

MASTERARBEIT

Konzeption und praktische Umsetzung einer
binauralen Livesession im Kontext der modernen Popmusik



HOCHSCHULE
DER MEDIEN

FAKULTÄT ELECTRONIC MEDIA
AUDIOVISUELLE MEDIEN (M.ENG)

Konzeption und praktische Umsetzung
einer binauralen Livesession im Kontext
der modernen Popmusik

Masterarbeit

im Studiengang
Audiovisuelle Medien (M.Eng.)

vorgelegt von

Marian Hepp

Matr.-Nr.: 35424

am 28. Februar 2020

an der Hochschule der Medien Stuttgart

Erstprüfer: Prof. Oliver Curdt

Zweitprüfer: Prof. Jens-Helge Hergesell

I. Ehrenwörtliche Erklärung

„Hiermit versichere ich, Marian Hepp, ehrenwörtlich, dass ich die vorliegende Masterarbeit mit dem Titel: „Konzeption und praktische Umsetzung einer binauralen Livesession im Kontext der modernen Popmusik“ selbstständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden.

Ich habe die Bedeutung der ehrenwörtlichen Versicherung und die prüfungsrechtlichen Folgen (§ 26 Abs. 2 Bachelor-SPO (6 Semester), § 24 Abs. 2 Bachelor-SPO (7 Semester), § 23 Abs. 2 Master-SPO (3 Semester) bzw. § 19 Abs. 2 Master-SPO (4 Semester und berufsbegleitend) der HdM) einer unrichtigen oder unvollständigen ehrenwörtlichen Versicherung zur Kenntnis genommen.

Stuttgart, 28. Februar 2020

II. Kurzfassung

3D Audio ist ein immer wichtiger werdender Aspekt in der Audiotechnik. Die Binauraltechnik hat dabei eine Sonderrolle inne, da sie das einzige Kopfhörerbasierte immersive Wiedergabeverfahren ist. Da Kopfhörer als Wiedergabemedium für Audioinhalte immer beliebter werden, gehört sie daher zu den zukunftssträchtesten Ansätzen im Kontext von 3D Audio. In der modernen Musikproduktion konnte sich die Binauraltechnik allerdings noch nicht etablieren. Hier dominiert das herkömmliche Stereowiedergabeverfahren. Das noch recht junge, an Relevanz gewinnende Format der Livesession bietet hier ein gutes Einsatzgebiet für Immersive Audio, da es sich klanglich bewusst von klassischen Studioproduktionen absetzen soll.

Diese Arbeit untersucht, anhand einer praktischen Umsetzung einer binauralen Livesession mit einer Band aus dem musikalischen Kontext der Popmusik, welche Schwierigkeiten und Problemstellungen bei einer solchen Produktion auftreten. Im Zuge dessen wird sowohl die ausführliche Planung und Konzeption, die Aufnahme an sich, sowie die Postproduktion detailliert dargelegt. Der Autor zeigt eine mögliche Herangehensweise auf, um eine binaurale Livesession erfolgreich umzusetzen. Das Ziel ist dabei, eine Lösung zu finden, die die aus der herkömmlichen Stereophonie stammenden Hörgewohnheiten der Hörer bedient und trotzdem eine immersive Klangästhetik aufweist, welche einen klaren Mehrwert gegenüber jener einer gewöhnlichen Stereoproduktion liefert.

Schlüsselworte:

3D Audio, Binaural, Livesession, Kopfhörer

III. Abstract

3D audio is an increasingly important aspect in audio technology. Binaural technology plays a special role in this, since it is the only headphone-based immersive production and playback process. As headphones are becoming increasingly popular as a playback medium for audio content, it is one of the most promising approaches in the context of 3D audio. However, binaural technology has not yet established itself in modern music production. The conventional stereo playback method dominates here. The still young format of the live session, which is gaining in relevance, offers a good area of application for immersive audio, since it should deliberately set itself apart from classic studio productions.

Using a practical implementation of a binaural live session with a band from the musical context of pop music, this work examines the difficulties and problems that arise with such a production. In the course of this, both the detailed planning and conception, the recording itself and the post-production are presented in detail. The author shows a possible approach to successfully implement a binaural live session. The aim is to find a solution that serves the listening habits that come from conventional stereophony and yet has an immersive sound aesthetic that delivers clear added value compared to that of a conventional stereo production.

Keywords:

3D Audio, binaural, livesession, headphones

IV. Danksagung

Ich möchte mich herzlich bei allen Personen bedanken, die zur Umsetzung dieser Arbeit und der Realisierung der Livesession beigetragen haben.

Dabei sind vor allem Oliver Szöke, Tobias Bosch, Lennart Bloch und Jonas Kivikoski von Foxframes zu nennen, welche sich sensationell um den visuellen Part der Session gekümmert haben und tatkräftig viel Engagement, Zeit und Equipment beigesteuert haben.

Zu Dank verpflichtet bin ich auch Fabian Glück, Bastian Kilper und Lukas Klotzbach von meiner Band Into The Fray, die ebenfalls viel Zeit und Mühe investiert haben.

Darüber hinaus danke ich auch dem Verein Alte Kirche Rulfingen e.V. sowie der Gemeinde Rulfingen für die Möglichkeit, die Produktion dort so unkompliziert umzusetzen. Dabei sind insbesondere Martina Eisele, Holger Stark und Manfred Moll zu nennen.

Besonderen Dank gilt auch meiner Familie, die sich vor Ort um Verpflegung und Unterkunft gekümmert hat und mir so den Rücken und den Kopf für das Wesentliche freigehalten hat. An dieser Stelle ist zu erwähnen, dass dies nicht nur im Rahmen der Session der Fall war, sondern während des gesamten Verlaufs meines Studiums. Vielen Dank für diesen nicht selbstverständlichen Rückhalt.

Abschließend möchte ich auch meinen beiden Prüfern, Prof. Oliver Curdt und Prof. Jens-Helge Hergesell, sowie Jörg Bauer für viele lehrreiche Vorlesungen, Workshops und Erfahrungen während meiner Zeit an der HdM danken.

V. Inhaltsverzeichnis

I.	Ehrenwörtliche Erklärung.....	I
II.	Kurzfassung.....	II
III.	Abstract.....	III
IV.	Danksagung.....	IV
V.	Inhaltsverzeichnis.....	V
VI.	Glossar	VII
1.	Einleitung.....	1
1.1.	<i>Motivation und Zielsetzung</i>	1
1.2.	<i>Vorgehen</i>	3
2.	Das Medium Livesession.....	4
2.1.	<i>Begriffserklärung</i>	4
2.2.	<i>Entstehung und Relevanz in der Musikindustrie</i>	5
2.3.	<i>Aktuelle Trends und Entwicklungen</i>	10
3.	Die Binauraltechnik	13
3.1.	<i>Räumliches Hören</i>	14
3.1.1.	<i>Auditive Richtungslokalisierung</i>	15
3.1.2.	<i>Auditive Entfernungswahrnehmung</i>	22
3.1.3.	<i>Selektives Hören</i>	24
3.1.4.	<i>Im-Kopf-Lokalisierung</i>	25
3.2.	<i>Funktionsprinzip</i>	25
3.2.1.	<i>Head Related Transfer Function</i>	26
3.2.2.	<i>Binaural Room Impulse Response</i>	28
3.2.3.	<i>Diffus-und Freifeldentzerrung von Kopfhörern</i>	28
3.3.	<i>Binaurale Mikrofonsysteme</i>	30
3.4.	<i>Binauralsynthese</i>	35
4.	Planung und Vorbereitung der Livesession	38
4.1.	<i>Technische Überlegungen</i>	38
4.1.1.	<i>Angestrebte auditive Ästhetik</i>	38
4.1.2.	<i>Einfluss der visuellen Komponente</i>	42
4.1.3.	<i>Erarbeitung eines binauralen Konzepts</i>	44
4.2.	<i>Die Auswahl der Künstler</i>	48

4.3.	<i>Die Wahl des Aufzeichnungsorts.....</i>	50
4.4.	<i>Vorproduktion mit den Künstlern.....</i>	55
4.5.	<i>Finanzierung.....</i>	59
5.	Aufnahme der Livesession	62
5.1.	<i>Signalkette, Mikrofonierung und Soundcheck.....</i>	62
5.2.	<i>Aufnahmestrategie und Herangehensweise.....</i>	70
5.3.	<i>Die Zusammenarbeit mit dem Filmteam.....</i>	71
6.	Postproduktion	73
6.1.	<i>Schnitt und Take-Auswahl.....</i>	73
6.2.	<i>Signalbereinigung</i>	75
6.3.	<i>Einrichten der Mixing-Sessions.....</i>	76
6.4.	<i>Mixing und Mastering.....</i>	80
6.4.1.	<i>Kunstkopf und Raumsignale</i>	81
6.4.2.	<i>Vocals.....</i>	81
6.4.3.	<i>Schlagzeug & Perkussion</i>	83
6.4.4.	<i>Gitarren.....</i>	84
6.4.5.	<i>Stage-Piano</i>	86
6.4.6.	<i>E-Bass und Synthesizer.....</i>	86
6.4.7.	<i>Summenbearbeitung und Mastering</i>	88
7.	Fazit und Ausblick	89
8.	Abbildungsverzeichnis	92
9.	Quellenverzeichnis	94
10.	Anhang.....	102

VI. Glossar

AAX	Avid Audio Extension
ADAT	Alesis Digital Audio Tape
AU	Audio Unit
BBC	British Broadcasting Corporation
BRIR	Binaural Room Impulse Response
DAW	Digital Audio Workstation
dB	Dezibel, logarithmische Verhältnisgröße zur Angabe von Pegeln und Pegeldifferenzen
DI	Direct Injection
EP	Extended Play
EQ	Equalizer
FFT	Fast Fourier Transformation
HRIR	Head Related Impulse Response
HRTF	Head Related Transfer Function
Hz	Hertz
ILD	Interaural Level Differences
ITD	Interaural Time Differences
kHz	Kilohertz
LTI	Linear Time Invariant
ms	Millisekunde
µs	Mikrosekunde
MPEG-H	Audioformat / Übertragungsstandard der Moving Picture Experts Group
Panning	Anordnung von Audioobjekten im räumlichen Panorama
RTAS	Real Time Audio Suite
VR	Virtual Reality
VST	Virtual Studio Technology

1. Einleitung

Die Relevanz von 3D-Formaten in der Audiobranche steigt weiter an. Durch neue Einsatzgebiete im Bereich der Virtual Reality und eine entsprechend große Nachfrage sowie neue technische Möglichkeiten, sind auch in die Jahre gekommene Wiedergabeverfahren, wie die Binauraltechnik oder Ambisonics, wieder lukrativ geworden und werden weiter erforscht und verbessert. In vielen Bereichen lösen neue Technologien aus dem Feld des Immersive Audio daher bereits die herkömmliche Stereophonie ab. Mit MPEG-H wurde sogar bereits ein offener universell einsetzbarer Audiostandard entwickelt, der immersive Inhalte auf verschiedensten Endgeräten wiedergeben kann.¹ Es ist daher anzunehmen, dass im Laufe der nächsten Jahre vorhandene Schwachstellen und Hürden, die den Weg zum massentauglichen Durchbruch bislang erschweren, weiter überwunden und dezimiert werden können.

Der regelrechte „Hype“ um 3D-Audio ist auch im Kontext der Musikproduktion zu spüren, da zunehmend vereinzelte Pionier-Projekte mit immersiven Wiedergabeverfahren realisiert werden. Allerdings handelt es sich hier oft um prestigeträchtige Pilotprojekte mit großem Marketingpotential. Von einem klaren Trend weg von den alten Stereo-Gewohnheiten, kann also absolut keine Rede sein. Im Bereich der Audiotechnik stellt die herkömmliche Stereophonie seit den 1960er Jahren den absoluten Standard für die Wiedergabe von Musik dar. Dennoch scheinen die 3D-Audio-Wiedergabeverfahren diesen Vorsprung langsam aufzuholen.

1.1. Motivation und Zielsetzung

Unter allen 3D-Audio-Wiedergabesystemen nimmt die Binauraltechnik eine Sonderrolle ein. Im Gegensatz zu den lautsprecherbasierten Alternativen funktioniert die binaurale Wiedergabe ausschließlich über Kopfhörer. Dadurch ist dieses Wiedergabeverfahren im Grunde für jeden Endkonsumenten verfügbar, der einen Kopfhörer besitzt und stellt technisch gesehen kaum Ansprüche an diesen. Laut einer Studie aus dem September

¹ Vgl. Fraunhofer-Institut für Integrierte Schaltungen IIS. (2020)

2019 besitzen in Deutschland 71 Prozent der Menschen Kopfhörer². Außerdem hat sich der Umsatz durch Kopfhörer in Deutschland seit dem Jahr 2012 ungefähr verdoppelt und steigt stetig an.³ Demzufolge ist anzunehmen, dass immer mehr Menschen in Deutschland Kopfhörer zur Musikwiedergabe nutzen, und dieser Trend weiter anhält. Die derzeitige Ausgangssituation scheint sich also sehr zu Gunsten der Binauraltechnik zu entwickeln. Dennoch spielt diese in der Popmusik keine Rolle. Dies liegt wohl hauptsächlich daran, dass sich binaurale Mischungen eher weniger für die Wiedergabe durch Lautsprechersysteme eignen. Da Musik immer möglichst kompatibel mit allen Wiedergabesystemen sein soll, damit sie potentiell auch in einem Club oder als Soundtrack für eine TV-Werbung oder einen Film genutzt werden kann, stellt dies einen nicht unerheblichen Nachteil dar. Darüber hinaus haben die meisten Musikkonsumenten im Bereich der Pop- und Rockmusik sehr ausgeprägte Hörgewohnheiten, von denen sie nur ungerne abrücken. Aus diesem Grund werden herkömmliche, druckvolle und direkte Stereomischungen häufig räumlichen, immersiven binauralen Mixen vorgezogen, vor allem da die Binauraltechnik bisher meist mit wahrnehmbaren Klangverfärbungen einhergeht.⁴

Diese Ansprüche haben die Musiker und Musikkonsumenten allerdings in erster Linie an klassische Studioproduktionen, also Alben, EPs und Singles, welche früher in Form von CDs oder Vinylplatten und heutzutage immer häufiger durch Streaming-Dienste konsumiert werden. Bei Musikformaten, die davon abweichen, also beispielsweise Konzertmitschnitte oder Livesessions, ist von einer deutlich höheren Toleranz seitens der Musiker und Endverbraucher gegenüber neuen, experimentellen Ansätzen auszugehen. Dies liegt daran, dass es sich hier um Formate handelt, die bereits bestehende Studiosongs live in einer alternativen, ungeschliffenen Version präsentieren. Hier ist also in der Regel sogar gewünscht, dass sich die Mischung von der Studioproduktion unterscheidet und einen musikalischen oder klanglichen Mehrwert bietet.

Für die Binauraltechnik scheint dies ein sehr gutes Einsatzgebiet und möglicherweise auch ein geeigneter Zugang zur modernen Popmusik zu sein, um sich dort auch

² Vgl. Bitkom Research. (2019).

³ Vgl. Statista GmbH. (2019). Kopfhörer

⁴ Vgl. Pfeifer. (2018). S.76f

langfristig und massentauglich zu etablieren. Aus diesem Grund soll im Rahmen dieser wissenschaftlichen Arbeit eine binaurale Livesession im Kontext der Popmusik konzipiert und umgesetzt werden. Hierbei sollen sowohl die erforderlichen organisatorischen Schritte als auch die technische Planung und Realisierung des Projekts nachvollziehbar dargelegt und erläutert werden. Dabei sollen vor allem Schwierigkeiten und Problemstellungen aufgeworfen, sowie Lösungsansätze gefunden und getestet werden.

1.2. Vorgehen

Im ersten Teil der Arbeit werden zunächst die relevanten wissenschaftlichen Grundlagen behandelt. Dabei wird zuerst das Medium der Livesession ausführlich erläutert. Im Zuge dessen wird zunächst der Begriff hergeleitet und definiert, anschließend wird die Entstehung und die Relevanz der Livesession in der Musikindustrie dargelegt. Abschließend werden aktuelle Trends und Entwicklungen aufgezeigt.

Das zweite Grundlagenkapitel behandelt die Binauraltechnik. Dabei wird zunächst deren Basis, das räumliche Gehör des Menschen, behandelt. Im Anschluss wird das technische Funktionsprinzip der Binauraltechnik genauer untersucht. Zum Abschluss des Kapitels werden binaurale Mikrofonsysteme und die Binauralsynthese beleuchtet. und anschließend die Binauraltechnik detailliert beschrieben.

Im zweiten Teil dieser wissenschaftlichen Arbeit folgt dann der praktische Teil in Form der Konzeption und Umsetzung der binauralen Livesession. Dabei werden die drei wesentlichen Produktionsschritte Vorbereitung, Aufnahme und Postproduktion ausführlich dargelegt und erläutert. Abschließend folgt ein Fazit, in dessen Rahmen das Ergebnis und der Prozess der durchgeführten Produktion kritisch hinterfragt und analysiert wird.

Da in dieser wissenschaftlichen Arbeit der Autor auch der ausführende Toningenieur bei der Produktion der binauralen Livesession ist, beziehen sich diese beiden Begriffe im Kontext der Erläuterungen und Ausführungen auf dieselbe Person. Ist also von dem Verfasser, dem Toningenieur oder dem Audioingenieur die Rede, ist stets der Autor dieser Arbeit gemeint.

2. Das Medium Livesession

In dem folgenden Kapitel wird das Medium Livesession genauer untersucht und beleuchtet. Im Zuge dessen folgt zunächst eine Begriffserklärung und Herleitung. Anschließend wird die Entstehung und die Rolle der Livesession in der modernen Musikindustrie ins Auge gefasst. Zum Abschluss werden schließlich aktuelle Trends und Entwicklungen hinsichtlich des Mediums aufgezeigt.

2.1. Begriffserklärung

Der Begriff „Livesession“ setzt sich aus den beiden englischsprachigen Worten „live“ und „session“ zusammen. Der erstgenannte Ausdruck lässt sich dabei mit den Adjektiven „lebendig“, „lebend“ oder „aktiv“ übersetzen, kann jedoch in anderem Kontext auch als „stromführend“, „unter Spannung“ stehend oder „spannungsführend“ ausgelegt werden. Im deutschen Sprachgebrauch hat der Begriff „live“ zwei Bedeutungen: Einerseits steht er im Bereich des Rundfunks und des Fernsehens für eine Übertragung als Direktsendung, andererseits kann er eine reale Anwesenheit, beispielsweise einer Person oder einer Personengruppe, verdeutlichen.⁵ Der Begriff „session“ lässt sich aus dem Englischen als „Sitzung“, „Versammlung“, „Tagung“ oder auch „Behandlung“ ins Deutsche übersetzen. Der Ausdruck „Session“ findet sich jedoch bereits seit einiger Zeit auch im deutschen Vokabular und wird im Duden als „sich über einen längeren Zeitraum erstreckende Tagung, Sitzungsperiode“ definiert.⁶

Der musikalisch geprägte Begriff „Livesession“, oftmals auch „Live-Session“ oder „Live Session“, hat in der Musikkultur zweierlei Bedeutungen. Ursprünglich diente er als ein Synonym für die sogenannte Studio Session, zu der sich ein loser Verbund von Musikern zum gemeinsamen Musizieren in einem Tonstudio trifft. Es handelt sich dabei um einen kreativen Prozess, dessen Ziel es ist, einen qualitativ hochwertigen und künstlerisch wertvollen Output zu generieren, durch den die beteiligten Künstler einen finanziellen

⁵ Vgl. Bibliographisches Institut GmbH. (2019). *live*.

⁶ Bibliographisches Institut GmbH. (2019). *Session, die*.

Mehrwert gewinnen können. Im Gegensatz zu der potentiell bekannteren Jam Session, bei der sich ebenfalls Musiker zum gemeinsamen Musizieren treffen, findet dies hier also nicht in einem zwanglosen Rahmen statt, sondern dient in der Regel der Sicherung des finanziellen Lebensunterhalts der Beteiligten.

Angesichts der Bedeutung der beiden Wortkomponenten „live“ und „session“ ist dieser Begriff für diesen Prozess also schlüssig herzuleiten: Es handelt sich um eine Versammlung von Musikern in einem Studio, mit dem Ziel, durch reale Anwesenheit und das damit verbundene Agieren und Reagieren aufeinander, gemeinsam einen kreativen Synergieeffekt zu erzielen.

Darüber hinaus beschreibt die Bezeichnung „Livesession“ in der Musikkultur allerdings auch die audiovisuelle Aufzeichnung einer musikalischen Live-Performance. Dieser Vorgang findet in der Regel ohne Publikum vor Ort statt und unterscheidet sich so klar von einem klassischen Konzertmitschnitt. Die musikalische Darbietung dient also in erster Linie dem Erstellen des audiovisuellen Endprodukts und wird erst später einem breiten Publikum zugänglich gemacht. Dies wird in der Regel durch Videoplattformen wie YouTube oder Vimeo, oder auch durch soziale Netzwerke wie Facebook oder Instagram bewerkstelligt.

Der Begriff lässt sich folglich auch hier schlüssig herleiten: Die Wortkomponente „live“ steht für Livemusik und signalisiert dem späteren Betrachter, dass die Künstler bei der Aufzeichnung, wie bei einem Konzert, in Echtzeit gemeinsam performt und miteinander interagiert haben. Das Wort „session“ beschreibt auch hier schlicht das Zusammenkommen der Musiker, in diesem Fall auch mit den Ton- und Videoschaffenden, die das audiovisuelle Endprodukt technisch realisieren.

Beide Bedeutungen für den Begriff lassen sich also logisch nachvollziehen und legitimieren. Sofern nicht anders angegeben, ist in dieser Arbeit jedoch stets von der zweitgenannten Form der Livesession die Rede.

2.2. Entstehung und Relevanz in der Musikindustrie

Die Livesession ist ein noch recht junges Medium in der Musikkultur. Dieser Umstand liegt allerdings weniger daran, dass die Idee und das Konzept dahinter besonders

innovativ oder bahnbrechend waren, sondern vielmehr an zwei entscheidenden Faktoren, die die Musikindustrie in der Folge nachhaltig beeinflussten und prägten.

Vor dem Einzug der Digitaltechnik in der Audio- und Videotechnik war es schlichtweg sehr kostenintensiv und aufwendig, eine solche Livesession umzusetzen. Der Fokus von Künstlern und Labels lag daher in audioteknischer Hinsicht ganz klar auf der Produktion von Tonträgern, die für den physischen Verkauf vorgesehen waren. Videotechnisch war es wiederum deutlich sinnvoller, Musikvideos oder Konzertmitschnitte zu realisieren, da diese potentiell im Fernsehen liefen oder als Konzertfilm vertrieben werden konnten. Folglich wurden dafür auch alle verfügbaren finanziellen Ressourcen aufgewendet.

Da eine Livesession vermutlich weder im Fernsehen noch im Kino - den damals vorherrschenden visuellen Massenmedien - die Chance auf eine Ausstrahlung gehabt hätte, machte es also rein finanziell gesehen absolut keinen Sinn, viel Geld in die Produktion eines solchen audiovisuellen Live-Videos zu stecken.

Nachdem die Digitaltechnik in den 1990er-Jahren sowohl in der Audio- als auch in der Videotechnik, Stück für Stück die Analogtechnik ablöste, sorgte dies dafür, dass die Produktionskosten für auditive und visuelle Werke sanken. Darüber hinaus wurden in den beiden Bereichen zunehmend massentaugliche Produkte für den breiten Markt entwickelt, die es auch semiprofessionellen Ton- und Filmschaffenden ermöglichte, relativ professionelle Ergebnisse zu erzielen. Dadurch eröffneten sich auch für viele unabhängige Künstler ohne großes finanzielles Budget zahlreiche Möglichkeiten.⁷

Die zweite technische und gesellschaftliche Errungenschaft, die den Weg für das Medium der Livesession ebnete, war das Internet. Mit dem Durchbruch des World Wide Web entstand ein neuer, immer wichtiger werdender Eckpfeiler der modernen Musikindustrie. Durch soziale Netzwerke, wie Myspace, Facebook oder Instagram, vor allem aber durch Videoplattformen, wie YouTube, taten sich in den 2000er-Jahren aus dem Nichts wichtige neue Vermarktungskanäle für Künstler und Musiklabels auf.⁸

Vor allem YouTube wuchs im Zuge dessen sehr schnell und wurde zur relevantesten Plattform für Musikvideos, da diese hier jederzeit nach Bedarf angeschaut werden konnten. Dieses Forum stellte herkömmliche große Musiksender, wie VIVA und MTV, in

⁷ Vgl. Westreicher. (2015). S. 116

⁸ Vgl. Anastasiadis. (2019). S. 1

den Schatten.⁹ Darüber hinaus war es nun jedem Menschen mit einem kostenfreien Account möglich, Inhalte auf der Plattform kostenlos hochzuladen. Nun war also nicht mehr der große Einfluss eines finanzstarken Musiklabels oder Vertriebs nötig, um die eigene Musik, inklusive Video, der ganzen Welt zur Verfügung zu stellen.

Die zwei essentiellen Faktoren, die dem Medium Livesession den Weg ebneten, waren also die Durchbrüche der Digitaltechnik und des Internets. Dadurch wurden alte Strukturen aufgebrochen, und neue Produktions- und Vermarktungswege eröffneten sich, die es der breiten Masse an Künstlern ermöglichte, professionell zu arbeiten und sich selbst zu vermarkten.¹⁰ Durch die angesprochenen sinkenden Kosten für eine professionelle, audiovisuelle Produktion und den neuen Vertriebskanälen für Musikvideos aller Art, wurden so in den 2000er-Jahren die ersten professionellen Livesessions realisiert und veröffentlicht.

Vorreiter und Pionier war dabei im Jahr 2002, bereits vier Jahre vor der Gründung von YouTube, der US-amerikanische Medienkonzern AOL, ehemals America Online, mit der Videoreihe „sessions@AOL“. Die Videos waren zunächst nur Kunden des Unternehmens vorbehalten, wurden später aber auf deren Homepage jedem frei zur Verfügung gestellt. Bis zur Einstellung des Formats im Jahr 2012 wurden zahlreiche Künstler diverser Genres live aufgezeichnet. Darunter Bands, wie Coldplay oder Fleetwood Mac, aber auch Solokünstler, wie Lady Gaga oder Kanye West.¹¹ Die Intention hinter der Videoreihe war dabei eine engere Kundenbindung durch exklusive Live-Performances und Interviews, sowie die Verknüpfung der Marke AOL mit einem populären, kunstaffinen Image.

Der große Durchbruch für das Format der Livesession begann allerdings erst gegen Ende der 2000er-Jahre. Ein wichtiger Faktor waren hierbei US-amerikanische Radiosender wie KEXP oder NPR (National Public Radio), welche mit eigenen YouTube-Kanälen damit begannen, regelmäßig Sessions mit Musikern, die in ihrem Studio zu Gast waren, auch visuell aufzuzeichnen und hochzuladen. Heute zählen sie mit jeweils mehreren Millionen Abonnements und jeweils annähernd einer Milliarde Videoaufrufen

⁹ Vgl. Anastasiadis. (2019). S. 84

¹⁰ Vgl. Anastasiadis. (2019). S. 1ff

¹¹ Vgl. Wikipedia. (2019). *Sessions@AOL*.

zu den erfolgreichsten Plattformen im Segment der Livesessions.^{12 13} Ein Vorläufer dieser Konzepte stellten die sogenannten „Peel Sessions“ des BBC-Moderators und Radio-DJs John Peel dar. Im Zuge dessen lud dieser ab Beginn der 1960er-Jahre prominente Musiker in sein BBC-Studio ein und nahm mit ihnen Livesessions auf, welche im Radio ausgestrahlt wurden.¹⁴ Dabei handelte es sich jedoch folglich lediglich um auditive Mitschnitte ohne Videomaterial.

Auch die sozialen Netzwerke sorgten durch ihre steigende Relevanz im Gesellschafts- und Sozialverhalten der Menschen, vor allem der jungen Erwachsenen, für einen Bedarf nach neuen Formaten und Inhalten. Vielen Künstlern wurde im Zuge dessen recht schnell bewusst, dass sie mittels Social Media eine enge Fanbindung erreichen und unkonventionell und schnell neue Anhänger gewinnen konnten. Durch soziale Netzwerke konnten sie sich künftig nahbar zeigen und eine gewisse Intimität suggerieren.¹⁵ In dieses derzeit wichtiger denn je erscheinende Konzept passt auch das Format der Livesession sehr gut, da die Künstler sich dort sehr authentisch und echt präsentieren können. Darüber hinaus bekommt der Fan von ihnen eine ehrliche Live-Performance, ganz intim und ohne Publikum.

Nicht unerheblich ist außerdem auch, viele Inhalte in der Rückhand zu haben, mit der man als Künstler seinen Feed auf den sozialen Netzwerken bedienen kann. Lässt man zu lange nichts von sich hören, wird man von den Fans schnell abgeschrieben und gerät in die Bedeutungslosigkeit. Adam Alpert, der Manager des erfolgreichen US-amerikanischen DJ-Duos The Chainsmokers formulierte dies folgendermaßen: “To stay hot, you have to deliver to your fans, and they consume quickly, [...] It’s not just about delivering music. It’s about delivering all types of content.”¹⁶ Aufgrund des recht geringen finanziellen und kreativen Aufwands, im Vergleich zur Produktion eines Musikvideos oder eines völlig neuen Songs im Studio, bieten sich Livesessions daher ideal an, um als Musiker schnell hochwertige Inhalte zu produzieren. Neben Interviews und Artikeln, für die jedoch ein kooperativer Partner bei der Presse erforderlich ist, etablierte sich die Livesession so schnell als wichtiges Vermarktungstool.

¹² Vgl. YouTube. (2020). *KEXP*

¹³ Vgl. YouTube. (2020). *NPR Music*

¹⁴ Vgl. Anastasiadis. (2019). S. 84

¹⁵ Vgl. Anastasiadis. (2019). S. 91-104

¹⁶ Vgl. Alpert. (2015).

Eine weitere nachhaltig wirkende Entwicklung in der Musikindustrie ist das sich wandelnde Hörverhalten der Musikkonsumenten. Kommerzielle Musik-Streamingdienste, beispielsweise das 2006 gegründete schwedische Unternehmen Spotify, bieten eine riesige Auswahl an Musik online gegen einen festen monatlichen Abonnement-Preis an. Wer das Angebot kostenlos wahrnehmen möchte, muss Werbeanzeigen zwischen den Songs hören.¹⁷ Pro Stream bekommen die Künstler dabei Geld vom Anbieter ausgeschüttet. In Deutschland funktioniert das Konzept, die Nutzerzahlen der Streamingdienste steigen stetig an (siehe Abb. 1) und die Musikindustrie erlebt einen kleinen wirtschaftlichen Aufschwung. Durch diese Entwicklung verzeichnete der Umsatz mit Kaufmusik beispielsweise 2013 in Deutschland nach 15 Jahren erstmals wieder einen Anstieg, anstatt weiter zu stagnieren.¹⁸

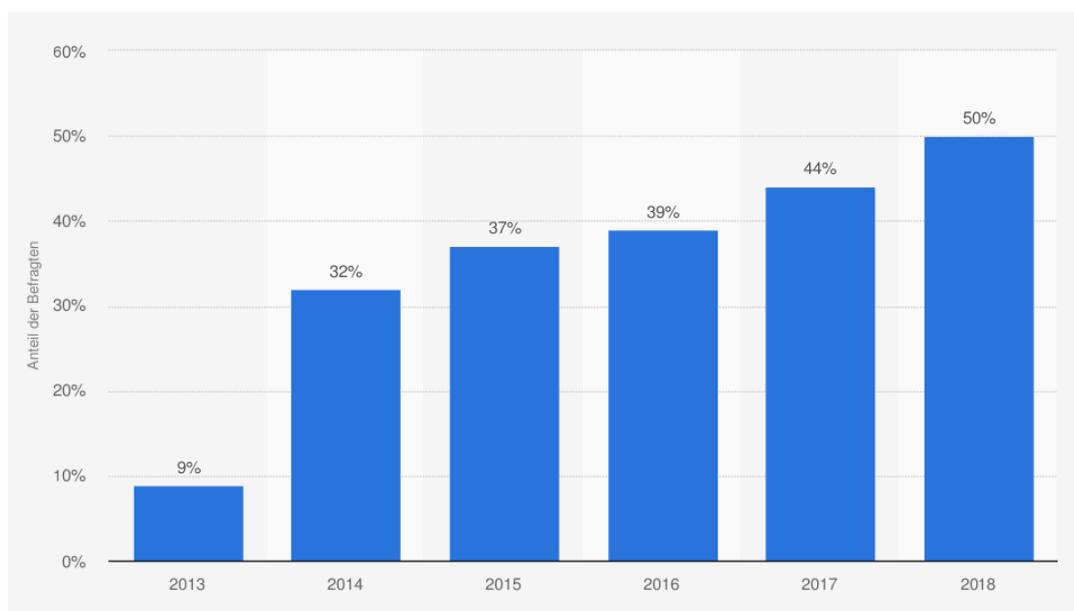


Abb. 1: Anteil der befragten Internetnutzer, die Musikstreaming-Dienste nutzen, in Deutschland in den Jahren 2013 bis 2018 ¹⁹

Durch die Streamingdienste änderte sich das Hörverhalten der Konsumenten in den letzten Jahren deutlich. Immer mehr Menschen konsumieren Musik nicht mehr in Form von durchgängigen Alben, sondern durch Playlisten, die aus einzelnen Songs vieler

¹⁷ Vgl. Spotify AB. (2020).

¹⁸ Vgl. Lischka. (2014).

¹⁹ Bitkom Research. (2018).

unterschiedlicher Künstler bestehen. Diese Playlists werden in der Regel vom Streamingdienst kuratiert und zusammengestellt. Nach Angaben der DiMA (Digital Media Association) im März 2018, gaben mehr als die Hälfte der befragten Nutzer an, dass Playlists für sie Alben ersetzen.²⁰

Ein echter Bezug zum Künstler geht dabei häufig verloren, da viele Musikkonsumenten zwar einzelne Songs kennen, sich aber selten mit der Gesamtheit des kreativen Schaffens eines Musikers auseinandersetzen. Folglich produzieren viele Künstler und Plattenfirmen zunehmend einzelne für sich stehende Songs.²¹ Ziel ist es dabei in der Regel, in einer hörerreichen Playlist platziert zu werden. Ist dies der Fall, soll aus dem Titel maximal viel Profit geschlagen werden. Hier bieten sich Livesessions als ideales Mittel an, um den Song zu vermarkten, alternative Versionen zur Verfügung zu stellen und sich als Künstler hinter dem Song zu präsentieren.

2.3. Aktuelle Trends und Entwicklungen

Beobachtet man derzeitige Trends im Bereich der Livesession, so fällt auf, dass neben etablierten Anbietern des Segments, wie die bereits angesprochenen Radiosender oder auch langjährig aktive YouTube-Kanäle, zunehmend auch finanzstarke Akteure in den Markt drängen.

So baut das österreichische Unternehmen Red Bull seine Markenpräsenz im Musikmarkt unter anderem durch Livesessions, welche teils auf der eigenen Webseite, als auch auf YouTube veröffentlicht werden, weiter aus. Darüber hinaus haben auch Major Labels das Potential des Formats erkannt. Die Plattenfirma Sony Music publiziert beispielsweise auf deren YouTube-Kanal „Filtr Germany“ neben Musikvideos zunehmend auch Livesessions ihrer Künstler.

Außerdem ist auch der Trend zu erkennen, dass immer mehr Musikmagazine und Tonstudios das Medium Livesession zur Selbstvermarktung nutzen.

Sinnbildlich für das steigende Interesse an Livesessions ist darüber hinaus der YouTube-Kanal „COLORS“ der in Berlin ansässigen COLORSxSTUDIOS GmbH. Seit seiner

²⁰ Vgl. Digital Media Association. (2018).

²¹ Vgl. Stern.de. (2018)

Gründung im Februar 2016 konnte der Channel über vier Millionen Follower aufbauen und circa 1,2 Milliarden Klicks auf seinen Videos verbuchen.²² Damit übertrifft er die statistischen Werte zahlreicher anderer langjährig aktiver, großer Livesession-Kanäle innerhalb kurzer Zeit und etabliert sich als derzeit größte Plattform für dieses Format. Ein weiterer Trend, der sich abzeichnet, ist die örtliche Verschiebung vieler Livesessions von klassischen Tonstudiumgebungen hin zu außergewöhnlichen Locations. So lassen sich auf Plattformen wie YouTube mittlerweile Videos in verschiedensten räumlichen Begebenheiten finden. Angefangen beim Stuttgarter Marmorsaal, über Hochhausdächer in Bangkok bis hin zur bolivianischen Salzwüste Salar de Uyuni (siehe Abb. 2) sind dabei scheinbar keine Grenzen gesetzt. Der Grund dafür ist offensichtlich, dass den Konsumenten auch visuell etwas geboten werden soll. Da dieser Trend scheinbar sehr positiv aufgefasst wird, scheint es für die meisten Zuschauer dabei zweitrangig zu sein, dass dies in der Regel auf Kosten der Audioqualität geschieht.



Abb. 2: Livesession des Künstlers FKJ in der Salzwüste Salar de Uyuni²³

Der nächste logische Schritt in der Entwicklung der Livesession ist nun das Live-Streaming. Vereinzelt wird dies auch schon praktiziert, allerdings kann man hier noch nicht von einem klaren Trend sprechen. Ein großes Ausrufezeichen setzte in diesem Kontext die britische Band Coldplay, welche zum Release ihres im November 2019

²² Vgl. YouTube. (2020). COLORS

²³ Cercle. (2019)

erschienenen Albums „Everyday Life“ eine halbstündige Livesession der neuen Songs von einem Dach in Jordanien live auf YouTube übertrug. Ein solches Unterfangen stellt natürlich eine enorme logistische Herausforderung dar, zumindest wenn man ein gewisses Maß an Qualität gewährleisten will, und wird so zumindest vorläufig nur großen, finanzstarken Akteuren vorbehalten bleiben. Dennoch gibt es vereinzelt auch kleinere Künstler, welche diesen Ansatz mit weniger Aufwand und kreativen Lösungen umsetzen.

3. Die Binauraltechnik

In dem folgenden Kapitel wird das dreidimensionale Wiedergabeverfahren der Binauraltechnik genauer untersucht. Dieses beruht auf den Prinzipien des räumlichen Gehörs des Menschen. Dabei wird die Tatsache ausgenutzt, dass im Schalldruckverlauf an den Trommelfellen alle für das räumliche Hören einer auditiven Umgebung benötigten Information enthalten sind.²⁴

Eberhard Sengpiel beschrieb die theoretische Grundlage der Binauraltechnik dabei wie folgt: „Am Eingang des Gehörs liegen zwei Signale, nämlich der modulierte Schalldruck am jeweiligen Trommelfell. Wenn diese beiden Signale mit unendlich kleinen Mikrofonen am Gehörgangseingang aufgenommen und genau an dieser Stelle über Kopfhörer wiedergegeben werden, dann müsste das Originalhörereignis mit der richtigen Klangfarbe und dem richtigen Raumeindruck wieder erscheinen.“²⁵ Folglich verfolgt die Binauraltechnik das Ziel, die am Trommelfell anliegenden Ohrsignale des Menschen künstlich zu reproduzieren. Im Zuge dessen werden spezielle Mikrofonsysteme und Software eingesetzt. Es kommt also nicht zur Synthese eines ganzen Schallfeldes, sondern zur Synthese der am Trommelfell anliegenden Schalldruckpegel.²⁶

Das Verfahren selbst wird bereits seit den 1970er-Jahren im Rundfunkbereich eingesetzt und funktioniert in der Regel nur mittels einer Wiedergabe über Kopfhörer. Dies unterscheidet die Binauraltechnik von den anderen dreidimensionalen Wiedergabeverfahren, die allesamt lautsprecherbezogen sind. Aufgrund dieser Sonderstellung wird die Binauraltechnik auch in Zukunft eine wichtige Rolle im Bereich der 3D-Audiotechnik spielen.

Im Laufe des Kapitels wird zunächst genauer auf das räumliche Hören des Menschen eingegangen, und anschließend wird das Funktionsprinzip der Binauraltechnik untersucht. Des Weiteren werden binaurale Mikrofonsysteme und der Prozess der Binauralsynthese beleuchtet.

²⁴ Vgl. Weinzierl. (2008). S.671

²⁵ Sengpiel. (1995).

²⁶ Vgl. Weinzierl. (2008). S.671

3.1. Räumliches Hören

Für den Menschen wahrnehmbare Schallwellen treten im alltäglichen Umfeld im Normalfall als Luftdruckschwankungen auf, die das Ohr als adäquaten Reiz wahrnimmt.²⁷ Diese Änderungen des Luftdrucks wandelt es in einem Schallfeld in Nervenreize um, die das Gehirn weiterverarbeiten kann.²⁸ Die Anatomie dieses menschlichen Sinnesorgans spielt dabei eine entscheidende Rolle hinsichtlich der Fähigkeit des räumlichen Hörens.

Spricht man vom räumlichen Hören oder auch Richtungshören, geht es - wie der zweite Begriff schon vermuten lässt - darum, die Richtung, sowie die Entfernung einer Schallquelle lokalisieren zu können. Darüber hinaus bezeichnet der Begriff auch die Befähigung des Menschen, durch das Ohr Informationen zur Größe des Raums, in dem er sich befindet, zu gewinnen.²⁹

Eine essenziell wichtige Rolle übernimmt dabei das Außenohr. Es besteht aus der Pinna, auch Ohrmuschel genannt, und dem Gehörgang. Es nimmt die Schallwellen auf und leitet diese an das Trommelfell im Mittelohr weiter.³⁰ Dort wird die Schallenergie in mechanische Schwingung übersetzt, die im Innenohr in elektrische Impulse umgewandelt und schließlich vom Gehirn verarbeitet werden. Das Frequenzspektrum der eintreffenden Schallenergie wird auf dem Weg durch die Ohrmuschel und den Gehörgang dabei in einem so großen Maß verzerrt, dass das Gehirn diese Änderungen registriert und daraus Informationen zur Lokalisation eines Schallereignisses gewinnen kann. Die Pinna und der Gehörgang bilden im Zuge dessen als Einheit ein extrem detailliertes Resonanzsystem, welches Verzögerungen von bis zu 300 μ s erzeugen kann, welche das Spektrum entsprechend beeinflussen. Entscheidend für das Ausmaß der Verzögerung ist dabei die Richtung, aus der die Schallwellen auf das Außenohr treffen.³¹ Das spektral veränderte Signal, das dann auf das Trommelfell trifft, wird auch als Ohrsignal bezeichnet. Da beim binauralen Hören Schall mit zwei Ohren wahrgenommen wird, ist folglich von Ohrsignalen die Rede.

²⁷ Vgl. Zenner. (2006). S.287

²⁸ Vgl. Webers. (2007). S. 93

²⁹ Vgl. Fleck. (2020)

³⁰ Vgl. Blauert. (1974). S.4

³¹ Vgl. Blauert. (1974). S. 55

Im Zuge der folgenden Erläuterungen zum Räumlichen Hören wird die Terminologie nach Jens Blauert verwendet. Dieser ordnet Begriffe mit dem Anhang „Hör-“ klar wahrnehmungsbedingten Termini zu und solche mit dem Anhang „Schall-“ im Gegenzug physikalischen Ausdrücken.³²

Am Beispiel der Ohrsignale lässt sich diese Terminologie gut veranschaulichen: Werden diese nur leicht abgeändert, ist das Gehirn nicht mehr in der Lage, das Schallereignis, also das reale, physikalisch messbare Ereignis, richtig zu verorten. Folglich stimmt das vom Menschen wahrgenommene Hörereignis, also dem Ereignis, welches das Gehirn aus den vorhandenen Informationen rekonstruiert, nicht mit dem tatsächlichen Schallereignis überein.

Es kann also zu einer Diskrepanz zwischen Schall- und Hörereignis kommen. Ausschlaggebende Faktoren hierfür können beispielsweise spezielle Klangcharakteristiken einer Schallquelle oder akustische Eigenschaften eines Raumes sein, die die menschliche Wahrnehmung täuschen. Logischerweise ist der Mensch daher bei der Lokalisation eines Schallereignisses im Alltag häufig auf zusätzliche Informationen anderer Sinnesorgane, wie den Augen, angewiesen, um eine schnelle und zuverlässige Verortung zu gewährleisten.

3.1.1. Auditive Richtungslokalisierung

Hinsichtlich der nachfolgenden Erklärungen zur auditiven Richtungslokalisierung von Schallereignissen, wird auch im folgenden Abschnitt ein enger Bezug zu Blauerts Erläuterungen bezüglich des räumlichen Hörens präsent sein. Aus diesem Grund wird das von ihm genutzte kopfbezogene Koordinatensystem miteinbezogen, welches zur einfacheren Veranschaulichung dienen soll. (siehe Abb. 3)

Es besteht aus einer Horizontal-, einer Median- und einer Frontalebene. Die Medianebene bildet dabei die Symmetrieebene. Sie unterteilt den Kopf dementsprechend in eine rechte und eine linke Seite und liegt genau auf der Blickachse. Die Horizontalebene liegt orthogonal zur Medianebene und befindet sich genau auf der interauralen Achse, also jener, die durch beide Ohren verläuft. Sowohl die Median-

³² Vgl. Blauert. (1974). S. 2

als auch die Horizontalebene beinhalten im Zuge dessen auch Höheninformationen, die Horizontalebene im Gegensatz dazu nicht. Sie verläuft sowohl auf der Blick- als auch auf der interauralen Achse und teilt den Kopf so auf der Höhe der Ohren in einen oberen und einen unteren Teil.³³

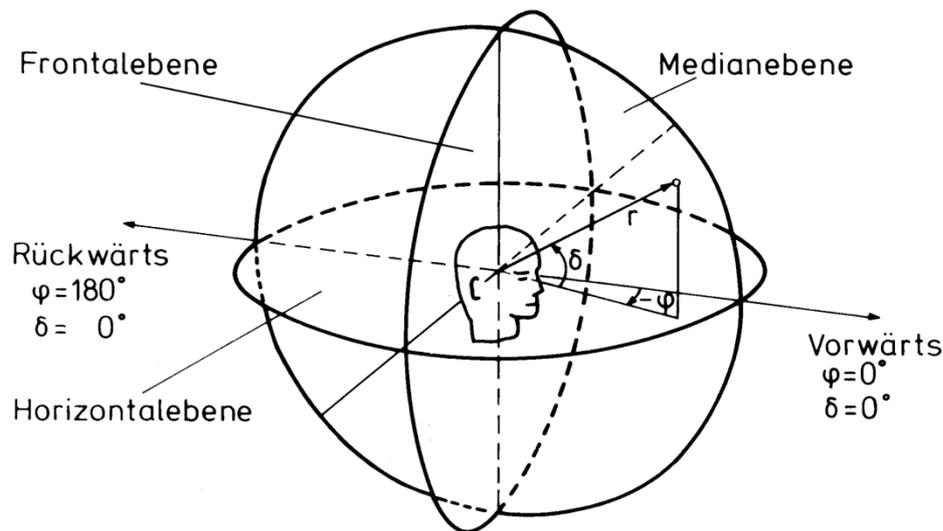


Abb. 3: Kopfbezogenes Koordinatensystem³⁴

Um genauer auf die auditive Richtungslokalisierung einzugehen, wird zunächst der Unterschied zwischen den Termini Lokalisation und Lokalisationsschärfe unterteilt. Unter dem erstgenannten Begriff versteht Jens Blauert dabei im Wesentlichen die Zuordnung eines Hörereignisses zu einem bestimmten Schallereignis. Genauer gesagt definiert er die Lokalisation als ein „Zuordnungsgesetz oder -regel [...] zwischen dem Ort eines Hörereignisses (z. B. bezüglich Richtung und/oder Entfernung) und einem bestimmten Merkmal oder bestimmten Merkmalen eines Schallereignisses oder eines anderen, mit dem Hörereignis korrelierten Ereignisses. [...]“.³⁵ Der Terminus der Lokalisationsunschärfe steht seiner Auffassung nach wiederum für die bei der Lokalisation mögliche Genauigkeit. Sie beschreibt die „Kleinste Änderung eines bestimmten Merkmals oder bestimmter Merkmale des Schallereignisses oder eines anderen, mit dem Hörereignis korrelierten Ereignisses, die gerade zu einer

³³ Vgl. Blauert. (1974). S. 11f

³⁴ Blauert. (1974). S. 11

³⁵ Blauert. (1974). S. 30

Ortsänderung des Hörereignisses (z. B. Bezüglich Richtung und/oder Entfernung) führt. Die Lokalisationsunschärfe ist eine Eigenschaft der Lokalisation.“³⁶

Die Basis für die Lokalisation einer Schallquelle beim binauralen Hören stellt hauptsächlich der Vergleich der beiden Ohrsignale dar. Erwähnenswert ist dabei, dass für den Menschen die horizontale Verortung evolutionär bedingt schon immer eine größere Relevanz hatte. Aus diesem Grund ist diese, gegenüber der Lokalisation auf der Medianebene, deutlich sensibler und besser ausgeprägt.³⁷ Der Mensch nutzt dabei zur Verortung eines Schallereignisses auf der Horizontalebene in der Regel zwei physiologisch bedingte Gegebenheiten: Die interauralen Pegeldifferenzen, kurz ILD (Interaural Level Difference), und die interauralen Laufzeitunterschiede, kurz ITD (Interaural Time Difference).³⁸ Beide beruhen dabei auf der Entfernungsdifferenz der Schallquelle zu den beiden Ohren. Ist der Abstand zu beiden Ohren gleich groß, treten weder ILDs noch ITDs auf. In diesem Fall muss sich die Schallquelle folglich genau auf der Medianebene befinden.

Die interauralen Pegeldifferenzen ergeben sich, wie der Name schon vermuten lässt, durch unterschiedliche Schalldruckpegel an den beiden Ohren. Das schallabgewandte Ohr befindet sich dabei im sogenannten Schallschatten des Kopfes. (siehe Abb. 4) Dieser entsteht durch die durch den Kopf verursachte Dämmung, da dieser sich zwischen der Schallquelle und dem entsprechenden Ohr befindet.³⁹ Folglich empfängt dieses Ohr einen wesentlich geringeren Schallpegel als das der Schallquelle nähere Ohr. Diese Differenz des Schallpegels wertet das Gehirn aus und kann so Informationen zur Lokalisation auf der Horizontalebene gewinnen. Der Grad der Absorption durch den Kopf hängt dabei von der Frequenz ab. Je kleiner die Frequenz ist, desto größer wird die Wellenlänge und je größer die Wellenlänge im Vergleich zu dem Hindernis ist - also in diesem Fall dem Kopf - desto größer werden die auftretenden Beugungseffekte.⁴⁰ Aus diesem Grund ist es Schallwellen, deren Frequenz unter 300 Hz liegt, möglich, sich um den Kopf herum zu beugen. Die Intensitätsunterschiede der Ohrsignale nehmen mit

³⁶ Blauert. (1974). S. 30

³⁷ Vgl. Guski. (1996). S. 273

³⁸ Vgl. Steckmann. (2018). S. 53

³⁹ Vgl. Friesecke. (2014). S. 135

⁴⁰ Vgl. Webers. (2007). S. 76

steigender Frequenz zu.⁴¹ Dementsprechend ist die Lokalisationsschärfe bei tieferen Frequenzen deutlich geringer als bei höheren Frequenzen.

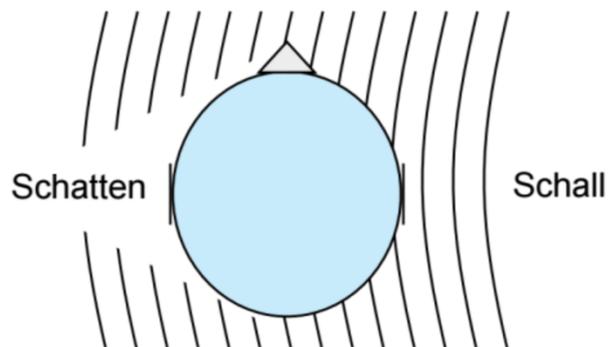


Abb. 4: Interaurale Pegeldifferenzen durch Abschattung des Kopfes⁴²

Die interauralen Laufzeitunterschiede sind für die Richtungslokalisierung auf der Horizontalebene ebenfalls sehr wichtig. Befindet sich eine Schallquelle unterschiedlich weit entfernt von den beiden Ohren, kommt es zu einem zeitlichen Versatz der beiden Ohrsignale, die das Gehirn auswerten und zur Richtungslokalisierung nutzen kann. (siehe Abb. 5)

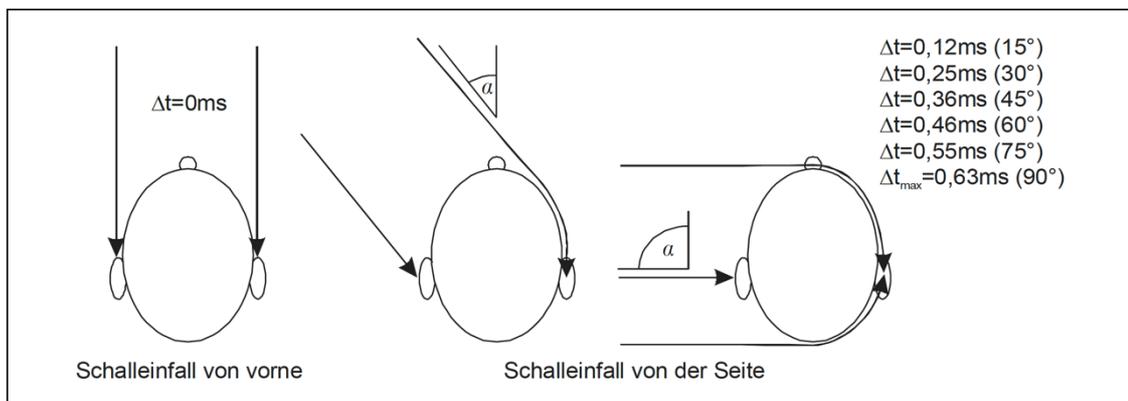


Abb. 5: Lokalisation durch Laufzeitunterschiede⁴³

Je größer der Unterschied der Entfernung zur Schallquelle ist, desto größer fallen diese Laufzeitunterschiede aus.⁴⁴ Erzeugt eine Schallquelle seitlich des Kopfes also ein

⁴¹ Vgl. Dickreiter. (2014). S. 130

⁴² Butz & Krüger. (2017). S. 27

⁴³ Friesecke. (2014). S. 135

⁴⁴ Vgl. Friesecke. (2014). S. 134f

Schallereignis, so treffen dessen Schallwellen zunächst auf das der Quelle zugewandte Ohr. Auf das zweite Ohr treffen sie erst mit einem deutlich wahrnehmbaren zeitlichen Unterschied. Dieser zeitliche Versatz lässt sich näherungsweise mit der folgenden Formel berechnen:

$$\Delta t = \frac{d \cdot \pi}{c} \cdot \frac{\alpha}{360^\circ}$$

Der Laufzeitunterschied des Schalls zwischen den beiden Ohren Δt wird also durch den Ohrabstand d , die Schallgeschwindigkeit c und den Einfallswinkel α bestimmt.⁴⁵

Die Lokalisationsunschärfe beträgt dabei ungefähr 2 – 3 Grad und ergibt sich aus der Tatsache, dass der Mensch ab einem zeitlichen Versatz von circa 30 μs dazu in der Lage ist, die Position einer neuen Schallquelle zu bestimmen. Der geringste wahrnehmbare Laufzeitunterschied beträgt sogar lediglich 10 μs .

Der maximal mögliche Laufzeitunterschied beträgt bei exakt seitlichem Schalleinfall auf den Kopf, entlang der interauralen Achse, ungefähr 630 μs . Bei seitlichem Schalleinfall ist die Lokalisationsunschärfe darüber hinaus deutlich größer und liegt bei 12 bis 18 Grad. Dies liegt daran, dass der Schall zum schallabgewandten Ohr zwei Wege nimmt: Einmal über die Stirn und einmal über den Hinterkopf.⁴⁶

Befindet sich eine Schallquelle genau auf der Medianebene und hat dementsprechend den gleichen Abstand zu beiden Ohren, ergeben sich weder interaurale Pegelunterschiede noch interaurale Laufzeitunterschiede. Ähnlich problematisch wird es vermeintlich, wenn zwei Schallquellen aus unterschiedlichen Richtungen der Horizontalebene, beispielsweise aus einem Winkel von 30 Grad und einem Winkel von 150 Grad, ausgehend von der Blickachse, auf die Ohren treffen. Da der zeitliche Versatz sowie die Druckdifferenz der beiden Ohrsignale hier rein theoretisch nämlich bei beiden Schallereignissen gleich wären, müsste es dem Menschen eigentlich schwerfallen, die Signale zu unterscheiden, obwohl eins von rechts-hinten (150 Grad) und eins von rechts-vorne (30 Grad) kommt. Das ist aber nicht so.

⁴⁵ Vgl. Friesecke. (2014). S. 134

⁴⁶ Vgl. Friesecke. (2014). S. 134f

Bei der Richtungslokalisierung auf der Medianebene spielen die Pinna und der Gehörgang durch das gemeinsam gebildete Resonanzsystem eine elementare Rolle. Je nachdem aus welchem Winkel die Schallwellen auf die Ohrmuschel treffen, werden diese vom Resonanzsystem des Außenohrs über das gesamte Frequenzspektrum unterschiedlich beeinflusst.

Mit seinen Blauertschen Bändern, auch richtungsbestimmende Bänder genannt, visualisiert Jens Blauert, welche Frequenzbereiche beim Eintreffen aus verschiedenen Richtungen durch das Außenohr maßgeblich verzerrt werden.

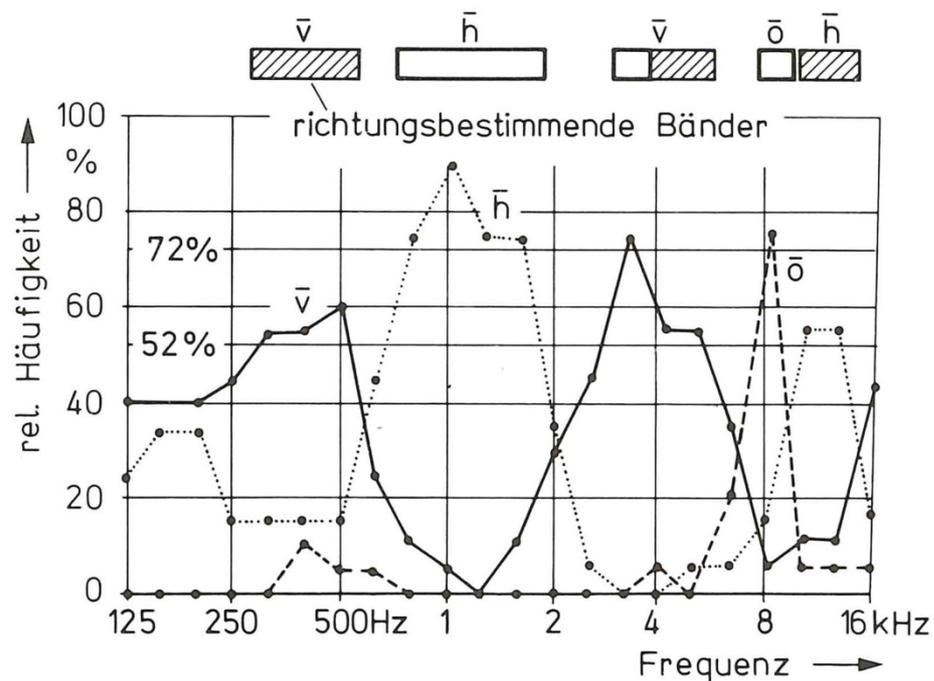


Abb. 6: Richtungsbestimmende Bänder nach Blauert⁴⁷

Dementsprechend kann man schlussfolgern, dass die Filterwirkung des Außenohrs hinsichtlich der auditiven Richtungslokalisierung auf der Medianebene eine elementare Rolle spielt. Sehr wichtig ist jedoch, dass die Signale für diese Form der Richtungsbestimmung ein sehr breites Frequenzspektrum haben sollten, damit die Wirkung des Resonanzsystems eine maximal effektive Wirkung erzielen kann. Ist dies der Fall, kann die Lokalisationsschärfe auf der Medianebene bis zu 4 Grad betragen. Darüber hinaus kann die Hörerfahrung eines Menschen diese weiter verbessern. Kennt

⁴⁷ Blauert. (1974). S. 33

man ein typisches Schallereignis und dessen spektrale Eigenschaften beispielsweise schon gut, fällt die entsprechende Richtungslokalisierung deutlich einfacher.⁴⁸

Ein weiterer wichtiger Faktor zur Richtungslokalisierung auf der Medianebene sind außerdem kleine Mikrobewegungen, die der Mensch intuitiv und unterbewusst mit seinem Kopf macht. Das Ohrsignal ändert sich dadurch logischerweise leicht und liefert so zusätzliche Informationen, die die Unterscheidung von Schallsignalen von vorne oder hinten verbessern. Die Richtungslokalisierung wird dadurch also verbessert und stabilisiert.⁴⁹

In der Theorie werden bei der auditiven Richtungslokalisierung auf der Horizontalebene also eher interaurale Pegeldifferenzen und Laufzeitunterschiede genutzt, wohingegen auf der Medianebene das Resonanzsystem des Außenohrs und entsprechende spektrale Verzerrungen zielführend sind. In der Realität ist es allerdings nicht möglich, klar zu bestimmen, welchen Bezugspunkt das Gehirn zur auditiven Richtungsbestimmung eines Schallereignisses nutzt. Vielmehr ist es stets ein Zusammenspiel aller zur Verfügung stehenden Informationen, welches schlussendlich zu einer bestmöglichen und zuverlässigen Verortung führt. Aufgrund der Tatsache, dass es beispielsweise bei Sinustönen mit einer Wellenlänge, die ein oder mehrfach dem Unterschied der Entfernung der beiden Ohrsignale zur Schallquelle entspricht, kein Phasenversatz auftritt, oder im Gegenzug bei niedrigen Frequenzen keine ausreichenden Intensitätsunterschiede auftreten, ist es auch absolut situations- und signalbedingt, welche Informationen genutzt werden.⁵⁰

Tritt der Fall auf, dass aus unterschiedlichen Richtungen mit zeitlichem Abstand von wenigen Millisekunden gleich mehrere Schallereignisse lokalisiert werden müssen, so greift der Haas-Effekt, welcher auch als Gesetz der ersten Wellenfront bekannt ist. Dieser besagt, dass die dabei früher auf das Trommelfell auftreffenden Schallwellen in diesem Fall richtungsbestimmend sind, auch wenn die nachfolgenden Schallwellen einen höheren Schallpegel haben.⁵¹

⁴⁸ Vgl. Blauert. (1974). S. 84f

⁴⁹ Vgl. Weinzierl. (2008). S. 676f

⁵⁰ Vgl. Webers. (2007). S. 124

⁵¹ Vgl. Blauert. (1974). S. 163

3.1.2. Auditive Entfernungswahrnehmung

Ein weiterer wichtiger Aspekt des räumlichen Hörens ist die auditive Entfernungswahrnehmung, welche im Gegensatz zur auditiven Richtungslokalisation auch weitestgehend monaural funktioniert, also mit nur einem Ohr. Dabei wendet der Mensch drei verschiedene Ansätze an.

In geschlossenen Räumlichkeiten ist naheliegend, dass das Gehirn das Verhältnis zwischen Direkt- und Diffusschall auswertet, um dadurch Schlüsse auf die Entfernung eines Schallereignisses zu ziehen.⁵² Je höher der Anteil des Direktschalls ist, desto näher wird die Schallquelle dabei lokalisiert. Überwiegt im Gegenzug der Diffusschallanteil, so wird die Schallquelle weiter entfernt verortet.

Aufgrund der Tatsache, dass im Freifeld kaum diffuser Schall vorhanden ist, wird hier eine andere akustische Gegebenheit genutzt. Im Gegensatz zu tieferen Frequenzen werden höhere Frequenzen durch die Luft stärker gedämpft.⁵³ Dementsprechend kommt es auf dem Weg der Schallwellen von der Quelle zum Empfänger zu spektralen Klangfarbenunterschieden, da die hohen Frequenzanteile schneller abnehmen. Dies wirkt sich allerdings erst auf größere Distanzen in einem deutlich wahrnehmbaren Maß aus.

Der dritte Ansatz zur auditiven Entfernungswahrnehmung basiert auf dem $1/r$ -Gesetz, welches verdeutlicht, dass der Schalldruck bei einer Verdoppelung des Abstands zur Schallquelle um 6 dB abnimmt.⁵⁴ Das menschliche Gehirn macht sich diese Gegebenheit zunutze, um Schallquellen entfernungstechnisch zu verorten. Die Hörerfahrung spielt hier jedoch eine sehr wichtige Rolle, da nur dem Hörer bekannte Schallereignisse mittels dieses Verfahrens in Relation gesetzt und die entsprechenden Entfernung somit korrekt lokalisiert werden können. Handelt es sich um ein unbekanntes Schallereignis, so kann der Mensch nicht zwischen nah und leise, sowie entfernt und laut unterscheiden.⁵⁵ Dies gilt allerdings im Grunde für alle drei Ansätze, die der Mensch zur auditiven Entfernungswahrnehmung nutzt. Im Falle einer unbekannt

⁵² Vgl. Webers. (2007). S. 125

⁵³ Vgl. Sengpiel. *Dämpfung der Luft bei hohen Frequenzen*

⁵⁴ Vgl. Sengpiel. *Schalldruck p und das reziproke Abstandsgesetz $1/r$*

⁵⁵ Vgl. Webers. (2007). S. 125

Schallquelle kann die Hörereignisentfernung also schnell von der tatsächlichen Schallereignisentfernung abweichen.⁵⁶

Eine weitere potentielle Fehlerquelle bei der auditiven Entfernungslokalisation liegt dem inkonsistenten Lautstärkeempfinden des Menschen über das gesamte Frequenzspektrum hinweg zugrunde. Die Fletcher-Munson-Kurven visualisieren dieses Phänomen, und man kann an ihnen klar erkennen, dass die Kurven mit zunehmendem Lautstärkepegel immer flacher und vermeintlich ausgeglichener werden. (siehe Abb. 7) Die wahrgenommene Klangfarbe eines bekannten Schallereignisses ändert sich also mit ansteigender Lautstärke, und besonders der Anteil der tiefen Frequenzen wird dadurch als lauter empfunden.⁵⁷ Dies kann zu einer ungenaueren Entfernungslokalisation führen.

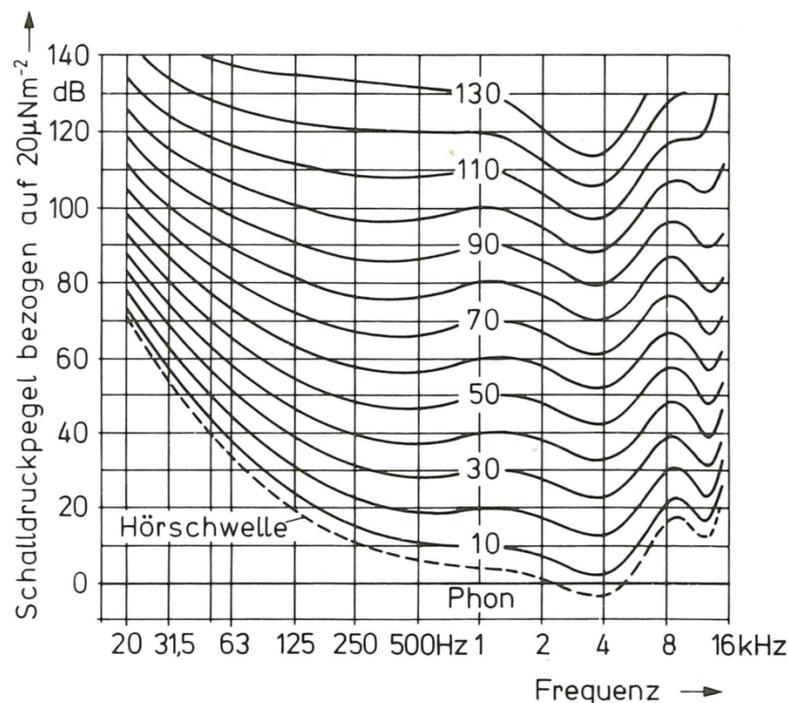


Abb. 7: Kurven gleicher Lautstärke nach Fletcher und Munson (1933)⁵⁸

⁵⁶ Vgl. Blauert. (1974). S. 125

⁵⁷ Vgl. Friesecke. (2014). S. 118

⁵⁸ Blauert. (1974). S. 98

3.1.3. Selektives Hören

Der Mensch ist dazu in der Lage, aus einer Gruppe von mehreren Schallquellen eine bestimmte davon herauszuhören. Da uns eben diese Begabung auch ermöglicht, auf einer lauten Party mit vielen Schallereignissen unseren Gesprächspartner gut zu verstehen und zu selektieren, wird dieses Phänomen auch Cocktailparty-Effekt genannt. Andere Konversationen, die Musik und sonstige Geräusche werden dabei als auditive Szene zusammengefasst. Diese Geräuschkulisse wird nicht bewusst wahrgenommen. Durch die selektive Reizverarbeitung kann der Mensch aus dieser auditiven Szene ein Schallereignis identifizieren und von den restlichen Schallereignissen separiert wahrnehmen.⁵⁹

Im Rahmen der sogenannten auditiven Szenenanalyse, auch Wahrnehmungsorganisation (engl.: perceptual organization) genannt, wird dabei jegliches eintreffende Signal analysiert und entsprechend der akustischen Eigenschaften anderen Signalen als zusammengehörig zugeordnet oder davon separiert.⁶⁰

Dabei wird zwischen der simultanen und der sequentiellen Gruppierung unterschieden. Treten mehrere Signale simultan auf, werden fünf Reizparameter zur Gruppierung genutzt: Ein gemeinsames Ein- und Aussetzen des Signals, eine gemeinsame Grundfrequenz, eine gemeinsame räumliche Herkunft, kohärente Amplituden- oder Frequenzmodulationen und die spektrale Form, also die entsprechende Klangfarbe des Signals.⁶¹

Das Ziel der sequentiellen Gruppierung ist das Zusammenfügen aufeinanderfolgender Schallereignisse, beispielsweise mehrere Töne zu einer Melodie, oder, falls sie nicht zusammenpassen, die Separation auf verschiedene Geräuschquellen. Das zentrale Konzept stellt dabei die „auditory stream segregation“, kurz „streaming“ genannt, dar. Im Zuge dessen werden zusammenhängend empfundene Folgen von Schallereignissen als „stream“ bezeichnet. Werden im Gegenzug mehrere getrennte Ereignisfolgen

⁵⁹ Vgl. Hagendorf. (2011). S. 134

⁶⁰ Vgl. Weinzierl. (2008). S. 72

⁶¹ Vgl. Weinzierl. (2008). S. 72

unterschiedlichen Geräuschquellen zugeordnet, nennt man dies „stream segregation“.⁶²

3.1.4. Im-Kopf-Lokalisation

Das Phänomen der Im-Kopf-Lokalisation tritt üblicherweise bei der Audio-Wiedergabe über Kopfhörer auf und ergibt sich aus der Tatsache, dass dieser dabei genau auf den beiden Ohrmuscheln positioniert ist. Die Rolle des Außenohrs bei der auditiven Wahrnehmung wird also im Grunde komplett ausgehebelt, da dieses vom Raum isoliert wird. Das Gehirn erhält logischerweise keine verwertbaren Informationen zur Richtungslokalisierung, wie sie bei natürlichen Schallwellen vorhanden wären, da die Schallwandler des Kopfhörers und somit auch das Schallereignis direkt an den Ohren auf der interauralen Achse wiedergegeben werden. Das vom Gehirn rekonstruierte Hörereignis kann daher keinen bekannten Mustern zugeordnet werden und wird folglich falsch verortet. Darüber hinaus kommt es aufgrund der klaren Separation des linken und rechten Kanals zu dem jeweiligen Ohr, im Gegensatz zu einer Wiedergabe über Lautsprecher, zu keinem Übersprechen auf das jeweils andere Ohr. Auch dies widerspricht den Prinzipien des binauralen, natürlichen Hörens.⁶³

In Folge dessen wird das Hörereignis nicht außerhalb des Kopfes, sondern zwischen den Ohren auf der interauralen Achse wahrgenommen. Wo genau das Signal verortet wird, hängt dabei davon ab, wie groß der Signalanteil des Schallereignisses auf den beiden Kanälen des Kopfhörers ist.

3.2. Funktionsprinzip

Für die Übersetzung des natürlichen, räumlichen Hörens des Menschen in eine künstlich reproduzierbare Umgebung sind die Ohrsignale elementar wichtig. Sie beinhalten alle Informationen, die das Gehirn zur auditiven Richtungs- und Entfernungswahrnehmung braucht. (siehe 3.1.) Sie stellen den Schlüssel zur artifiziellen Synthese des räumlichen

⁶² Vgl. Weinzierl. (2008). S. 72

⁶³ Vgl. Dickreiter. (2014). S. 133

Hörens dar und lassen sich mithilfe der „Head Related Transfer Function“ und der „Binaural Room Impulse Response“ beschreiben. Durch die Binauralisierung lässt sich mit ihrer Hilfe ein beliebiges trockenes Signal mittels Faltungsalgorithmen zu einem binauralen Signal umrechnen. (siehe 3.4.) Außerdem spielt die „Head Related Transfer Function“ auch bei der Entwicklung von binauralen Mikrofonsystemen eine entscheidende Rolle. (siehe 3.3.)

3.2.1. Head Related Transfer Function

Die „Head Related Transfer Function“, kurz HRTF, ist im Deutschen auch als Außenohrübertragungsfunktion oder kopfbezogene Übertragungsfunktion bekannt. Sie „beschreibt für einen bestimmten Schalleinfallswinkel die Schallübertragung aus dem Freifeld an einem Punkt am Eingang des Gehörgangs (Ohrkanal).“⁶⁴

Bei der auditiven Wahrnehmung des Menschen wird dabei zwischen dem distalen und dem proximalen Schallreiz unterschieden. Der Erstgenannte ist dabei der von der Schallquelle, also dem Sender ausgehende Reiz, wohingegen Letzterer im Gegenzug der beim Trommelfell, also im Grunde der bei dem Empfänger ankommende Reiz ist. Der Zusammenhang dieser beiden Schallreize lässt sich durch Mikrofonaufnahmen im Ohrkanal bestimmen. Anschließend kann man das Frequenzspektrum des distalen Schallreizes mit dem des proximalen Schallreizes vergleichen und kann so die Außenohrübertragungsfunktion bestimmen.⁶⁵

Die HRTF hängt dabei in erster Linie von den physiologischen Eigenschaften des Menschen ab. Durch Beugungen und Reflexionen an Körperteilen, wie der Ohrmuschel, aber auch am Kopf und dem Torso, erfährt der Schall eine spektrale Filterung, durch die der Mensch in der Lage ist, ein Schallereignis zu lokalisieren (siehe 3.1.1.). Da diese körperlichen Merkmale bei jedem Menschen ähnlich, aber doch einzigartig sind, variiert diese natürliche Filterwirkung, die von der HRTF beschrieben wird, auch bei jedem Menschen. Folglich hat jede Person eine individuell einzigartige Außenohrübertragungsfunktion.⁶⁶ (siehe Abbildung 8)

⁶⁴ Sengpiel. (1995).

⁶⁵ Vgl. Hellbrück. (2004). S. 151f

⁶⁶ Vgl. Sengpiel. (1995).

Um eine ideale authentische Wiedergabe einer binauralen Aufnahme gewährleisten zu können, sollte man also die individuelle HRTF des Hörers nutzen. Andernfalls riskiert man eine mitunter mangelhafte Lokalisation.⁶⁷ Dennoch ist der Mensch auch dazu in der Lage, mit fremden Ohren oder einer generalisierten HRTF räumlich zu hören und dies durch einen Lerneffekt und Erfahrungen zu verbessern. Dabei kann es allerdings vor allem auf der Medianebene schnell zu einer Verwechslung zwischen Vorne und Hinten kommen.⁶⁸

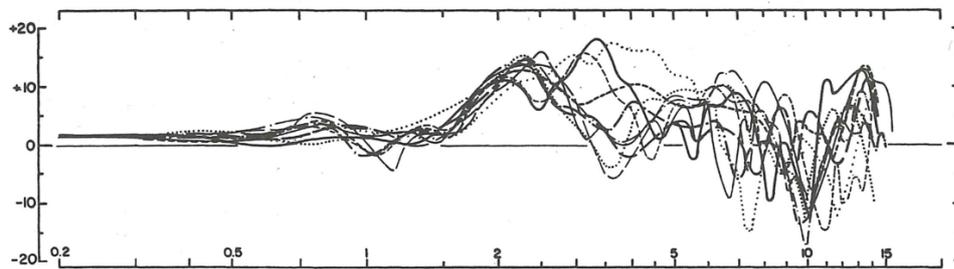


Abb. 8: Außenohrübertragungsfunktion gemessen an zehn Versuchspersonen, frontaler Schalleinfallswinkel⁶⁹

Da eine Übertragungsfunktion grundsätzlich ein Ausgangssignal in Abhängigkeit vom Eingangssignal im Frequenzbereich darstellt, und das Äquivalent im Zeitbereich eine Impulsantwort ist, berechnet man die HRTF aus der sogenannten HRIR. Diese sogenannte „Head-Related Impulse Response“ wird für jedes Ohr gemessen und stellt das Signalverhalten im Zeitbereich dar. Um die HRTF zu errechnen, muss man lediglich eine Fourier-Transformation der HRIR durchführen.⁷⁰

Aufgrund der Tatsache, dass die HRTF ausschließlich die akustische Wirkung der geometrischen Einflüsse in reflexionsarmen Räumen beschreibt, beträgt die typische Länge der HRIR, also der HRTF dargestellt im Zeitbereich, lediglich wenige Millisekunden.⁷¹

⁶⁷ Vgl. Møller. (1995). S. 300f

⁶⁸ Vgl. Laumann. (2016). S.12

⁶⁹ Görne. (2006). S 104

⁷⁰ Vgl. Hellmich. (2012). S. 14

⁷¹ Vgl. Weinzierl. (2008). S. 673

3.2.2. Binaural Room Impulse Response

Im Gegensatz zu einer „Head-Related Impulse Response“ wird eine „Binaural Room Impulse Response“, kurz BRIR, nicht im Freifeld, sondern in einem akustisch wirksamen Raum gemessen. Sie enthält, neben den im Freifeld relevanten Informationen, zusätzlich auch Reflexionsmuster des Raums, die Rückschlüsse auf die Position der Schallquelle und des Empfängers geben.⁷²

Sie ermöglicht folglich die künstliche Rekonstruktion eines räumlichen Eindrucks, der auch auf Kopfhörern wiedergegeben werden kann.

Da BRIRs mindestens die Länge der entsprechenden Nachhallzeit T_{60} haben und damit meist deutlich länger als HRIRs sind, steigt der Rechenaufwand bei der Binauralisierung dabei um mehrere Größenordnungen.⁷³

3.2.3. Diffus-und Freifeldentzerrung von Kopfhörern

Da bei der Wiedergabe über Kopfhörer die Schallwandler direkt am Trommelfell sitzen, und die spektral verfärbende Wirkung des Außenohrs, welche durch die HRTF beschriebene wird, somit ausgehebelt wird, würde ein wiedergegebenes Signal verfärbt und seltsam klingen. Außerdem kann es zu einer Im-Kopf Lokalisation kommen. (siehe 3.1.4) Dies liegt daran, dass die spektralen Veränderungen durch das Außenohr vom Gehirn nicht als Klangfarbenmerkmal wahrgenommen werden, sondern lediglich zur Verortung einer Schallquelle dienen. Man spricht dabei von der sogenannten inversen Filterung der Außenohrübertragungsfunktion.⁷⁴

Um diese Problematik zu lösen, werden die Signale von dem Kopfhörer spektral so entzerrt, dass der Verlauf des Schalldrucks am Trommelfell dem einer HRTF gleicht. Würde diese sogenannte Kopfhörer-Übertragungsfunktion eine Kopfübertragungsfunktion im Freifeld mit einer Schalleinfallrichtung exakt von vorne genau reproduzieren, so würde die inverse Filterung stattfinden und ein wiedergegebenes Monosignal würde durch den Kopfhörer theoretisch vorne lokalisiert

⁷² Vgl. Weinzierl. (2008). S. 673

⁷³ Vgl. Weinzierl. (2008). S. 673

⁷⁴ Vgl. Theile. (1984). S. 4

werden. Man spricht hierbei von einem freifeldentzerrten Kopfhörer. Da dabei allerdings insbesondere bei der Wiedergabe stereofoner Signale Ohrsignale entstehen, die nicht zu den typischen interauralen stereofonen Merkmalen passen, wendet das Gehirn hier nicht die inverse Filterung an, und es kommt im Zuge dessen zu wahrnehmbaren Klangverfärbungen.⁷⁵

Die Lösung für dieses Problem stellt eine Diffusfeldentzerrung des Kopfhörers mit einem richtungsneutralen Bezug dar. „Damit ist ein Kopfhörer-Übertragungsmaß, das die Ankopplung des Kopfhörers an den Ohrkanaleingang ohne lineare Verzerrungen sicherstellt, physikalisch definiert: Das Kopfhörer-Übertragungsmaß muss übereinstimmen mit demjenigen Außenohr-Übertragungsmaß, welches in einem richtungsneutralen Schallfeld gemessen wird.“⁷⁶

Die Kopfhörer-Übertragungsfunktion entspricht hier also nicht einer HRTF aus einer bestimmten Richtung im Freifeld, sondern Durchschnittswerten aus Übertragungsfunktionen aller Richtungen im Diffusfeld. Daraus ergibt sich eine größtmögliche klangliche Neutralität bei der Wiedergabe von raumbezogenen Signalen mittels Kopfhörer. Aus diesem Grund ist diese Form der Entzerrung aktuell die Standardlösung für das Entzerrungsproblem bei Kopfhörern.⁷⁷

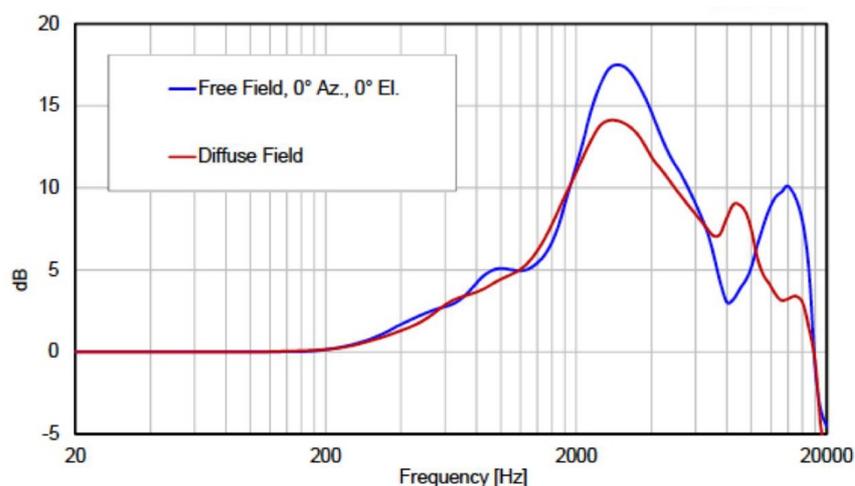


Abb. 9: Beispiel einer Freifeld-HRTF und Diffusfeld-HRTF⁷⁸

⁷⁵ Vgl. Dickreiter. (2014). S.350

⁷⁶ Dickreiter. (2014). S. 350f

⁷⁷ Vgl. Dickreiter. (2014). S. 351f

⁷⁸ Schröder. (2018).

3.3. Binaurale Mikrofonsysteme

Mikrofone zur Aufnahme binauraler Signale existieren schon seit vielen Jahrzehnten. Bereits 1930 entwickelte F.A. Firestone, ein Wissenschaftler der Michigan University, das vermeintlich erste Kunstkopfmikrofon aus einer Schaufenster-Wachsbüste.⁷⁹ Dabei handelt es sich um die wohl bekannteste und weitest verbreitetste Form des binauralen Mikrofons.



Abb. 10: Kunstkopfmikrofon KU 100 des Herstellers Neumann ⁸⁰

Das Prinzip hinter dem Kunstkopfmikrofon ist sehr einfach und logisch nachvollziehbar. Wie der Name schon vermuten lässt, handelt es sich um ein Stereomikrofon, welches einem menschlichen Kopf nachempfunden ist. (siehe Abb. 10) Anstelle der beiden Trommelfelle befinden sich die beiden Mikrofonkapseln. Dadurch entsprechen die beiden Mikrofonsignale hinsichtlich Richtcharakteristik sowie Laufzeit- und Pegeldifferenzen menschlichen Ohrsignalen. In der Theorie ist es somit möglich, ein durch den Kunstkopf aufgenommenes Schallfeld bei der späteren Wiedergabe exakt zu reproduzieren. (siehe Abb. 11)

Der Hörer hat also im Idealfall den Eindruck, selbst im Raum zu sein. Dies schließt dementsprechend auch das räumliche Verorten von im Schallfeld aufgenommenen Schallereignissen mit ein.⁸¹ Das dafür mitverantwortliche spektrale Resonanzsystem, welches das Außenohr in Kombination mit dem Kopf bildet, sowie die interauralen

⁷⁹ Vgl. Brech & Paland. (2015). S. 153

⁸⁰ Georg Neumann GmbH. (2018). *KU 100*

⁸¹ Vgl. Dickreiter. (2014). S. 345

Laufzeit- und Pegeldifferenzen werden also bereits bei der Aufnahme durch die anatomische Kopfform des Mikrofons berücksichtigt und reproduziert. (siehe 3.1.)

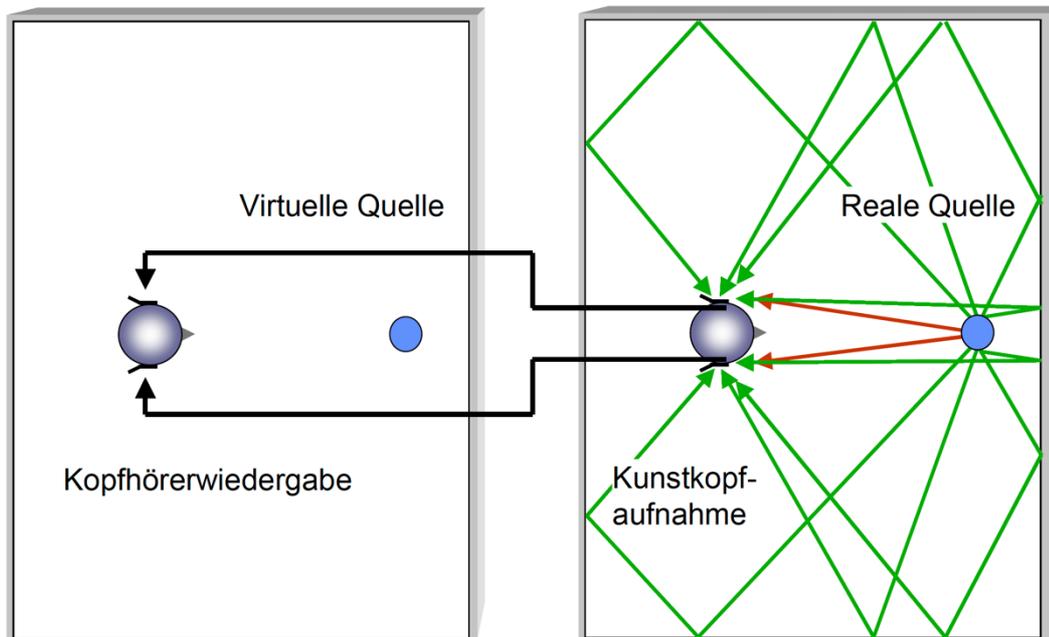


Abb. 11: Veranschaulichung der Kunstkopfstereofonie ⁸²

Ein Kunstkopfmikrofon hat dementsprechend eine eigene Kopfübertragungsfunktion, die von der Form des nachgebildeten Kopfes abhängt. Kunstköpfe variieren im Zuge dessen sehr stark hinsichtlich der Genauigkeit der nachgebildeten menschlichen Anatomie. Es gibt sehr detaillierte Modelle, bei denen sogar die Schultern und der Torso mit modelliert wurden, aber auch sehr rudimentäre Ansätze, bei denen lediglich eine kopffartige Form berücksichtigt wird. Um gute Ergebnisse zu gewährleisten, wird oftmals auch die Form der menschlichen Pinna aus Silikon nachgebildet. In der Regel kommen dabei hinsichtlich der Kopf- und Ohrmuschelform allgemeine Mittelwerte zum Einsatz, um im Endeffekt auch eine möglichst gemittelte Kopfübertragungsfunktion für das Mikrofon zu generieren. Dies entspringt der Intention, dass Inhalte, die mit dem Kunstkopf aufgezeichnet werden, später von möglichst vielen Menschen richtig verortet werden können, da eine große Äquivalenz der gemittelten HRTF des Mikrofons zur individuellen HRTF des Hörers besteht.

⁸² Theile. (2002).

Kunstkopfmikrofone werden neben dem Einsatz von Hörbuch- und Musikproduktionen hauptsächlich in der modernen Messtechnik eingesetzt. Im Zuge dessen gibt es eine Vielzahl unterschiedlicher Modelle, die meisten sind allerdings nicht im freien Handel erhältlich. Im Onlineshop von Europas größtem Versandzentrum für Musikinstrumente und Audiozubehör, einem der weltweit führenden Händler in diesem Bereich, sind beispielsweise nur zwei verschiedene Modelle verfügbar.⁸³ Dabei handelt es sich neben dem kostengünstigen Modell „Dummy Head John“ der Firma Soundman, welches allerdings ohne integrierte Mikrofone geliefert wird, um den „KU 100“ des Herstellers Neumann.⁸⁴ Dieser ist eines der Nachfolgemodelle des „KU 80“, eines der ersten kommerziell genutzten Kunstkopfmikrofone, welches Anfang der 1970er-Jahre auf den Markt kam.⁸⁵ Das Angebot an erhältlichen Kunstkopfmikrofonen ist also sehr gering. Neben den Kunstkopfmikrofonen gibt es noch weitere binaurale Mikrofonsysteme. Viele Mikrofonhersteller bieten beispielsweise kleine kompakte Mikrofonkapseln an, welche man wie In-Ear-Kopfhörer in den Gehörgang steckt und so die Möglichkeit hat, Aufnahmen mit der eigenen HRTF zu machen. Der eigene Kopf wird also im Grunde zu einem Kunstkopfmikrofon umgewandelt. Da diese Systeme in ihrer Bauweise allerdings meist dennoch zu groß sind, um die Mikrofonkapsel direkt vor dem Trommelfell aufzuzeichnen und häufig auch das spektrale Resonanzsystem des Außenohrs beeinflussen, entstehen im Zuge dessen selten perfekte Ergebnisse.



Abb. 12: Sennheiser AMBEO Smart Headset ⁸⁶

⁸³ Vgl. Gerl. (2018)

⁸⁴ Vgl. Thomann GmbH. (2020)

⁸⁵ Vgl. Georg Neumann GmbH. (2018). *KU 80*

⁸⁶ Sennheiser electronic GmbH & Co. KG. (2020). *AMBEO Smart Headset*

Diese Mikrofonsysteme sind allerdings eher weniger für die Produktion binauraler Aufzeichnungen, welche für eine breite Masse gedacht sind, ausgelegt. Dies liegt an dem Umstand, dass bei der Aufnahme die individuelle HRTF des Mikrofonträgers mit aufgenommen wird, und es sich dabei nicht um eine gemittelte Außenohrübertragungsfunktion handelt. Das aufgenommene Signal wird also in erster Linie bei der Person gute Ergebnisse liefern, welche es bei der Aufnahme getragen hat, und ist somit eher für Endverbraucher und Hobby-Aufnahmen für den Eigenbedarf geeignet.

Eine weitere Mikrofontechnik, welche zunehmend genutzt wird, um binaurale Aufnahmen zu generieren, ist das Soundfield-Mikrofon, welches ebenfalls bereits seit den 1970er-Jahren existiert. Es basiert auf einem mathematischen Ansatz der Schallfeldabtastung auf einer kugelsymmetrischen Oberfläche und liefert Mikrofon-signale, die für eine Wiedergabe im Ambisonics-Verfahren ausgelegt sind.⁸⁷ Dieses 3D-Audioverfahren wird, ähnlich wie auch die Binauraltechnik, wieder vermehrt eingesetzt, da für Aufnahmen einerseits nur ein recht kompaktes Mikrofonsystem nötig ist und damit andererseits fast alle Wiedergabeformate bedient werden können.⁸⁸ Da ein Soundfield-Mikrofon allerdings kein binaurales Signal aufnimmt, und daher erst eine Umwandlung erfolgen muss, um ein solches zu generieren, stellt diese Mikrofontechnik einen Sonderfall dar und gehört eigentlich nicht zu den binauralen Mikrofonsystemen. Da binaurale Aufnahmen dennoch zunehmend mit solchen Mikrofonen produziert werden, sollte diese Mikrofontechnik dennoch nicht unerwähnt bleiben.

Die Mikrofonkapseln, die bei binauralen Mikrofonsystemen verwendet werden, sind stets Druckempfänger, da dieses Empfängerprinzip dem des menschlichen Trommelfells entspricht. Aus diesem Grund ist es bei der Entwicklung eines Kunstkopfmikrofons sinnvoll, eine Freifeld- oder Diffusfeldentzerrung durchzuführen. „Ersteres bedeutet hier, dass der Kunstkopf bei frontaler Beschallung den gleichen Frequenzgang aufweist wie ein freifeldentzerrtes Messmikrofon. Für Signale aus anderen Richtungen ergeben sich deutliche klangliche Verfärbungen. Da Kunstköpfe meist in überwiegend diffusem Schallfeld verwendet werden, werden bevorzugt

⁸⁷ Vgl. Weinzierl. (2008). S. 591

⁸⁸ Vgl. Zotter & Franz. (2019). S. ix

diffusfeldentzerrte Kunstköpfe eingesetzt.“⁸⁹ Dadurch entstehen bei der transauralen Wiedergabe der Kunstkopfaufnahmen weniger Klangverfärbungen, und es ergibt sich so eine Kompatibilität zur Lautsprecherwiedergabe. Aufgrund der Tatsache, dass heutzutage auch Kopfhörer in der Regel diffusfeldentzerrt sind (siehe 3.2.3.), kann so außerdem eine optimale Anpassung zwischen dem Kunstkopfmikrofon und dem Kopfhörer gewährleistet werden.⁹⁰

Da durch binaurale Mikrofonsysteme binaurale Signale erzeugt werden, eignen sich diese eher nicht zur Wiedergabe über Lautsprecher. Selbst wenn man eine klare Signaltrennung des linken Lautsprechers auf das linke Ohr und umgekehrt gewährleisten könnte, würde das erneute Durchlaufen des Außenohrs und dessen spektrale Filterwirkung zu Klangverfärbungen und einer mangelnden Lokalisation führen. Um eine authentische Wiedergabe zu ermöglichen, sollten also Kopfhörer verwendet werden. Dies gilt allerdings für alle binauralen Signale und wurde bereits bei der Einleitung des Kapitels erwähnt.

Ein weiterer Nachteil binauraler Mikrofonsysteme ist, dass in der Regel kein Headtracking eingesetzt werden kann, da diese bei der Aufnahme eine statische Position haben. Daraus resultiert ein kopfbezogenes Klangbild. Aus diesem Grund wird die Lokalisation auf der Medianebene erschwert, da die Mikrobewegungen, die der Hörer unterbewusst macht, keinen Effekt haben.⁹¹ (siehe 3.1.1.)

Bei der Postproduktion von Aufnahmen mittels binauraler Mikrofonsysteme besteht, verglichen mit herkömmlichen Aufnahmen, die für eine gewöhnliche Stereoproduktion bestimmt sind, nur wenig Handlungsspielraum. Aufgrund der Tatsache, dass die Kopfübertragungsfunktion bereits mit auf das Signal aufgenommen wurde, sollte man den Frequenzgang im Nachhinein nach Möglichkeit nicht mehr beeinflussen. Sowohl Equalizer als auch Kompressoren oder Effektgeräte sollten daher nicht mehr angewandt werden, da diese den Frequenzgang des Signals verändern und so unter Umständen die inverse Filterung der HRTF verhindern können. Eine räumliche Verortung wäre dann nicht mehr möglich. Lautstärkeveränderungen können darüber hinaus zu einer beeinträchtigten Entfernungswahrnehmung der aufgezeichneten Schallereignisse

⁸⁹ Weinzierl. (2008). S. 359

⁹⁰ Vgl. Dickreiter. (2014). S. 346

⁹¹ Vgl. Weinzierl. (2008). S. 676f

führen. Das Schneiden zwischen verschiedenen Kunstkopfsignalen, beispielsweise zwischen mehreren Durchläufen einer musikalischen Darbietung, stellt kein Problem dar, solange die Position des Kunstkopfmikrofons und der von ihm aufgezeichneten Schallquellen im Raum identisch bleiben.

3.4. Binauralsynthese

Die Binauralsynthese beschreibt den Prozess, wenn Ohrsignale nicht direkt durch ein spezielles Mikrofonsystem aufgenommen, sondern durch eine digitale Signalverarbeitung synthetisiert werden. Dieser Vorgang beruht auf dem Umstand, dass es sich bei der Schallübertragung von einem beliebigen Punkt im Raum bis zu dem Trommelfell um ein LTI-System (englisch: linear time-invariant system), also um ein lineares zeitinvariantes System handelt. Mittels einer mathematischen Faltung lässt sich aus diesem Grund ein ausgegebenes Signal einer Schallquelle $s(t)$ durch die BRIR, welche die Schallübertragung vom Standort der Schallquelle Ohr repräsentiert, zu einem Ohrsignal $g(t)$ umrechnen. Dieses synthetisierte Signal enthält alle durch das Übertragungssystem, beziehungsweise die Impulsantwort $h(t)$ induzierten auditiven Charakteristika.⁹² Folglich kann durch die Binauralsynthese ohne den Einsatz spezieller dafür vorgesehener Mikrofone ein binaurales Signal generiert werden, welches bei einer Wiedergabe über Kopfhörer ein im Raum stattfindendes Hörereignis simuliert.

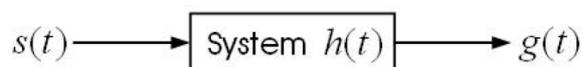


Abb. 13: LTI-System im Zeitbereich ⁹³

Bei der herkömmlichen diskreten mathematischen Faltung werden die beiden zu verrechnenden Bestandteile, also das Ausgangssignal der Schallquelle und die Impulsantwort, beispielsweise die BRIR des linken Ohrs, Sample für Sample miteinander multipliziert. Dadurch entsteht ein enormer Rechenaufwand, der eine

⁹² Vgl. Weinzierl. (2008). S. 671

⁹³ Keinath & Tarnow. (2009). S. 9

Echtzeitberechnung extrem zeitaufwendig macht. Aus diesem Grund bedient man sich bei der Binauralsynthese in der Regel der schnellen Faltung, welche auf dem Umstand beruht, dass eine Faltung im Zeitbereich einer Multiplikation im Frequenzbereich entspricht. Die beiden Bestandteile der Faltung werden also zunächst durch eine schnelle Fourier-Transformation, kurz FFT (englisch: fast fourier transformation), aus der Darstellung im Zeitbereich in die Darstellung im Frequenzbereich transformiert. Nun können beide Spektren ganz einfach multipliziert werden. Das Ergebnis dieser Rechnung stellt nun die Fourier-Transformierte des Ausgangssignal des LTI-Systems dar. Dieses kann nun durch eine umgekehrte Fourier-Transformation errechnet werden, und man erhält somit das binauralisierte Signal.⁹⁴

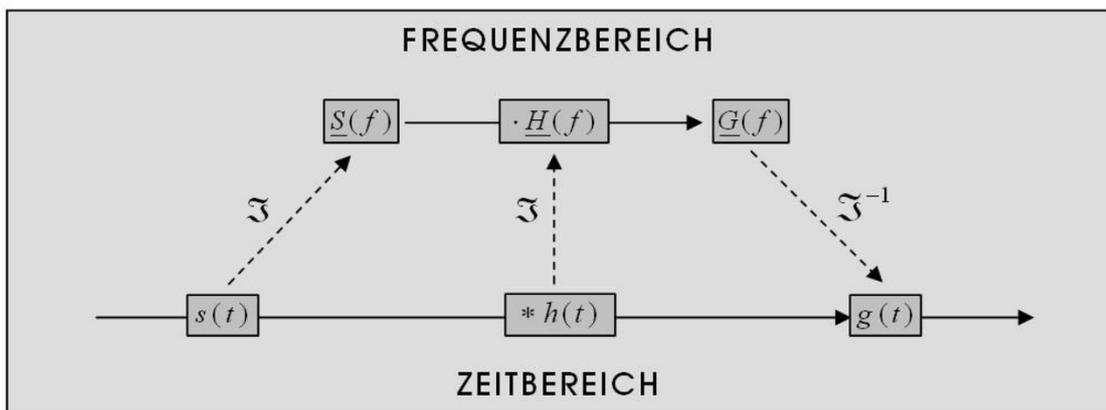


Abb. 14: Schnelle Faltung ⁹⁵

Die Tatsache, dass der Rechenaufwand bei der Binauralsynthese stets sehr geringgehalten werden soll, hängt hauptsächlich damit zusammen, dass im Zuge dessen häufig auch ein Headtracking-System zum Einsatz kommt. In diesem Fall muss die Berechnung in Echtzeit durchgeführt werden, um Verzögerungen zu vermeiden. Neigt der Hörer den Kopf nach links, so muss das Eingangssignal synchron zu der Bewegung mit der entsprechenden Impulsantwort für den sich stetig ändernden Schalleinfallswinkel verrechnet werden, um ein authentisches Hörereignis zu gewährleisten. Aus dieser Problematik ergibt sich auch, dass stets ein sehr großer Datensatz an entsprechenden BRIRs aus vielen verschiedenen Schalleinfallsrichtungen

⁹⁴ Vgl. Keinath & Tarnow. (2009). S. 12

⁹⁵ Keinath & Tarnow. (2009). S.12

vorhanden sein muss, um eine richtungsdynamische Binauralsynthese zu erzeugen. Da hierbei in der Regel nicht alle exakten Schalleinfallrichtungen abgedeckt werden können, kommt es dabei meist zu einer Interpolation zwischen mehreren Impulsantworten. Dies stellt allerdings auch einen zusätzlichen Rechenaufwand dar. Zahlreiche Software-Unternehmen bieten Programme an, mit denen eine Binauralsynthese in einer DAW durchgeführt werden kann. Dabei handelt es sich um Plug-Ins, welche in üblichen Formaten wie VST oder AU angeboten werden. Manche Programme bieten darüber hinaus auch Standalone-Versionen an, die auch eigenständig ohne DAW funktionieren. In „Logic Pro X“, der DAW des US-amerikanischen Technologieunternehmens Apple, ist sogar bereits ein Tool zur Binauralisierung integriert, allerdings bietet dieser binaurale Panner nicht allzu viele Einstellungsmöglichkeiten.⁹⁶

⁹⁶ Vgl. White. (2019)

4. Planung und Vorbereitung der Livesession

In dem folgenden Kapitel wird der Planungs- und Denkprozess hinsichtlich der binauralen Livesession genauer beleuchtet. Im Zuge dessen werden zunächst die technischen Überlegungen detailliert dargelegt. Dabei wird insbesondere auf die angestrebte auditive Ästhetik, die Erarbeitung des binauralen Konzepts und darüber hinaus auch auf den Einfluss der visuellen Komponente hinsichtlich der Planung der Livesession genauer eingegangen. Anschließend wird die Auswahl der Künstler und der Location schlüssig erläutert. Zum Abschluss des Kapitels wird dann der Prozess der Vorproduktion mit den Musikern sowie die Finanzierung erläutert.

4.1. Technische Überlegungen

4.1.1. Angestrebte auditive Ästhetik

Verschafft man sich einen Überblick des aktuellen Angebots an Livesessions, so fällt schnell auf, dass es dabei sehr unterschiedliche Ansätze gibt. Auffällig ist, dass sich verhältnismäßig viele dieser Aufnahmen recht schnell als vermeintlich nicht live eingespielte Sessions entlarven lassen. Mit einem geschulten Auge und einem technisch und musikalisch versierten Gehör, kann man dies oft schnell erkennen. Das lässt sich hauptsächlich an Diskrepanzen zwischen dem auditiven und dem visuellen Inhalt feststellen, also folglich beispielsweise durch eine fehlende Synchronisation der beiden Wahrnehmungsebenen. Für einen tontechnisch bewanderten Menschen ist auch klar, dass man als Sänger nicht lautstark mit rund vier Zentimetern Abstand in ein Großmembran-Kondensator Mikrofon singen kann, ohne dabei deutlich wahrnehmbare Popp-Geräusche zu verursachen. Hier wurde also die Aufnahme von Audio und Video offensichtlich getrennt voneinander vorgenommen. Da die Künstler die Performance bei der Videoaufnahme nicht exakt reproduzieren konnten, entstanden im Zuge dessen wahrnehmbare Diskrepanzen bei der Synchronisation zwischen Bild und Ton. Dem allgemeinen, weniger geschulten Musikhörer bleibt dies allerdings vermutlich oft verborgen, und die Illusion einer Live-Performance gelingt dennoch.

Der Sinn einer solchen Vorgehensweise ist sicherlich eine bestmögliche Audio-Performance und Qualität durch einen Zusammenschchnitt mehrerer Durchläufe und womöglich sogar dem Einsatz des Overdub-Verfahrens zur Vermeidung von Übersprechen unter den Signalen. So lässt sich schnell die Qualität einer herkömmlichen Studioproduktion erreichen. Darüber hinaus lassen sich durch dieses Vorgehen nötige Kompromisse zwischen den Bild- und den Tonschaffenden vermeiden. Das Setup kann so, hinsichtlich Aufstellung und Mikrofonierung, nach der Aufnahme des Audiomaterials einfach auf die Bedürfnisse des Filmteams umgestellt werden. Der Poppchutz am Großmembran-Mikrofon, der für die Audioaufnahme elementar wichtig ist, um ein verwertbares Gesangssignal bei naher Besprechung zu gewährleisten, wird dann beispielsweise einfach abgenommen, damit das Gesicht des Sängers oder der Sängerin im Bild gut zu sehen ist. Rein qualitätstechnisch ist dies also absolut nachvollziehbar, jedoch leidet darunter eindeutig die Authentizität. Inwieweit das dann noch den Sinn einer Livesession erfüllt, ist fraglich.

Oftmals lässt sich jedoch auch ein komplett gegensätzlicher Ansatz erkennen. Der französische Musikblog La Blogothèque aus Paris, ein sehr populärer Anbieter von Livesessions, produziert mit den Formaten „A Take Away Show“ und „One To One“ beispielsweise audiovisuelle Produktionen mit einer sehr rohen, authentischen Ästhetik. Bei dem erstgenannten Format geht ein kleines Produktionsteam mit den eingeladenen Künstlern ins Freie und nimmt spontan eine Session auf. Dies kann während einem Spaziergang sein - es wird dann während dem Gehen aufgenommen - oder aber auch spontan in einem Park oder einer anderen inspirierenden Örtlichkeit. Oftmals spielen die Künstler dabei sehr heruntergebrochene, rohe Versionen ihrer Songs, und es wird in der Regel lediglich mit einem an der Filmkamera angebrachten Mikrofon aufgenommen.

Bei dem Format „One To One“ dagegen wird ein einzelner unwissender Mensch mit verbundenen Augen in einen Raum geführt und einem Künstler gegenübergestellt, der für diese Person dann einen Song in sehr intimer Atmosphäre performt. Auch hier wird technisch gesehen ein relativ geringer Aufwand betrieben: Es gibt mehrere Kameraeinstellungen, und auditiv wird die Performance in der Regel mit nur einem Mikrofon aufgezeichnet, da es sich ebenso meist um sehr einfach gehaltene, rohe Darbietungen handelt.

Es lassen sich also technisch sehr anspruchsvoll gestaltete und aufwendige Aufnahmen finden, an denen eine hohe audiovisuelle Qualität klar über der Authentizität steht, und umgekehrt auch sehr einfach gehaltene Produktionen, an denen eine Momentaufnahme und der spontane Live-Gedanke an erster Stelle steht. Natürlich gibt es jedoch auch viele Ansätze, die auf einem Kompromiss dieser beiden Extreme basieren.

Betrachtet man in diesem Kontext die Binauraltechnik etwas genauer, so gestaltet sich die Suche nach Livesessions zunächst recht schwierig. Die wenigen Produktionen, die man dabei findet, bewegen sich hauptsächlich im Bereich der Klassik oder des Jazz. Auch einige Rock-, Pop- und Folktaufnahmen lassen sich finden, allerdings wurden diese nicht selten mit Kunstkopf-Eigenkonstruktionen realisiert und bieten dementsprechend nicht die beste klangliche Qualität. Es finden sich also verhältnismäßig nur wenige qualitativ hochwertige Livesessions, bei denen die Binauraltechnik eingesetzt wird. Auffällig ist darüber hinaus auch, dass dabei fast immer ein Kunstkopf-Mikrofon verwendet wird. Folglich haben all diese Aufnahmen auch eine sehr ungeschliffene, indirekte Klangästhetik, da der Kunstkopf mit verhältnismäßig großem Abstand zu den einzelnen Instrumenten positioniert werden muss, um alle ausgewogen abzubilden. Die Signale weisen also einen recht hohen Raumanteil auf, bieten dafür aber eine authentische Abbildung der Räumlichkeit und der einzelnen Positionen im Raum und zum Kunstkopf.

Livesessions, bei denen die Binauralsynthese zum Einsatz kommt, gibt es dagegen nur sehr selten. Ein Beispiel hierfür stellt eine Produktion des Unternehmens Sennheiser zur Vermarktung ihres „AMBEO VR Mic“ dar, einem Ambisonic-Mikrofon, das speziell für räumliche 360 Grad-Aufnahmen entwickelt wurde.⁹⁷ Im Zuge dessen wurde eine Livesession mit der Schweizer Band EgoPusher realisiert. Dabei kam neben dem beworbenen Mikrofon auch das Binauralisierungstool „AMBEO Orbit“ des Unternehmens zum Einsatz, welches genutzt wurde, um eine Wiedergabe der direkt abgenommenen Einzelsignale im Raum zu simulieren.⁹⁸ Da die Künstler elektronische Synthesizer-Klänge mit akustischen Instrumenten, wie Schlagzeug oder Violine,

⁹⁷ Vgl. Sennheiser electronic GmbH & Co. KG. (2020). *AMBEO VR Mic*

⁹⁸ Vgl. YouTube. (2019). EgoPusher "Flake" Binaural Version with Sennheiser AMBEO VR Mic | Sennheiser

vereinen, bot sich das musikalische Setup hierfür sehr gut an, da neben den akustischen Instrumenten, die von dem Mikrofon eingefangen werden, auch einzelne, nicht im Raum hörbare Signale, wie DI-Signale von Synthesizern, vorhanden waren, die sich dadurch sehr gut für den Einsatz der Binauralsynthese eignen.

Vergleicht man diese Livesession mit anderen rein Kunstkopf basierten Produktionen, so wirkt diese deutlich definierter und direkter. Sie nähert sich aus soundästhetischer Sicht allerdings auch sehr an eine gewöhnliche Stereo Livesession an. Das liegt allerdings auch an der sehr konservativen Verteilung der Instrumente und Signale im Raum. Im Grunde spielt sich dabei alles vor dem Hörer ab, und lediglich vereinzelte Effekte wie Hall oder Delays verteilen sich um den Hörer herum.

Angesichts der Tatsache, dass die Hörgewohnheiten der Menschen, die maßgeblich auf das herkömmliche Stereoformat und die damit verbundenen Konventionen hinsichtlich Mix und Klangästhetik geprägt sind, eine starke Rolle bei der Bewertung einer musikalischen Aufnahme spielen, muss dies nicht unbedingt ein Nachteil sein. Wichtig ist jedoch, dass man es schafft, die dritte Dimension wirklich für den Hörer zu öffnen und ein immersives, umhüllendes Sounderlebnis zu produzieren.

Generell verspricht die Kombination von einem binauralen Mikrofonsystem mit der Binauralsynthese soundästhetisch also eine interessantere Lösung, als eine Aufnahme lediglich mit einem Mikrofon, um das alle Instrumente positioniert werden. Um eine binaurale Livesession im Vergleich zu einer herkömmlichen zweidimensionalen Mischung konkurrenzfähig zu gestalten, scheint dies der beste Weg zu sein, da man auch in der Postproduktion durch die Binauralsynthese noch viele Möglichkeiten in der Klanggestaltung hat und so eine auditiv hochwertige, definierte und druckvolle binaurale Mischung kreieren kann.

Das Ziel ist also ganz klar, die Vorteile der herkömmlichen Stereophonie hinsichtlich Punch und Direktheit nach Möglichkeit zu erhalten und mit der Räumlichkeit der Binauraltechnik zu verbinden, um so eine hochwertige dreidimensionale Sounderfahrung für den Hörer zu generieren, die zwar die typischen Hörgewohnheiten bedient, diese jedoch zusätzlich um eine dritte Dimension sinnvoll erweitert.

4.1.2. Einfluss der visuellen Komponente

Ein wichtiger Aspekt, den man als ausführender Toningenieur bei der Planung einer Livesession nicht außer Acht lassen sollte, ist der Einfluss des visuellen Parts der Produktion. Man sollte in den organisatorischen Gedankengängen immer darauf achten, auch die bildlichen Faktoren und eventuellen Reibepunkte mit dem Filmteam im Auge zu behalten.

Dies beginnt bereits bei der Wahl des Equipments und der Mikrofonierung. Für das Ohr des Audioengineers irrelevante Faktoren, wie Kabelfarben oder das Aussehen eines Mikrofons, führen dabei nicht selten zu Diskussionen mit dem Kameramann oder dem Regisseur, da jene Aspekte die visuelle Ästhetik maßgeblich beeinflussen und oftmals potentiell verschlechtern. Dasselbe gilt für den Einsatz von Akustik-Modulen zur saubereren Signaltrennung oder Poppeschützen.

Darüber hinaus ist auch die Aufstellung der Musiker ein Faktor, bei dem es zu Streitpunkten kommen kann. Dabei ist aus tontechnischer Sicht besonders das Übersprechen der einzelnen Signale untereinander ein ausschlaggebender Punkt. Fordert das mitwirkende Filmteam also beispielsweise aus bildästhetischen Gründen eine klassische Liveaufstellung, bei der der Schlagzeuger direkt hinter dem Sänger positioniert ist, hat man als Toningenieur unter Umständen ein enormes Übersprechen des Drumsets auf das Vocal-Signal und somit im Mix enorme Schwierigkeiten, die Stimme sauber herauszuarbeiten.

Das Zusammenspiel zwischen Bild und Ton sorgt also bereits bei herkömmlichen Stereo Livesessions öfters für Differenzen unter den Beteiligten. Bei Kunstkopf-Aufnahmen ist man aus tontechnischer Sicht noch einmal deutlich weniger flexibel, was die Aufstellung der Musiker angeht, da alle entsprechend so positioniert werden müssen, dass die Signale der unterschiedlichen Instrumente ausbalanciert aufgezeichnet werden und direkt aus der richtigen Richtung auf den Kunstkopf treffen.

Bei der audiovisuellen Realisation einer Livesession ist es also elementar wichtig, dass die beteiligten Ton- und Bildschaffenden kompromissbereit an die Sache heran gehen und ein bestmöglicher Weg für das Endprodukt gewählt wird. Ob dabei die visuelle oder die auditive Komponente Vorrang hat, sollte vorab geklärt sein.

Der Einfluss des Bildes hat allerdings nicht nur negative Folgen für den Toningenieur. Durch visuelle Eindrücke können neue Elemente im musikalischen Arrangement besser vorgestellt und die auditive Szenerie visualisiert werden. Der Hörer kann sich im Zuge dessen neben den Ohren also auch auf die Augen verlassen. Durch die Kombination dieser beiden Sinne ist es dadurch auch für den Audioengineer einfacher, ein schlüssiges auditives Abbild der Performance nachzubilden.

Speziell was die Binauraltechnik angeht, hilft die visuelle Komponente darüber hinaus auch eine Verwechslung von Signalen auszuschließen, die von vorne beziehungsweise hinten kommen. Sind keine visuellen Anhaltspunkte vorhanden, werden Schallereignisse meist nicht vorne lokalisiert, da außerhalb des Sichtfelds befindliche Schallquellen in der Regel hinter dem Menschen liegen.⁹⁹ Sieht der Konsument also die Musiker vor sich im Video, werden die Signale vermutlich eher vorne lokalisiert, als bei einer Wiedergabe der Livesession ohne visuelle Komponente.

Einen interessanten Einsatz der Binauraltechnik in Kombination mit einer visuellen Ebene stellt das gemeinsame Projekt „Sound and Vision“ des Regisseurs Chris Milk und des Musikers Beck Hansen dar. Es handelt sich um eine Livesession, bei der der Sänger Beck auf einer zentralen Bühne mit seiner Gitarre den Song „Sound and Vision“ von David Bowie performt. Unterstützt wird er dabei von einem großen Orchester und Chören, die ihn umringen. Durch eine 360 Grad-Kamera und ein spezielles binaurales Mikrofonsystem, welches im Grunde aus einem Kunstkopf mit acht Ohren, also vier in verschiedene Richtungen orientierte Ohrpaaren besteht, ist es dabei möglich, mittels Headtracking eine richtungsdynamische audiovisuelle 360 Grad-Wiedergabe zu realisieren. Der Hörer kann also virtuell neben Beck auf der Bühne stehen und mit einer Drehung des Kopfes auch die entsprechende Hörperspektive dynamisch verändern. Das entsprechende Signal wird dabei in Echtzeit aus den vier Stereosignalen des Spezialmikrofons interpoliert.¹⁰⁰

Da dies allerdings einen enormen technischen und kostenintensiven Aufwand zur Folge hätte, da es sich beispielsweise bei dem genutzten Mikrofonsystem um eine Spezialanfertigung handelte, stellt diese Produktion eine absolute Ausnahme dar und

⁹⁹ Vgl. Arnold. (2013).

¹⁰⁰ Vgl. Katzif. (2020).

verkörpert wohl auch eher ein immersives Gesamtkunstwerk als eine typische Livesession. Da außerdem kaum einem Musikhörer ein Headtracking-System zur Verfügung steht, wird im Falle der in dieser Arbeit behandelten Produktion, eine binaurale Livesession ohne Headtracking bevorzugt.

4.1.3. Erarbeitung eines binauralen Konzepts

Hinsichtlich der Klangästhetik ist das Ziel der Livesession, einen druckvollen, direkten Klang zu gewährleisten, die allgemeinen stereobasierten Hörgewohnheiten zu bedienen und dennoch eine dritte räumliche Dimension zu kreieren, welche dadurch einen Mehrwert hinsichtlich der gewöhnlichen Stereophonie darstellt. (siehe 4.1.1.) Um dies technisch umzusetzen, ist ein durchdachtes Konzept nötig.

Die Schwierigkeit darin besteht hauptsächlich in dem Spagat zwischen der angestrebten sehr direkten, druckvollen Ästhetik und der umhüllenden Räumlichkeit, die erzielt werden soll. Letzteres würde sich am besten schlichtweg durch eine Kunstkopfaufnahme realisieren lassen, da man dann als Hörer einen authentischen auditiven Raumeindruck gewinnt, und die Instrumente entsprechend ihres tatsächlichen Abstandes zum Kunstkopfmikrofon lokalisieren würde. Es ergibt sich eine schlüssige Abbildung der auditiven Szenerie, die allerdings einen sehr rohen ungeschliffenen Klang hat und kaum Spielraum hinsichtlich einer Nachbearbeitung in der Postproduktion zulässt.

Will man einen direkten, druckvollen Sound, führt kein Weg an der Polymikrofonierung vorbei. Um einen möglichst hohen Direktschallanteil und wenig Übersprechen der anderen Instrumente zu gewährleisten, muss man die Mikrofone dabei möglichst nah an den einzelnen Schallquellen positionieren. Dies hat den Vorteil, dass man beim Mixing eine große Bearbeitungsfreiheit hat, wenn die Signale möglichst gut isoliert von den anderen Klangelementen aufgenommen wurden. Dadurch kann man die Charakteristika der einzelnen Instrumente sehr gut herausarbeiten und einen transparenten, kompakten Mix gewährleisten.

Kombiniert man die Polymikrofonierung mit der Binauralsynthese, so lässt sich bei einer sauberen Signaltrennung ein recht glaubhafter dreidimensionaler Mix erstellen. Da man

im Zuge dessen allerdings durch die uneingeschränkte Freiheit hinsichtlich Signal-levelling, Effekt-Prozessierung und der Anordnung der einzelnen Elemente um den Hörer schnell die Grenzen einer authentischen auditiven Szenerie überschreitet, wirken diese Mischungen oft wenig realistisch. Bei einer reinen Audioproduktion kann dies ein artifizieller Ansatz sein, der durchaus akzeptiert wird. Schließlich bilden herkömmliche Stereo-Mixe oftmals auch kein natürliches Abbild einer musikalischen Live-Performance, sondern spielen im Gegenzug mit den Grenzen der technischen Möglichkeiten. Im Zusammenhang mit einer Livesession ist es allerdings aufgrund der visuellen Informationen, die der Zuhörer zu der auditiven Szenerie erhält, wichtig, diese Glaubhaftigkeit der Mischung im Hinterkopf zu behalten. Panning-Fahrten des Vocal-Signals um den Zuhörer herum sollten also beispielsweise eher vermieden werden, wenn der Sänger im Bild eine statische Position beibehält und sich nicht auch im Bild bewegt. Dies würde den Zuhörer andernfalls einerseits verwirren und andererseits den Bezug zwischen Audio und Video trennen, was zu einem unglaublichen Eindruck der Live-Performance führen würde.

In Relation mit dem Ziel, gängige Hörgewohnheiten der Musikkonsumenten zu bedienen, um Nachteile gegenüber einer herkömmlichen Stereo-Session zu vermeiden, führt dies wiederum zu erschwerten Bedingungen, um eine immersive, dreidimensionale Räumlichkeit zu erzeugen. Dies liegt daran, dass sich bei der herkömmlichen Stereophonie die musikalischen Instrumente ausschließlich vor dem Hörer oder zwischen den beiden Ohren, je nachdem ob die Wiedergabe über Lautsprecher oder Kopfhörer erfolgt, auffächern. Dies wird vom Hörer aufgrund der daraus resultierenden Hörgewohnheiten auch so erwartet. Werden im Song dominante Instrumente hinter dem Hörer positioniert, kann dies auf Dauer als unangenehm empfunden werden. Es ist also ratsam, die einzelnen Instrumente vor dem Hörer zu positionieren.

Da eine Aufstellung der Musiker in einem 360 Grad-Winkel um die Hörerposition herum daher nicht empfehlenswert ist, muss man stattdessen auf andere Mittel zurückgreifen, um einen immersiven, umhüllenden Effekt zu erzielen. Naheliegender ist hierbei, den Diffusschall im Raum zu nutzen, der aus allen Richtungen auf den Hörer trifft und somit auch einhüllend wirken kann. Um dies in einem ausreichend wahrnehmbaren Maß gewährleisten zu können, ist es wichtig, in einem Raum aufzunehmen, welcher eine

große Nachhallzeit hat. Eine akustisch optimierte Studioumgebung, in der der Diffusschall mittels Absorber und Diffusoren auf ein Minimum reduziert wurde, eignet sich hierbei nicht.

Im Zuge dessen empfiehlt sich der Einsatz eines Kunstkopfmikrofons, um den natürlichen Schalleindruck im Raum während der Performance möglichst realitätsnah aufzuzeichnen. Essenziell ist dabei natürlich, dass auch echte akustische Instrumente genutzt werden, die das Schallfeld im Raum ausreichend anregen.

Es ist also naheliegend, für die Aufnahme der Livesession ein Kunstkopfmikrofon für das Basissignal und eine Polymikrofonierung der einzelnen Instrumente für die Binauralisierung in der Postproduktion zu nutzen. Durch die Ergänzung dieser beiden Ansätze sollte es möglich sein, die gewünschte Klangästhetik zu erreichen. Das Basissignal liefert dabei einen authentischen Raumeindruck der Performance und sorgt für einen immersiven Effekt durch den Diffusschallanteil im Aufnahmebereich, wohingegen man die binauralisierten Einzelsignale später im Mix als Stützschnale hinzunehmen kann und so einen großen Einfluss auf die Klanggestaltung der einzelnen Elemente hat.

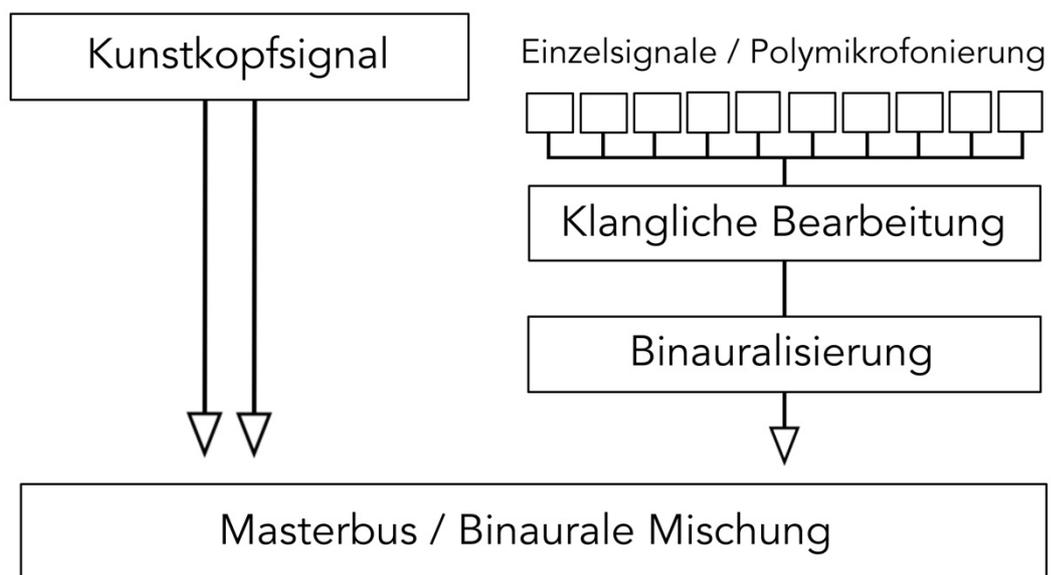


Abb. 15: Binaurales Konzept der Livesession ¹⁰¹

¹⁰¹ Eigene Darstellung

Dies hat zur Folge, dass auf dem finalen Mix der Livesession parallel zwei verschiedene HRTFs genutzt werden: Einerseits die HRTF des Kunstkopfmikrofons und andererseits die der Binauralisierungssoftware. Da es sich allerdings bei beiden um gemittelte Außenohrübertragungsfunktionen handelt, sollte die Lokalisation bei der Wiedergabe dennoch bei jedem Hörer recht gut funktionieren. Eine fehlerhafte Lokalisation, vor allem eine Verwechslung zwischen vorne und hinten, ist natürlich nicht auszuschließen. Dies stellt aber ein generelles Problem der Binauraltechnik dar, welches in diesem Fall in Kauf genommen werden muss. (siehe 3.2.1.)

Um beim Mixing neben dem Hall ein weiteres Mittel zur Realisierung eines immersiven Effekts zu haben, ist es darüber hinaus wichtig, isolierte Signale, die bei der Aufnahme nicht im Raum stattfinden, zur Verfügung zu haben. Dies können Samples, Synthesizer oder auch DI-Signale jeglicher Art sein. Da diese nicht direkt von dem Kunstkopfmikrofon mit einer räumlichen Information aufgenommen werden, können diese Signale somit mittels der Binauralsynthese frei im zur Verfügung stehenden Raum verteilt werden. Folglich auch seitlich oder hinter dem Hörer. Darüber hinaus kann man diese Elemente auch stark prozessieren, da auf ihren Signalen kein Übersprechen anderer Instrumente enthalten ist. Mittels Delays und ähnlichen räumlichen Effekten kann man somit also auf eine abstraktere Art und Weise eine Räumlichkeit und einen immersiven Effekt aufbauen. Wichtig ist dabei, dass diese Signale eher von ergänzender Natur sein sollten. Die klangliche Grundstruktur des Songs sollte sich nach wie vor vor dem Hörer abspielen. Tragende Rhythmusinstrumente, wie die Bassdrum oder der Bass, sowie der Leadgesang sollten hierfür also keine Verwendung finden.

Zusammenfassend gesagt, kristallisiert sich also der Ansatz heraus, dass eine Kombination aus der Aufnahme mit einem Kunstkopf sowie einer Polymikrofonierung und der anschließenden Binauralsynthese im Mix den potentiell besten und vielseitigsten Weg darstellt, um auch in der Postproduktion noch viele Freiheiten zu haben. Für die dreidimensionale Immersion und Räumlichkeit sollen dabei maßgeblich der Hall sowie binaural synthetisierte, isolierte Einzelsignale genutzt werden. Im Zuge dessen ist es wichtig, in einem sehr lebendigen Raum mit einer recht großen Nachhallzeit aufzunehmen. Darüber hinaus wäre es von Vorteil, wenn die Künstler sowohl akustische Instrumente nutzen, die im Raum zu hören sind, sowie auch elektronische Elemente, die raumunabhängig aufgezeichnet werden können.

4.2. Die Auswahl der Künstler

Die Auswahl der Künstler ist für die Vorbereitung einer Livesession ein essentiell wichtiger Faktor. Sowohl die Instrumentierung als auch die Aufstellung und die Anzahl der Bandmitglieder spielen in den Planungsprozessen eine wichtige Rolle. Aufgrund der angestrebten Soundästhetik und dem geplanten binauralen Konzept stellte sich heraus, dass diese dabei nach Möglichkeit gewisse Kriterien erfüllen müssen. Sie sollten sowohl akustische Instrumente als auch elektronische Signale nutzen. (siehe 4.1.3.) Da sich diese Arbeit darüber hinaus im Kontext der Popmusik bewegt, sollten die musikalischen Akteure sich außerdem in diesem Genre bewegen.

Das branchenübliche Vorgehen ist, dass die Musiker als Kunden in Erscheinung treten und die Livesession in Auftrag geben. In diesem Fall müsste man sich als beteiligter Toningenieur den musikalischen und klanglichen Gegebenheiten, die die Band mitbringt, unterordnen. Aus diesem Grund war es hier also als Audioengineer definitiv sinnvoll, die beteiligten Künstler auf Grundlage der benannten Kriterien auszuwählen und nicht umgekehrt, um ein bestmögliches Ergebnis zu gewährleisten, auch wenn dies nicht der typischen Vorgehensweise entspricht.

Die Wahl fiel dabei auf die Stuttgarter Indie Pop-Band „Into The Fray“. Aufgrund der Tatsache, dass der Verfasser dieser Arbeit selbst Teil der Band ist, war es naheliegend, diese als musikalischen Akteur in das Projekt zu integrieren. Es ist zwar untypisch, dass der Toningenieur gleichzeitig auch als Musiker in Erscheinung tritt, jedoch ergaben sich aus dieser Konstellation einige nicht unwesentliche Vorteile im Vergleich zur Arbeit mit fremden Musikern. Einerseits waren die Songs und deren tiefergehenden Strukturen und die unterschiedlichen Elemente dem Verfasser und ausführenden Toningenieur bereits bekannt, was eine gute logistische und kreative Planung im Vorfeld erleichterte, andererseits konnte er darüber hinaus auch in Proben und in der Vorproduktion maßgeblich einen musikalischen Einfluss nehmen, um die Songs so hinsichtlich Arrangement und Instrumentierung zusätzlich für die binaurale Livesession zu optimieren.

Die dadurch bedingte größere Belastung des Verfassers bei der Aufnahme der Session, aufgrund der Doppelbesetzung als Toningenieur und Musiker, schien angesichts dieser Vorteile akzeptabel. Darüber hinaus ist zwar anzunehmen, dass diese beiden Rollen im

branchenüblichen Normalfall strikt voneinander getrennt werden, tatsächlich kommt es in der Praxis allerdings häufiger vor, dass die Rollengrenzen im Laufe der Produktion verschwimmen oder sogar die Künstler selbst aus Kostengründen für die tontechnische Betreuung verantwortlich sind. Dieser Fall tritt vor allem auf, wenn Künstler eine Livesession selbst produzieren und folglich nicht die audiovisuelle Infrastruktur eines darauf spezialisierten Partners nutzen können.

Zudem erfüllt die Band die relevanten Kriterien und hätte sich daher auch ohne die besagten erleichternden Umstände als musikalischer Partner für die Livesession angeboten.



Abb. 16: Into The Fray ¹⁰²

„Into The Fray“ wurde 2015 von dem Sänger und Gitarristen Klotzbach als Solo-Projekt gegründet. In den folgenden Jahren entwickelte sich das Projekt zu einer Band, und die Genre-Stilistik wandelte sich von Indie Folk hin zu einem ambienten Indie Pop, mit Einflüssen verschiedenster musikalischer Richtungen. Aktuell besteht die Band neben dem bereits erwähnten Sänger und Gitarristen aus dem Lead-Gitarristen Fabian Glück, dem Schlagzeuger Bastian Kilper und dem Bassisten Marian Hepp, dem Verfasser dieser Arbeit. Neben den akustischen Grundelementen Gesang, Gitarre, Bass und

¹⁰² Pressebild, Foto: Florian Müller

Schlagzeug, nutzt die Band jedoch auch zunehmend elektronische Elemente in ihren Arrangements. Somit werden sowohl die Genre-Vorgabe als auch die musikalischen Kriterien erfüllt.

4.3. Die Wahl des Aufzeichnungsorts

Der Aufzeichnungsort wird bei Livesessions in den letzten Jahren zunehmend wichtiger. Dabei zeichnet sich der Trend ab, dass Künstler und Produzenten vermehrt von akustisch optimierten Studioumgebungen hin zu ausgefallenen Orten ausweichen. Der Grund hierfür liegt in dem visuellen Mehrwert und einem interessanteren Setting für den Zuschauer. (siehe 2.3.) Auch im Zuge dieser Arbeit soll diese Entwicklung nicht unbeachtet bleiben.

Aus tontechnischer Sicht bietet eine akustisch optimierte Studioumgebung natürlich auf den ersten Blick viele Vorteile. Einerseits ist bereits eine funktionierende tontechnische Aufnahmeinfrastruktur gegeben, andererseits sind Aufnahmeräume in Tonstudios in der Regel aus akustischer Sicht sehr vielseitig konstruiert und bieten dadurch durch eine wohlüberlegte Positionierung der unterschiedlichen Schallquellen viele Möglichkeiten zur Einflussnahme der Klangästhetik.¹⁰³ Wird im Overdubbing-Verfahren gearbeitet, so kann man dabei nach und nach alle Instrumente individuell im Raum positionieren. Nimmt man jedoch ein musikalisches Ensemble live auf, folglich mit mehreren Schallquellen in einem Raum, kann dies schnell zu Problemen führen, vor allem, wenn zusätzlich ein visueller Mitschnitt gemacht wird. Die Aufnahmeräume vieler Studios bieten dafür nicht den nötigen Platz, da sie häufig nicht dafür ausgelegt und konzipiert worden sind. Große Studios mit entsprechend dimensionierten Aufnahmeräumen sind im Gegenzug meist sehr kostenintensiv.

Darüber hinaus haben Aufnahmeräume in Studios in der Regel eine recht kurze Nachhallzeit. „Sinnvolle Nachhallzeiten liegen in Abhängigkeit von Raumvolumen und Nutzung im Bereich zwischen 0,1 Sekunden für sehr kleine Räume für die Aufnahme von einzelnen Instrumenten und Sängern und etwa 0,8 Sekunden für mittelgroße

¹⁰³ Vgl. Friesecke. (2014). S. 88

Aufnahmeräume.“¹⁰⁴ Dies würde bei der Umsetzung des aufgestellten binauralen Konzepts zu Problemen führen, da der Diffusschall im Raum dabei eine entscheidende Rolle bei der Generierung eines immersiven, umhüllenden Klangs einnehmen soll. (siehe 4.1.3.)

Folglich spricht nicht nur der Trend, hin zu ausgefallenen visuellen Räumlichkeiten und die damit verbundene positive Aufnahme durch das Publikum, gegen die Aufnahme in einem Tonstudio, sondern vielmehr auch das erarbeitete binaurale Konzept und damit ein potentiell weniger immersives Ergebnis.

Fasst man diese beiden Punkte zusammen, wird dementsprechend ein visuell ansprechender Raum mit einer großen Nachhallzeit gesucht. Eine naheliegende Lösung stellen dabei sakrale Versammlungsstätten wie Kirchen oder Kapellen dar. Beim Bau dieser Gebäude wurde häufig unter Aufwendung großer finanzieller Mittel auf eine dem Auge schmeichelnde visuelle Ästhetik geachtet. Darüber hinaus wurden sie für viele Besucher ausgelegt, haben hohe Decken und besitzen dementsprechend ein großes Raumvolumen. Dieses variiert je nach Größe und kann Werte von circa 2000 bis zu 200 000 m³ haben. Auch die Architektur und Ausstattung im Raum fällt je nach Baustil sehr unterschiedlich aus und hat so einen großen Effekt auf die Akustik im Raum.¹⁰⁵ Hinzu kommen sehr schallharte und reflektierende Wände, da hierbei in der Regel Stein verbaut wurde.

Betrachtet man die Formel zur Berechnung der Nachhallzeit von Wallace Clement Sabine, so lässt sich aus diesen Umständen eine große Nachhallzeit herleiten.

$$RT = 0,163 \cdot \frac{V}{S \cdot a}$$

RT steht dabei für die Nachhallzeit in Sekunden, V für das Raumvolumen in m³, S für die Oberfläche des Raums in m² und a für den mittleren Absorptionskoeffizienten aller Flächen des Raums.¹⁰⁶ Da das Raumvolumen, also der Zähler des Bruchs, bei Kirchen, wie bereits angesprochen, sehr groß ist und der mittlere Absorptionskoeffizient

¹⁰⁴ Weinzierl. (2008). S. 293

¹⁰⁵ Vgl. Dickreiter. (2003). S. 33

¹⁰⁶ Vgl. Friesecke. (2014). S. 60

aufgrund der schallharten, reflektierenden Wände und Böden in der Regel eher geringer ausfällt, ergeben sich so schnell recht hohe Nachhallzeiten für sakrale Gebäude.

Die Wahl des Aufzeichnungsortes fiel im Rahmen der dargelegten Überlegungen auf eine kleine Dorfkirche im oberschwäbischen Raum. Dabei handelt es sich um die sogenannte „Alte Kirche“ in Rulfingen, im Landkreis Sigmaringen. Sie wurde im 15. Jahrhundert im spätgotischen Stil erbaut. Da die Kirchengemeinde des Ortes 1974 eine neue Kirche erhielt, wurde die alte Kirche profaniert und die religiösen Artefakte sowie die Kirchenbänke entfernt. In den folgenden Jahren stand das Gebäude leer und blieb ungenutzt, entwickelte sich aber ab 1997 zu einer vielseitig genutzten Kulturstätte, in der seither Konzerte, Varietés, Lesungen, Ausstellungen, Kabarette und Comedy-Veranstaltungen stattfinden.¹⁰⁷

Aufgrund der Profanierung und dem jahrelangen Leerstand hat die Kirche viel von ihrer religiösen Wirkung und ihrem alten Glanz eingebüßt. Dies ist im Falle der Livesession allerdings ein entgegenkommender Umstand, da die Band Into The Fray sich nicht in eine bestimmte religiöse Richtung positionieren will. Die Alte Kirche in Rulfingen bietet sich daher optimal an, um einerseits die typische architektonische und klangliche Ästhetik einer Kirche zu nutzen, andererseits aber klare Assoziationen zum christlichen Glauben zu vermeiden. Darüber hinaus ist es aufgrund der Tatsache, dass sich dort keine Kirchenbänke mehr befinden, möglich, die Musiker mit ausreichend Platz mitten in der Kirche zu positionieren, was eine deutlich größere Flexibilität ermöglicht.

Gotische Kirchen haben in der Regel ein verhältnismäßig großes Raumvolumen und aufgrund von vielen steinernen Oberflächen eine sehr geringe Schallabsorption im Bereich der tiefen und mittleren Frequenzen. Dieser Umstand führt meist zu sehr langen Nachhallzeiten und einer eher dunkleren Klangfärbung des Halls.¹⁰⁸

Anhand der Baupläne der alten Kirche lässt sich ein grobes Gesamtvolumen des Innenraums von circa 1729 m³ errechnen. Dieses ergibt sich aus dem groben Raumvolumen des Altarraums von circa 450 m³ und dem des eigentlichen Saals von 1279 m³. Es ist also im Kontext kirchentypischer Dimensionen als sehr gering

¹⁰⁷ Vgl. Alte-Kirche e.V. (2020).

¹⁰⁸ Vgl. Dickreiter. (2003). S. 33

einzuordnen. Die Fläche des Bodens und der Decke beträgt jeweils ungefähr 231 m², die Gesamtfläche aller Wände beträgt dagegen circa 488 m². Die Gesamtoberfläche des Raumes ist dementsprechend ungefähr 950 m². Da die Baupläne sehr alt sind und viele zur genauen Berechnung notwendigen Maßangaben fehlen, handelt es sich hierbei um stark vereinfachte, grobe Berechnungen. Von einer Kalkulation der Nachhallzeit mittels der Formel nach Sabine wird aus diesem Grund abgesehen.

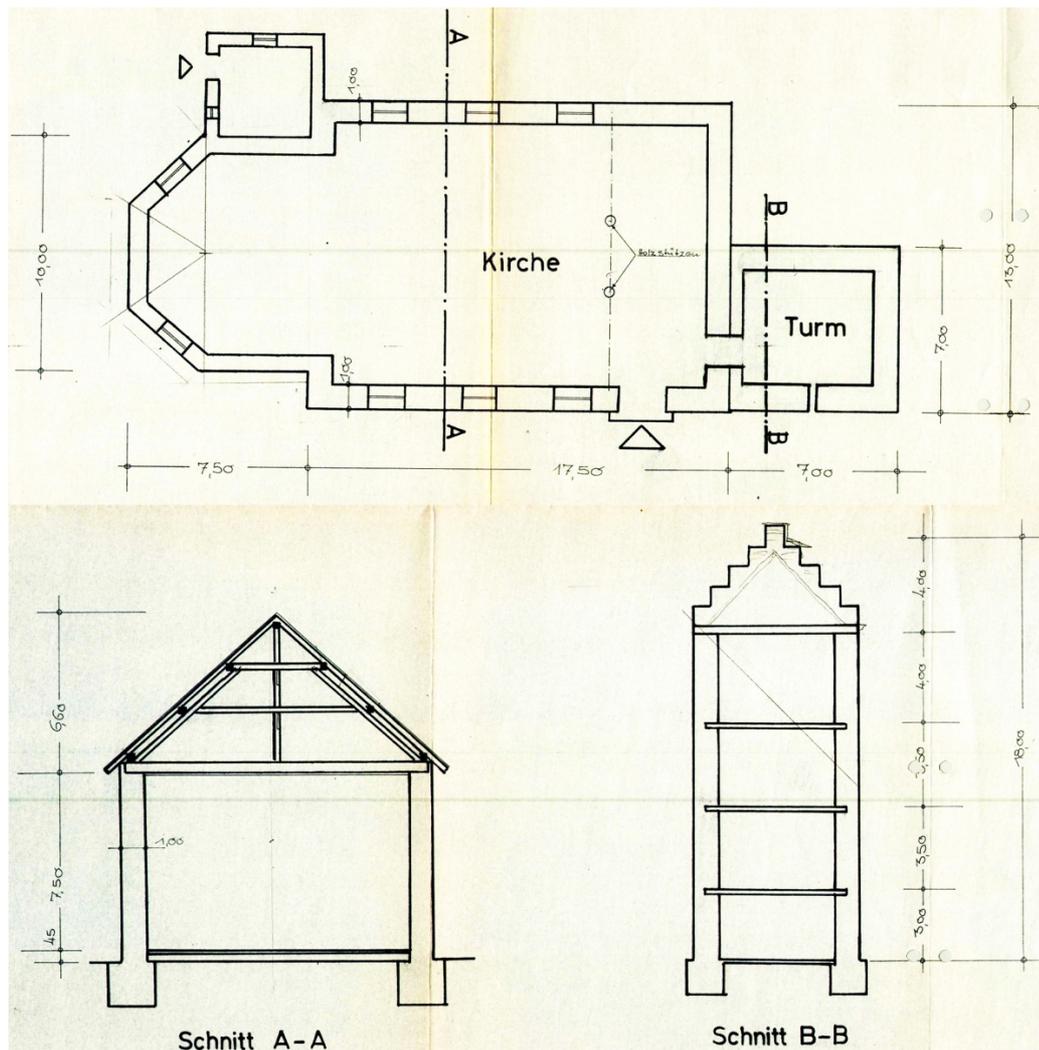


Abb. 17: Baupläne der Alten Kirche in Rulfingen ¹⁰⁹

Aufgrund der fehlenden Informationen hinsichtlich Materialien, Fensterflächen, der Bühnenhöhe im Altarraum, sowie von vorhandenen Holzflächen auf dem Saalboden, wären die daraus resultierenden Ergebnisse zu ungenau und nicht repräsentativ.

¹⁰⁹ Stadtbauamt Mengen. (1978).

Anhand der Baupläne und des groben Raumvolumens lässt sich dennoch sagen, dass die Kirche verhältnismäßig recht klein ist und so eine Nachhallzeit bietet, die zwar ausreichend lang ist, um einen umhüllenden Effekt zu erzielen, aber dennoch ein recht definiertes Klangbild zulässt. Bei einer Begehung der Location vor der tatsächlichen Produktion erhärtete sich dieser Eindruck. Die Nachhallzeit des Innenraums der alten Kirche liegt schätzungsweise zwischen 1,5 und 2 Sekunden. Da es sich lediglich um eine kurze technische Einweisung handelte, war es leider aus zeittechnischen Gründen nicht möglich eine Impulsantwort aufzunehmen.



Abb. 18: Innenraum der Alten Kirche Rulfingen ¹¹⁰

Besonders hinsichtlich der Wirkung des Schlagzeugs im Raum würden zu lange Nachhallzeiten zu einem undefinierten, diffusen Klangbild führen. Da gotische Kirchen eine Nachhallzeit von bis zu 12 Sekunden oder mehr erreichen können, stellt die ausgewählte Kirche eher ein Sonderfall dar, da sie mit ihrer geschätzten Nachhallzeit

¹¹⁰ Foto: Foxframes

eher im Bereich klassischer Konzertsäle liegt.¹¹¹ Die Alte Kirche in Rulfingen schien daher im Vorfeld als eine gute Wahl für die Aufzeichnung der Livesession.

4.4. Vorproduktion mit den Künstlern

Im Rahmen der Vorproduktion arbeitet das audiovisuelle Produktionsteam eng mit den musikalischen Akteuren zusammen, um bei der Aufnahme vor Ort einen reibungslosen und detailliert organisierten Ablauf gewährleisten zu können. Aus visueller Sicht geht es dabei in erster Linie um Absprachen hinsichtlich gestalterischer Elemente, wie der Lichtführung oder welche Farben im Fokus stehen sollen, und um den generellen visuellen Look des Endprodukts. Auf Grundlage der Resultate dieser Kommunikation wird dann entsprechend benötigtes Equipment geliehen und ein Konzept ausgearbeitet.

Aus audiatechnischer Sicht gestaltet sich die Vorproduktion etwas komplexer und mit einem engeren Kontakt zu den Künstlern. Hinsichtlich der binauralen Livesession fing dieser Prozess bereits bei der Songauswahl an. Into The Fray hat im Herbst 2019 ihre letzte EP „alas“ veröffentlicht. Davon mussten zwei Songs ausgewählt werden, die gute klangliche Voraussetzungen mitbringen.

Im Zuge dessen wurden einige Songs schon aufgrund ihrer lauten, druckvollen Ästhetik schnell aussortiert. Da diese sehr auf verzerrten Gitarrensensoren und aggressiven Drums basieren, wäre ihre klangliche Wirkung sehr schwierig umzusetzen gewesen, da solche Signale durch die Binauralisierung bzw. durch die räumliche Aufnahme mit einem Kunstkopf deutlich an Direktheit und Punch verlieren. Außerdem hätten solche lauten Schallquellen auch ein deutlich extremeres Übersprechen der unterschiedlichen Signale zueinander im Raum zur Folge gehabt. Aufgrund der Tatsache, dass der Sänger Lukas Klotzbach beispielsweise oft sehr leise singt, hätte dies unter Umständen zu enormen klanglichen Einbußen, bis hin zur Unbrauchbarkeit des Vocal-Signals führen können.

Es kamen also nur die Songs der EP in Frage, die einen balladeskeren, ruhigeren Charakter haben. Die Wahl fiel dabei einerseits auf den Titel „Bloom“, ein sehr melancholischer Track, der sich dynamisch langsam aufbaut und in seiner Studioversion

¹¹¹ Vgl. Dickreiter. (2003). S. 33-45

eine Vielzahl an Instrumenten verschiedenster Art beinhaltet und andererseits auf den Song „In Dreams“, eine etwas ruhigere Ballade.

Letzterer bietet aufgrund seiner klassischen Besetzung von Gesang, Gitarre, Bass und Drums etwas weniger Freiraum bei der Binauralisierung im Mix und muss daher hauptsächlich über den Raumklang einen immersiven Effekt erzielen. Da es sich jedoch um die aktuelle Single der Künstler handelt, drängten diese darauf, diesen Song performen zu dürfen. Dies bietet im Endeffekt jedoch eine gute Möglichkeit, um zu prüfen, ob ein binauraler Raumeindruck allein eine ausreichend umhüllende auditive Wirkung erzielen kann.

Nach der Auswahl der Songs wurden diese in Koproduktion zwischen Toningenieur und Band weiter für eine Performance bei der binauralen Livesession optimiert. Im Zuge dessen wurde vom Rhythmusgitaristen und Sänger Lukas Klotzbach eine Akustik-Gitarre statt einer E-Gitarre genutzt, um akustisch nicht auf einen sperrigen Verstärker angewiesen zu sein und eine andere Klangfarbe ins Spiel zu bringen. Fabian Glück, der Lead-Gitarrist, spielte stattdessen nicht wie üblich über einen, sondern über zwei Verstärker. Da er für seinen Sound sehr viele stereofähige Delay- und Hall-Effekte nutzt, stellte dies kein Problem dar. So repräsentiert Klotzbach mit seiner Akustikgitarre im Panorama vor dem Hörer die Mitte, während sich Fabian Glück mit seinen sphärischen E-Gitarren-Klängen im Mix links und rechts positioniert. Dadurch bleibt ein ausgeglichenes Klangbild bestehen.

Darüber hinaus wurde mithilfe der zur Verfügung stehenden Effektpedale ein System erdacht, mit dem Fabian Glück durch einen Fußschalter seine Verstärker stumm schalten kann. Ein Monosignal seiner Leadgitarre wird dennoch durch eine DI-Box nach seinem Effekt-Pedalboard abgegriffen. Dies erlaubt eine bewusste Steuerung dessen, was im Raum wiedergegeben und somit vom Kunstkopf aufgezeichnet wird und was nicht. Da Fabian Glück durch das sogenannte Swelling und Shredding teilweise sehr sphärische Klangwände produziert, die sich für die Wiedergabe im Raum sehr gut eignen und das Schallfeld gut anregen, er aber teilweise auch akzentuierte, rhythmische Melodien spielt, die sich hervorragend zur räumlichen Verteilung und Automation mittels der Binauralsynthese anbieten, ist so beides möglich. Mit einem einfachen Fußtritt ist er in der Lage, sein Gitarrensignal von der räumlichen Wiedergabe zu isolieren.

Ein weiterer problematischer Faktor stellte trotz der Auswahl ruhigerer Songs das Drumset dar. Aufgrund der gängigen Hörgewohnheiten soll sowohl das Schlagzeug sowie der Gesang vor dem Hörer positioniert sein. Da beide musikalischen Elemente auch im Raum zu hören und somit lokalisierbar sind, bedeutete das folglich, dass sich Sänger Lukas Klotzbach vor dem Kunstkopf positionieren muss und Drummer Bastian Kilper sich zentral dahinter aufbaut, um eine Verortung beider Schallquellen von vorne zu gewährleisten. Die Positionen von Bassist Marian Hepp und Lead-Gitarrist Fabian Glück sind im Grunde flexibel, da ihre Instrumente an ihrer Position nicht als Schallquelle in Erscheinung treten. Sie werden entweder durch Verstärker im Raum an anderer Position lokalisiert werden oder gar nicht im Raum hörbar sind. Daraus ergibt sich die folgende Aufstellung:

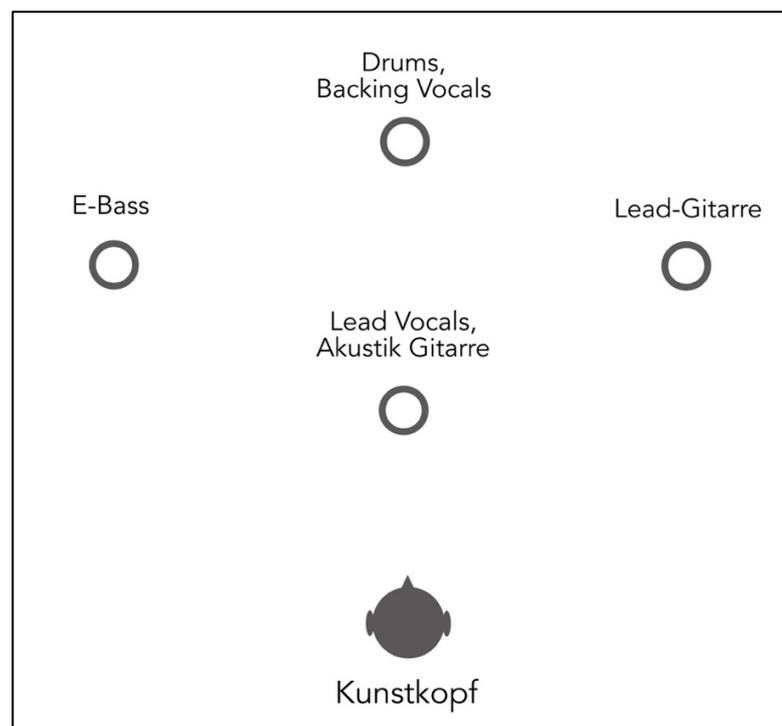


Abb. 19: Aufstellung "In Dreams" ¹¹²

Hinsichtlich des Signalübersprechens des Schlagzeugs auf das Vocal-Signal ist dieser Aufbau logischerweise sehr kontraproduktiv. Hinzu kommt der Umstand, dass bei der Vorproduktion im Proberaum nicht klar evaluiert werden konnte, wie dicht das

¹¹² Eigene Darstellung

Schlagzeug im Rücken des Sängers stehen muss, um auf dem Kunstkopf-Signal einen ausreichenden Direktschallanteil zu haben. Dies würde sich erst in der Location vor Ort ergeben.

Aufgrund dieser Faktoren einigte man sich darauf, dass der Schlagzeuger Bastian Kilper Hot Rods aus Kunststoff verwendet. Dies sind spezielle Drumsticks, die im Vergleich zu herkömmlichen Sticks ein leiseres Spiel ermöglichen. Dadurch kann die Lautstärke des Schlagzeugs im Raum so gering wie möglich gehalten werden, ohne dass die Transienten zu verwaschen klingen und sich nicht mehr klar durchsetzen.

Bei dem Titel „Bloom“ ist die Aufstellung der Musiker an sich zwar fast identisch, allerdings ist die Instrumentierung etwas anders. Die Rollen von Lukas Klotzbach und Fabian Glück sowie deren Setups ändern sich dabei nicht. Allerdings spielt Drummer Bastian Kilper hier sowohl Piano als auch Perkussion und Becken, Bassist Marian Hepp spielt dagegen neben dem E-Bass auch einen Synthesizer und ebenfalls ein Becken. Um dies möglich zu machen, tauschen Lead-Gitarrist Fabian Glück und der Bassist Hepp bei diesem Song den Platz und es ergibt sich die folgende Aufstellung:

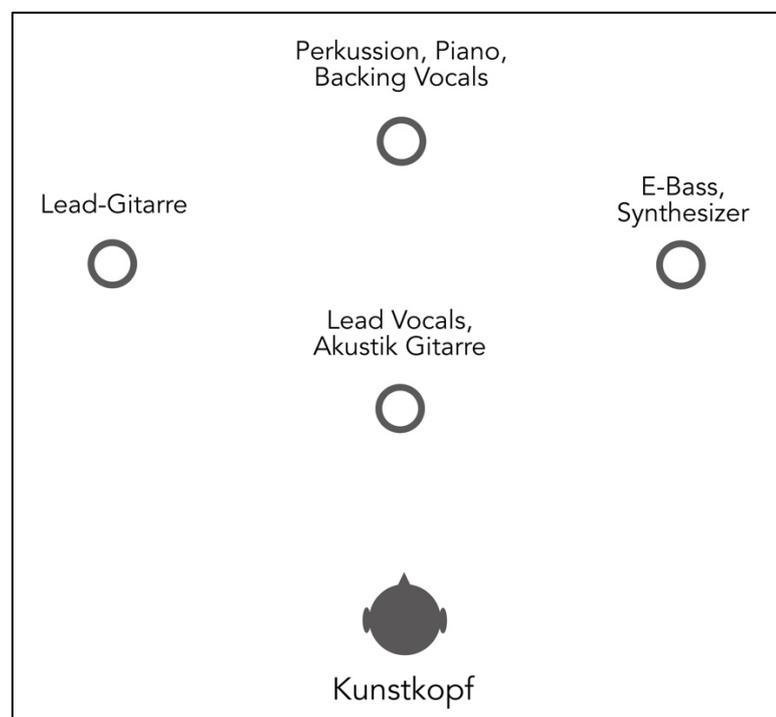


Abb. 20: Aufstellung "Bloom" ¹¹³

¹¹³ Eigene Darstellung

Nachdem die Arrangements, die Instrumentierung und auch die Aufstellung final festgelegt wurden, war es nun möglich, einen genauen Spur- und Mikrofonplan anzulegen und sich Gedanken um die Aufnahmeinfrastruktur zu machen. Auch dieser Planungsprozess fand im Rahmen der Vorproduktion statt, wird jedoch aus strukturellen Gründen erst im nächsten Kapitel genauer ausgeführt. (siehe 5.1.)

4.5. Finanzierung

Obwohl der Fokus dieser Arbeit klar auf die audioteknische Organisation und Durchführung der Livesession gerichtet ist, soll dennoch auch die Finanzierung angesprochen und beleuchtet werden. Dieser Aspekt stellt ebenso ein wichtiges Rahmenelement bei der Planung einer Livesession dar und sollte daher nicht unberücksichtigt bleiben.

Da eine Livesession in der Regel nicht direkt verkauft werden kann, sondern maximal durch sehr viele Klicks auf Videoplattformen wie YouTube Geld generieren kann, dient sie in erster Linie zu Marketing- und Promotionszwecken. (siehe 2.2.) Für Musiker stellt dieses Format also in erster Linie eine Investition dar, aus der man, zumindest aus rein finanzieller Sicht, keinen effektiven Mehrwert herausziehen kann. Aus diesem Grund versuchen viele Künstler mit externen Partnern zusammen zu arbeiten, die mit ihren Sessions einerseits eine große Reichweite durch eine bestehende große Gefolgschaft erreichen können und andererseits ein Produktionsteam zur Verfügung stellen, welches die Session technisch umsetzt. Dadurch entfallen die hohen Produktionskosten für die Musiker. Da diese renommierten Anbieter allerdings nur ungern mit unbekanntem, kleineren Künstlern zusammenarbeiten, sondern im Gegenzug lieber namhafte Musiker mit einer großen Nachfrage zu sich einladen, bleibt diesen unbekannteren Künstlern in der Regel nur übrig, die Produktion selbst zu realisieren.

Aus diesem Status Quo heraus, wurde auch die im Rahmen dieser Arbeit realisierte Livesession geplant, da der Verfasser, wie bereits erwähnt, gleichzeitig auch Teil der Band ist. Da nur ein minimales Finanzvolumen zur Verfügung stand, kann man hier also von einer Low-Budget-Produktion sprechen. Im Zuge dessen sollten alle erforderlichen Ressourcen, wie die Location, das Filmteam, die Leihe von Equipment, aber auch die

Unterkunft und die Verpflegung des Teams möglichst wenige Kosten verursachen, das Ergebnis sollte dennoch möglichst professionell ausfallen.

Für die visuelle Umsetzung der Livesession konnte die Stuttgarter Filmproduktionsfirma Foxframes gewonnen werden, ein der Band Into The Fray nahestehendes Kollektiv von Filmemachern. Diese übernahmen gegen eine geringe, mehr symbolische Aufwandsentschädigung die komplette Videoproduktion inklusive der Leihe von benötigtem Licht- und Kameraequipment sowie die Fahrtkosten. Da das Unternehmen auch ein Tonstudio betreibt, wurde darüber hinaus ein Großteil des bei der Session genutzten Audio-Equipments von ihnen zur Verfügung gestellt. Die restliche Audio-Ausrüstung wurde aus Privatbeständen des Verfassers dieser Arbeit und der Bandmitglieder von Into The Fray bereitgestellt. Das Kunstkopfmikrofon wurde von der Hochschule der Medien Stuttgart ausgeliehen.

Der Aufzeichnungsort, die Alte Kirche, durfte dank des Vereins „Alte-Kirche e.V.“ und der Gemeinde Rulfingen kostenlos über den Zeitraum eines gesamten Wochenendes genutzt werden. Es fielen hierbei also keinerlei Mietkosten an. Alle Beteiligten wurden darüber hinaus kostenlos in einem dem Verfasser nahestehenden Privathaushalt untergebracht.

Zusammengefasst kann man also sagen, dass die Kostenaufstellung, maßgeblich durch das Aufbringen unentgeltlicher Leistungen seitens bandnaher dritter Akteure, sehr geringgehalten werden konnte.

Kostenpunkt	Kosten
Aufwandsentschädigung Foxframes	300,00
Spritkosten PKW 1	31,00
Spritkosten PKW 2	24,00
Leihgebühr Hazer	45,00
Verpflegung	164,00
Verbrauchsmaterial	28,00
<i>gesamt</i>	592,00

Abb. 21: Kostenaufstellung der Livesession ¹¹⁴

¹¹⁴ Eigene Darstellung

Die anfallenden Kosten betragen sich also insgesamt auf rund 600 Euro. Gemessen an üblichen Produktionskosten für vergleichbare Dreharbeiten ist dies ein extrem geringer Betrag. Es sollte allerdings angemerkt werden, dass bei dieser Produktion sowohl der Audioengineer als auch das Filmteam unentgeltlich gearbeitet haben und auch die Unterkunft und die Location kostenlos zur Verfügung gestellt wurden. Davon ist im Normalfall nicht auszugehen, dementsprechend ist dieser Gesamtkostenbetrag auch keinesfalls als branchenüblich, sondern vielmehr als Produkt zahlreicher wohlwollender Unterstützer zu bewerten.

5. Aufnahme der Livesession

In dem folgenden Kapitel werden die Prozesse und Abläufe während der Aufnahme der Livesession genauer erläutert. Dabei wird zunächst die vor Ort genutzte Signalkette, die Mikrofonierung sowie der Soundcheck ausführlich dargelegt. Im Anschluss werden die generelle Aufnahmestrategie und die Herangehensweise beschrieben. Zum Schluss des Kapitels wird dann die Zusammenarbeit mit dem Filmteam und in diesem Kontext explizite Reibungspunkte und Diskussionen beleuchtet.

5.1. Signalkette, Mikrofonierung und Soundcheck

Bei der Aufnahme der Livesession war es aufgrund der Tatsache, dass als Aufzeichnungsort eine alte Kirche und kein audiotechnisch gut ausgestattetes Tonstudio gewählt wurde wichtig, eine gute Aufnahme-Infrastruktur zu gewährleisten. Es musste also ein möglichst kompaktes, mobiles Setup konzipiert werden, welches zuließ, ausreichend viele Kanäle aufzunehmen und auch vier autarke Monitorwege für die Musiker auszugeben. An dieser Stelle hätte sich auf jeden Fall der Einsatz eines digitalen Mischpults angeboten, da jedoch keines mit qualitativ zufriedenstellenden Vorverstärkern zur Verfügung stand, und die Mietkosten für ein hochwertiges Modell zu hoch gewesen wären, wurde ein anderer Weg eingeschlagen.

Die Wahl fiel dabei auf eine Kombination mehrerer Geräte mit hochwertigen Preamps und Analog-Digital-Wandlern, die in Kombination genutzt wurden. Die Basis stellte dabei das „Apollo Twin Duo MKII“ dar, welches kaskadiert mit einem „Apollo 8p“ genutzt wurde. Bei beiden Geräten handelt es sich um vollwertige Thunderbolt-Audiointerfaces des Herstellers Universal Audio, welche zusammen zehn Mikrofonvorverstärker haben. Um die für die Livesession zur Verfügung stehende Anzahl der Preamps auf 18 zu erhöhen, wurden via ADAT zusätzlich acht weitere Vorverstärker des Geräts „Esotar MP-8D“ der Marke Alto genutzt. Zur Aufnahme der einzelnen Spuren wurde die DAW „Pro Tools 11“ des Herstellers Avid und ein MacBook Pro verwendet. Die Signale wurden dabei mit einer Auflösung von 24 Bit und einer Samplerate von 48 kHz aufgenommen.



Abb. 22: Das mobile Aufnahmesetup vor Ort ¹¹⁵

Im Rahmen der Vorproduktion (siehe 4.4.) ergab sich im Austausch mit den Musikern bei der Probe ein grober Überblick über die zu erwartenden Signale der jeweiligen Songs. Somit war vorab bereits klar, dass 18 Kanäle ausreichen würden, um die Livesession zu realisieren.

Für die Umsetzung der Session wurden drei Tage ins Auge gefasst. Da das Filmteam erst am zweiten Tag anreiste, wurde so sichergestellt, dass an dem Vortag genügend Zeit vorhanden war, um einen ausgiebigen Soundcheck durchzuführen. Vor Ort angekommen musste zunächst einmal das im Proberaum erdachte Setup hinsichtlich Aufstellung und Position im Raum auf die Alte Kirche übertragen werden. Dabei wurden verschiedenste Positionen unter Berücksichtigung klanglicher und raumakustischer Aspekte ausprobiert. Maßgeblich war hier in erster Linie das Drumset, da dieses das lauteste musikalische Element darstellte und somit den größten Einfluss auf das Schallfeld im Raum nahm. Die Wahl fiel letztlich, auch unter der Berücksichtigung von visuellen Aspekten, auf eine relativ zentrale Position vor dem Altarraum.

¹¹⁵ Foto: Foxframes

Im Anschluss erfolgte die Mikrofonierung. Der Song „In Dreams“ besteht dabei lediglich aus den vier Elementen Schlagzeug, Bass, Gitarre und Gesang. Hier nimmt vor allem die Abnahme des Drumsets viele Kanäle ein, nämlich insgesamt sieben. Darüber hinaus benötigt auch die Lead-Gitarre von Fabian Glück durch die Abnahme der beiden Verstärker sowie die Abnahme des DI-Signals drei Channels. Die beiden Gesangsspuren, der E-Bass und die Akustik-Gitarre fallen dagegen mit insgesamt vier Spuren eher weniger ins Gewicht. Beachtet man abgesehen davon noch die zwei Stereo-Signale, die durch den Kunstkopf und zwei zusätzlich genutzte Raummikrofone entstehen, so ergibt sich eine Kanalliste mit genau 18 Signalen.

Nr.	Signal	Preamp / Input	Mikrofon
01	Kunstkopf L	Apollo Twin – In 1	Neumann KU 100
02	Kunstkopf R	Apollo Twin – In 2	Neumann KU 100
03	Bassdrum In	MP-8D – In 1	Shure Beta 91A
04	Bassdrum Out	MP-8D – In 2	Audix D6
05	Snare Top	MP-8D – In 3	Audix D1
06	Snare Bottom	MP-8D – In 4	Audix D2
07	Tom	MP-8D – In 5	Audix D4
08	Overhead L	MP-8D – In 6	Neumann KM184
09	Overhead R	MP-8D – In 7	Neumann KM184
10	Vocal Bastian	MP-8D – In 8	Sennheiser MD441
11	Akustik Gitarre	Apollo 8d – In 1	Neumann U87
12	E-Gitarre DI	Apollo 8d – In 2	DI-Signal
13	E-Gitarre Amp L	Apollo 8d – In 3	Sennheiser MD421
14	E-Gitarre Amp R	Apollo 8d – In 4	Shure SM57
15	Vocal Lukas	Apollo 8d – In 5	Shure SM7B
16	E-Bass	Apollo 8d – In 6	DI-Signal
17	Room L	Apollo 8d – In 7	Sennheiser MKH 20
18	Room R	Apollo 8d – In 8	Sennheiser MKH 20

Abb. 23: Kanalliste "In Dreams" ¹¹⁶

Die Auswahl der Mikrofone war im Zuge dessen maßgeblich von dem vorhandenen Equipment abhängig. Bei den Drums kamen für die Kessel hauptsächlich dynamische Mikrofone der D-Serie des Herstellers Audix zum Einsatz. Aufgrund der Tatsache, dass

¹¹⁶ Eigene Darstellung

diese speziell für den Einsatz am Schlagzeug entwickelt wurden, bieten sie den Vorteil, dass sie mittels Klemmen direkt an der Trommel angebracht werden können und außerdem recht kompakt konstruiert sind. Bezüglich einer visuellen Aufzeichnung bedeutet dies ein optisch ästhetischeres Bild durch weniger Stative und eine optisch dezente Mikrofonie. Außerdem überzeugen die Mikrofone auch durch einen druckvollen Klang und für das Schlagzeug vorteilhafte Richtcharakteristiken. Für die Bassdrum wurde zusätzlich das Grenzflächenmikrofon „Beta 91A“ des Herstellers Shure verwendet, um durch dessen spezielle Klangcharakteristik später im Mix mehr Einfluss nehmen zu können. Als Overhead-Mikrofone dienten zwei „KM184“ von Neumann, welche im AB-Verfahren über dem Drumset positioniert wurden. Die Wahl fiel dabei bewusst auf die Nieren-Richtcharakteristik, um den Raum so gut es geht auszublenden und ein möglichst direktes, ausgewogenes Gesamtklangbild des Schlagzeugs einzufangen, welches später binauralisiert und dadurch mit einem synthetisch-generierten Raumeindruck versehen werden kann. Auf die Mikrofonierung der Hi-Hat wurde bewusst verzichtet, um die Zahl der Kanäle gering zu halten.

Für die Abnahme der beiden Gitarrenverstärker wurden mit dem Sennheiser „MD421“ und dem Shure „SM57“ zwei für diesen Zweck recht typische dynamische Mikrofone verwendet. Hinsichtlich der Akustik-Gitarre fiel die Entscheidung auf eine Mono-Mikrofonierung mit einem Neumann „U87“ Großmembranmikrofon. Das Doppelmembrankapselsystem war dabei auf eine Nieren-Richtcharakteristik geschaltet und auf den Steg, allerdings recht nah am Schallloch der Gitarre, ausgerichtet. Eine Stereomikrofonierung hätte hier wenig Sinn gemacht, da das Signal später binauralisiert und frontal vorne angeordnet wird. Für diesen Zweck ist ein Monosignal deutlich besser geeignet.

Für die Aufnahme der Stimmen fiel die Wahl in beiden Fällen auf ein dynamisches Mikrofon. Um das Übersprechen des Schlagzeugs auf das Gesangssignal möglichst gering zu halten, wurde bei Drummer Bastian Kilper mit dem Sennheiser „MD441“ ein Mikrofon mit einer Supernieren-Richtcharakteristik verwendet. Für Hauptsänger Lukas Klotzbach wurde das Shure „SM7B“ verwendet. Durch den aufsteckbaren, dezenten Windschutz können Popp-Laute vermieden werden, ohne dass ein großer, visuell unästhetischer Poppschutz genutzt werden muss.

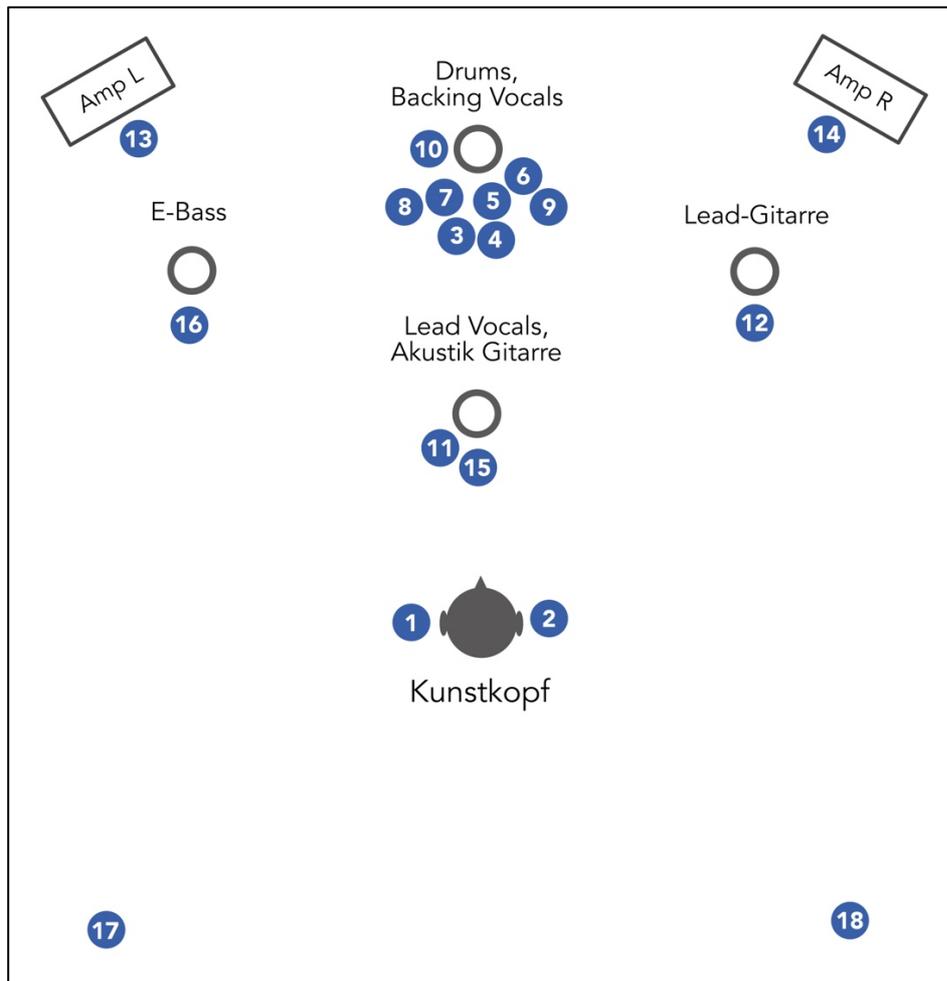


Abb. 24: Skizzierte Verteilung der Kanäle bei dem Song "In Dreams" ¹¹⁷

Als Kunstkopf-Mikrofon diente das Modell „KU100“ des Herstellers Neumann. Dabei handelt es sich um eines der gängigsten und hochwertigsten Modelle des Marktes. (siehe 3.3.) Besonders die gemittelte HRTF dieses Modells funktioniert bei den meisten Menschen sehr gut und gewährleistet eine gute räumliche Lokalisation.

Vor Ort kristallisierte sich schnell heraus, dass sich die Hauptstimme auf dem Kunstkopfsignal nicht gegen das Schlagzeug und den Rest der Band durchsetzen kann. Das Mikrofon hätte im Zuge dessen wenige Zentimeter entfernt von dem Gesicht des Sängers positioniert werden müssen, was einerseits die visuelle Ästhetik zerstört und andererseits eine direkte Mikrofonierung der Stimme platztechnisch sehr schwer gestaltet hätte. Daher kam der Entschluss auf, den „KU100“ etwas weiter weg zu positionieren, um zu gewährleisten, dass die Balance von den Gitarren und dem

¹¹⁷ Eigene Darstellung

Schlagzeug ausgewogen ist. Die Stimme wurde in diesem Kontext außen vorgelassen und musste später durch die Binauralsynthese des Einzelsignals in den Vordergrund gestellt werden. Letztlich stand das Kunstkopfmikrofon circa 1,5 m entfernt frontal vor dem Sänger.

Die Entscheidung, neben dem Kunstkopf zwei zusätzliche Raummikrofone aufzustellen, liegt in einer größeren Flexibilität im Mix begründet. Einerseits könnte man diese bei einem herkömmlichen Stereomix anstelle des Kunstkopfsignals nutzen, andererseits könnten sie durch die Binauralsynthese als Stützsingale fungieren, um die Räumlichkeit des Kunstkopf-Signals zu unterstützen. Hierfür wurden zwei „MKH 20“ des Herstellers Sennheiser verwendet. Dabei handelt es sich um ein Hochfrequenz-Kondensatormikrofon mit einer Kugelcharakteristik, welches nach Bedarf mittels einer schaltbaren Diffusfeldkorrektur für den Einsatz außerhalb des Hallradius optimiert werden kann. Damit ist es bestens als Raummikrofon geeignet.

Nr.	Signal	Preamp / Input	Mikrofon
01	Kunstkopf L	Apollo Twin – In 1	Neumann KU 100
02	Kunstkopf R	Apollo Twin – In 2	Neumann KU 100
03	Bassdrum	MP-8D – In 1	Audix D6
04	Pandeiro	MP-8D – In 2	beyerdynamic TG D57
05	Mono Overhead	MP-8D – In 3	Neumann KM184
06	-	MP-8D – In 4	-
07	Stage Piano L	MP-8D – In 5	DI-Signal
08	Stage Piano R	MP-8D – In 6	DI-Signal
09	Synthesizer	MP-8D – In 7	DI-Signal
10	Vocal Bastian	MP-8D – In 8	Sennheiser MD441
11	Akustik Gitarre	Apollo 8d – In 1	Neumann U87
12	E-Gitarre DI	Apollo 8d – In 2	DI-Signal
13	E-Gitarre Amp L	Apollo 8d – In 3	Sennheiser MD421
14	E-Gitarre Amp R	Apollo 8d – In 4	Shure SM57
15	Vocal Lukas	Apollo 8d – In 5	Shure SM7B
16	E-Bass	Apollo 8d – In 6	DI-Signal
17	Room L	Apollo 8d – In 7	Sennheiser MKH 20
18	Room R	Apollo 8d – In 8	Sennheiser MKH 20

Abb. 25: Kanalliste "Bloom" ¹¹⁸

¹¹⁸ Eigene Darstellung

Die Kanalliste des Titels „Bloom“ zeigt, dass die Instrumentierung hier etwas anders als bei dem Song „In Dreams“ war. Anstelle eines ganzen Drumsets kamen hier verschiedene Perkussion-Elemente zum Einsatz. Eine Cajon ersetzte dabei die Bassdrum und ein Pandeiro trat an die Stelle der Snare und der Tom. Außerdem wurde auch ein Shaker sowie ein einzelnes Becken genutzt. Neben diesen perkussiven Elementen spielte Schlagzeuger Bastian Kilper hier auch ein Stagepiano des Herstellers Nord. Das eingesetzte Klang-Preset sollte dabei dem akustischen Klang eines Klaviers entsprechen. Bassist Marian Hepp spielte neben dem E-Bass auch einen Synthesizer sowie am Ende des Songs zusätzlich auch das Becken. Die Rollen von Fabian Glück und Lukas Klotzbach sowie deren Setups blieben unverändert. Die Kanalliste änderte sich also nicht in allen Belangen.

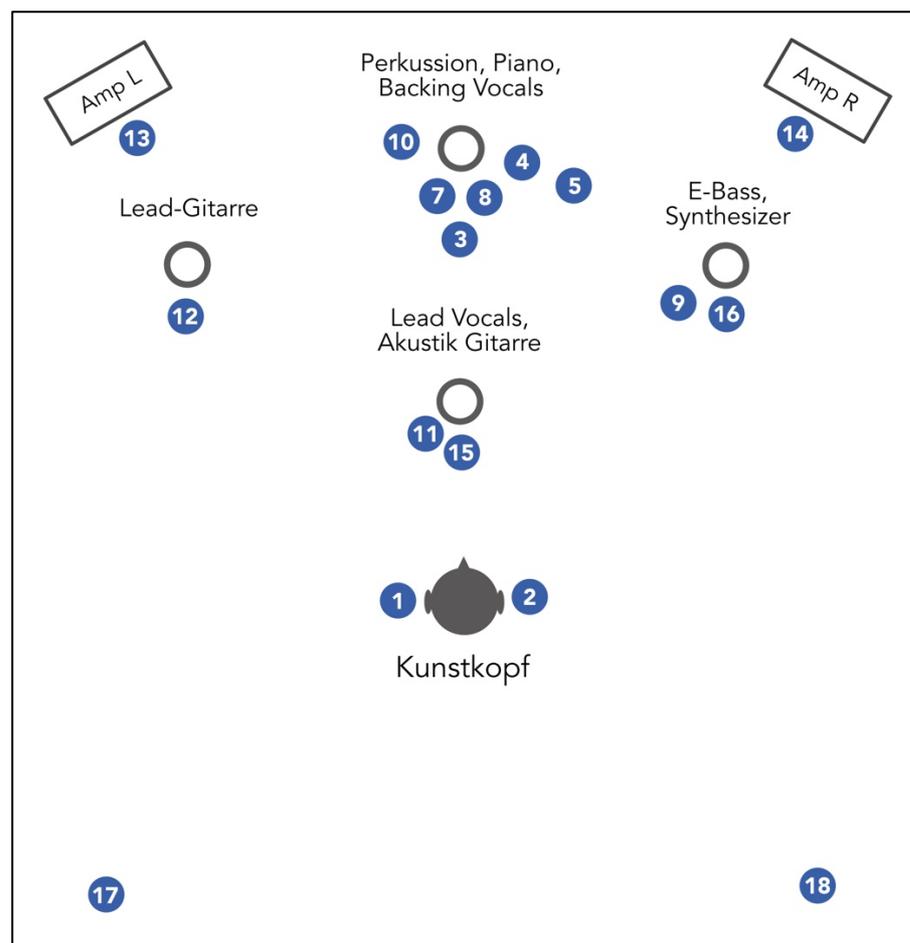


Abb. 26: Skizzierte Verteilung der Kanäle bei dem Song "Bloom" ¹¹⁹

¹¹⁹ Eigene Darstellung

Hinsichtlich der Mikrofonierung wurde für die Cajon das Audix „D6“, welches vorher als Bassdrum-Mikrofon diente, genutzt. Für die Abnahme des Pandeiro wurde ein darunter geklemmtes TG D57 des Herstellers beyerdynamic verwendet. Da das daraus resultierende Signal allerdings sehr basslastig und eher dumpf klingt, sollte dies später im Mix unbedingt in Kombination mit dem Overhead-Signal genutzt werden.

Da alle Overhead-relevanten Schallquellen bei dem Song „Bloom“ recht kompakt auf einer Seite des Drummers angesiedelt sind, fiel die Entscheidung hier im Gegensatz zu dem anderen Titel auf eine Mono-Mikrofonierung. Eine stereofone Abnahme hätte hier wenig Sinn gemacht, da man aufgrund der dichten Anordnung von Shaker, Pandeiro und Becken je nach Aufstellung entweder zwei sehr ähnliche Overhead-Signale oder ein sehr unausgewogenes Stereobild erhalten hätte. Beides wäre für die spätere Binauralisierung eher unvorteilhaft gewesen. Da alle Instrumente räumlich von ungefähr dem gleichen Punkt ausgingen, ist ein Mono-Overhead-Signal, welchem später beim Mixing eine bestimmte Position im virtuellen Raum zugeordnet wird, deutlich effektiver. Das Stagepiano und der Synthesizer wurden jeweils mithilfe von DI-Boxen aufgenommen. Da es sich bei den Synthesizer-Klängen um Bass-Signale handelt, genügte hier ein Mono-Signal.

Nachdem die Mikrofonierung erfolgreich abgeschlossen war, einige Audio-Probendurchläufe aufgenommen und alle Einzelsignale auf ihre Qualität überprüft wurden, musste gewährleistet werden, dass allen Musikern ein ausgewogenes personalisiertes Monitorsignal zur Verfügung steht. Da alle Künstler In-Ear-Systemen ausgerüstet sind, welche auch bei Konzerten der Band zum Einsatz kommen, war das nötige Equipment hierfür bereits bandseitig vorhanden. Über die Outputs der Universal Audio Interfaces konnte in der DAW „Pro Tools“ mittels AUX-Wegen für jeden Musiker ein individualisierter In-Ear-Mix generiert werden. Da die Binauralsynthese aufgrund der benötigten großen Prozessor-Leistung und damit einhergehenden Latenzen erst in der Postproduktion zum Einsatz kommen sollte, handelte es sich hierbei um herkömmliche Stereomischungen. Dies hatte den Vorteil, dass die Künstler mit dem von Live-Konzerten gewohnten Monitor-Sound performen konnten. Das Kunstkopf-Signal wäre darüber hinaus einerseits zu indirekt und wenig durchsetzungsfähig gewesen und hätte andererseits im Raum stumme Instrumente, die über eine DI-Box abgenommen wurden, nicht berücksichtigt.

Beim Soundcheck wurde zunächst „In Dreams“ und anschließend „Bloom“ vorbereitet. Bei der späteren Aufnahme war die Song-Reihenfolge umgekehrt, da man so das Setup von „Bloom“ nach dem Check einfach belassen konnte.

5.2. Aufnahmestrategie und Herangehensweise

Nachdem das Filmteam den Aufbau des Lichtsetups sowie die Vorbereitung des Equipments abgeschlossen hatte, konnte mit den Aufnahmen für die Livesession begonnen werden. Im Vorfeld wurde sich darauf geeinigt, dass jeder der beiden Songs mehrmals nacheinander durchgespielt werden muss. Dies lag in der Tatsache begründet, dass nur mit einer Kamera gedreht wurde. Es wurden also systematisch alle Kameraeinstellungen und -winkel, die der Regisseur für erforderlich hielt, in mehreren Takes abgearbeitet. In der Postproduktion wurden diese dann zu einem ästhetischen visuellen Gesamtwerk zusammengefügt. Eine authentische Livesession hinsichtlich einer One-Take-Aufnahme war also aus videotechnischer Sicht gar nicht möglich, ohne dass man sich auf eine statische totale Kameraeinstellung hätte einigen müssen. Um dies zu vermeiden und ein abwechslungsreiches, professionelles Video mit vielen Perspektiven zu kreieren, entschied man sich daher für das erläuterte Vorgehen.

Aufgrund der Tatsache, dass die Musiker also in jedem Fall mehrere Durchgänge der Songs für das Filmteam spielen mussten, führte dies zu dem Entschluss, diese Herangehensweise auch auf den auditiven Aspekt der Session zu übertragen. Auch hier wurden also mehrere Takes aufgenommen, aus denen später eine Audiofassung geschnitten werden konnte. Um dabei ein stetig gleichbleibendes musikalisches Tempo und damit ein reibungsloses Schneiden zwischen Durchläufen sowie eine Synchronität zwischen Bild und Ton verschiedener Takes sicherzustellen, hörten die Künstler auf ihrem Monitor-Mix auch ein Klick-Signal. Da die Musiker allerdings ohnehin auch bei Konzerten auf einen Klick spielen, stellte dieses Vorgehen keine Umstellung für diese dar. Wichtig war allerdings, dass die Musiker stets exakt gleich performten und keine musikalischen Variationen oder ähnliches einbauten.

Um Ermüdung, Stress und damit verbundene qualitative Einbußen zu verhindern, wurde zeittechnisch für jeden Song ein ganzer Tag eingeplant, um ausreichend Pausen und

genügend Schlaf gewährleisten zu können. Insgesamt wurden pro Song zwischen 10 und 15 Durchläufe aufgenommen, um alle gewünschten Kameraeinstellungen abzarbeiten. Dies führte sowohl bei den Musikern als auch bei dem Kameramann zu einer schnellen Ermüdung und Anstrengung. Die großzügige Zeitplanung erwies sich aus diesem Grund während der Produktion als sehr positiv und zielführend. Darüber hinaus hatte das anberaumte große Zeitfenster auch visuelle Gründe. „Bloom“ sollte bei Dunkelheit und „In Dreams“ bei Tag aufgenommen werden, um zwei verschiedene ästhetische Looks einzufangen. Der erste Titel wurde daher am Abend des zweiten Tages aufgenommen, der Zweite dann am Vormittag des Folgetags bei Helligkeit.

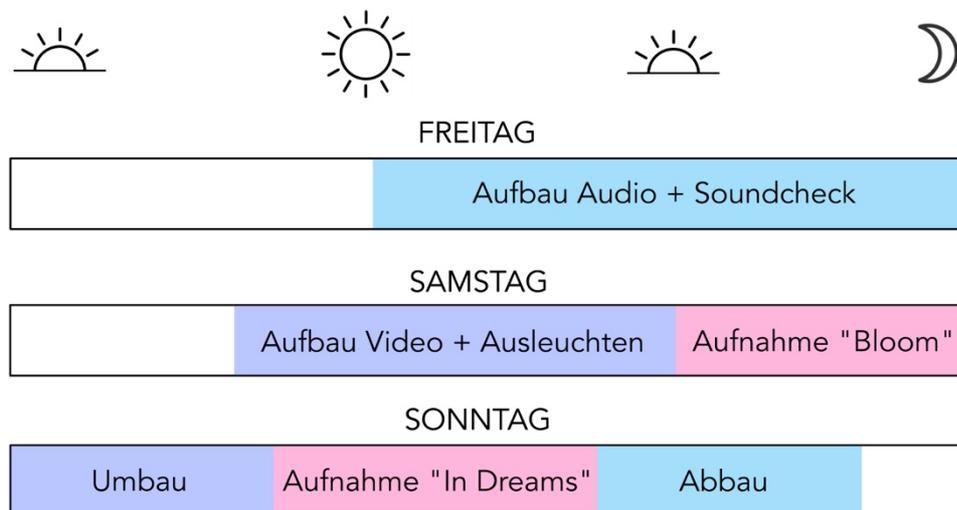


Abb. 27: Visuelle Darstellung des Zeitplans ¹²⁰

5.3. Die Zusammenarbeit mit dem Filmteam

Durch den Einfluss der visuellen Komponente mussten während der Produktion aus tontechnischer Sicht einige Kompromisse eingegangen werden. Diese bewegten sich allerdings in einem absehbaren Rahmen und waren somit präventiv gut zu lösen.

In diesem Kontext war die Mikrofonierung ein relevanter Faktor. Bereits im Vorhinein wurde hierbei durch eine gezielte Wahl des Equipments, beispielsweise hinsichtlich der

¹²⁰ Eigene Darstellung

Schlagzeug-Abnahme, auf eine optisch dezente Umsetzung geachtet. Klangliche Abstriche mussten bei der Auswahl des Gesangsmikrofons von Sänger Lukas Klotzbach gemacht werden. Hierbei kam zunächst ein „U47“ Großmembran-Kondensatormikrofon des Herstellers Neumann zum Einsatz, welches sehr gut mit der Stimme harmonierte und ein hervorragend aufgelöstes Klangbild lieferte. Da der dafür nötige Poppchutz visuell allerdings als sehr störend empfunden wurde, da er das Gesicht größtenteils verdeckte, fiel die Wahl nach Absprache mit den Musikern und dem Filmteam letztlich auf das „SM7B“ von Shure. (siehe 5.1.)

Ein weiterer zu beachtender Aspekt im Kontext der Zusammenarbeit zwischen Video und Audio, war das simultane Aufnehmen des visuellen und auditiven Materials. Da die Mikrofone, vor allem der Kunstkopf natürlich auch hörbare Schritte und Bewegungsgeräusche mitaufnahmen, bestand stetig das Risiko, dass das Filmteam bei der Bewegung durch den Raum auf der Aufnahme hörbare Störgeräusche verursacht. Abgesehen davon stellte der Kameramann besonders bei Nahaufnahmen ein Hindernis für den Schall dar, wenn er sich zwischen dem Kunstkopf und der Schallquelle befand. Ist die effektive Fläche des Hindernisses, also die des Kameramanns, größer als die Wellenlänge, so bildet sich ein Schallschatten, in dem kein Direktschall zu hören ist.¹²¹ Befand sich der Kunstkopf also in dem Schallschatten des Kameramanns, so äußerte sich dies bei höheren Frequenzen mit einer verhältnismäßig kleinen Wellenlänge folglich durch eine indirektere, dumpfere Klangästhetik der Schallquelle auf dem Kunstkopf-Signal.

Dieser Umstand hatte zur Folge, dass aus auditiver Sicht nur die Takes bedenkenlos genutzt werden konnten, bei denen der Kameramann eine Totale gefilmt hat. Bei diesen Durchläufen hatte dieser nämlich einerseits einen recht großen Abstand zu den Musikern, um alle im Bild zu haben, und andererseits behielt er eine statische Position und bewegte sich nicht durch den Raum. Aus diesem Grund wurden bei den ersten vier bis fünf Durchläufen des jeweiligen Songs genau solche Einstellungen eingefangen. Aufgrund der Tatsache, dass Musiker sowieso nur einige Takes spielen können, bevor die Formkurve der Performance stetig abfällt, konnten so die besten musikalischen Durchgänge benutzt werden, und dieser Umstand stellte kein Problem dar.

¹²¹ Vgl. Friesecke. (2014). S. 32

6. Postproduktion

Nach der Vorproduktion und der eigentlichen Produktion der Livesession, welche den Aufnahmeprozess an sich darstellt, folgt die Postproduktion. Dieser Vorgang ist bei der Umsetzung des binauralen Konzepts nicht weniger wichtig als die Aufnahme vor Ort selbst. Aufgrund der angewandten Aufnahmestrategie mit mehreren musikalischen Durchläufen (siehe 5.2.) ist hier der erste Schritt der Schnitt eines Mastertakes für jeden Song durch eine Auswahl aus den einzelnen aufgenommenen Takes. Anschließend werden der Prozess der Bereinigung einzelner Signale, sowie die Session-Struktur, welche beim Einrichten festgelegt wird, genauer erläutert. Abschließend wird der kreative Prozess des Mixings und das dabei angewandte Vorgehen detailliert dargelegt. Im Zuge dessen wird auch der Vorgang des finalen audiatechnischen Produktionsschritts, dem Mastering, genauer erläutert.

6.1. Schnitt und Take-Auswahl

Bevor sich die Auswahl der besten Takes für den finalen Audioschnitt herauskristallisierte, musste zunächst das Vorgehen mit dem Filmteam abgesprochen werden. Im Fall der Session für diese Arbeit wurde im Zuge dessen abgemacht, dass ein finaler Audiotake angefertigt wird, auf den das Video-Material synchron geschnitten wird. Die Audio- und Videotakes werden also losgelöst voneinander behandelt, um bei beiden Aspekten maximal viel Bearbeitungsfreiraum zu haben. Aus diesem Grund bestanden bei dem Schnitt audioseitig keine Grenzen und es konnten die musikalisch besten Durchläufe genutzt werden.

Insgesamt standen für „Bloom“ fünf und für „In Dreams“ sechs Takes zur Auswahl, die verwendet werden konnten. Da diese musikalisch alle auf einem sehr guten Level waren, reduzierte dies den Arbeitsaufwand enorm, da nicht so viel Material durchgehört werden musste und dennoch alle Songabschnitte auf einem der Durchläufe überzeugend festgehalten wurden. Da schon während der Aufnahme der Livesession stets Notizen zu den Takes gemacht wurden, um später bei der Postproduktion einen

guten Überblick zu haben, war es recht einfach, den vermeintlich Besten als Basistake und somit als Grundlage für jeden der beiden Songs auszuwählen.

Aufgrund der Tatsache, dass sich alle Schallquellen und Mikrofone bei der Aufnahme im selben Raum und zusätzlich noch recht nah beieinander befanden, war auf den meisten Signalen allerdings ein großes Übersprechen anderer Instrumente vorhanden. Eine Ausnahme bildeten hier natürlich musikalische Elemente, die lediglich mit einem DI-Signal abgenommen wurden. Da auf dem Kunstkopfsignal, welches im Mix die Soundbasis darstellen soll, auch alle im Raum hörbaren Elemente abgebildet sind, war es sinnvoll, dass alle simultan hörbaren Signale stets von demselben Take sind. Wenn also geschnitten wurde, dann wurde spurenübergreifend zwischen den Takes geschnitten.

Der Mastertake des Songs „Bloom“ besteht in erster Linie aus dem fünften Take, welcher als Basis für den Schnitt gewählt wurde. Insgesamt gibt es während des Songs lediglich einen Schnitt. Dieser befindet sich ganz am Ende des Songs, in dem lauten Part der dynamischen Steigerung. Da hier bei dem Shredding-Einsatz der Lead-Gitarre ein Störgeräusch zu hören ist, welches vermutlich durch ein nicht ganz sauberes Greifen des Bundes entstand, musste hier auf das Ende eines alternativen Durchlaufs, nämlich dem des vierten, zurückgegriffen werden.

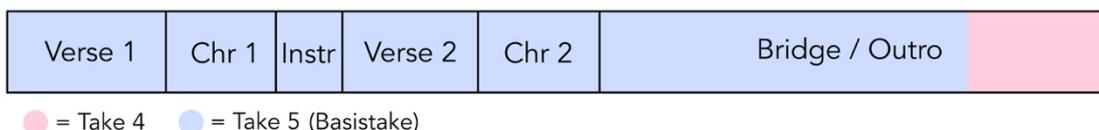


Abb. 28: Songablauf und Schnitte bei "Bloom" ¹²²

Bei dem Titel „In Dreams“ wurde im Laufe des Songs zweimal geschnitten. Bei dem Großteil des Songs ist dabei der dritte Durchlauf, welcher als Basis diente, zu hören. Da die ruhige Bridge, welche zu Beginn lediglich aus Gesang und Akustikgitarre besteht, bei dem fünften Take stimmlich überzeugender und akzentuierter war, wurde bei der Bridge dieser Durchlauf reingeschnitten. Der letzte Chorus stammt wieder von dem Basistake. Da die Drums auf dem fünften Take bei dem ersten Chorus hinsichtlich des

¹²² Eigene Darstellung

Timings besonders sauber gespielt waren, hätte sich hier angeboten, diesen auch aus diesem Take zu entnehmen. Da es sich allerdings immer noch um eine Live-Aufnahme handelt und die Performance anderer Instrumente darunter tendenziell eher gelitten hätte, wurde dies nicht berücksichtigt.

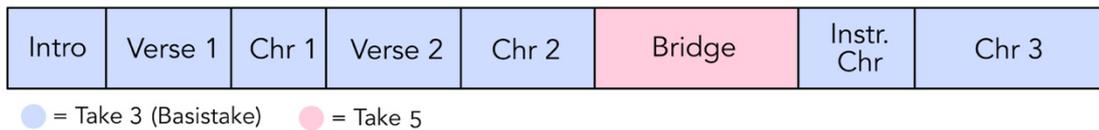


Abb. 29: Songablauf und Schnitte bei "In Dreams" ¹²³

Nachdem die beiden Mastertakes in der DAW „Pro Tools“ fertig zusammengeschnitten waren, wurden die Einzelspuren gerendert, da beim Mixing eine andere DAW zum Einsatz kam.

6.2. Signalbereinigung

Um sicher zu stellen, dass das Rendering der Spuren fehlerlos vonstatten ging, wurden diese vor dem Einrichten der Mixing-Sessions noch einmal gegengehört. Dabei wurde auch auf potentielle Störgeräusche, wie starkes Rauschen oder Brummen, geachtet, um diese gegebenenfalls nach Möglichkeit zu entfernen.

Bereits bei der Aufnahme der Livesession fiel bei einem der Gitarrenverstärker ein Netzbrummen im Bereich von 100 Hz auf, welches vor Ort nicht vollständig eliminiert werden konnte. Im Raum und somit auch auf dem Kunstkopfsignal war dieses Brummen kaum wahrzunehmen und somit irrelevant, auf dem Einzelsignal des Verstärkers trat dieses jedoch sehr störend in Erscheinung. Aus diesem Grund wurde das Netzbrummen mit dem Reparatur-Programm „RX 7“ des Software-Herstellers iZotope und dem darin inkludierten „De-Hum“-Modul spektral herausgerechnet. Dabei wurden nicht nur die Grundfrequenz von 100 Hz sondern auch die Oberschwingungen weitestgehend

¹²³ Eigene Darstellung

entfernt, ohne dabei hörbare Artefakte oder Klangeinbußen auf dem Gitarren-Signal zu generieren.

Abgesehen von diesem Signal wiesen alle weiteren Aufnahmespuren eine gute Qualität auf und konnten ohne weitere Bereinigung beim Mixing genutzt werden. Auch das Übersprechen konnte durch den ausgiebigen Soundcheck vor der Aufnahme gut unter Kontrolle gehalten werden. Dennoch wurde der Versuch unternommen, das Übersprechen der Instrumente auf das Lead-Vocal-Signal in der Postproduktion weiter zu reduzieren, da dieses später im Mix bei beiden Songs eine sehr dominante Rolle einnehmen wird. Hierfür wurde das „Music Rebalance“-Modul von „RX 7“ genutzt, welches das eingegebene Audiomaterial analysiert und in Vocals, Bass, Percussion und weitere Instrumente unterscheidet. Jedes dieser Elemente kann daraufhin hinsichtlich der Lautstärke angepasst werden.¹²⁴ Um das Vocal-Signal zu bereinigen, wurden daher alle Elemente außer die Stimme herausgezogen. Das Ergebnis wies allerdings bei den Lead-Vocal-Signalen beider Songs stark hörbare Artefakte und Klangeinbußen auf. Aus diesem Grund wurden letztlich doch die Originalsignale mit dem Übersprechen verwendet.

Neben den hier angesprochenen Aktivitäten zur Bereinigung der Signale wurden darüber hinaus natürlich weitere Schritte unternommen, wie die Filterung durch Hoch- und Tiefpassfilter, der Einsatz von Gate-Effekten oder das Freischneiden von einzelnen Signalen. Diese Bearbeitungsprozesse wurden allerdings erst während dem Mixing durchgeführt, um die klanglichen Auswirkungen im musikalischen Kontext besser einschätzen zu können.

6.3. Einrichten der Mixing-Sessions

Da das Mixing in der DAW Live 10 des in Berlin sitzenden Unternehmens Ableton durchgeführt wurde, mussten die gerenderten Einzelspuren der geschnittenen Mastertakes pro Song in einer neuen, übersichtlichen Session angelegt werden. Die Hauptgründe hierfür sind gewohnheitsbedingte Faktoren und Mixing-Routinen des Verfassers, der es gewohnt ist, in „Live“ zu arbeiten und zu mischen. Um ein optimales

¹²⁴ Vgl. iZotope, Inc. (2020)

Ergebnis zu gewährleisten, sollte sichergestellt werden, dass mit gewohnten Tools und Plug-Ins gearbeitet werden konnte. Da der fertige binaurale Mix darüber hinaus als einfaches Stereo-File gerendert wird, reichen die Export-Möglichkeiten dieser DAW hier vollkommen aus.

Bei der Suche nach einer geeigneten Software für die Binauralisierung der Einzelsignale standen letztlich zwei Programme in der engeren Auswahl. Dabei handelte es sich einerseits um den Spatial Audio Designer des Unternehmens New Audio Technology von Tom Ammermann und andererseits um dearVR Pro des Düsseldorfer Anbieters Dear Reality. Beide Tools verbindet eine übersichtliche grafische Oberfläche und eine einfache Einbindung in die DAW durch die Plug-In-Formate AAX, AU und VST, sowie im Falle des Spatial Audio Designers auch noch durch RTAS.

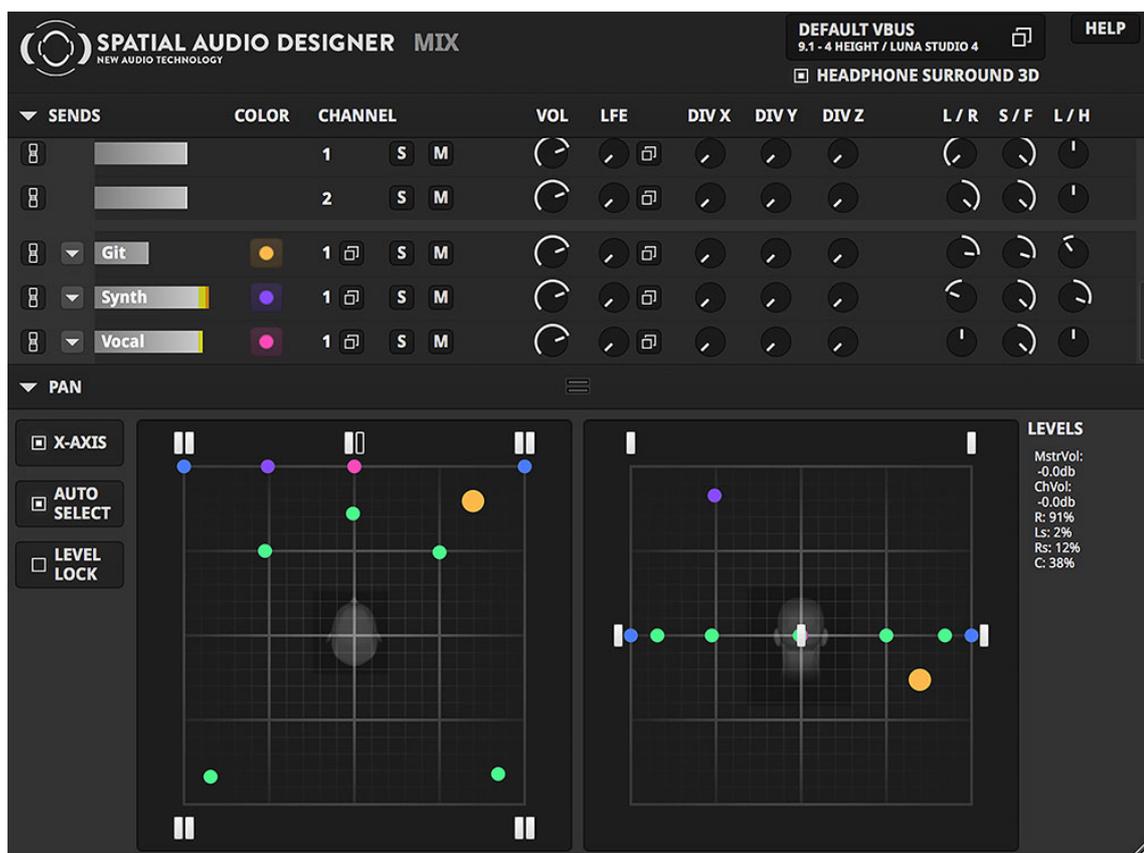


Abb. 30: Grafische Oberfläche des Spatial Audio Designers ¹²⁵

¹²⁵ New Audio Technology GmbH. (2020).

Die Wahl fiel letztlich auf das Programm dearVR Pro. Die Gründe hierfür lagen vor allem an dem Umstand, dass bei dem Spatial Audio Designer lediglich Lautsprecheranordnungen binauralisiert werden können.¹²⁶ Dabei stehen verschiedene Räumlichkeiten zur Auswahl, die allerdings hauptsächlich im Studiobereich angesiedelt sind und teils eine färbende Klangcharakteristik aufweisen. Klangobjekte, die zwischen den virtuellen Lautsprechern positioniert werden, klingen als Phantomschallquelle mitunter weniger überzeugend. Das Tool von Dear Reality bietet im Gegenzug über 45 sehr unterschiedliche Raum-Presets, welche auf Faltung und algorithmischen Berechnungen basieren. Aufgrund der großen Auswahl, welche vom Innenraum eines Autos bis zur Kathedrale reicht, stellte dies bei der Imitation des Raums der Alten Kirche die bessere Auswahl dar. Sehr vorteilhaft ist außerdem die Möglichkeit, die Reflektionen im Raum durch verschiedene Parameter zu beeinflussen.¹²⁷



Abb. 31: Grafische Oberfläche von dearVR Pro¹²⁸

Auch hinsichtlich der Einbindung in die DAW bietet dearVR Pro für das Mixing der Livesession mehr Vorteile als der Spatial Audio Designer. Bei Letzterem sendet man durch einzelne sogenannte Send-Module in den unterschiedlichen Kanälen alle Einzelsignale zu einem Mix-Modul, in welchem die Binauralisierung all dieser Signale zentral geregelt und als Stereosignal ausgegeben wird. Da in Live 10 insertierte Plug-

¹²⁶ Vgl. New Audio Technology GmbH. (2020).

¹²⁷ Vgl. Dear Reality GmbH. (2020)

¹²⁸ Dear Reality GmbH. (2020)

Ins, also auch die entsprechenden Send-Module, stets Pre-Fader eingesetzt werden, kann folglich die Balance der unterschiedlichen binauralisierten Signale ausschließlich in dem Mix-Modul festgelegt werden. Bei dearVR Pro ist dies einfacher gelöst. Hier bindet man das Programm mittels eines Plug-Ins ein, welches in jede Einzelspur geladen wird. Die Binauralisierung findet also auf jedem Kanal separat statt, und jeder Kanal gibt ein binaurales Signal aus. Dies nimmt zwar mehr CPU in Anspruch als das System des Spatial Audio Designers, sollte im Rahmen der Session allerdings in einem unproblematischen Rahmen bleiben. Mit dem Programm von Dear Reality bleibt der Workflow während dem Mixing also sehr übersichtlich, da der herkömmliche, gewohnte Signalfluss der DAW intakt bleibt.

Aufgrund der Tatsache, dass binauralisierte Signale möglichst nicht mehr bearbeitet werden sollten, um das dabei erzeugte künstliche Ohrsignal intakt zu halten und eine inverse Filterung durch das Gehör zu gewährleisten, muss das dearVR Pro-Plug-In logischerweise immer am Ende des jeweiligen Kanalzugs, nach einer potentiellen Bearbeitung, die das Spektrum des Signals beeinflussen könnte, eingesetzt werden. Aus diesem Grund macht es bei der Strukturierung der Mixing-Sessions für die beiden Songs keinen Sinn, Signale für die klangliche Bearbeitung in Gruppen zusammen zu legen, da dies die bereits erfolgte Binauralisierung möglicherweise wieder zerstören würde. Obwohl die Summenbearbeitung in Gruppen bei dem Erstellen herkömmlicher Stereo-Mischungen durchaus üblich ist, beispielsweise bei der Kompression von Drums,¹²⁹ stehen hier beim Mixing also im Grunde alle Elemente für sich und laufen nach der Binauralisierung direkt in den Master und werden summiert.

Um die Sessions übersichtlich zu halten, wurden in der Vorbereitung dennoch Gruppen angelegt. Allerdings nicht um die beinhalteten Signale in Summe zu bearbeiten, sondern um die Spuren zu sortieren und eine gute Grundordnung zu schaffen und zu bewahren. Da alleine für die Lead-Gitarre beispielsweise drei verschiedene Signale vorlagen und darüber hinaus bereits vor dem Mixing zu erahnen war, dass im späteren Prozess einige Signale, beispielsweise die Lead-Vocals, auf mehrere Spuren aufgeteilt werden müssen, um verschiedene Bearbeitungsketten zu nutzen, erwies sich dies schlichtweg als logischer Schritt. Auch Stereosignale, welche in Form von zwei

¹²⁹ Vgl. Owsinski. (2007). S. 88f

Monosignalen rausgerendert wurden, wurden in Gruppen zusammengefasst. In Kombination mit einer guten Farbcodierung blieben die beiden Mixing-Sessions so trotz einer Vielzahl an Spuren stets übersichtlich und gut strukturiert.



Abb. 32: Farbliche Session-Kodierung und Anordnung des Songs "Bloom" ¹³⁰

6.4. Mixing und Mastering

In dem folgenden Abschnitt wird der Vorgang des Mixing genauer dargelegt und erklärt. Generell wurde bei der Erarbeitung des binauralen Konzepts festgelegt, dass die allgemeinen Hörgewohnheiten der Musikkonsumenten bedient werden sollen, um keine Nachteile gegenüber herkömmlichen Stereomischungen zu erzeugen. (siehe 4.1.3) Dies muss besonders beim Mixing beachtet werden. Da allerdings die Positionen der einzelnen Instrumente bereits bei der Aufnahme festgelegt und mit dem Kunstkopf festgehalten wurde, gab es hier einen guten Anhaltspunkt, auf dem aufgebaut werden konnte. Bei dem gesamten Mixing-Prozess wurde dabei stets mit den finalen Videotakes gearbeitet, um zu gewährleisten, dass Ton und Bild miteinander harmonieren.

Im vorherigen Abschnitt wurde im Rahmen der Ausführungen hinsichtlich der Session-Einrichtung darauf hingewiesen, dass eine Summenbearbeitung in Gruppen bei diesem Projekt wenig Sinn macht. (siehe 6.2.) Aus diesem Grund werden in diesem Abschnitt alle musikalischen Elemente in der Mischung songübergreifend einzeln behandelt. Abschließend wird die finale Summenbearbeitung der Mischung auf dem Masterbus erläutert.

¹³⁰ Eigener Screenshot

6.4.1. Kunstkopf und Raumsignale

Die Ausgangsbasis bei dem Erstellen der Mischungen der beiden Songs stellte jeweils das Kunstkopfsignal dar. Es diente im Mixing-Prozess als Hauptsignal und als exakte Abbildung der akustischen Ereignisse im Raum, zu welchem nach und nach die Einzelsignale der Instrumente hinzugenommen wurden, um ein schlüssiges Gesamtbild zu kreieren. Gleichzeitig definiert es auch die Position des Hörers im Raum von welcher ausgegangen wird. Da es sich bereits um ein binaurales Signal handelte, fand hier keine weitere Klangbearbeitung statt, um die korrekte Lokalisation nicht durch spektrale Veränderungen zu beeinflussen. Aus diesem Grund wurde bereits bei der Aufnahme versucht, durch die Positionierung des Kunstkopfs im Raum und zu den Schallquellen ein möglichst ausgeglichenes Klangbild zu erhalten. (siehe 5.1.)

Um den räumlichen Klang des Kunstkopfsignals weiter zu verstärken und besonders den Raumklang von hinten zu unterstützen, wurden die beiden aufgenommenen Mono-Raumsignale mithilfe des dearVR Pro-Plug-Ins links und rechts hinter dem Hörer positioniert. Dies funktioniert im Mix erstaunlich gut und macht einen spürbaren Unterschied hinsichtlich der Immersion der Mischung aus. Als Inspiration für diesen Ansatz diente ein üblicher Mixing-Ansatz bei Surround-Produktionen. Hier wird das räumliche Ambiente oftmals auf die Surround-Kanäle hinter dem Hörer gemischt, um einen größeren umhüllenden Effekt zu generieren, ohne dabei die Instrumente um den Hörer zu positionieren.¹³¹

6.4.2. Vocals

In Popmusik-Produktionen stellt die Stimme in der Regel das wichtigste musikalische Element dar und sollte daher so früh wie möglich in den Mixing-Prozess mit einbezogen werden, um sicherzustellen, dass die Vocals noch genügend Platz im Mix haben, um sich gut durchzusetzen.¹³² Aus diesem Grund wurden bei der Erstellung der Mischungen

¹³¹ Vgl. Owsinski. (2007). S. 155

¹³² Vgl. Owsinski. (2007). S. 41

die Lead-Vocal-Signale des Sängers Lukas Klotzbach als erstes musikalisches Element integriert und ausgearbeitet.

Im Zuge dessen fand zunächst eine standardmäßige Bearbeitung statt, wie sie auch bei herkömmlichen Stereoproduktionen vollzogen worden wäre. Dabei wurden zunächst die scharfen S-Laute durch einen De-Esser am Anfang der Bearbeitungskette kontrolliert. Anschließend wurde eine Entzerrung des Frequenzgangs durch einen parametrischen, sauber klingenden Equalizer durchgeführt. Hierbei kam unter anderem ein Lowcut-Filter zum Einsatz, um das Übersprechen im Bassbereich zu minimieren. Außerdem wurden die Frequenzen um 230 Hz relativ breitbandig angehoben, um der Stimme etwas mehr Kraft zu verleihen. Darüber hinaus wurden mithilfe eines High-Shelf-Filters die Höhen dezent angehoben, um einen offeneren Vocal-Sound zu kreieren und die Konsonanten etwas mehr zu betonen.¹³³ Abschließend wurde die Hauptstimme dann noch komprimiert, um dem sehr dynamischen Gesangsstil des Sängers Lukas Klotzbach etwas entgegenzuwirken. Hier war ein subtiles Vorgehen sehr wichtig, da man durch eine zu starke Kompression und das vorhandene Übersprechen potentiell schnell negative Resultate bei den Klangbildern der anderen Instrumente erzeugt. Dies liegt vor allem daran, dass die Lead-Vocals das lauteste und prominenteste Element in der Mischung sind und darauf befindliches Übersprechen dementsprechend sehr laut im Mix wahrgenommen wird.

Im Falle der aufgenommenen Live-Session betraf dies in erster Linie das Schlagzeug, beziehungsweise die Perkussion-Elemente. Bei dem Song „Bloom“, welcher in dieser Hinsicht dezenter und leiser instrumentiert ist, taten sich im Zuge dessen weniger Probleme auf als bei dem Titel „In Dreams“. Dort konnte aus diesem Grund etwas stärker komprimiert werden.

Anschließend wurde das dearVR Pro-Plug-In in die Signalkette eingefügt, um die Binauralisierung durchzuführen. Dabei wurde die Hauptstimme, ungefähr entsprechend der realen Gegebenheiten bei der Aufnahme, direkt vor dem Hörer positioniert. Als Raumeinstellung wurde, wie bei den meisten anderen binauralisierten Signalen auch, das Raum-Preset „Chapel“ verwendet, welches im Kontext des Videos und auch aufgrund der Hörerfahrung vor Ort bei der Aufnahme am authentischsten klang.

¹³³ Vgl. Owsinski. (2007). S. 66

Um die starke Dynamik der Vocals, vor allem bei dem Song „Bloom“, zu kontrollieren, wurden außerdem mehrere Vocal-Tracks angelegt und Automationen erstellt. Auf weitere räumliche Effekte bei der Bearbeitung, wie Delays oder Reverbs, wurde aufgrund der angestrebten natürlichen Ästhetik aufgrund der Live-Charakteristik verzichtet. Der Raumhall, welcher durch das dear VR Pro-Plug-In realisiert wird, reicht hierbei vollkommen aus. Die einzige Ausnahme bildet der laute Schlussteil des Songs „Bloom“. Hier ist sehr dezent ein Echo der Hauptstimme abwechselnd links und rechts hinter dem Hörer zu vernehmen. Dies bietet einen zusätzlichen immersiven Effekt, ohne aufdringlich zu wirken und schien daher als förderlich für die Mischung.

Die Backing-Vocals von Schlagzeuger Bastian Kilper wurden nach demselben Vorgehen, allerdings erst zu einem deutlich späteren Zeitpunkt, im Mix untergebracht. Da diese nur sehr unscheinbar im Mix wahrnehmbar sind, stellte dies, abgesehen von dem Übersprechen des Schlagzeugs auf diesem Signal, keine ernstzunehmenden Probleme bei der Umsetzung dar.

6.4.3. Schlagzeug & Perkussion

Das Schlagzeug stellt bei dem Song „In Dreams“ die rhythmische Basis dar und wurde dementsprechend beim Mixing direkt nach den Vocals in den Mix integriert. Dabei wurden zunächst die beiden Overhead-Signale frontal links und rechts vor dem Hörer positioniert. Die Anordnung entsprach dabei ungefähr der optimalen Position eines imaginären Stereolautsprecherpaars im Raum. Die virtuellen Positionen der beiden Overheadsignale und die des Hörers bilden also ein gleichseitiges Dreieck.

Der daraus resultierende Klang entspricht nicht dem natürlichen Höreindruck bei der Aufnahme vor Ort, sondern wirkt im Gegenzug eher breiter und ähnelt dem gewohnten Klang eines Schlagzeugs aus einer herkömmlichen Stereomischung. Dies dient dazu, die Hörgewohnheiten zu bedienen und eine differenziertere Wahrnehmung der später hinzugefügten Einzelsignale des Schlagzeugs über ein breiteres Panorama gewährleisten zu können. Alternativ hätte man auch, wie bei dem Song „Bloom“, ein Mono-Overhead-Mikrofon nutzen können, das zentral vor dem Hörer positioniert wird. Dies hätte allerdings zur Folge gehabt, dass das komplette Drumset dann sehr kompakt

und punktuell von vorne wahrgenommen wird. Dies wurde angesichts der recht einfach gehaltenen Instrumentierung des Songs „In Dreams“ nicht als sinnvoll erachtet. Hier soll das Schlagzeug recht groß und dominant wahrgenommen werden.

Die Einzelsignale wurden zunächst alleinstehend mit Kompressoren, Equalizer und in manchen Fällen auch mit Gates bearbeitet, um ihnen den nötigen Druck zu verleihen, sich im Mix durchzusetzen. Daraufhin wurden die einzelnen Elemente mit dem dearVR Pro-Plug-In entsprechend ihrer vermeintlichen Position auf Basis der Overhead-Signale angeordnet. Da viele Signale durch die Binauralisierung deutlich an Druck verloren haben, mussten dann im Anschluss noch einmal Anpassungen an der vorangegangenen Bearbeitung durchgeführt werden.

Als etwas problematisch gestaltete sich hier beispielsweise das Snare-Signal bei „In Dreams“. Entweder rutschte das Signal zu weit in den Hintergrund und setzte sich nicht durch, oder es war zu dominant und integrierte sich nicht in den Mix. Letzteres trat dann ein, wenn man versuchte, ohne eine Binauralisierung zum Ziel zu kommen.

Die Bassdrum, beziehungsweise die Cajon, wurde im Prozess des Mixings nicht binauralisiert. Die Signale wurden also nicht verräumlicht, sondern direkt wiedergegeben und werden somit bei der Wiedergabe im Kopf lokalisiert. Dies liegt daran, dass „[...] tieffrequente Klangquellen wie Bassdrum und Bass durch die Binauralisierung ihre Kompaktheit verlieren und für ein druckvolles Klangerlebnis nicht verräumlicht werden sollten. Die erwähnte Im-Kopf-Lokalisation kann an dieser Stelle einen klanglichen Vorteil darstellen.“¹³⁴ So behalten die Drums einen gewissen Punch und eine gute Durchsetzungsfähigkeit. Besonders beim Song „Bloom“, bei dem es nur einige wenige Perkussion-Elemente gibt, ist es sehr wichtig, dass die Cajon, welche im Grunde das rhythmische Kernstück des Titels ist, einen druckvollen Klang hat.

6.4.4. Gitarren

Da die Akustikgitarre mit einem Mikrofon und nicht durch einen verbauten Tonabnehmer aufgenommen wurde, war auf dem Signal verhältnismäßig viel Übersprechen vorhanden, welches zunächst mittels eines Lowcut-Filters bestmöglich

¹³⁴ Pfeifer. (2018). S. 50

entfernt wurde. Wegen diesem Übersprechen konnte hier, ähnlich wie bei den Vocals, nur sehr dezent komprimiert werden. Aufgrund der Tatsache, dass Lukas Klotzbach sehr dynamisch spielt, wurden für die Akustikgitarre daher mehrere Spuren angelegt, und es wurden Automationen gefahren, um der Dynamik so besser entgegenwirken zu können. Bei der Binauralisierung wurde die Akustikgitarre am selben Punkt positioniert wie die Stimme. Allerdings wurde die Höhe, entsprechend den realen Bedingungen, etwas nach unten angepasst.

Aufgrund der Tatsache, dass die Lead-Gitarre sowohl durch zwei stereofon verkabelte Verstärker als auch durch eine Mono-DI-Box nach dem Effekt-Pedalboard aufgenommen wurde (siehe 4.4.), ergaben sich hier beim Mixing viele Freiräume. Durch diese Vorgehensweise und das bewusste Stummschalten der Amps während der Aufnahme durch Gitarrist Fabian Glück ist die Lead-Gitarre nur in ausgewählten Passagen im Raum und somit auf dem Kunstkopfsignal zu hören. Die beiden Mikrofonsignale wurden dabei zunächst durch Equalizer und Kompressor hinsichtlich Dynamik und Frequenzgang angepasst. Anschließend wurden die Signale durch das dearVR Pro-Plug-In entsprechend der tatsächlichen Positionen der Verstärker im Raum binauralisiert und verteilt. Die Signalanteile, in denen die Lead-Gitarre nicht über die Verstärker gespielt wurde, wurden durch Schnitte entfernt.

Die Lead-Gitarren-Elemente, die lediglich auf dem DI-Signal präsent sind, konnten im Gegenzug durch die Binauralsynthese frei im Raum verteilt werden. Dabei kommt hauptsächlich die Swelling-Spieltechnik zum Einsatz, bei der im Grunde kein klarer Anschlag zu hören ist, sondern sphärische „Klangwolken“ durch Reverb- und Delay-Effekte erzeugt werden. Diese eignen sich aufgrund ihrer unauffälligen Klangcharakteristik sehr gut für eine dynamische Verteilung im Raum durch eine Automation. Bei den entsprechenden Song-Passagen bewegt sich dieses Signal um den Hörer herum und fördert so die Immersion, ohne zu aufdringlich und unglaubwürdig zu wirken.

Bei dem Titel „Bloom“ spielte Glück darüber hinaus nach dem ersten Chorus ein sehr statisches Pattern, bei dem er stets dieselbe Note sehr abgehackt und akzentuiert anschlug. Da hier anfangs keine Vocals zu hören sind, eignet sich diese Stelle sehr gut, um dem Hörer die Immersion der Mischung sehr offensichtlich vorzuführen. Die Lead-Gitarre wurde dabei beim Prozess der Binauralisierung mittels Automationen um den

Kopf des Hörers herumbewegt. Dieses Element ist auch am Ende des Songs durch die Wiedergabe eines Looper-Pedals, das Fabian Glück vor Ort nutzte, noch einmal zu hören.

6.4.5. Stage-Piano

Das Stage-Piano kam lediglich bei dem Song „Bloom“ zum Einsatz. Beim Mixing wurden die beiden Einzelsignale, welche bei der Aufnahme der Session durch eine Stereo-DI-Box aufgezeichnet wurden, links und rechts von dem Hörer positioniert. Da die Perkussion-Elemente bei diesem Song auch aufgrund des Mono-Overhead-Signals alle zentral vor dem Hörer positioniert sind, legt sich das Piano so im Panorama darum herum und liefert einen großflächigen, offenen Klang.

Es bestand auch die Überlegung, lediglich ein Monosignal des Stage-Pianos zu nutzen und dieses links oder rechts von dem Hörer zu positionieren. Da das Instrument allerdings von Schlagzeuger Bastian Kilper zentral hinter dem Sänger Lukas Klotzbach gespielt wurde und bei der Aufnahme auch dort zu sehen ist, wurde dieser Ansatz schnell verworfen.

Bearbeitet wurde das Stage-Piano-Signal nur sehr dezent mit einem Equalizer und einem Kompressor, da kein Übersprechen vorhanden war und die beim Spiel entstandene Anschlagsdynamik nicht zerstört werden sollte. Abgesehen davon handelte es sich um bereits gut klingende Samples, die lediglich etwas in den Höhen angehoben werden mussten.

6.4.6. E-Bass und Synthesizer

Der E-Bass wurde vor Ort bei der Aufnahme durch eine DI-Box aufgezeichnet. Im Mix standen so alle Türen offen und es bestand nicht die Gefahr, dass der Bass auf dem Kunstkopfsignal und den Raummikrofonen unwiderruflich und undefiniert abgebildet wurde. Probleme durch Übersprechen oder Raumresonanzen im tieffrequenten Bereich spielen in der Postproduktion so keine Rolle.

Beim Mixing wurde das Signal zunächst durch eine virtuelle Verstärker-Simulation prozessiert und klanglich angepasst. Dabei kam eine Simulation des Amp-Modells „SVT 4 Pro“ des Herstellers Ampeg zum Einsatz. Anschließend wurde ergänzend ein Equalizer und ein Kompressor verwendet, um den Klang des E-Basses weiter zu optimieren.

Bei dem Song „Bloom“ kam neben dem E-Bass auch ein Software-Synthesizer zum Einsatz, welcher bei der Aufnahme durch ein Audio-Interface ausgegeben und durch eine DI-Box aufgenommen wurde. Dieses Synth-Signal spielt sich ebenso im tieffrequenten Bereich ab und ist besonders im Subbass sehr präsent. Da das Signal schon bei der Klangsynthese optimiert wurde, erfolgte hier bei der Mischung keine weitere klangliche Bearbeitung.

Hinsichtlich der Binauralisierung stellte sich bei beiden Bass-Signalen dieselbe Frage wie bei der Bassdrum. Durch eine Verräumlichung würden unter Umständen deutliche Einbußen der Kompaktheit und Durchsetzungsfähigkeit der Signale riskiert.¹³⁵ Da der Bass als Fundament im Mix eine tragende Rolle spielt, wurde daher zunächst der Ansatz verfolgt, die Signale, ohne eine Binauralsynthese in der Mischung zu etablieren. Bei dem recht dumpfen Synth-Signal funktionierte dies auch recht gut, da es sich um ein verhältnismäßig schmalbandiges Signal handelt, welches lediglich im Bass-Bereich präsent ist. Bei dem E-Bass, dessen Klang sich im Gegenzug aus einem recht breitbandigen Frequenzspektrum zusammensetzt, funktionierte dies allerdings nicht zufriedenstellend. Die Klarheit der Mischung ging dadurch verloren, und andere Elemente waren nicht mehr ausreichend differenziert wahrnehmbar, wohingegen der E-Bass zu aufdringlich im Vordergrund stand. Daher wurde dieses Signal bei beiden Songs verräumlicht und zentral vor dem Hörer, allerdings etwas nach unten versetzt, positioniert. Dieser Ansatz lieferte einen akzeptablen Klang, obwohl der zu erwartende klangliche Effekt eintrat. Angesichts der Live-Ästhetik der Produktion schienen diese Abstriche hinsichtlich der Durchsetzungsfähigkeit des Signals jedoch als annehmbar, um das Signal authentisch in den Mix zu integrieren.

Im ersten Teil des zweiten Chorus wurde darüber hinaus auch das Synth-Signal binauralisiert, da der Einsatz des Instruments hier zu dominant und präsent

¹³⁵ Vgl. Pfeifer. (2018). S. 50

wahrgenommen wurde. Durch die Verräumlichung konnte dieser Umstand entschärft werden.

6.4.7. Summenbearbeitung und Mastering

Das Mastering beschreibt einen Vorgang in der Audioproduktion, der in der Regel nach dem Mixing erfolgt und den Zweck erfüllt, die betreffende Mischung für die Vervielfältigung zu optimieren. Dabei können sowohl technische Fehler und Mängel behoben, aber auch klanggestalterische und ästhetische Änderungen vollzogen werden.¹³⁶ Dies geschieht in der Regel durch eine Bearbeitung der finalen Mastersumme der Mischung. Da die Lautstärke bei der Bewertung zweier Songs oder Mischungen eine essentielle Rolle spielt, wird diesem Faktor beim Mastering bereits seit den 1950er-Jahren besondere Aufmerksamkeit gewidmet.¹³⁷

Um die Mixe der Live-Session gegenüber anderen Produktionen konkurrenzfähig zu halten, wurde auch hier ein Mastering durchgeführt. Da es sich um einen binauralen Mix handelt, traten dabei allerdings einige Probleme auf, die beachtet werden mussten. Einerseits legen die richtungsbestimmenden Bänder nach Jens Blauert nahe, dass der Frequenzgang bei der Bearbeitung eines binauralen Signals nicht verzerrt werden sollte, um eine korrekte Lokalisation gewährleisten zu können. Andererseits wirkt sich auch eine Einengung der Dynamik nachteilig darauf aus.¹³⁸ Darüber hinaus sollten auch die Phasen einzelner Frequenzbänder nicht durch Filter beeinflusst werden, da auch dies den dreidimensionalen Höreindruck mindern könnte.

Die finale Summenbearbeitung auf dem Masterbus wurde daher auf das Wesentliche reduziert und diente ausschließlich dazu, die Mischung lauter zu machen. Es wurde keine ästhetische Klangbearbeitung durch EQs oder Multiband-Kompressoren durchgeführt, um das binaurale Signal intakt zu halten. Etwaige klanggestalterische Änderungen wurden direkt im jeweiligen Kanal des Einzelsignals, noch vor der Verräumlichung, realisiert. Auf der Mastersumme kamen letztlich lediglich ein Kompressor und ein Limiter zum Einsatz, um die Lautheit der Mischungen zu steigern.

¹³⁶ Vgl. Dickreiter. (2014). S. 880ff

¹³⁷ Vgl. Owsinski. (2007). S. 113

¹³⁸ Vgl. Theile. (2002).

7. Fazit und Ausblick

Im Rahmen dieser Arbeit wurden zunächst das Medium der Livesession und die Binauraltechnik als Grundlagen ausführlich erläutert, um eine verständliche Ausgangssituation für die folgenden Ausführungen zu gewährleisten. Im anschließenden praktischen Teil der Arbeit wurde dann die Konzeption und Umsetzung der binauralen Livesession detailliert dargelegt.

Dabei traten besonders bei der Planung im Vorhinein wichtige Fragen auf, die es zu klären galt. Hier ist vor allem die Entscheidungsfrage zu nennen, welcher klangliche Ansatz mit der Session verfolgt wird. Wenn ein möglichst realistisches Abbild des Schallfelds bei der Aufnahme reproduziert werden soll, ist dies meist mit verhältnismäßig indirekten Signalen verbunden, und die Mischung wirkt potentiell deutlich weniger druckvoll und unkontrollierter. Im Gegenzug erscheinen die Räumlichkeit und die Immersion der Produktion in diesem Fall deutlich authentischer und glaubhafter. Versucht man sich an die Hörgewohnheiten aus der herkömmlichen Stereophonie zu halten und vernachlässigt dabei die realistischen Raumanteile und Lautstärkeverhältnisse, so verhält sich dies genau gegensätzlich.

Da das erklärte Ziel vorab war, eine Livesession zu generieren, die die Hörgewohnheiten bedient und somit konkurrenzfähig zu herkömmlichen Stereoproduktionen ist, fiel die Wahl auf den zweitgenannten Ansatz. Um dennoch einen großen Mehrwert durch die immersiven Möglichkeiten der Binauraltechnik zu erzielen, musste daher ein Konzept erdacht werden, welches dies gewährleistet. Die Wahl fiel dabei auf eine Kombination aus Kunstkopfmikrofon und einer Polymikrofonierung, deren Einzelsignale in der Postproduktion binauralisiert werden konnten.

Dabei sollte die Immersion in erster Linie durch eine lebendige Räumlichkeit des Aufzeichnungsortes realisiert werden, welche den Hörer in Form von Nachhall und Reflektionen umgibt. Da die Band lediglich aus vier Musikern bestand und alle Instrumente tragend und dominant in dem musikalischen Klangbild der Songs vertreten waren, konnten diese darüber hinaus nicht hinter dem Hörer positioniert werden, ohne bestehende Hörgewohnheiten zu verletzen und somit Irritationen auszulösen. Daher

bestand die Konvention, dass alle Instrumente im Wesentlichen vor dem Hörer angeordnet werden mussten.

Das Ergebnis wirkt – so die subjektive Meinung des Autors - zwar überzeugend und immersiv, allerdings wird es Laien vermutlich schwerfallen, die entstandenen Mischungen auf Anhieb als dreidimensional zu identifizieren. Hätten sie die Möglichkeit zwischen einer herkömmlichen Stereoverision und der binauralen Version hin und her zu schalten, würden der Mehrwert und der immersive Aspekt schnell ersichtlich werden. Der Anspruch sollte jedoch sein, dass die Hörer dies auch ohne einen direkten Vergleich als dreidimensionale, binaurale Mischung einordnen können.

Dass dies vermeintlich nicht so ist, liegt vor allem daran, dass der Raumklang als maßgebliche dreidimensionale Komponente hinsichtlich seiner Klangcharakteristik sehr unscheinbar ist, und der Fokus der Hörer daher immer vorne auf den Instrumenten und beim Sänger liegt. Die einzige Ausnahme bildet hier die zweite Strophe des Songs „Bloom“. Hier wandert die Lead-Gitarre, die in dieser Passage sehr dominant ist, durch eine Automation um den Kopf des Hörers und generiert einen gewissen Aha-Effekt. Um eine noch bessere Immersion zu erreichen, sollten mehr demonstrative Signale wie dieses im Arrangement etabliert werden.

Solche Effekte stehen allerdings stets im Konflikt mit der visuellen Komponente der Livesession. Dies war auch ausschlaggebend dafür, dass in dieser Hinsicht bei den beiden Songs der Produktion nicht offensiver gemischt wurde. Es sollte unbedingt vermieden werden, dass Bild und Ton zu Irritationen beim Betrachter der Session führen, da dies die Immersion zerstören würde. Im Zuge dessen ist man bei einer Livesession also durch die visuelle Komponente gezwungen, konservativer zu mischen, als beispielsweise bei ausschließlich auditiven binauralen Musikmischungen. Haben die Musiker in visueller Hinsicht eine feste Position vor dem Hörer, sollten die Signale dementsprechend auch aus dieser Richtung wahrzunehmen sein.

Da im Rahmen der Popmusik stets von kleinen Besetzungen auszugehen ist, bei denen alle Musiker eine tragende Rolle spielen und dementsprechend vor dem Hörer positioniert sein sollten, wäre hier bei künftigen Produktionen ein Ansatz, zusätzliche Musiker in das musikalische Arrangement zu integrieren. Diese sollten allerdings eher ergänzend und untermalend agieren. Hierbei würden sich beispielsweise Streicher oder Bläser eignen, die um den Hörer herum positioniert werden können. Darüber hinaus

eignen sich elektronische Signale wie Synthesizer oder abstrakte Samples sehr gut für solche experimentellen Effekte, da der Hörer diesen keine spezifische akustische Klangquelle im Raum zuordnet. Dementsprechend werden Signale dieser Art weniger kritisch wahrgenommen, was eine authentische, räumliche Lokalisation betrifft.

Eine weitere Möglichkeit, die Immersion einer solchen binauralen Livesession noch weiter zu verbessern, wäre die Implementierung von einem Headtracker. In Kombination mit einer 360 Grad-Kamera könnte die auditive Szenerie vor Ort sehr gut repräsentiert werden und das Klangbild würde sich entsprechend der Kopfhaltung des Hörers anpassen. Durch die Möglichkeit der aktiven Interaktion würde dieser Effekt weiter verstärkt werden. Da dies allerdings durch eine VR-Brille erfolgen müsste, um ein gutes Ergebnis zu erzielen, würden die technischen Ansprüche an den Endverbraucher ansteigen und neue Hürden schaffen.

Abschließend lässt sich sagen, dass es durchaus möglich ist, eine binaurale Livesession zu produzieren, die sowohl klassische Hörgewohnheiten aus der herkömmlichen Stereophonie bedient als auch eine dreidimensionale Komponente bietet. Allerdings stellt dies stets einen Spagat zwischen diesen beiden Aspekten dar, da diese bisher nicht Hand in Hand gehen und sich eher widersprechen. Dennoch lassen sich auf Basis klanglicher Kompromisse in eine der beiden Richtungen auch mit den aktuellen technischen Möglichkeiten überzeugende Ergebnisse erzielen. Trotzdem gilt es die mögliche Immersion und die Lokalisationsschärfe der Binauraltechnik weiter zu verbessern und Klangverfärbungen zu minimieren. Ein Schlüsselbaustein scheint hier eine effiziente Lösung für die Messung und Einbindung der individuellen HRTF des Hörers zu sein.

In den nächsten Jahren wird die Binauraltechnik mit Sicherheit weiterentwickelt und verbessert. Für die Zukunft ist es wichtig, die breite Masse an Hörern weiter an das Thema 3D-Audio heran zu führen und in diesem Bereich ein Bewusstsein und Erfahrungswerte auf Seiten der Endverbraucher zu schaffen. So ist es unter Umständen möglich, die seit Jahrzehnten bestehenden Hörgewohnheiten langsam aufzubrechen und auch die Musikproduktion gegenüber neuen Formaten zu öffnen.

8. Abbildungsverzeichnis

Abb. 1: Anteil der befragten Internetnutzer, die Musikstreaming-Dienste nutzen, in Deutschland in den Jahren 2013 bis 2018	9
Abb. 2: Livesession des Künstlers FKJ in der Salzwüste Salar de Uyuni.....	11
Abb. 3: Kopfbezogenes Koordinatensystem	16
Abb. 4: Interaurale Pegeldifferenzen durch Abschattung des Kopfes	18
Abb. 5: Lokalisation durch Laufzeitunterschiede	18
Abb. 6: Richtungsbestimmende Bänder nach Blauert	20
Abb. 7: Kurven gleicher Lautstärke nach Fletcher und Munson (1933)	23
Abb. 8: Außenohrübertragungsfunktion gemessen an zehn Versuchspersonen, frontaler Schalleinfallswinkel	27
Abb. 9: Beispiel einer Freifeld-HRTF und Diffusfeld-HRTF	29
Abb. 10: Kunstkopfmikrofon KU 100 des Herstellers Neumann	30
Abb. 11: Veranschaulichung der Kunstkopfstereofonie	31
Abb. 12: Sennheiser AMBEO Smart Headset	32
Abb. 13: LTI-System im Zeitbereich	35
Abb. 14: Schnelle Faltung	36
Abb. 15: Binaurales Konzept der Livesession	46
Abb. 16: Into The Fray	49

Abb. 17: Baupläne der Alten Kirche in Rulfingen	53
Abb. 18: Innenraum der Alten Kirche Rulfingen	54
Abb. 19: Aufstellung "In Dreams"	57
Abb. 20: Aufstellung "Bloom"	58
Abb. 21: Kostenaufstellung der Livesession	60
Abb. 22: Das mobile Aufnahmesetup vor Ort	63
Abb. 23: Kanalliste "In Dreams"	64
Abb. 24: Skizzierte Verteilung der Kanäle bei dem Song "In Dreams"	66
Abb. 25: Kanalliste "Bloom"	67
Abb. 26: Skizzierte Verteilung der Kanäle bei dem Song "Bloom"	68
Abb. 27: Visuelle Darstellung des Zeitplans	71
Abb. 28: Songablauf und Schnitte bei "Bloom"	74
Abb. 29: Songablauf und Schnitte bei "In Dreams"	75
Abb. 30: Grafische Oberfläche des Spatial Audio Designers	77
Abb. 31: Grafische Oberfläche von dearVR Pro	78
Abb. 32: Farbliche Session-Kodierung und Anordnung des Songs "Bloom"	80

9. Quellenverzeichnis

Anastasiadis, M. (2019).

Social-Pop-Media - Zur Mediatisierung von Popmusik und digitaler Musikpraxis in sozialen Medien. Wiesbaden: Springer Verlag.

Blauert, J. (1974).

Räumliches Hören. Stuttgart: S. Hirzel Verlag.

Brech, M., & Paland, R. (2015).

Kompositionen für hörbaren Raum. Bielefeld: transcript Verlag.

Clement, M., Schusser, O., & Papies, D. (2009).

Ökonomie der Musikindustrie. Wiesbaden: Springer Gabler.

Dickreiter, M. (2003).

Mikrofon-Aufnahmetechnik. Stuttgart, Leipzig: S. Hirzel Verlag.

Dickreiter, M., Dittel, V., & Wöhr, M. (2014).

Handbuch der Tonstudioteknik, Band 1 & 2 (8. Auflage). Berlin, Boston: Walter de Gruyter GmbH.

Fedtke, T. (2007).

Kunstkopftechnik - Eine Bestandsaufnahme. ACUSTICA (2007), Vol. 93.

Friesecke, A. (2014).

Die Audio-Enzyklopädie - Ein Nachschlagewerk für Tontechniker. Berlin: Walter de Gruyter GmbH.

Guski, R. (1996).

Wahrnehmen - ein Lehrbuch. Stuttgart: W. Kohlhammer GmbH.

Hagendorf, H., Krummenacher, J., Müller, H.-J., & Schubert, T. (2011).

Wahrnehmung und Aufmerksamkeit. Berlin, Heidelberg: Springer Verlag.

Hellbrück, J., & Kellermeier, W. (2004).

Hören. Hogrefe-Verlag: Göttingen.

Hellmich, M. (2012).

Erstellung einer Datenbank von kopfbezogenen Impulsantworten. Ilmenau:

Technische Universität Ilmenau.

Keinath, D., & Tarnow, C. (2009).

Entwicklung und Erprobung einer Methode zur Integration binauraler

Raumimpulsantworten bei der Mischung von Mehrspurproduktionen. Detmold:

Hochschule für Musik Detmold.

Laumann, K. (2016).

Über Methoden zur Qualitätsbeurteilung von virtuellen Kopfhörern. München:

Technische Universität München, Lehrstuhl für Mensch-Maschine Kommunikation.

Møller, H., Sørensen, M. F., Hammershøi, D., & Jensen, C. (1995).

Head-related transfer functions of human subjects. *Journal of the Audio Engineering Society*, 43(5), 300-321.

Moore, B. (2004).

An Introduction to the Psychology of Hearing, 5. Auflage. San Diego, London:

Emerald.

Owsinski, B. (2007).

Mischen wie die Profis - Das Handbuch für Toningenieur. München: GC Carstensen Verlag.

Pfeifer, M. (2018).

Evaluierung des produktionstechnischen Potenzials und konsumentenorientierten Mehrwerts von Binaural-3D-Audio im Kontext von Popmusik. Stuttgart: Hochschule der Medien, Fakultät Electronic Media

Steckmann, A. (2018).

Eintauchen in virtuelle Räume. Hamburg: Diplomics Verlag GmbH.

Wörle, A. (2006).

Optimierung eines Blind Upmix Binaural Systems für die Kopfhörerwiedergabe. Graz: Technische Universität Graz.

Webers, J. (2007).

Handbuch der Tonstudioteknik für Film, Funk und Fernsehen (9. Auflage). München: Franzis Verlag.

Weinzierl, S. (2008).

Handbuch der Audiotechnik. Berlin: Springer Verlag.

Westreicher, N. (2015).

Das Phantastische im Musikvideo. Wien: LIT Verlag.

Zenner, H. (2006).

Hören. In F. Schmidt, & H. Schaible, Neuro- und Sinnesphysiologie. Berlin, Heidelberg: Springer Verlag.

Zotter, F., & Frank, M. (2019).

Ambisonics - A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality. Cham: Springer Verlag.

Internetquellen

(zuletzt auf ihre Verfügbarkeit geprüft am 25.02.2020)

Alpert, A. (2015).

Unless you're Adele, you have no business releasing album tracks all at once. Von qz.com: <https://qz.com/536000/unless-youre-adele-you-have-no-business-releasing-album-tracks-all-at-once/> abgerufen

Alte-Kirche e.V. (2020).

Kirchengeschichte. Von alte-kirche.info: <http://www.alte-kirche.info/kirchengeschichte.php> abgerufen

Arnold, A. (2013). 40 Jahre Kunstkopf. Ein noch nicht abgeschlossenes Kapitel der Hörspielgeschichte. Von br.de: <https://www.br.de/mediathek/podcast/artmix-galerie/40-jahre-kunstkopf-ein-noch-nicht-abgeschlossenes-kapitel-der-hoerspielgeschichte/31837> abgerufen

Bibliographisches Institut GmbH. (2019).

live. Von dudен.de: <https://www.duden.de/rechtschreibung/live> abgerufen

Bibliographisches Institut GmbH. (2019).

Session, die. Von dudен.de: https://www.duden.de/rechtschreibung/Session_Periode_Abschnitt_Tagung abgerufen

Bitkom Research. (2018).

Anteil der befragten Internetnutzer, die Musikstreaming-Dienste nutzen, in Deutschland in den Jahren 2013 bis 2018. Von statista.com: <https://de.statista.com/statistik/daten/studie/685445/umfrage/nutzung-von-musikstreaming-in-deutschland/> abgerufen

Bitkom Research. (2019).

Sind Sie im Besitz von Kopfhörern? Von statista.com:

<https://de.statista.com/statistik/daten/studie/1047606/umfrage/umfrage-zum-besitz-von-kopfhoerern-in-deutschland/> abgerufen

Dear Reality GmbH. (2020).

DearVR Pro. Von dearvr.com: <https://www.dearvr.com/products/dearvr-pro> abgerufen

Digital Media Association. (2018).

A MIDIA Research Report March 2018. Von dima.org: <https://dima.org/wp-content/uploads/2018/04/DiMA-Streaming-Forward-Report.pdf> abgerufen

Fleck, C. (2019). Räumliches Hören.

Von hoerhelfer.de: <https://www.hoerhelfer.de/Richtungshoeren-raeumliches-Hoeren/> abgerufen

Fraunhofer-Institut für Integrierte Schaltungen IIS. (2020).

MPEG-H Audio. Von iis.fraunhofer.de:

<https://www.iis.fraunhofer.de/de/ff/amm/rundfunk-streaming/mpegh.html> abgerufen

Georg Neumann GmbH. (2018).

KU 80. Von neumann.com: <https://de-de.neumann.com/ku-80> abgerufen

Gerl, M. (2018).

Vom kleinen Musikladen zum europaweit größten Online-Shop . Von

sueddeutsche.de: <https://www.sueddeutsche.de/bayern/unternehmen-thomann-vom-kleinen-musikladen-zum-europaweit-groesste-online-shop-1.3940519> abgerufen

iZotope, Inc. (2020).

Music Rebalance. Von izotope.com:

<https://www.izotope.com/en/products/rx/features/music-rebalance.html> abgerufen

Katzif, M. (2020).

BECK'S 'SOUND & VISION'. Von likemgmt.com:

<https://www.likemgmt.com/projects#/sound-vision/> abgerufen

Lischka, K. (2014).

Wie das Streaming-Geschäft den Musikmarkt verändert. Von

spiegel.de: <https://www.spiegel.de/netzwelt/web/download-daemmerung-wie-streaming-den-musikmarkt-veraendert-a-944652.html> abgerufen

New Audio Technology GmbH. (2020).

Spatial Audio Designer. Von newaudiotechnology.com:

<https://newaudiotechnology.com/products/spatial-audio-designer/> abgerufen

Sengpiel, E. (1995).

Kopfbezogene Übertragungsfunktion HRTF. Von sengpielaudio.com:

<http://www.sengpielaudio.com/KopfbezogeneUebertragungsfunktionHRTF.pdf>
abgerufen

Sengpiel, E. (1998).

Transaural Stereo - Kunstkopf-Stereofonie über Lautsprecher. Von sengpielaudio.com:

<http://www.sengpielaudio.com/TransauralStereo.pdf> abgerufen

Sengpiel, E. (kein Datum).

Dämpfung der Luft bei hohen Frequenzen. Von sengpielaudio.com:

<http://www.sengpielaudio.com/Rechner-luft.htm> abgerufen

Sengpiel, E. (kein Datum).

Schalldruck p und das reziproke Abstandsgesetz $1/r$. Von sengpielaudio.com:

<http://www.sengpielaudio.com/Rechner-abstandsgesetz.htm> abgerufen

Sennheiser electronic GmbH & Co. KG. (2020).

AMBEO VR Mic. Von sennheiser.com: <https://de-de.sennheiser.com/mikrofon-3d-audio-ambeo-vr-mic> abgerufen

Spotify AB. (2020).

Spotify Premium. Von spotify.com: <https://www.spotify.com/de/premium/> abgerufen

Statista GmbH. (2019).

Kopfhörer. Von statista.com:

<https://de.statista.com/outlook/15010600/137/kopfhoerer/deutschland> abgerufen

Stern.de. (2018). Darum veröffentlicht Cros Plattenfirma keine Alben mehr. Von <https://www.stern.de/kultur/musik/darum-veroeffentlicht-cros-plattenfirma-keine-alben-mehr-7974238.html> abgerufen

Theile, G. (1984).

Sind „Klangfarbe“ und „Lautstärke“ vollständig determiniert durch das Schalldruckpegel-Spektrum am Trommelfell? München: Institut für Rundfunktechnik. Von [hauptmikrofon.de: https://hauptmikrofon.de/theile/1984-2_Trommelfell-Spektrum-vs-Klangfarbe_DAGA-1984.pdf](https://hauptmikrofon.de/theile/1984-2_Trommelfell-Spektrum-vs-Klangfarbe_DAGA-1984.pdf) abgerufen

Theile, G., Wittek, H., & Reisinger, M. (2002).

Wellenfeldsynthese-Verfahren: Ein Weg für neue Möglichkeiten der räumlichen Tongestaltung. München: Institut für Rundfunktechnik. Von [hauptmikrofon.de: https://hauptmikrofon.de/theile/2002-3_WFS_22.TMT-2002.pdf](https://hauptmikrofon.de/theile/2002-3_WFS_22.TMT-2002.pdf) abgerufen

Thomann GmbH. (2020).

Von [thomann.de: https://www.thomann.de/](https://www.thomann.de/) abgerufen

White, P. (2019). Binaural Panning In Logic Pro. Von [soundonsound.com: https://www.soundonsound.com/techniques/binaural-panning-logic-pro](https://www.soundonsound.com/techniques/binaural-panning-logic-pro) abgerufen

Wikipedia. (2019).

Sessions@AOL. Von <https://en.wikipedia.org/wiki/Sessions@AOL> abgerufen

YouTube. (2019).

Egopusher "Flake" Binaural Version with Sennheiser AMBEO VR Mic | Sennheiser. Von https://www.youtube.com/watch?v=jk1LZVzNu54&feature=emb_title abgerufen

YouTube. (2020).

COLORS. Von

<https://www.youtube.com/channel/UC2Qw1dzXDBAZPwS7zm37g8g/about> abgerufen

YouTube. (2020).

KEXP. Von <https://www.youtube.com/user/kexpradio/about> abgerufen

YouTube. (2020).

NPR Music. Von <https://www.youtube.com/user/nprmusic/about> abgerufen

10. Anhang

Die Anhänge zu dieser Masterarbeit können dem beiliegenden Datenträger entnommen werden.

Dieser beinhaltet:

- die digitale Version der schriftlichen Ausarbeitung im PDF-Format
- die beiden fertigen Videos der Livesession im MP4-Format
- die finalen Audiomischungen im WAV-Format (24 Bit, 48 kHz)
- alle Einzelspuren der beiden Mastertakes der Songs (24 Bit, 48 kHz)
- die Ableton Live Sets der beiden Audiomischungen
- alle Abbildungen als Bilddateien
- zusätzliche Impressionen von der Aufnahme vor Ort (Fotos: Foxframes)