messung in die neue Regie wurde ebenfalls umgesetzt. Gerade hierbei ergaben sich diverse Schwierigkeiten bezüglich des Routings, die vor allem auf eine mangelnde Flexibilität seitens des Realisers, was die Möglichkeiten zur Kanalbelegung betrifft, zurückzuführen sind. Die Frage nach der Möglichkeit verschiedene Räume mit dem Realiser abzubilden, konnte ebenfalls beantwortet werden. Hier zeigt sich, dass dies mit einigen wenigen Anpassung der Messeinstellungen möglich ist.

Die Auswertung des Vergleichs der realen mit den virtuellen Raumimpulsantworten ergab grundsätzlich eine gute Übereinstimmung zwischen den beiden Varianten. Unter Berücksichtigung der Effekte des Signal-Processings, verlaufen die Frequenzgänge weitgehend ähnlich, ebenso wie die Impulsantworten. Einige dieser Effekte traten unerwartet auf. Besonders hervorzuheben sind hier die Generierung eines 2dB Pegel-Offsets zwischen dem linken und rechten Ohr durch die Kopfhörerentzerrung, sowie variierende Latenzen zwischen Regie und Realiser, die bei den verschiedenen Messdurchgängen noch einmal unterschiedlich ausfallen. Alles in allem kann jedoch festgehalten werden, dass der Realiser rein messtechnisch gute Reproduktionen der binauralen Raumimpulsantworten und Frequenzgänge liefert.

Die konkrete Arbeit in der Regie war von einigen wenigen Herausforderungen geprägt. Die meisten dieser Probleme traten am Realiser selbst auf, da dieser sich nach einiger Zeit oft aufhing oder während einiger Messversuche abstürzte. Ebenso gestaltete sich die Bedienung des Realiser über Fernbedienung und eingebauten Bildschirm etwas schwierig.

Ein nächster Schritt wäre die Klärung der beschriebenen Abweichungen in den Messergebnissen. Möglicherweise ist hier die Verwendung der flatEQ-Kurve praktikabler, als die Verwendung der autoEQ-Kurve. Ebenfalls bleibt die Frage offen, woher die unterschiedlichen Latenzen zwischen den verschiedenen Durchgängen kommen. Möglicherweise könnte in beiden Fällen ein Update der Firmware des Realisers bereits Abhilfe schaffen. Gleichmaßen könnte das Thema durch die wahrnehmungstechnische Verifizierung der Messergebnisse weitergeführt werden. Dabei müsste die Frage geklärt werden, ob bzw. inwieweit die oben beschriebenen Abweichungen vom realen Abhörraum die Wahrnehmung beeinflussen. Da der Realiser die RIRs jedoch bereits relativ genau reproduziert, wäre es naheliegend, dass die Simulation auch wahrnehmungsseitig akkurat ist. Um dies

zu erörtern könnte ein vergleichender Hörtest durchgeführt werden. Speziell an der HdM wäre ferner eine Messung in der Regie B ein möglicher nächster Schritt. Die Besonderheit dort ist, dass es neben einem 9.1.4 Lautsprecher Setup auch eine Sennheiser Ambeo-Soundbar vorhanden ist. Diese erzeugt die Höhen- und Surround-Kanäle durch Reflexionen an den Wänden. Da sie jedoch über keinen PCM-Eingang verfügt, kann der Realiser nicht direkt angeschlossen werden und damit die interne Generation der Sweeps nicht genutzt werden. Diese müssen folglich extern erzeugt und in die Soundbar eingespeist werden. Grundsätzlich unterstützt der Realiser die Möglichkeit der externen Generation der Mess-Sweeps, doch ist diese Methode um einiges schwerer umzusetzen. Im Manual des Realisers gibt es hierzu eine detaillierte Anleitung.

Abschließend bleibt Folgendes zu sagen: Die Binauraltechnik ist und bleibt ein spannendes Thema. Es ist dabei sehr faszinierend, was mit den entsprechenden Tools alles möglich ist und wie komplex das fast schon bestechend einfache Prinzip der Binauralsynthese doch wiederum ist. Das Hören ist letzten Endes ein vielschichtiger Vorgang und es braucht viel Expertise und Fachwissen, um das entsprechende Equipment zu entwickeln, zu bedienen und einzusetzen. Bezugnehmend auf die Motivation aus Kapitel 1.1, ist es schwer eventuelle Lösungen auf ein triviales Niveau herunterzubrechen. Letzten Endes blickt man mit Spannung auf zukünftige Entwicklungen oder ist möglicherweise sogar selbst an diesen beteiligt.

Literatur – und Quellenverzeichnis

Algazi, V., Duda, R., Thompson, D. & Avendano, C. (2002). The CIPIC HRTF database. In *Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No.01TH8575)*. <https://doi.org/10.1109/aspaa.2001.969552>

Armstrong, C., Thresh, L., Murphy, D. & Kearney, G. (2018). A Perceptual Evaluation of Individual and Non-Individual HRTFs: A Case Study of the SADIE II Database. *Applied Sciences*, 8(11), 2029. <https://doi.org/10.3390/app8112029>

Aural ID. (o. D.). Genelec. Abgerufen am 30. Juli 2025, von <https://www.genelec.com/aural-id>

Blauert, J. (1996). Spatial Hearing. In *Spatial hearing : the psychophysics of human sound localization* (Revised edition, S. 1–489). The MIT Press. http://direct.mit.edu/books/oa-monograph/chapter-pdf/2056429/f011001_9780262268684.pdf (Ursprünglich veröffentlicht 1974)

Blauert, J. & Braasch, J. (2008). Räumliches Hören. In S. Weinzierl (Hrsg.), *Handbuch der Audiotechnik* (S. 87). <https://doi.org/10.1007/978-3-540-34301-1>

Bomhardt, R., De La Fuente Klein, M. & Fels, J. (2016). A high-resolution head-related transfer function and three-dimensional ear model database. *Proceedings Of Meetings On Acoustics*, 050002. <https://doi.org/10.1121/2.0000467>

Denk, F., Ernst, S. M. A., Ewert, S. D., Kollmeier, B. & Heeren, J. (2018). *The Oldenburg Hearing Device (OIHeaD) HRTF Database* [Database Documentation]. https://uol.de/f/6/dept/mediphsyk/ag/mediphsyk/download/paper/denk/OIHeaD-HRTF_doc_v1.0.3.pdf?v=1676456071

Dolby Atmos Music - Musik wie nie zuvor. (o. D.). Dolby. Abgerufen am 30. Juli 2025, von https://www.dolby.com/de/musik/?gad_source=1&gad_campaignid=21838397744&gbraid=0AAAAACoWuPrG29E9Wb31rS3y5ErbWbKG3&gclid=EAlaIqobChMIjOienOLkigMVaXJBah1mGDa5EAAYASAAEgIeIFD_BwE

Duraiswami, R., Zotkin, D. N. & Gumerov, N. A. (2004). Interpolation and Range Extrapolation of HRTFs [Conference Paper]. In *Proceedings the IEEE Conference of Acoustics, Speech, and Signal Processing*. <https://www.researchgate.net/publication/236024149>

Farina, A. (2000). Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique. *Journal Of The Audio Engineering Society*. https://zenway.ru/uploads/01_12/134-aes00.pdf

Fraunhofer IIS. (o. D.). *MPEG-H Audio*. Fraunhofer-Institut für Integrierte Schaltungen IIS. Abgerufen am 28. Juli 2025, von <https://www.iis.fraunhofer.de/de/ff/amm/rundfunk-streaming/mpegh.html>

Friesecke, A. (2014). *Die Audio-Enzyklopädie: Ein Nachschlagewerk für Tontechniker* (2. Aufl.). De Gruyter.

Georg Neumann GmbH. (2020). *KU 100 Bedienungsanleitung*. Abgerufen am 6. August 2025, von https://www.neumann.com/de-de/serviceundsupport_de/downloads?productName=KU+100

Hammershøi, D. & Møller, H. (2005). Binaural Technique: Basic Methods for Recording, Synthesis, and Reproduction. In J. Blauer (Hrsg.), *Communication Acoustics* (S. 223–254). IEEE Computer Society Press.

Hess, W. (2012). Head-Tracking Techniques for Virtual Acoustics Applications [Convention Paper 8782]. In *Journal of the Audio Engineering Society*. <https://aes2.org/publications/elibrary-page/?id=16524>

IRCAM. (o. D.). [Messaufbau der Listen HRTF Datenbank]. Listen HRTF Database. Abgerufen am 27. Mai 2025, von http://recherche.ircam.fr/equipes/salles/listen/system_protocol.html

Kapralos, B. (2003). *Auditory Perception and Virtual Environments* [Ph.D Qualification Exam Document]. York University.

Li, S. & Peissig, J. (2020). Measurement of Head-Related Transfer Functions: A Review. *Applied Sciences*, 10(14), 5014. <https://doi.org/10.3390/app10145014>

Lindau, A. (2014). *Binaural resynthesis of acoustical environments: Technology and perceptual evaluation*. <https://doi.org/10.14279/depositonce-4085>

Menolotto, P. (2022). Meeting Report: August 2022 – John Mulcahy; Room EQ Wizard (REW) software. In *Audio Engineering Society - Melbourne Section*. Abgerufen am 10. Juli 2025, von <https://aesmelbourne.org.au/aug2022-mtg-report/>

Møller, H., Sørensen, M. F., Hammershøi, D. & Jensen, C. B. (1995). Head-related transfer functions of human subjects. *Journal Of The Audio Engineering Society*, 43(5), 300–321.

Müller, S. & Massarani, P. (2001). Transfer-Function Measurement with Sweeps. *Journal Of The Audio Engineering Society*, 49(6), 443–471. <https://www.aes.org/e-lib/browse.cfm?elib=10189>

Smyth, S. (2021). Realiser A16 User Manual. In *Smyth Research*. Abgerufen am 20. Juni 2025, von https://smyth-research.com/wp-content/uploads/2021/10/A16_v2.0_user_manual_20_Oct_2021.pdf

Sound mit personalisiertem 3D-Audio für AirPods und Beats erleben. (2025, 23. April). Apple Support. Abgerufen am 30. Juli 2025, von <https://support.apple.com/de-de/102596>

Weinzierl, S. (2008). Aufnahmeverfahren. In S. Weinzierl (Hrsg.), *Handbuch der Audiotechnik* (S. 586–589). <https://doi.org/10.1007/978-3-540-34301-1>

Weinzierl, S. & Slavik, K. M. (2008). Wiedergabeverfahren. In S. Weinzierl (Hrsg.), *Handbuch der Audiotechnik* (S. 671–673). <https://doi.org/10.1007/978-3-540-34301-1>

Werner, S., Klein, F., Mayenfels, T. & Brandenburg, K. (o. D.). *A summary on acoustic room divergence and its effect on externalization of auditory events*. 8th International Conference On Quality Of Multimedia Experience, Lissabon, Portugal. <https://doi.org/10.1109/qomex.2016.7498973>

Wo kann ich Inhalte in Dolby Atmos anschauen? (o. D.). Dolby. Abgerufen am 30. Juli 2025, von <https://www.dolby.com/de/experience/home-entertainment/articles/wo-kann-ich-inhalte-in-dolby-atmos-anschauen/>

Zhang, W., Samarasinghe, P. N., Chen, H. & Abhayapala, T. D. (2017). Surround by Sound: A Review of Spatial Audio Recording and Reproduction [Journal-article]. *Applied Sciences*, 532. <https://doi.org/10.3390/app7050532>

Zotkin, D., Duraiswami, R. & Davis, L. (2004). Rendering Localized Spatial Audio in a Virtual Auditory Space. *IEEE Transactions On Multimedia*, 6(4), 553–564. <https://doi.org/10.1109/tmm.2004.827516>