

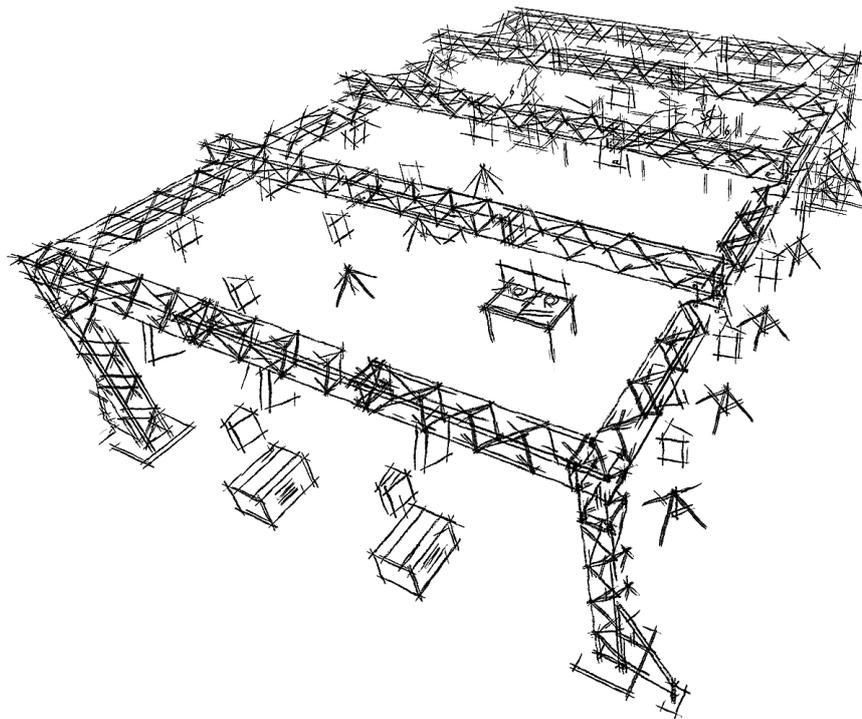
# Masterarbeit

Hochschule der Medien  
Fakultät Electronic Media  
Studiengang Elektronische Medien

## **Konzeption und praktische Umsetzung einer 3D-Audio-Live-Produktion**

---

Untersuchung und Entwicklung mit den Schwerpunkten im technischen Design  
und dreidimensionaler Live-Mischungen im Kontext populärer Musik



# Masterarbeit

Hochschule der Medien  
Fakultät Electronic Media  
Studiengang Elektronische Medien

## **Konzeption und praktische Umsetzung einer 3D-Audio-Live-Produktion**

---

Untersuchung und Entwicklung mit den Schwerpunkten im technischen Design  
und dreidimensionaler Live-Mischungen im Kontext populärer Musik

vorgelegt von:           Benedikt Maile (Matrikel-Nr. 28962)  
                                  Daniel Schiffner (Matrikel-Nr. 28487)

vorgelegt am:            18. Mai 2016

Erstprüfer:                Prof. Oliver Curdt (HdM Stuttgart)

Zweitprüfer:             Dr. Helmut Wittek (Schoeps GmbH)



## **Eidesstattliche Erklärung**

Hiermit erklären wir, Benedikt Maile und Daniel Schiffner, an Eides statt, dass wir die vorliegende Masterarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt haben. Die Stellen unserer Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen sind, haben wir in jedem Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht. Dasselbe gilt sinngemäß für Tabellen und Abbildungen. Diese Arbeit hat in dieser oder einer ähnlichen Form noch nicht im Rahmen einer anderen Prüfung vorgelegen.

Stuttgart, den 17. Mai 2016

Stuttgart, den 17. Mai 2016

-----  
Benedikt Maile

-----  
Daniel Schiffner

## **Abstract**

The possibilities of three-dimensional speaker configurations evolve continuously and have become an essential part of the film industry, the home entertainment and virtual reality. Even in terms of live-events, a trend can be observed which demands for the next dimension in audio performances.

This work examined opportunities for a live concert performance of popular music on a 3D speaker configuration with reference to a practical implementation, technical requirements, standards and opportunities of three-dimensional speaker configurations were analyzed. In the process a three-dimensional audio work environment was created that served as a development environment for three-dimensional live sound and as a showcase of the technical approach for an implementation. During pre-production an approach for mixing techniques was developed which corresponds to desired intention while played back on the three-dimensional speaker configurations. As a result methods of mixing, possibilities in three-dimensional space as well as their effect and plausibility were discovered.

With the attained knowledge, a three-dimensional audio live production environment was developed which enabled the realisation of a three-dimensional audio live concert with the band Eau Rouge. Accompanying particular production phases, creative and technical possibilities for each phase were presented and evaluated. Based on the practical experience, an outlook of the potential of three-dimensional live performances was developed.

Keywords: 3D Audio, live concert, surround

## **Kurzfassung**

Die Möglichkeiten dreidimensionaler Lautsprecherkonfigurationen entwickeln sich kontinuierlich weiter und sind bereits Teil der Filmindustrie, des Home-Entertainments und der virtuellen Realität geworden. Auch im Veranstaltungsbereich ist eine Entwicklung zu beobachten, welche die nächste Dimension in Audio-Darbietungen bedienen möchte.

Diese Arbeit untersucht anhand einer praktischen Umsetzung, wie ein Live-Konzert populärer Musik auf einer dreidimensionalen Lautsprecherkonfiguration aufgeführt werden kann. Dabei wurden die technischen Anforderungen, Standards und Möglichkeiten analysiert. Eine 3D-Audio-Arbeitsumgebung wurde geschaffen, die im Rahmen der Vorproduktion als Entwicklungsumgebung für dreidimensionale Live-Mischungen und technische Ansätze der Umsetzung diente. Während der Vorproduktion wurde ein mischtechnischer Ansatz entwickelt, der auf dreidimensionalen Lautsprecherkonfigurationen der gewünschten Intention entsprach. Dabei wurden Methoden der Mischung, gestalterische Möglichkeiten im dreidimensionalen Raum sowie deren Wirkung und Plausibilität untersucht.

Aus dieser Erkenntnis wurde eine 3D-Audio-Live-Produktions-Umgebung entwickelt, mit welcher ein 3D-Audio-Live-Konzert der Stuttgarter Band Eau Rouge umgesetzt werden konnte. Anhand der einzelnen Produktionsabschnitte wurden kreative und technische Möglichkeiten aufgezeigt und bewertet. Mit der gewonnenen Praxiserfahrung wurde ein Ausblick über das Entwicklungspotential dreidimensionaler Live-Performances erarbeitet.

Schlüsselworte: 3D Audio, Live-Konzert, Rund-um-Klang

## Danksagung

Wir möchten ganz besonders folgenden Sponsoren, Herstellern, Firmen und der Hochschule für die vielseitige Unterstützung beim Entstehen dieser Produktion und Arbeit danken.

Unser Dank gilt auch allen Professoren, technischen Mitarbeitern und Helfern die uns im Rahmen dieser Produktion unterstützt haben.

Weiter danken wir Freunden und Familie für Ihre Unterstützung während der Produktion und Erstellung dieser Arbeit.

Sponsoren, Hersteller und Firmen:

Mixtown Veranstaltungstechnik, Direct.Out Technologies, d&b Audiotechnik, Wireworx, Iosono Sound, NTP Technology, SMM, RME, Media Online, Crystal Sound Veranstaltungstechnik, Orangetec, Steinberg, Schoeps Mikrofone, Auro 3D Technologies, Plugin Alliance, BW Bank, Mevis.TV, Mplan.

Professoren, technische Mitarbeiter und Helfer:

Matthias Adler, Dominik Arnold, Jörg Bauer, Jonas Baumbach, Guido Benkart, Peter Bollinger, Matthias Bürgel, Prof. Oliver Curdt, Ismail Daif, Arline Scarleth Duarte, Jan Ehrlich, Florian Gundert, Hdm Haustechnik, Tim Heumesser, Rainer Kampe, Christian Kittel, Kim Knödler, Christian Knoll, Andreas Koch, Jürgen Kössinger, Axel Kühlem, Jan Langhammer, Kerstin Lauer, Axel Lembke, Winnie Leyh, Peter Marquardt, Timo Maul, Steffen Mühlhöfer, Lukas Müller, Ulrich Otte, Marcel Schlegel, Ida Schneider, Heiko Schulz, Martin Schwöri, Alexandra Seidel, Tobias Van, Mikael Vest, Courtney Wachtel, Michael Werner, Helmut Wittek, Elena Zondler, Ralf Zuleeg.

# Inhalt

<b>Eidesstattliche Erklärung</b> .....	<b>I</b>
<b>Abstract</b> .....	<b>II</b>
<b>Kurzfassung</b> .....	<b>III</b>
<b>Danksagung</b> .....	<b>IV</b>
<b>Inhalt</b> .....	<b>V</b>
<b>Glossar</b> .....	<b>VIII</b>
<b>1 Einleitung</b> .....	<b>1</b>
1.1 Ausgangspunkt der Arbeit .....	1
1.2 Motivation für 3D-Audio.....	2
1.3 Ziele von 3D-Audio .....	3
1.4 Schwerpunkt der Arbeit.....	4
<b>2 Grundlagen 3D Audio</b> .....	<b>5</b>
2.1 Nomenklatur .....	5
2.2 Kanalbasiertes Audio .....	7
2.3 Objektbasiertes Audio .....	7
2.4 Kopfbezogenes Koordinatensystem .....	8
2.5 Räumliches Hören.....	9
2.5.1 Lokalisation.....	9
2.5.2 Monaurale Faktoren.....	9
2.5.3 Binaurale Faktoren .....	10
2.6 Entfernung und Tiefe .....	12
2.7 Klangfarbe und Räumlichkeit.....	13
2.8 Prinzipien der klanglichen Reproduktion .....	13
2.8.1 Mehrkanal-Sterefonie .....	13
2.8.2 Vector Base Amplitude Panning.....	13
2.8.3 Ambisonics .....	15
2.8.4 Wellenfeldsynthese.....	18
2.8.5 Binauraltechnik.....	23

<b>3</b>	<b>3D-Audio-Wiedergabesysteme .....</b>	<b>26</b>
3.1	Auro 3D .....	26
3.1.1	Verteilung der Schallfelder .....	29
3.1.2	Kanalübersprechen.....	30
3.1.3	Mischpult-Integration .....	30
3.1.4	Auro 3D Authoring Tools.....	31
3.2	IoSono .....	35
3.2.1	Verteilung der Schallfelder .....	35
3.2.2	Kanalübersprechen.....	36
3.2.3	Mischpult-Integration .....	36
3.2.4	IoSono-Core .....	37
3.2.5	IoSono Control Unit.....	38
3.2.6	Spatial Audio Workstation.....	46
<b>4</b>	<b>Der Selbstversuch: Vorbereitung des 3D-Konzerts .....</b>	<b>50</b>
4.1	Mischung.....	53
4.1.1	Allgemeines .....	53
4.1.2	Grundlagen Wahrnehmung .....	57
4.1.3	Signal-Bearbeitung als stilistisches Mittel.....	58
4.1.4	Räumliche Klangbearbeitung.....	63
4.1.5	Einfluss der Dreidimensionalität auf die Bearbeitung .....	64
4.1.6	Das System und seine gestalterischen Möglichkeiten .....	66
4.1.7	Hörgewohnheiten .....	66
4.1.8	Dreidimensionales Mischen als gestalterisches Mittel.....	67
4.1.9	Immersion und Dreidimensionalität.....	69
4.1.10	Räumlichkeit & Plausibilität .....	69
4.1.11	Wahrnehmung im Kontext der Liveproduktion.....	70
4.2	Die Musik.....	73
4.2.1	Musikalische Auswahlkriterien.....	73
4.2.2	Die Band: „Eau Rouge“ .....	73
4.2.3	Instrumentierung .....	74
4.3	Die Vorproduktion.....	75
4.3.1	Vorbereitung.....	75

4.3.2	Technisches Setup der Vorproduktion .....	77
4.3.3	Kontroll- & Interaktionsschnittstellen .....	85
4.3.4	Zuspieler .....	89
4.3.5	Live-Inhalt.....	92
4.3.6	Liveproduktion in Iosono.....	100
4.3.7	Systemvorstellung: Vorproduktion der Liveproduktion in Iosono.....	100
<b>5</b>	<b>Der Selbstversuch: Umsetzung des 3D-Konzerts .....</b>	<b>103</b>
5.1	Die Mischung: exemplarisch beschrieben am Song „Circles“ .....	103
5.1.1	Panorama .....	103
5.1.2	Arrangement .....	104
5.2	Veranstaltungsort .....	112
5.3	Technisches Setup der Produktion .....	114
5.3.1	Lautsprecherkonfiguration .....	114
5.3.2	Iosono Lautsprecher-Setup.....	123
5.3.3	Iosono Systemeinmessung.....	125
5.3.4	Signalkette .....	128
5.3.5	Gefährdungsanalyse .....	129
5.4	Positionierung der Band .....	129
5.5	Monitoring .....	132
<b>6</b>	<b>Aufzeichnung &amp; Abspielen .....</b>	<b>133</b>
<b>7</b>	<b>Schlussfolgerung und Ausblick .....</b>	<b>134</b>
<b>8</b>	<b>Refrenzen .....</b>	<b>136</b>
<b>9</b>	<b>Abbildungsverzeichnis .....</b>	<b>141</b>
<b>10</b>	<b>Anhang .....</b>	<b>144</b>

## Glossar

AD/DA	Analog Digital / Digital Analog
ATSC	Advanced Television System Committee
Daemon	Programm das im Hintergrund abläuft und bestimmte Dienste zur Verfügung stellt.
dB	Dezibel, logarithmische Verhältnisgröße zur Angabe von Pegeln und Pegeldifferenzen.
dB <sub>FS</sub>	Dezibel full scale
dB <sub>SPL</sub>	absolute Pegelangabe für den Schalldruckpegel mit einem Bezugswert von $0 \text{ dB}_{\text{SPL}} = 2 \times 10^{-5} \text{ Pa}$
dB <sub>U</sub>	absolute Pegelangabe eines Spannungspegels mit einem Bezugswert von $0 \text{ dB}_{\text{U}} = 0,775 \text{ V}$
Digilink	Standard zum Austausch von Audio- und Steuerdaten der Firma Avid
DPI	Deep Packet Inspection
DSP	Digitaler Signalprozessor: durch die Architektur und Anbindung der DSP Systeme können sehr geringe Latenzen erzielt werden.
FIR	Finite Impulse-Response
HE	Höheneinheit (1HE = 44,45 mm oder 1 ¾ Zoll)
IIR	Infinite Impulse-Response
INR	Impulse-Response-Noise-Ratio
LFE	Low frequency effects channel
LTC	Longitudinal (oder Linear) Timecode
LWL	Lichtwellenleiter
MADI	Multichannel Audio Digital Interface
MM	Multimode
MIDI	Musical Instrument Digital Interface

MTC	MIDI Timecode
PCIe	Peripheral Component Interconnect Express ist ein Standard zur Verbindung von Peripheriegeräten mit dem Chipsatz eines Hauptprozessors.
Phantomschallquelle	Virtuelle Schallquelle, die zwischen zwei Lautsprechern wahrgenommen wird
SMPTE	Society of Motion Picture and Television Engineers
SNR	Signal-Noise-Ratio
Sub-D	Engl. D-Subminiature; eine verbreitete Bauform eines Steckersystems für Daten- und Audioverbindungen.
Sweet Spot	Der optimale Hörort
Sweet Area	Der optimale Hörbereich
TCP/IP	Transmission Control Protocol / Internet Protocol
UDP	User Datagram Protocol
XLR-F / XLR-M	Industriestandard für elektrische Steckverbindung: Ohne Zusatz wird in dieser Arbeit von einer 3poligen Verbindung ausgegangen. Die Zusätze F und M definieren die Art der Steckverbindung. F steht für female und M für male.



# 1 Einleitung

## 1.1 Ausgangspunkt der Arbeit

Dreidimensionale Audioformate sind in der Lage akustische Ereignisse aus allen drei Dimensionen zu verarbeiten und zu reproduzieren. Eine Vielzahl an Wiedergabesystemen für dreidimensionales Audio sind bereits auf dem Markt und wurden in unterschiedlichen Anwendungen praktisch erprobt. Dreidimensionale Audioformate für Filmproduktionen können in speziell dafür ausgestatteten Kinosälen, über unterschiedliche Lautsprecherkonfigurationen erlebt werden. Ein gemeinsamer Standard hat sich dabei noch nicht etabliert.<sup>1</sup> Neben den in Filmproduktionen üblichen Formaten wie Dolby Atmos, Auro 3D oder Auro Max gibt es im Bereich der Musik auch mehrkanalige Produktionen die keinem Standard entsprechen und vielmehr als eine Art Klanginstallation produziert und reproduziert werden. Im Bereich der standardisierten Lautsprecherkonfigurationen für dreidimensionale Musikproduktionen ist momentan eine höhere Gewichtung des Auro 3D Formates zu beobachten.

Sowohl in Film-, Musik- und Live-Produktionen kommen auch Wiedergabesysteme zum Einsatz, die auf dem ursprünglichen Gedanken der Wellenfeldsynthese aufbauen. Diese Systeme werden seit Beginn der 90er Jahre entwickelt, erprobt und in Bereichen der Klanginstallation, Live-Beschallung und Auralisation eingesetzt. Zu aktuellen Entwicklungen zählt auch das Iosono System der gleichnamigen GmbH, einem Fraunhofer Spin-Off. Iosono hat eine Vielzahl an Referenzen in den Bereichen Event-, Club- und Museumsinstallationen sowie der virtuellen Realität und in Kinosälen.<sup>2</sup> Entwicklungen im Bereich von 3D-Audio-Live-Konzepten und deren Umsetzung im Kontext populärer Musik gibt es bisher nur wenige.

Standardisierungsinstitutionen wie die SMPTE oder die ATSC beschäftigen sich bereits mit kanalbasiertem und objektbasiertem dreidimensionalem Audio. Neben einigen videospezifischen Themen beschäftigen sich diese Standardisierungsversuche in erster Linie mit Immersion, Personalisierung, Interaktion und Übertragungsmöglichkeiten. 3D Audio unterstützende Formate und Codecs wie Dolby Atmos, DTS:X oder MPEG-H stehen in den Startlöchern und sind zukunftsorientiert design. In diesem Zusammenhang wird das Thema der Binauralisierung und der Ausblick eines eventuell bevorstehenden Kopfhörer-Entertainments stark diskutiert.

---

<sup>1</sup> vgl. SMPTE 2014: 9

<sup>2</sup> Iosono 2015a

## 1.2 Motivation für 3D-Audio

Mit der Einführung digitaler, hochauflösender Projektionssysteme in die Kinos nimmt auch der Anteil an Filmen mit dreidimensionalen Audioformaten zu. Die durch heutige Projektionssysteme realisierbare Bildqualität bietet ein ultimatives visuelles Kinoerlebnis in 3D. Auch der restliche Multi-Media -Markt bewegt sich zunehmend in Richtung dreidimensionaler visueller Inhalte.

Große Filmproduzenten wie George Lucas oder David Lynch sind sich einig, dass der Ton für mindestens 50% des emotionalen Filmerlebnisses verantwortlich ist.

*George Lucas: "Sound ist 50% des Filmerlebnisses und ich bin davon überzeugt, dass das Publikum vom Sound bewegt und erregt wird wie vom Bild."<sup>3</sup>*

*David Lynch: "Jedes Mal , wenn ich Sound höre, sehe ich Bilder. Dann fallen mir die Ideen zu: Es macht mich fast verrückt! Filme sind 50% visuell und 50% Sound."<sup>4</sup>*

Es wächst die Akzeptanz und das kreative Potential filmischer Produktionen eben durch eine Weiterentwicklung der dreidimensionalen Filmproduktionen, sowohl im Bild als auch im Ton. Ein weiterer interessanter Punkt liegt darin, dass viele evolutionsbedingte, überlebenswichtige Instinkte des Menschen von durch Umgebungsgeräusche ausgelösten Emotionen abgeleitet sind. Man nehme das Beispiel eines über einer Person abbrechenden Astes oder eines lauernden Raubtieres, das hinter einer Person im Gebüsch raschelt. Das Orten eines Geräusches, ein schnelles, instinktives Verarbeiten der Signale und damit zusammenhängende Reaktionsmuster bilden hier einen Vorteil. Das menschliche Gehör nimmt also – sofern nicht eingeschränkt – Tonsignale wahr, deren Quellen im umgebenden Raum lokalisiert werden. Diesem räumlichen Hören möchten 3D-Audio-Konzepte entsprechen.

Das Bedürfnis nach emotionaler Attraktion sowie Orientierung im visuellen Raum steigt kontinuierlich. Darum sollte auch in der auditiven Welt die dritte Raumdimension mit einbezogen werden um eine noch realistischere Wahrnehmung zu erzielen und das stetig steigende Verlangen nach emotionaler Attraktion zu befriedigen. Die Weiterentwicklung der Mehrkanalformate von 2.0 zu 5.1 hin zu dreidimensionalen Formaten wie 11.1 oder beliebig vielen Lautsprechern, ist der logische Schritt in Richtung immersiver Audioerlebnisse.

---

<sup>3</sup> Weiss 2015: 46

<sup>4</sup> Weiss 2015: 46

Neben dem Standpunkt des Konsumenten gibt es die entsprechenden Interessen der Industrie. Die erfolgreiche Vermarktung von dreidimensionalem visuellem Content in Film, Fernsehen und der Spieleindustrie sieht sowohl die professionelle, als auch semi-professionelle Audioindustrie als einen neuen Absatzmarkt. Die technischen Möglichkeiten entwickeln sich rasant weiter. Prozessoren werden performanter, kleiner und billiger. Algorithmen werden stets optimiert und auf das performante Processing angepasst. Die Vielzahl an Audiokanälen welche 3D-Audio mit sich bringt erfährt aus Sicht der Technik fast keine Restriktionen mehr und passt je nach Anwendungsfall bereits in herkömmliche Laptops, Smartphones und Tablets. Seit der Einführung von Smartphones und Tablets wurden auch Auswirkungen auf die Hörgewohnheiten beobachtet. Laut BITKOM (2013) konsumieren 93% der 14- bis 29-Jährigen Audio über Kopfhörer - meistens in Zusammenhang mit dem Smartphone.<sup>5</sup> Die Einführung und Optimierung von binauralem Content ist darum von großem Interesse.

### **1.3 Ziele von 3D-Audio**

Mit 3D-Audio wird in erster Linie eine natürlichere, daraus resultierend, besser empfundene räumliche Wiedergabe verfolgt. Der Ansatz einer dreidimensionalen Musikproduktion ermöglicht eine wesentlich höhere Informationsdichte, entsprechend bietet Sie ein erhöhtes kreatives Potenzial. Eine musikalische Szene kann beispielsweise differenzierter aufgebaut werden. In Abhängigkeit des Wiedergabesystems können größere Hörzonen geschaffen werden. Der Sweet Spot in Surround- und Stereo-Formaten, kann zur Sweet Area vergrößert werden und ermöglicht ein Eintauchen und Begehen der musikalischen Szene. Es werden immersive Erlebnisse geschaffen, welche das generelle Ziel von 3D Audio implizieren. Unabhängig in welchem Zweig der technischen Entwicklungen wir uns befinden, ob Videotechnik, Projektionstechniken oder eben Audiotechnik wird eine möglichst perfekte Abbildung der Realität angestrebt. Dabei werden bei 3D-Audio die zur Verfügung stehenden technischen Möglichkeiten genützt, um entweder virtuelle Klangwelten zu schaffen oder Klangwelten, die mit den Erlebnissen der realen Welt übereinstimmen. Darüber hinaus bestehen nahezu unbegrenzte kreative Möglichkeiten, um Immersionseffekte zu verstärken oder bisher noch nie da gewesene Klangerlebnisse zu erzeugen. 3D-Audio in Verbindung mit Interaktivität oder Intraaktivität kann sogar eine zusätzliche 3.5 oder 4. Dimension öffnen, welche den Konsumenten komplett eintauchen und in der immersiven Umgebung interagieren lässt.

---

<sup>5</sup> Bitkom 2013

## **1.4 Schwerpunkt der Arbeit**

In dieser Arbeit liegt der Focus auf den Möglichkeiten einer 3D-Audio-Live-Produktion im Bereich populärer Musik. Einzelne Produktions- und Planungsphasen werden von Beginn an in Bezug auf technische Anforderungen und Möglichkeiten sowohl in der Theorie, als auch am praktischen Beispiel analysiert. Neben dem Entwurf und der praktischen Umsetzung einer 3D-Audio-Live-Umgebung für kleine bis mittelgroße Veranstaltungsorte, liegt der zweite große Schwerpunkt auf der Entwicklung eines mischtechnischen Ansatzes, der auf dreidimensionalen Lautsprecherkonfigurationen der gewünschten Intention entspricht. Kreative Möglichkeiten im dreidimensionalen Raum wurden auf Wirkung und Plausibilität analysiert.

Die technischen Herausforderungen und Möglichkeiten wurden im Rahmen der Vorproduktionen erarbeitet. Mit den daraus gewonnenen Erkenntnissen wurde ein Konzept für ein 3D-Audio-Live-Konzept mit der Stuttgarter Band Eau Rouge entwickelt. Ein kleiner Ausblick über Interaktionsmöglichkeiten für Zuschauer und Musiker wurde am Beispiel eines praktischen Versuches verdeutlicht. An Hand der Zuhörerreaktionen und Hörerfahrungen sowie mit der gewonnenen Praxiserfahrung wurde ein Ausblick zum Entwicklungspotential dreidimensionaler Live-Performances aufgezeigt. Dabei wurden auch Teile der Produktion binaural aufgenommen und ausgewertet.

Nachdem in Kapitel 2 die Grundlagen für 3D-Audio, dem räumlichen Hören und die Prinzipien der klanglichen Reproduktion geklärt wurden, werden in Kapitel 3 die produktionsspezifischen Wiedergabesysteme vorgestellt und auf ihre technischen und akustischen Eigenschaften untersucht. In Kapitel 4 wird der Selbstversuch im Rahmen der Vorproduktion beschrieben. Daraus entwickelt sich ein mischästhetischer Ansatz und eine Misch-Methode für den dreidimensionalen Raum. Anhand der einzelnen Projektphasen der Vorproduktion werden unter anderem in Bezug auf die technische Konzeption und Herangehensweise, Probleme und Lösungsansätze erläutert. Nach der Konzeption und Umsetzung des 3D-Audio-Live-Konzertes in Kapitel 5 werden in Kapitel 6 Möglichkeiten der Aufzeichnung und Wiedergabe erläutert. Abschließend gibt es in Kapitel 7 einen Ausblick über das Entwicklungspotential von dreidimensionalen Audio-Produktion sowie eine Einschätzung für die Zukunft von 3D-Audio-Live-Konzerten.

## 2 Grundlagen 3D Audio

Als Basis für diese Arbeit werden Grundlagen der Audio- und Übertragungstechnik vorausgesetzt. Es ist zu beobachten, dass die Nomenklatur im Bereich 3D Audio neue Definitionen mit sich bringt. Diese wirken zum Teil verwirrend, da sie im Rahmen von Werbung, aber auch technischen Berichten häufig überspitzt oder falsch verwendet werden. Die folgenden Definitionen bauen auf den vorausgesetzten Grundlagen auf und bilden die Basis für die vorliegende Arbeit.

### 2.1 Nomenklatur

#### 3D Audio

Klang ( <i>engl. Sound</i> )	periodischer Schalldruckverlauf in der Luft, der vom Gehör wahrgenommen werden kann und somit hörbar ist.
Audio	eine elektrische Repräsentation von Sound Audio selbst kann nicht vom Gehör wahrgenommen werden und ist somit nicht hörbar.
Klangfeld ( <i>engl. Soundfield</i> )	akustischer Raum, welcher durch Reproduktion einer oder mehrerer Audio-Quellen erzeugt werden kann.
Immersive Soundfield ( <i>dt. mitreißendes Klangfeld</i> )	akustischer Raum, in dem Sound aus allen drei Dimensionen reproduziert werden kann.
Immersive Sound ( <i>dt. mitreißender Klang</i> )	der Sound, der innerhalb eines immersiven Klangfeldes hörbar wird.
Immersive Audio ( <i>dt. mitreißendes Audio</i> )	Audio, das mit der Absicht erstellt wird, in einem immersiven Klangfeld reproduziert zu werden.
Lautsprecherkonfiguration ( <i>engl. Soundfield Configuration</i> )	eine definierte Aufstellung und Konfiguration von Lautsprechern, die das gewünschte Klangfeld übermitteln können; Bsp.: Mono, Stereo, Surround oder 3D-Audio.

## Räumliches Hören

Lokalisation	Zuordnungsgesetz oder -regel (Operator) 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 <sup>6</sup> , Beispiele: Zuordnung von Hörereignisort und Schallquellenort; Zuordnung von Hörereignisrichtung und interauraler Schalldruckpegeldifferenz - auch Zuordnung von Hörereignisrichtung und Kopfbewegung. <sup>7</sup> Schallereignisse werden als Hörereignisse außerhalb des Kopfes lokalisiert. <sup>8</sup>
Im-Kopf-Lokalisation	Die Schallreize eines Schallereignisses sind so geartet, dass sie nicht einer möglichen "natürlichen" Schallquelle außerhalb des Kopfes zugeordnet werden können. Das Reizmuster ist für das Gehör völlig neuartig. Vom Gehör für die Lokalisation nicht auswertbare Signale werden für die Wahrnehmung durch die Im-Kopf-Lokalisation gekennzeichnet. <sup>9</sup>
Richtung	Richtung aus welcher ein Hörereignis wahrgenommen wird.
Distanz	der wahrgenommene Abstand zwischen Hörer und des wiedergegebenen Schallereignisses. <sup>10</sup>
Tiefe	die perspektivische Wahrnehmung einer wiedergegebenen Audio-Szene als Ganzes. <sup>11</sup>
Stabilität	der Grad bis zu welchem sich die Lokalisierung eines Hörereignisses mit der Zeit ändert.
Verortung	die räumliche Unterscheidung von Hörereignissen; <sup>12</sup> der Grad, bis zu welchem ein Hörereignis an exakt einer bestimmten Position wahrgenommen werden kann.

---

<sup>6</sup> vgl. Theile 1980

<sup>7</sup> vgl. Blauert 2014

<sup>8</sup> vgl. Dickreiter 2008

<sup>9</sup> ebenda

<sup>10</sup> vgl. Rumsey 2002

<sup>11</sup> ebenda

## 2.2 Kanalbasiertes Audio

Bei kanalbasiertem Audio handelt es sich um eine Anzahl an Audio-Kanälen welche die nötigen Informationen zur Wiedergabe des gewünschten Klangfeldes beinhalten. Dabei bezeichnet ein Kanal eine Auswahl an verschiedenen sequenzierten Audiosamples die zur Wiedergabe über einen einzelnen Lautsprecher oder über ein Lautsprecherarray vorgesehen sind.<sup>13</sup>

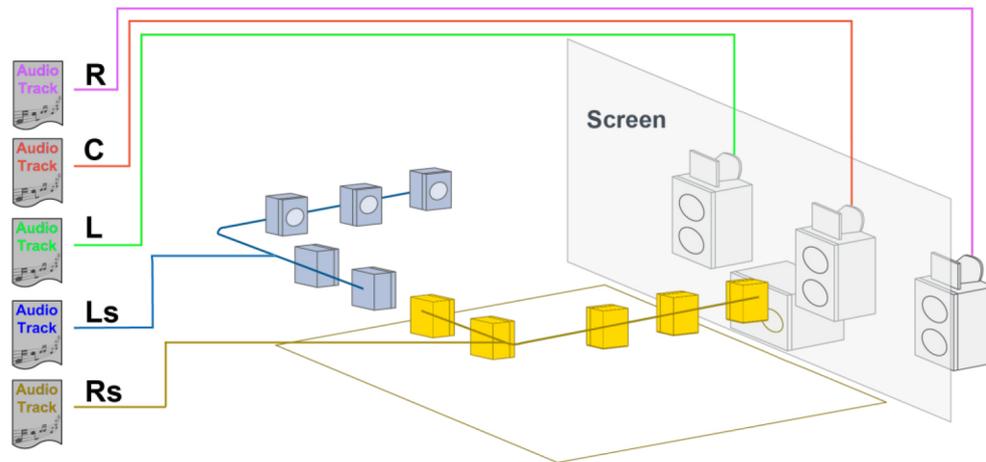


Abbildung 1: Kanalbasiertes Audio<sup>14</sup>

## 2.3 Objektbasiertes Audio

Objektbasiertes Audio umfasst eine Anzahl an Audio-Objekten, die die nötigen Informationen zur Wiedergabe des gewünschten Klangfeldes beinhalten. Ein Objekt ist dabei Audio in beliebiger Länge verknüpft mit Metadaten, die beschreiben, wie es im Klangfeld wiedergegeben werden soll. Metadaten beschreiben die Position, die Ausbreitung, die Bewegungsenergie und zusätzliche Rendering-Informationen. Objekte können im Klangfeld bewegt werden und an einer einzelnen Position oder über einen bestimmten Lautsprecher wiedergegeben werden.

<sup>13</sup> SMPTE 2014

<sup>14</sup> SMPTE 2014: 25

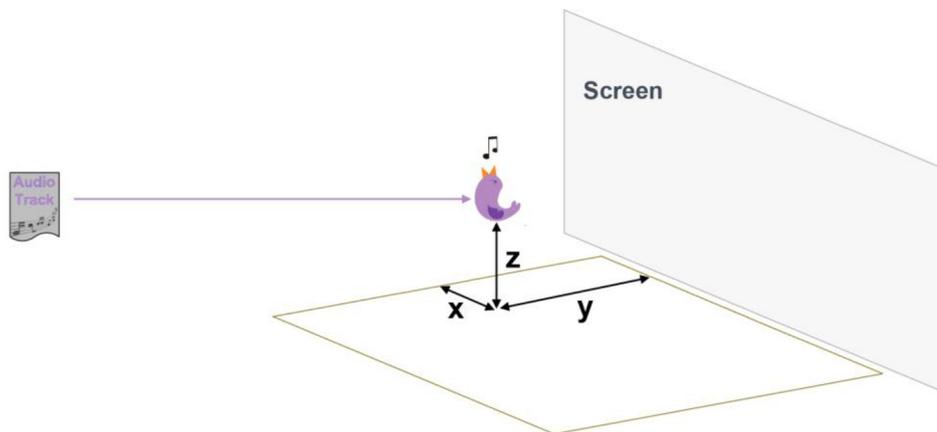


Abbildung 2: Objektbasiertes Audio<sup>15</sup>

## 2.4 Kopfbezogenes Koordinatensystem

Für die Beschreibung von Hörereignisorten oder -ausdehnungen wird vorzugsweise das kopfbezogene Polarkoordinatensystem verwendet. Dessen Ursprung befindet sich auf der interauralen Achse zwischen beiden Gehörkanaleingängen.

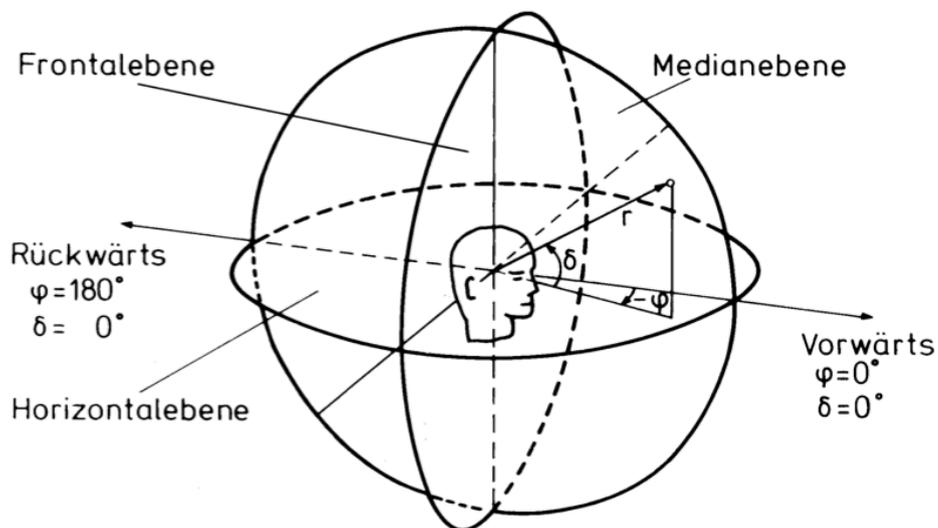


Abbildung 3: Kopfbezogenes Koordinatensystem<sup>16</sup>

<sup>15</sup> SMPTE 2014: 26

<sup>16</sup> Weinzierl 2008: 88

Im kopfbezogenen Koordinatensystem verläuft die Horizontalebene durch die interaurale Achse und der Unterkante der Augenhöhe. Die Frontalebene verläuft orthogonal zur Horizontalebene durch die interaurale Achse. Die Medianebene verläuft orthogonal sowohl zur Horizontal, als auch zur Frontalebene und teilt den fiktiven Kopf in zwei Hälften. Die Position einer Schallquelle wird mit den Polarkoordinaten Azimut  $\varphi$ , Elevation  $\delta$ , und Entfernung  $r$  dargestellt.<sup>17</sup> Die Position von Schallquellen oder Audio-Objekten kann neben dem kopfbezogenen Koordinatensystem auch mit dem kartesischen Koordinatensystem über die Koordinaten  $X, Y, Z$  beschrieben werden.

## 2.5 Räumliches Hören

Die Hörwahrnehmung der Position einer Schallquelle, also das Erkennen von Richtung und Entfernung einer Schallquelle, beruht auf vielen Faktoren und stellt einen komplexen Vorgang dar. Im Folgenden soll ein kurzer Überblick aus der Sicht der physikalischen Akustik sowie der Psychoakustik gegeben werden. Auf neurophysiologische Modelle wird in dieser Arbeit nicht näher eingegangen.

### 2.5.1 Lokalisation

Bei dem Vorgang der Lokalisation ordnet man einer wahrgenommenen Schallquelle die Parameter Azimut  $\varphi$ , Elevation  $\delta$  und Entfernung zu. Eine seitliche Auslenkung in der Horizontalebene wird mit dem Seitenwinkel  $\varphi$  angegeben (Azimut). Positionen in der Medianebene werden mit dem Erhebungswinkel  $\delta$  (Elevation) bezeichnet.<sup>18</sup> Die Wahrnehmung dieser Parameter gliedert sich dabei in monaurale und binaurale Faktoren und basiert auf verschiedenen Mechanismen, die jeweils auf unterschiedlichen physikalischen Eigenschaften der Schallquelle beruhen.

### 2.5.2 Monaurale Faktoren

Monaurale Faktoren beschreiben Änderungen, die von beiden Ohren gleich wahrgenommen werden. Die Änderung einer Schallquellenposition führt zu einer Änderung des Amplitudenspektrums der Außenohrübertragungsfunktion im vorwiegend hochfrequenten Bereich. Ausreichender Frequenzinhalt des Quellspektrums sowie eine gute Kenntnis der spektralen Beschaffenheit der Schallquelle sind demnach Voraussetzung, um diesen Faktor zur Positionswahrnehmung durch den Menschen nutzen zu können.<sup>19</sup>

---

<sup>17</sup> Weinzierl 2008

<sup>18</sup> vgl. Blauert 2014: 11

<sup>19</sup> vgl. Majdak 2002: 5

### 2.5.3 Binaurale Faktoren

Die binauralen Faktoren beschreiben Unterschiede in der Wahrnehmung zwischen dem der Schallquelle zugewandten, ipsilateralen, und abgewandten, kontralateralen, Ohr.

#### **Interaural Time Difference (ITD)**

*dt. Lokalisation durch interaurale Laufzeitunterschiede.*

Bei dieser Methode, werden vom Gehirn sowohl Laufzeitunterschiede als auch Hüllkurven beider Ohrsignale ausgewertet.<sup>20</sup> Diese ergeben sich beispielsweise durch Beugung der Schallwellen um den Kopf. Diese Methode ist für Frequenzen unterhalb 1,5 kHz wirksam, da die Wellenlänge größer als der Kopfdurchmesser ist. Aus einem durchschnittlichen Kopfdurchmesser von ca. 17 cm resultiert ein maximaler Laufzeitunterschied von 0,63ms. An Hand der kleinsten wahrnehmbaren Laufzeitdifferenz von 0,03ms kann man nach Dickreiter eine Lokalisationsgenauigkeit von ca. 3°-5° ableiten.<sup>21</sup>

#### **Interaural Level Difference (ILD)**

*dt. Lokalisation durch interaurale Intensitätsunterschiede.*

Frequenzen über 1,5 kHz werden durch den Kopf abgeschattet. Dadurch entstehen Pegelunterschiede zwischen dem ipsilateralen und dem kontralateralen Ohr.

#### **Cone of Confusion**

ITDs können in Anbetracht von Kurven gleicher interauraler Zeitdifferenzen,<sup>22</sup> Mehrdeutigkeit aufweisen, vor allem in Bezug auf Vorne-Hinten-Lokalisation. Mehrere Studien haben bereits bewiesen, dass die Vorne-Hinten Lokalisation bereits durch kleine Kopfbewegungen verbessert werden kann.<sup>23</sup>

---

<sup>20</sup> vgl. Blauert 2014: 132

<sup>21</sup> vgl. Dickreiter 2008: 106

<sup>22</sup> vgl. Whightman & Kistler 1997

<sup>23</sup> vgl. Perrett 1997; Macpherson 2013

## **Head Related Transfer Function (HRTF)**

*dt. Außenohr-Übertragungsfunktion.*

Bewegt sich eine Schallquelle auf der Medianebene, kommt es nahezu zu keiner ITD und ILD. Für die Richtungswahrnehmung bei Hörereignissen in der Medianebene ist die bei jedem Menschen individuelle HRTF verantwortlich. Im Frequenzbereich bezeichnet man diese Funktion als HRTF – auf Zeitbasis als *Head Related Impulse Response* (HRIR).

Die HRTF beschreibt die linearen Verzerrungen des Schallsignals am Eingang des Gehörgangs, die durch das Zusammenwirken der individuellen anatomischen Merkmale beim Hören entstehen.<sup>24</sup> Sie „[...] ist das Ergebnis von akustischer Abschattung, Beugung, Verzögerung, Resonanzen und Reflexionen durch Torso, Schulter, Kopf, Ohrmuschel, den Eingang in den Ohrkanal und den Ohrkanal selbst.“<sup>25</sup> Die HRTF ändert sich in Abhängigkeit der Kopfausrichtung und Position kontinuierlich.

Diese Übertragungsfunktionen können über Mikrofone in beiden Ohren einer Versuchsperson oder eines Kunstkopfes gemessen und aufgezeichnet werden. Es gibt aber auch die Möglichkeit, HRTFs über unterschiedliche Modelle zu formen, beispielsweise über das Kugelmodell von Blauert, die Zerlegung mit der Hauptkomponentenmethode nach Kistler and Wightman oder die individuelle Anpassung der HRTFs durch Time-Stretching.

## **Binaural Room Impulse Response (BRIR)**

Während HRTFs aufgenommene Impulsantworten in reflexionsarmen Räumen beschreiben, die nur die akustische Wirkung von Außenohr, Kopf und Torso charakterisieren, bezeichnen BRIR binaurale Impulsantworten, die in akustischen wirksamen Räumen gemessen wurden. Diese enthalten neben dem Freifeldanteil auch die für Quell- und Empfängerposition spezifischen Reflektionsmuster des Raumes. HRIRs haben Längen von lediglich wenigen Millisekunden und können ohne gravierende perzeptive Beeinträchtigung durch 72 Samples bei 48 kHz repräsentiert werden.<sup>26</sup> Dahingegen haben BRIRs eine Länge entsprechend der Nachhallzeit des aufgenommenen Raumes. Der Rechenaufwand zur Auralisation eingesetzter Faltungsalgorithmen steigt somit um mehrere Größenordnungen.<sup>27</sup>

---

<sup>24</sup> vgl. Blauert 2014: 62

<sup>25</sup> Weinzierl 2008: 586

<sup>26</sup> vgl. Hammershoi u. Moller 2005

<sup>27</sup> vgl. Weinzierl 2008: 673

## 2.6 Entfernung und Tiefe

Der Vergleich zwischen Bild und Ton zeigt, dass die dritte Dimension unterschiedlich definiert ist. Die Parameter einer 2D Darstellung bestehen aus Breite x Höhe, wohingegen 2D Audio mit den Parametern Breite x Tiefe beschrieben wird. Bei einer visuellen Darstellung entsteht die dritte Dimension über Tiefeninformation, wohingegen bei 3D-Audio die dritte Dimension mit der Höhe beschrieben wird. Das bedeutet mit den Worten von Theile, dass mit einer einfachen 2.0-Stereofonie keine zufriedenstellende Abbildung des natürlichen Raumeindrucks beziehungsweise der räumlichen Tiefe möglich ist.

Es wird lediglich eine Darstellung der räumlichen Perspektive vergleichbar mit der perspektivischen Darstellung in einem flächigen Bild erreicht. Bei 3/2-Stereofonie kann mit Hilfe der Surround-Kanäle die Abbildungsebene zwischen den vorderen Lautsprechern verlassen werden. Daraus ergeben sich Möglichkeiten für die Reproduktion des frühen Seitenschalls zur Darstellung der räumlichen Tiefe sowie des Nachhalls zur Darstellung des Raumeindrucks und der Umhüllung.<sup>28</sup>

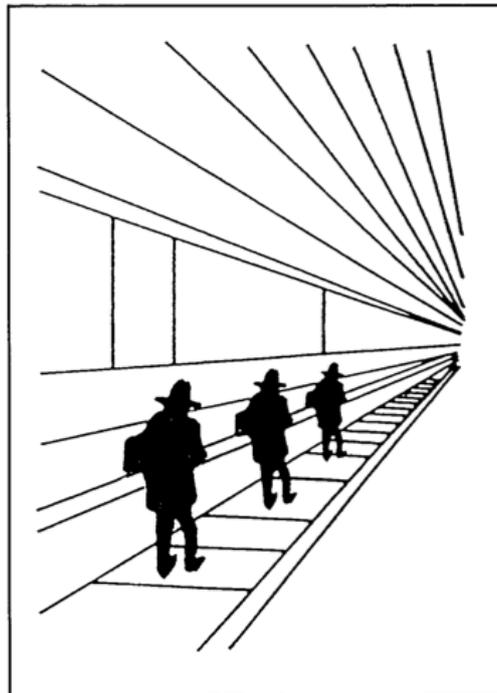


Abbildung 4: Die visuelle Perspektive kann mit der stereofonen Perspektive verglichen werden<sup>29</sup>

---

<sup>28</sup> vgl. Theile 1991: 762

<sup>29</sup> ebenda

## 2.7 Klangfarbe und Räumlichkeit

*”Timbre is that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar.”<sup>30</sup>*

Die Definition der *American Standard Acoustical Terminology* gibt Aufschluss über die Bedeutung der Klangfarbe bei der Unterscheidung von Klängen. Sie ist eines der am häufigsten genutzten Attribute, um Klang zu definieren. Die Fähigkeit, Klangfarbe einer virtuellen Schallquelle korrekt oder zumindest plausibel wiederzugeben, ist ein wichtiger Bestandteil eines jeden Schallreproduktionssystems. Nach Rumsey würde der Zuhörer dabei eher einen Kompromiss bei der räumlichen Genauigkeit als eine verminderte Klangfarbe in Kauf nehmen.<sup>31</sup>

## 2.8 Prinzipien der klanglichen Reproduktion

Im Bereich 3D Audio gibt es neben der Mehrkanalstereofonie und Binauraltechnik das Prinzip der Schallfeldrekonstruktion. Dazu zählen die Wellenfeldsynthese und das Ambisonicsformat. Alle bekannten räumlichen Tonübertragungsverfahren lassen sich darauf zurückführen oder stellen Mischformen dar. Je nach Anwendungsbereich können bestimmte Vorteile des Prinzips genutzt beziehungsweise Nachteile des Verfahrens umgangen werden.<sup>32</sup>

### 2.8.1 Mehrkanal-Stereofonie

Mehrkanal-Stereofonie basiert auf der Grundlage der kanalbasierten Lautsprecher-Stereofonie. Diese beruht im Prinzip auf den Gesetzmäßigkeiten der Lokalisation im überlagerten Schallfeld von zwei oder mehr Lautsprechern. Die Richtungsdarstellung erfolgt in der Abbildungsebene zwischen benachbarten Lautsprechern.<sup>33</sup>

### 2.8.2 Vector Base Amplitude Panning

Vector Base Amplitude Panning, kurz VBAP, wurde 1997 von Pulkki eingeführt. Das VBAP-Verfahren ist eine mathematische Formalisierung des traditionellen stereofonen Panorama-Potentiometers. Es berechnet die Signalpegel für jeweils benachbarte Lautsprecherpaare, im dreidimensionalen Fall für Lautsprechertripel, in Abhängigkeit von der gewünschten Abbildungsrichtung der

---

<sup>30</sup> ASA 1960

<sup>31</sup> vgl. Rumsey et al. 2005

<sup>32</sup> vgl. Theile et al 2002

<sup>33</sup> vgl. Theile 1981: 163

Schallquelle für beliebige zwei- beziehungsweise dreidimensional angeordnete Lautsprecherkonfigurationen. Im Gegensatz zu Ambisonics (Kapitel 2.8.3) oder Wellenfeldsynthese (Kapitel 2.8.4.) wird keine physikalische Synthese von realen oder virtuellen Schallfeldern angestrebt, sondern eine auf dem Phänomen der Summenlokalisierung beruhende Erzeugung von Phantomschallquellen.<sup>34</sup> Psychoakustische Grundlage ist das aus einem akustischen Modell für Schallausbreitung am Kopf abgeleitete *tangent law*<sup>35</sup>.

$$\frac{\tan \theta}{\tan \theta_0} = \frac{g_1 - g_2}{g_1 + g_2} \quad (1)$$

Es sagt den Hörereignisort der Phantomschallquelle (panning angle  $\theta$ ) aus den Verstärkungsfaktoren  $g_1$  und  $g_2$  der benachbarten Lautsprecher und dem Winkel  $\theta_0$  der Lautsprecherbasis voraus. (siehe Abbildung 5)

$$g_1^2 + g_2^2 = 1 \quad (2)$$

Mit Formel (2) wird die Leistungssumme der Lautsprechersignale konstant gehalten, wodurch eine konstante Summenlautheit erzeugt wird.

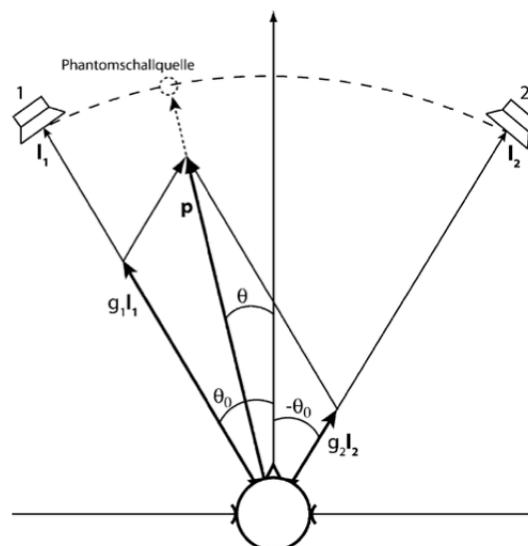


Abbildung 5: Lokalisationswinkel der Phantomschallquelle  $\theta$  und Versatzwinkel  $\theta_0$

Typischerweise ist  $\theta_0 = 30^\circ$  und  $|\theta| \leq \theta_0$ <sup>36</sup>

<sup>34</sup> vgl. Pulkki 2001

<sup>35</sup> vgl. Bennet 1985

<sup>36</sup> vgl. Weinzierl, 2008, S. 658

Der zur gewünschten Phantomschallquelle zeigende Vektor  $p$  lässt sich als Linearkombination der Lautsprechervektoren  $I_1$  und  $I_2$  konstruieren. Er lässt sich auf drei Dimensionen erweitern, wenn man als Vektorbasis ein Triplet aus drei Lautsprechervektoren  $I_1$ ,  $I_2$  und  $I_3$  und die zugehörige  $3 \times 3$ -Matrix verwendet.

Hörversuche zeigten, dass der wahrgenommene Hörereignisort im lateralen Bereich, wo sich die Lokalisationskurve nicht durch das *tangent law* beschreiben lässt, durch die so berechneten Vektoren nur ungenau annähern lässt. Auch innerhalb der Medianebene führt das reine Amplituden-Panning nur zu unscharfen Richtungsinformationen.<sup>37</sup>

### 2.8.3 Ambisonics

Ambisonics ist ein Aufnahme- und Wiedergabeverfahren zur Reproduktion von zwei-beziehungsweise dreidimensionalen Schallfeldern über eine variable Lautsprecheranordnung. Im Gegensatz zu Vector Based Amplitude Panning (VBAP), bei dem ein Quellsignal zwischen drei Lautsprechern aufgeteilt wird, tragen bei Ambisonics alle Lautsprecher zur Synthetisierung des Schallfeldes bei. Michael Gerzon entwickelte, basierend auf den Absätzen von Alan Blumlein, der bereits in den 30er Jahren gute Erfahrungen mit koinzidenten Mikrofonanordnungen gemacht hatte, zusammen mit seinen Studienkollegen Peter Carven und Richard Lee an der Oxford University den Ambisonic Ansatz.<sup>38</sup>

Der Vorteil von Ambisonics gegenüber anderen Surround-Technologien ist die Einführung eines von der Lautsprecherkonfiguration und der Anzahl an Quellsignalen unabhängigen Zwischenformats. Die Qualität der Schallfeldreproduktion hängt von der Ambisonics-Ordnung sowie von der Anordnung und Anzahl der entsprechenden Lautsprechern ab. Die Lautsprecherpositionen des Wiedergabesetups sind von der Produktion des Ambisonics-Schallfeldes unabhängig.

Der grundlegende Ansatz basiert hier auf dem Vergleich einer aufgenommenen Schallwelle (Referenzwelle) mit der durch einen Lautsprecher synthetisierten Schallwelle. Es wird davon ausgegangen, dass sowohl die aufgenommene Schallwelle als auch die abgestrahlte Schallwelle vom Lautsprecher einer ebenen Welle entspricht. Dies trifft dann zu, wenn die aufgenommene Schallquelle sowie die wiedergebenden Lautsprecher weit genug von der Abhörposition entfernt sind (Freifeld Bedingung).<sup>39</sup>

---

<sup>37</sup> Pulkki 2001

<sup>38</sup> vgl. Gerzon 1973

<sup>39</sup> vgl. Kronlacher 2012

Ambisonics Enkodierung, Manipulation und Dekodierung werden durch eine Reihe von einfachen Matrixmultiplikationen und Additionen ermöglicht.

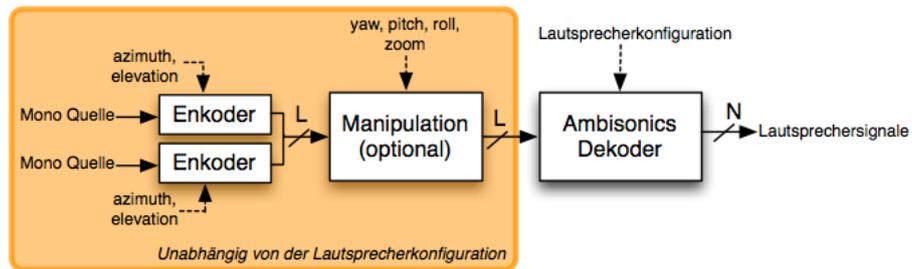


Abbildung 6: Ambisonic Blockdiagramm<sup>40</sup>

Nach der von Michael A. Garzon im Jahr 1970 entwickelten Theorie wird ein dreidimensionales Schallfeld in vier Kanäle (W, X, Y und Z) kodiert was auch als B-Format bezeichnet wird. Im B-Format können mit folgender Vorschrift Monoquellen kodiert (positioniert) werden:

$$W = \frac{1}{k} \sum_{i=1}^k s_i \left[ \frac{1}{\sqrt{2}} \right] \quad \text{Omnidirektionaler Anteil} \quad (3.1)$$

$$X = \frac{1}{k} \sum_{i=1}^k s_i [\cos(\phi_i) \cos(\theta_i)] \quad \text{nach X gerichteter Anteil} \quad (3.2)$$

$$Y = \frac{1}{k} \sum_{i=1}^k s_i [\sin(\phi_i) \cos(\theta_i)] \quad \text{nach Y gerichteter Anteil} \quad (3.3)$$

$$Z = \frac{1}{k} \sum_{i=1}^k s_i [\sin(\theta_i)] \quad \text{nach Z gerichteter Anteil} \quad (3.4)$$

$s_i$  bezeichnet die einzelnen Monoquellen welche an der gegebenen Position mit dem horizontalen Winkel  $\phi_i$  (Azimuth) und dem vertikalen Winkel  $\theta_i$  (Elevation) enkodiert wird.

<sup>40</sup> Kronlacher, 2012

Die grafische Darstellung der vier B-Format Kanäle sieht wie folgt aus:

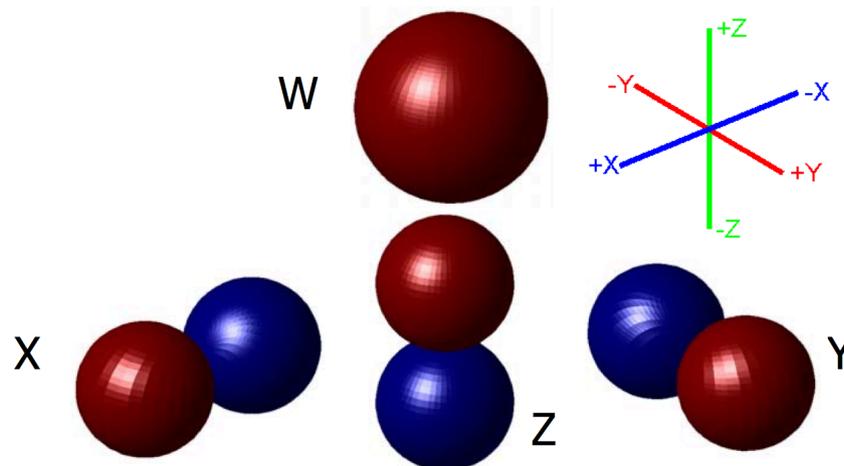


Abbildung 7: Erste Ordnung Ambisonics - W,X,Y und Z Kanal, dabei ist der In-Phase Anteil rot.<sup>41</sup>

Die Grafik verdeutlicht, dass für den W-Kanal ein Signal von einem omni-direktionalen Mikrofon aufgezeichnet werden kann. Für die Aufnahme der X, Y und Z Kanäle werden drei Mikrofone mit einer Acht als Richtcharakteristik benötigt, um ein reales Schallfeld aufzuzeichnen. In der idealen Umsetzung müssten diese vier Mikrofone gemeinsam im Ursprung des Koordinatensystems positioniert werden, was sich praktisch nur sehr schwer verwirklichen lässt. In der Praxis werden Ambisonics Aufnahmen üblicherweise im A-Format aufgenommen. Dieses Format setzt sich aus vier Mikrofonen mit Nierencharakteristik zusammen, welche tetraedisch angeordnet werden.<sup>42</sup> Durch einfache Summen- und Differenzbildung kann das A-Format in das B-Format übergeführt werden. Im B-Format kann ein Ambisonics Signal sehr einfach um seine Achsen rotieren und in der Richtungsdominanz (Zoom) verändert werden.<sup>43</sup>

$$W = 0.5 \cdot (A + B + C + D) \quad \text{Omnidirektionaler Anteil} \quad (4.1)$$

$$X = (A + C) - (B + D) \quad \text{nach X gerichteter Anteil} \quad (4.2)$$

$$Y = (A + B) - (C + D) \quad \text{nach Y gerichteter Anteil} \quad (4.3)$$

$$Z = (A + D) - (B + C) \quad \text{nach Z gerichteter Anteil} \quad (4.3)$$

#### Ambisonics Formatüberführung. A-Format zu B-Format

<sup>41</sup> vgl. Wiggins 2004

<sup>42</sup> vgl. Kronlacher 2012

<sup>43</sup> vgl. Wiggins 2004

Um das bestmögliche Ergebnis bei Aufnahmen zu erzielen, müssten die Mikrofon-Kapseln idealerweise koizident angeordnet sein. Da dies in der Praxis nicht möglich ist hat Gerzon Möglichkeiten entwickelt, die Klangverfärbungen auf Grund des räumlichen Abstandes der Mikrofonkapseln zu minimieren.<sup>44</sup>

Da Ambisonics in seiner ersten Ordnung nur eine beschränkte Genauigkeit in der Abbildung und Reproduktion von Schallfeldern bietet, wurde Anfang des Jahres 1990 gezeigt, dass der Ambisonics Ansatz auf höhere Ordnungen erweiterbar ist. Mit Higher Order Ambisonics (HOA) wird die Richtungsabbildung verbessert und der Sweet Spot vergrößert. Der Higher-Order-Ambisonics-Ansatz resultiert in einer Reihenentwicklung in Sphärischen Harmonischen.<sup>45</sup>

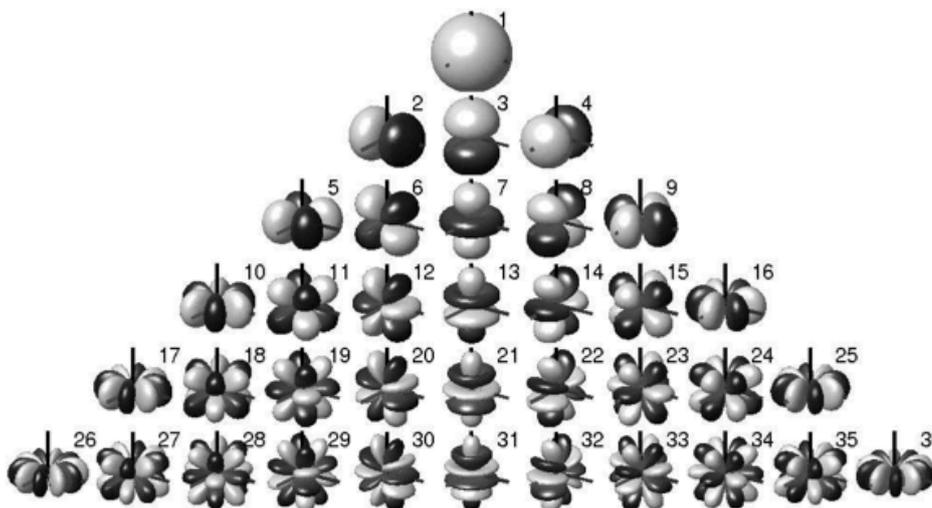


Abbildung 8: Sphärische Harmonische<sup>46</sup>

#### 2.8.4 Wellenfeldsynthese

Als objektbasiertes System bildet die Wellenfeldsynthese keine Phantomschallquellen zwischen benachbarten Lautsprechern ab, sondern strebt vielmehr eine reale physikalische Synthese eines räumlich ausgedehnten Schallfelds an. Dabei erzeugen Interferenzen von einer Vielzahl an dicht zueinander positionierten Lautsprechersignalen das re-synthesierte Schallfeld.

Die Basis der Wellenfeldsynthese liefert das nach Christian Huygens und Augustin Fresnel benannte Huygens-Fresnel-Prinzip. Grundlegende Phänomene der akustischen aber auch optischen Wellenausbreitung, wie Reflexion, Brechung und Beugung, können damit beschrieben

<sup>44</sup> vgl. Gerzon 1975

<sup>45</sup> vgl. Sontacchi 2003

<sup>46</sup> Nachbar 2011

werden. Bei der Wellenfeldsynthese kann jeder Punkt einer sich ausbreitenden Wellenfront als Ausgangspunkt einer neuen sich ausbreitenden Elementar- oder Sekundärwelle betrachtet werden.<sup>47</sup>

Die folgende Abbildung zeigt links das theoretische Huygens-Fresnel Prinzip und rechts wie es prinzipiell bei der WFS zu Anwendung kommt. Wenn aus dem Punkt S eine Kugelwelle ausgeht, bewegt sie sich fort, indem die in Schwingung versetzten Teilchen ihre Bewegungen auf ihre Nachbarpartikel übertragen. Dieser Vorgang setzt sich symmetrisch in alle Richtungen fort. Die blaue Wellenfront kann als Summe aller Elementarwellen (Punktschallquellen) an der Oberfläche O (rot) betrachtet werden. Durch Kenntnis der Wellenfront auf der Fläche O lässt sich der Schwingungszustand in einem Punkt P berechnen. Bei der Umsetzung der WFS ersetzt man die Punktquellen durch geeignete Lautsprecher und erzeugt dadurch die entsprechende Kugelwelle. Dadurch wird die Schallquelle S zu einer virtuellen Schallquelle. Der Hörer an Punkt P empfängt jedoch die gleiche Wellenfront welche die Schallquelle S ausstrahlen würde.<sup>48</sup>

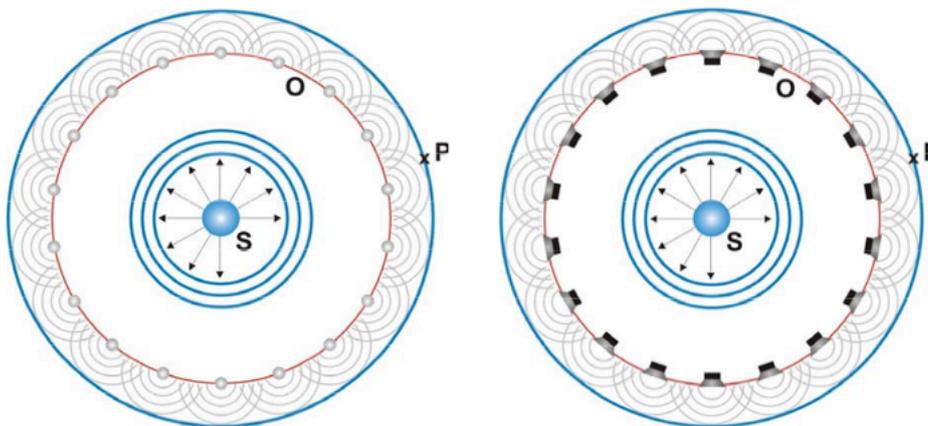


Abbildung 9: Huygen-Fresnel-Prinzip<sup>49</sup>

Das Kirchhoff-Helmholz-Integral liefert eine quantitative, mathematische Formulierung für die beschriebene Schallfeldsynthese nach dem Huygens-Fresnel-Prinzip.<sup>50</sup>

<sup>47</sup> vgl. Born & Wolf 1975

<sup>48</sup> vgl. Wittek 2007

<sup>49</sup> Theile et al. 2003

<sup>50</sup> vgl. Start 1997

Die physikalische Aussage des Kirchhoff-Helmholtz-Integrals beschreibt, dass ein Wellenfeld, das der akustischen Wellengleichung genügt, im Inneren eines Raumvolumens eindeutig durch die Bedingung auf dem Rand definiert ist:<sup>51</sup>

$$p_A = \frac{1}{4\pi} \int_S \left[ \underbrace{\left( p \frac{1+jkr}{r} \cos \varphi \frac{e^{-jkr}}{r} \right)}_{\text{Dipol}} + \underbrace{\left( j\omega \rho_0 v_n \frac{e^{-jkr}}{r} \right)}_{\text{Monopol}} \right] dS \quad (5)$$

Mit der *Formel (5)* als Anleitung zur Schallfeldsynthese ergibt sich der Schalldruck  $p_A(r)$  am Hörerplatz  $r_A$  als Summe von Schallsignalen auf der Oberfläche  $S$ .

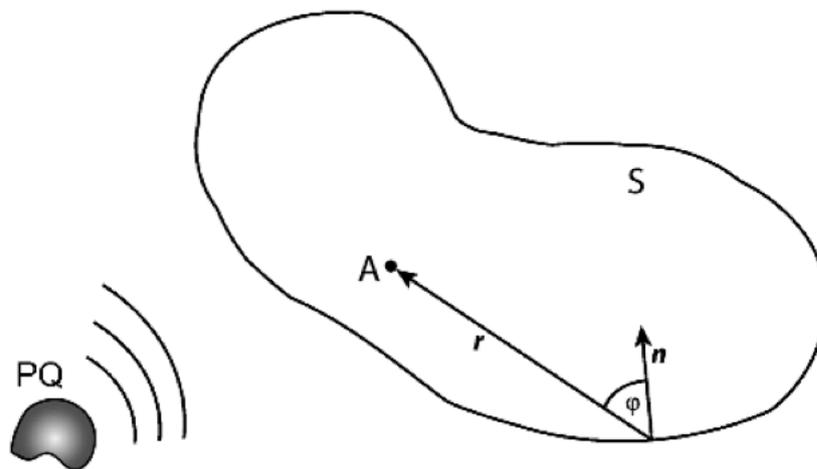


Abbildung 10: Kirchhoff-Helmholtz-Integral<sup>52</sup>

An einem beliebigen Punkt A (Hörerplatz) in einem Raumvolumen V ist der Schalldruck einer externen Primärquelle (PQ) durch Schalldruck und -schnelle auf der das Volumen umschließenden Oberfläche S gegeben (Kirchhoff-Helmholtz-Integral).<sup>53</sup> In der Wellenfeldsynthese wird die Oberfläche S durch das Lautsprecherarray gebildet. Schalldruck und -schnelle der abzubildenden Quelle lassen sich für reale Quellen durch Mikrofone messen (Druck- beziehungsweise Gradientenempfänger) oder für virtuelle Quellen berechnen. Für die Synthese virtueller

<sup>51</sup> vgl. Weinzierl 2008: 666

<sup>52</sup> Berkhout et al.1993

<sup>53</sup> vgl. Berkhout et al.1993

Schallfelder werden typischerweise Punktquellen oder ebene Wellen als besonders einfache Modelle der Wellenausbreitung angenommen.<sup>54</sup>

In praktischer Hinsicht wäre eine unmittelbare Umsetzung der WFS nach dem Kirchhoff-Helmholtz-Integral nahezu undenkbar. Mit seiner theoretischen Basis eignet es sich jedoch hervorragend, um das gewünschte Schallfeld zu berechnen. In der Praxis ist es nicht möglich unendlich große oder gar zweidimensionale Lautsprecherarrays mit derartig kleinem Lautsprecher-Abstand zu installieren. Hinzu kommen die akustischen Eigenschaften des realen Raumes, die Einfluss auf die virtuelle Akustik haben. In jedem Fall müssen Kompromisse eingegangen werden. Einige der Kompromisse gegenüber der exakten mathematischen Lösung werden in der Weiterentwicklung des zunächst von Berkhout im Jahr 1988 erarbeiteten Konzeptes 1993 von De Vries und Kollegen an der TU Delft formuliert. Neben dem Verzicht der Anordnung von Schallquellen auf einer geschlossenen Oberfläche zugunsten eines nur in einer Ebene angeordneten Arrays von Lautsprechern und dem Verzicht auf die Verwendung monopol- und dipolförmiger Quellen zugunsten einer meist näherungsweise monopolförmigen Charakteristik realer Lautsprecher nahm sich die Weiterentwicklung vor allem auch dem Grundproblem der Abmessungen realer Lautsprechersysteme an. Auf Grund des diskreten Abstands von realen Lautsprechern entsteht Spatial Aliasing. Ähnlich dem Abtasttheorem können durch die Interferenz von Lautsprechersignalen in endlichem Abstand nur Wellenformen bis zu einer bestimmten Grenzfrequenz exakt synthetisiert werden. Über folgende Formel lässt sich die Grenzfrequenz abschätzen und ist dabei abhängig vom Lautsprecherabstand  $\Delta x$ :

$$f_g = \frac{c}{2\Delta x} \quad (6)$$

Je nach Form und Ausbreitungsrichtung der synthetisierten Welle sowie Position des Hörers verhält sich die Grenzfrequenz anders.<sup>55</sup> Spatial Aliasing tritt für ebene Wellen, die sich senkrecht zu einem geraden Lautsprecherarray ausbreiten, erst oberhalb folgender Grenzfrequenz auf:

$$f_g = \frac{c}{\Delta x} \quad (7)$$

---

<sup>54</sup> vgl. Weinzierl 2008: 667

<sup>55</sup> vgl. Spors u. Rabenstein 2006; Leckschat u. Baumgartner 2005

Um Wellenfronten oberhalb einer Wellenlänge  $\lambda_g$  korrekt zu synthetisieren, ist daher in der Theorie folgender Lautsprecherabstand nötig:<sup>56</sup>

$$\Delta x < \frac{\lambda_g}{2} \quad (8)$$

Nach E. Start haben sich Lautsprecherabstände von 10 bis 30 cm für Signale mit üblicher spektraler Zusammensetzung, wie beispielsweise Sprache und Musik, als ausreichend für eine Wellenfeldsynthese erwiesen, ohne die Lokalisation gravierend zu beeinträchtigen.<sup>57</sup>

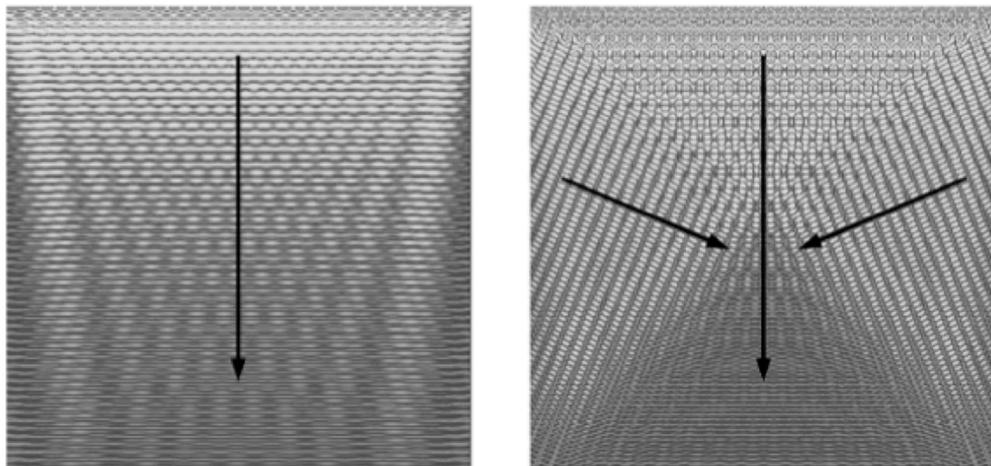


Abbildung 11: Aliasing-Artefakte

Aliasing-Artefakte für ein lineares Lautsprecherarray mit Lautsprecherabständen von  $\Delta_x = 30$  cm. Eine ebene Welle mit  $f = 1000$  Hz ( $\lambda = 34$  cm) wird fehlerfrei synthetisiert (links), während bei  $f = 1600$  Hz ( $\lambda = 21$  cm) bereits Artefakte in Form von Alias-Wellen entstehen, die sich schräg zur Originalwelle ausbreiten (rechts).<sup>58</sup>

Um perzeptive Beeinträchtigungen auf Grund der Aliasing-Artefakte zu minimieren, gibt es unter anderem den hybriden Ansatz zwischen WFS und Stereo. Die OPSI-Methode (Optimised Phantom Source Imaging in wavefield synthesis) verfolgt den Ansatz, unterhalb  $f_{alias}$ , WFS als System der klanglichen Reproduktion zu verwenden und oberhalb  $f_{alias}$  die Abbildung der virtuellen Quelle über stereophone Phantomquellen zu reproduzieren.<sup>59</sup> Einen ähnlichen Ansatz verfolgt auch der Isoono-Algorithmus.

<sup>56</sup> vgl. Weinzierl, 2008: 669

<sup>57</sup> vgl. Start 1997

<sup>58</sup> Weinzierl 2008: 670

<sup>59</sup> vgl. Wittek 2007

## 2.8.5 Binauraltechnik

Bei der Binauraltechnik handelt es sich in erster Linie um die Reproduktion der Ohrsignale, wie sie von der Kunstkopfstereofonie bekannt sind. Anstelle der Reproduktion eines geeigneten Schallfeldes am Wiedergabeort werden mit Hilfe eines Kunstkopfes die am Ort der Aufnahme wirksamen Ohrsignale aufgezeichnet und später über den Kopfhörer wiedergegeben. Idealerweise sind die reproduzierten Signale des Kunstkopfes identisch mit den Ohrsignalen, die der Hörer am Ort der Aufnahme empfangen würde. Sofern dies der Fall ist, korrespondiert das “reale” Schallereignis im Ursprungsraum mit dem reproduzierten “virtuellen” Hörereignis im Wiedergaberaum.<sup>60</sup>

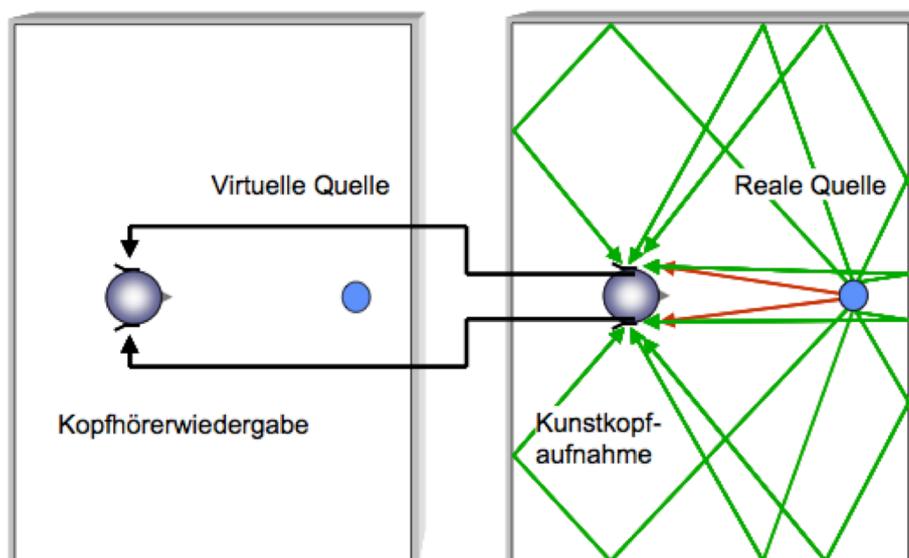


Abbildung 12: Kunstkopfstereofonie<sup>61</sup>

Neben der Aufnahme mit Hilfe eines Kunstkopfes lässt sich das binaurale Kopfhörersignal auch durch Faltung des am Senderort abgegebenen Signals mit den beiden kopfbezogenen Raumimpulsantworten des entsprechenden Raumes (BRIRs) erzeugen. Dieses Signal wird dann über Kopfhörer unmittelbar dem Ohrkanal zugeführt und enthält dann in der Regel alle durch das Übertragungssystem induzierten auditiven Merkmale.<sup>62</sup>

<sup>60</sup> vgl. Theile et al 2002

<sup>61</sup> ebenda

<sup>62</sup> vgl. Weinzierl 2008: 672

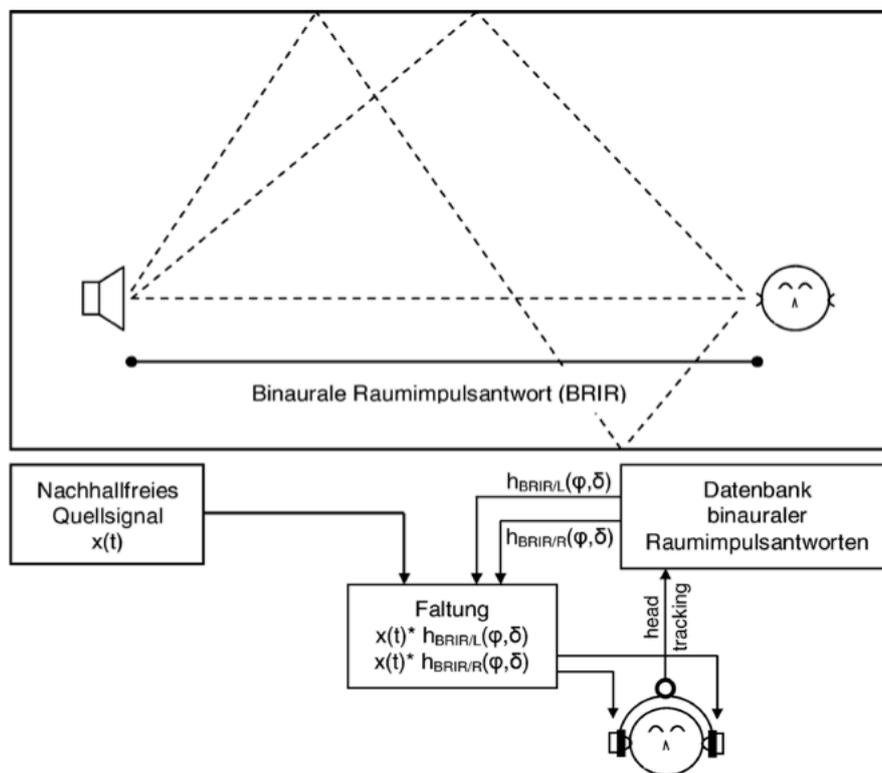


Abbildung 13: Binaurale Raumimpulsantwort

Ein binaurales Signalpaar wird durch Faltung eines nachhallfreien Quellsignals mit zwei zuvor gemessenen oder simulierten binauralen Raumimpulsantworten synthetisiert, die durch einen Headtracker entsprechend der Kopforientierung des Hörers aus einem Datensatz ausgewählt und dynamisch nachgeführt werden.<sup>63</sup>

Der praktische Einsatz der binauralen Technik bringt dabei umfangreiche Einschränkungen mit sich:

- Die Reproduktion der Ohrsignale muss sehr genau erfolgen, Unterschiede des individuellen Außenohres zum Kunstkopf-Außenohr verursachen bereits Beeinträchtigungen. Auch wirken Eingriffe wie Dynamikeinengung, Lautstärkeinstellung oder Störgeräusche bei der Wiedergabe nachteilig.<sup>64</sup>

<sup>63</sup> vgl. Weinzierl 2008: 672

<sup>64</sup> vgl. Theile et al. 2002

- Die Reproduktion der Ohrsignale muss über Kopfhörer erfolgen. Sogenannte Transauralisationsverfahren, welche für die Lautsprecherwiedergabe vorher eine inverse Filterung der Außenohrübertragungsfunktion durchführen, sind nicht präzise genug und schränken die Hörzone auf wenige Zentimeter ein.<sup>65</sup>
- Neuere Untersuchungen haben nachgewiesen, dass die Ohrsignale sogar an die momentane Kopfhaltung des Hörers angepasst werden müssen. Die dynamischen Ohrsignalmerkmale hatten bei Auswertung durch das Gehör eine bisher unterschätzte Bedeutung, insbesondere für die korrekte Lokalisation in der Medianebene.<sup>66</sup>
- Eine künstlerische Gestaltung des Klangbildes ist nur in sehr engem Rahmen möglich. Das gilt für die klangliche und räumliche Balance der Instrumente ebenso wie für die Darstellung des Raumes und der einhüllenden Atmosphäre. Die Abhängigkeit von den Gegebenheiten der Aufführung sind extrem groß und in den meisten Fällen unpraktikabel. Zudem stellt ein gelungenes Klangbild für die Kopfhörerwiedergabe immer eine Abbildung eines Hörerplatzes im Aufnahmeraum dar und ist nicht optimal für „wohzimmergerechte“ Lautsprecherwiedergabe (mangelnde künstlerische Kompatibilität).<sup>67</sup>

---

<sup>65</sup> vgl. ebenda

<sup>66</sup> vgl. Mackensen et al. 1998

<sup>67</sup> vgl. Heile et al 2002

### 3 3D-Audio-Wiedergabesysteme

Im Bereich der 3D-Audio Produktion gibt es eine überschaubare Anzahl an unterschiedlichen Formaten und Systemen. Sowohl objektbasierte als auch kanalbasierte Systeme sind dabei, sich auf dem Markt zu etablieren. Die Grundlage für die Verwirklichung des im Rahmen dieser Arbeit umgesetzten 3D-Audio-Live-Konzertes basieren während der Vorproduktionen auf dem Auro-3D- und während des finalen Live-Konzerts auf dem Iosono-System. Im Folgenden werden die Grundlagen und Arbeitsabläufe beider Systeme beschrieben.

#### 3.1 Auro 3D

Das kanalbasierte Auro-3D-Format bietet Lösungen mit Lautsprecheranordnungen von 8.0 bis 15.1. Mit der einfachsten Lösung, zwei quadrophonischen Grundaufstellungen, die übereinander angeordnet sind, würde man dem ITU 5.1-Standard nicht gerecht werden. Ergänzt man die 8.0 Aufstellung durch einen zusätzlichen Center- und LFE-Lautsprecher, so ergibt sich die kleinste Auro-3D-Lösung: Auro 3D 9.1. Daraus entsteht eine natürliche Energieverteilung von: fünf Lautsprechern vorne - vier hinten, vier Lautsprecher links - vier rechts, fünf Lautsprecher unten - vier oben. Diese Energieverteilung nimmt man als natürliches Klangbild wahr.

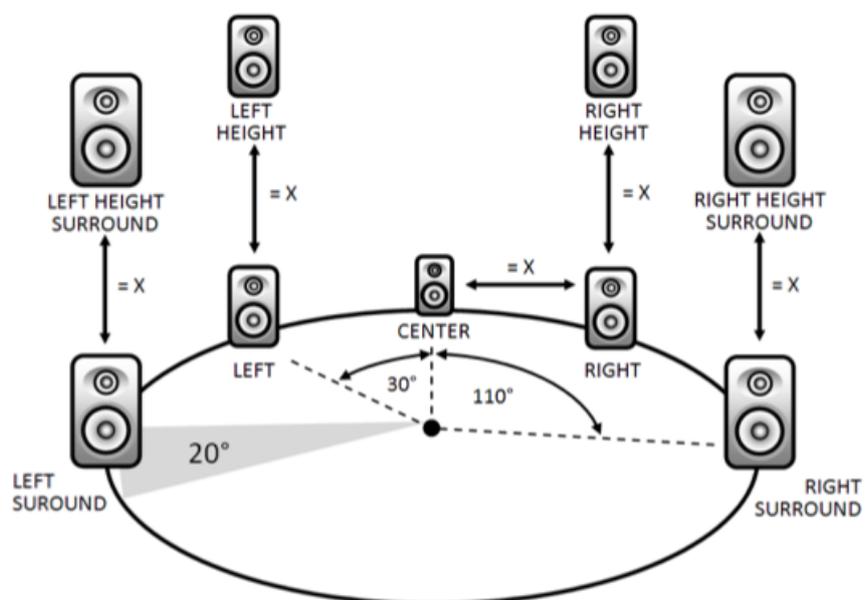


Abbildung 14: Lautsprecheranordnung für Auro 3D (9.1)

Darstellung ohne LFE; Die untere Anordnung der Lautsprecher folgt dem ITU-R BS.775 Standard. Der ideale Abstand zwischen den normal hohen und den hohen Lautsprechern entspricht ungefähr der halben Basisbreite (Basisbreite = Abstand zwischen Left und Right Lautsprecher)<sup>68</sup>

L	Left channel
C	Center channel
R	Right channel
LS	Left surround channel (rear or side)
RS	Right surround channel (rear or side)
LFE	Low Frequency Effect channel
HL	Height Left channel
HC	Height Center channel
HR	Height Right channel
HLS	Height Left surround channel
HRS	Height Right surround channel
T	Top channel

*Auro 3D Kanalbezeichnung*

---

<sup>68</sup> vgl. Baelen 2001: 14

Neben der klanglichen Ausgewogenheit ist die Kompatibilität zu standardisierten 5.1-Surround-Systemen nach ITU-R BS.775 und zu Stereo-Systemen von besonderer Bedeutung. Intensive Experimente und Hörvergleiche in den Galaxy Studios im belgischen Mol führten zu einer optimalen Lautsprecherkonfiguration. Für eine optimale räumliche Darstellung ergaben sich dabei Neigungswinkel der oberen Lautsprecher von 25-35 Grad bezogen auf den Sweet Spot.

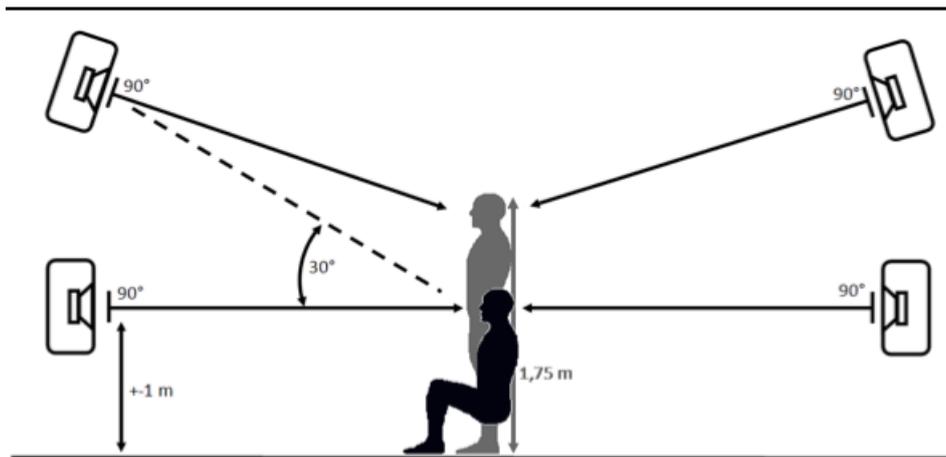


Abbildung 15: Seitenansicht der Auro-3D Anordnung

Seitliche Auro-3D-Lautsprecheranordnung, Darstellung ohne LFE. Die untere Lautsprecheranstellung entspricht dem ITU-Standard. Die Achsen der oberen Lautsprecher zeigen auf den Kopf des stehenden Zuhörers (grau) um eine optimale räumliche Darstellung zu erzielen.<sup>69</sup>

Ein Auro 3D Lautsprecher Setup kann mit bis zu 14 (13.1) Kanälen ergänzt werden. Daraus ergeben sich nach dem Standard folgende Möglichkeiten:

- Auro 3D 9.1 = 5.1 Surround + vier Höhenlautsprecher
- Auro 3D 10.1 = Auro 9.1 + Top-Channel (Deckenlautsprecher a.k.a. Voice of God)
- Auro 3D 11.1 = Auro 10.1 + Höhen-Center-Lautsprecher)
- Auro 3D 11.1b = 7.1 Surround + vier Höhenlautsprecher
- Auro 3D 13.1 = Auro 11.1b + Höhen-Center-Lautsprecher + Top-Channel

<sup>69</sup> vgl. Baelen 2011: 15

Man spricht bei Auro-3D-Lautsprecher setups auch von einer 3-Layer-Konfiguration. Dabei beschreibt der erste untere Layer die Standard 5.1 Surround-Ebene in Ohrhöhe. Die zweite Ebene mit 30 Grad Erhebung über den unteren Lautsprechern beschreibt den Höhen-Layer. Dieser besteht aus einer 5.0 oder 6.0 Lautsprecheranordnung. Die dritte Ebene beschreibt den Top-Layer, in der Fachliteratur auch „Stimme Gottes“ genannt.<sup>70</sup>

### 3.1.1 Verteilung der Schallfelder

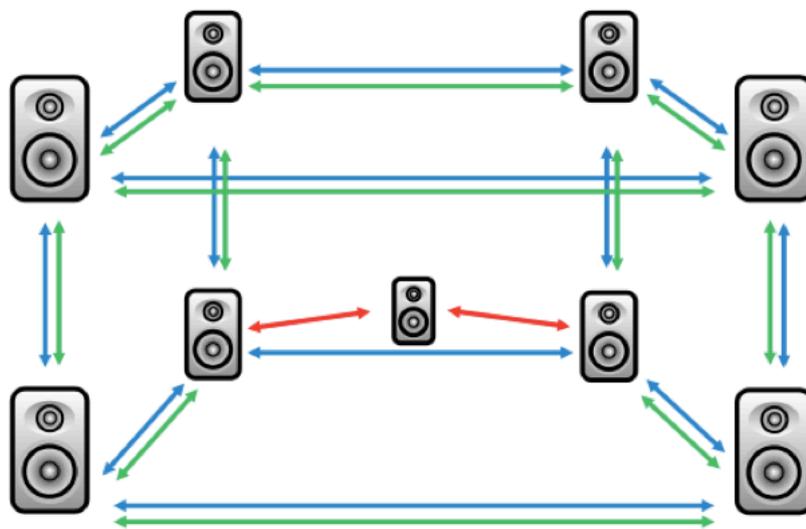


Abbildung 16: Mögliche Verteilung der räumlichen Attribute bei Auro-3D (9.1) Wiedergabe; Rot =Direktschall, Blau = Erstreflexionen, Grün = Nachhall

Auf der L-C-R-Basis wird der Direktschall der Schallquelle abgebildet. Auf den hinteren und oberen Lautsprechern wird Räumlichkeit in Form von dekorreliertem Diffusschall übertragen.<sup>71</sup>

Vor allem in Musikproduktionen ist die Aufgabe der unteren Frontlautsprecher in erster Linie eine korrekte Wiedergabe des frontalen Hörereignisses. Dabei wird eine möglichst genaue Lokalisation der aufgenommenen Instrumente angestrebt. Mit steigender Lautsprecheranzahl kommt es zunehmend zur Verfärbung durch Phantomschallquellen - die Abbildung von Direktschallquellen in diskret zugewiesenen Lautsprechern wird zunehmend schwieriger. Um dies zu vermeiden, sollte das Übersprechen zwischen Mikrofonen bei der Aufnahme möglichst gering sein.<sup>72</sup>

<sup>70</sup> Daele & Baelen 2012: 13

<sup>71</sup> vgl. Philipp 2012

<sup>72</sup> vgl. ebenda

### 3.1.2 Kanalübersprechen

Für gute räumliche Ergebnisse sollte der Diffusschallanteil möglichst unkorreliert und der Direktschallanteil klein sein. Dahingegen ist eine scharfe Abbildung zwischen zwei benachbarten Lautsprechern nur mit Signalen möglich, die wenig Laufzeitversatz und Pegelunterschied aufweisen. Kohärente Signale auf mehr als zwei Lautsprechern wirken sich zunehmend negativ auf die Lokalisationsschärfe und Klangfarbe aus, die auch stark von der Hörposition im Abhörraum abhängig ist. Die Gefahr von unerwünschtem Übersprechen (Crosstalk) zwischen den Mikrofonen steigt jedoch mit der Anzahl der Wiedergabekanäle. Übersprechen kann bei der Aufnahme durch große Laufzeiten – beispielsweise Groß-AB - oder durch optimierte Verfahren wie „OCT“ vermieden werden.<sup>73</sup>

### 3.1.3 Mischpult-Integration

Um Auro 3D kompatibles Audiomaterial abmischen zu können, sollte ein Mischpult in der Lage sein, Mehrkanal-Busse mit bis zu 14 Kanälen zu verwalten. Mit zwei speziellen Joysticks ausgestattet, kann das Mischpult Audiosignale innerhalb des dreidimensionalen Sound-Bereiches beliebig platzieren. Hersteller wie AMS-Neve produzieren mit ihren digitalen Mischpulten 88D und DFC Auro 3D fähige Konsolen. Im Broadcast-Bereich hat die Firma Lawo bereits erste Tests der 3D Audio Misch- und Monitoring-Integration in Ihrer MC<sup>2</sup>66-Serie durchgeführt und auf der NAB-Show 2016 vorgestellt.<sup>74</sup> Im Live-Bereich sind bisher noch keine einschlägigen Entwicklungen zu beobachten.



*Abbildung 17: AMS-Neve DFC Digital-Mixing-Konsole – 3D Panoramaregler  
Linke Hand: Bereich für das horizontale Panning der Audiosignale.  
Rechte Hand: Bereich für das vertikale Panning der Audiosignale*

<sup>73</sup> vgl. Theile & Wittek 2011; vgl. Philipp 2012; vgl. Nipkow 2011

<sup>74</sup> Lawo 2016

### 3.1.4 Auro 3D Authoring Tools

Auro Technologies bietet mit seinen Auro-3D Authoring Tools eine Reihe an Plug-ins, die sowohl dreidimensionales Panning, als auch dreidimensionale Up- und Downmixes sowie die entsprechende Encodierung ermöglichen. Im Hintergrund läuft dabei eine Auro-3D Host Application, die das Processing aller Audiostreams übernimmt.

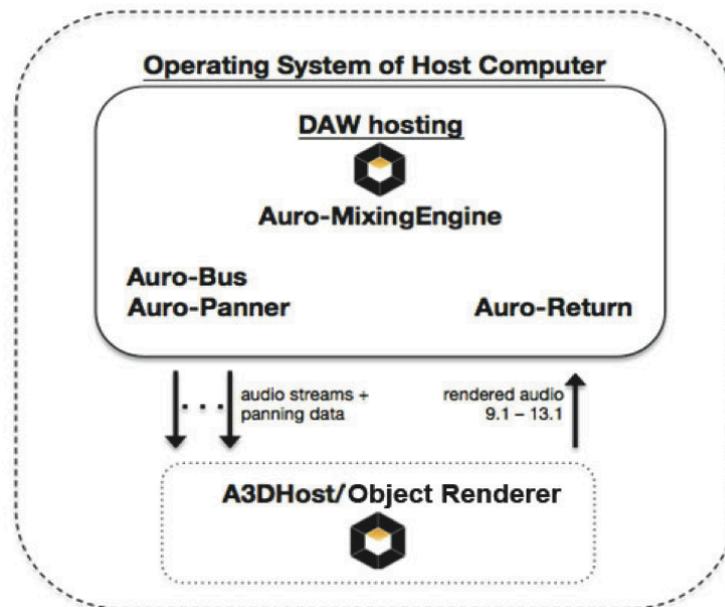


Abbildung 18: Auro 3D Blockdiagramm<sup>75</sup>

Das **Auro-Panner** Plug-in ermöglicht 3D-Panning und ersetzt dabei den DAW eigenen Panner, da die meisten DAWs zum gegenwärtigen Zeitpunkt noch kein dreidimensionales Panning unterstützen. Der Auro-Panner muss dabei in jeden Audio-Kanal, der im dreidimensionalen Raum bewegt werden soll inseriert werden. Die Informationen jedes einzelnen Auro-Panner werden im Anschluss zum Auro-Bus-Plug-in gesendet.

<sup>75</sup> Auro Technologies 2016: 26



Abbildung 19: Auro-Panner-Plug-in<sup>76</sup>

Das **Auro-Bus**-Plug-in sammelt alle Informationen der Auro-Panner auf den jeweiligen Audio-Spuren oder Sub-Gruppen. Alle verbundenen Auro-Panner können damit summiert und gelevelt werden. Im Weiteren besteht die Möglichkeit, Parameter für den automatischen Downmix vorzugeben.

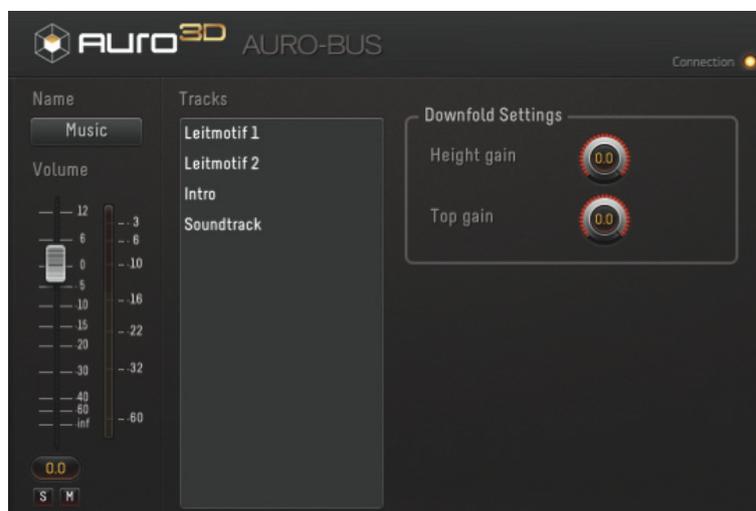


Abbildung 20: Auro-Bus-Plug-in<sup>77</sup>

<sup>76</sup> Auro Technologies 2016: 32

<sup>77</sup> Auro Technologies 2016: 41

In der Oberfläche des **Auro-Mix Engine**-Plug-ins wird zunächst die Auro-3D-Konfiguration vorgenommen (Auro 9.1, 11.1, 13.1). Weitere Mix-Parameter und Encodierungseinstellungen können festgelegt werden. Das Plug-in zeigt alle Pegel des Auro-3D-Mixes und gibt auf dem insertierten Kanal einen entsprechenden Downmix aus.



Abbildung 21: Auro-Mixing-Engine<sup>78</sup>



Abbildung 22: Auro-Return-Plug-in<sup>79</sup>

<sup>78</sup> Auro Technologies 2016: 12

<sup>79</sup> Auro Technologies 2016: 16

Das **Auro-Return**-Plug-in wird auf den DAW Auspielwegen inseriert. Es stellt der DAW die Kanäle des Auro 3D-Mixes zur Verfügung. Auf Grund der Kanal-Limitierung in der Bus-Struktur vieler DAWs müssen mehrere Multichannel-Tracks oder Einzelspuren erstellt werden, auf welchen das Auro-Return-Plug-in inseriert wird. Eine Möglichkeit dabei ist es in Pro Tools zwei 5.1 Spuren zu erstellen. Auf die erste wird das Auro-Return-Plug-in angewendet um den Höhen-Layer auszuspielen, auf den zweiten 5.1 Multitrack kann der untere Layer ausgespielt werden.

Über die **Auro-Settings**-Applikation können globale Einstellung, wie beispielsweise das Solo-Verhalten der Plug-ins oder hardware- und softwarespezifische Einstellungen bezüglich der verwendeten DAWs, vorgenommen werden. Eine sehr nützliche Funktion, die in der Plug-in-Version der Vor-Produktion leider noch nicht zur Verfügung stand, ist die Fader-Link Option. Diese gibt Solo-, Mute-, Lautstärke und XY-Panning-Informationen der Pro Tools eigenen Fader an die Auro-Plug-ins weiter. Eine Funktion, die das Arbeiten in großen Projekten wesentlich intuitiver gestaltet.

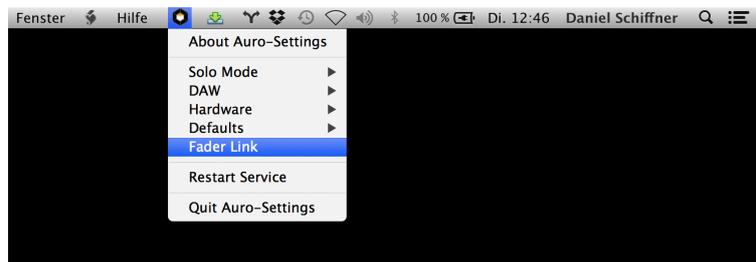


Abbildung 23: Auro-Settings - Fader Link<sup>80</sup>

Weitere Bestandteile der Auro Authoring Tools, wie Auro-Aux-Engine, Auro-DMix, Auro-Decoder, und Auro-Encoder, waren für diese Arbeit nicht relevant und werden nicht näher beschrieben.

---

<sup>80</sup> Screenshot

## **3.2 Iosono**

Iosono wurde 2004 als ein Spin Off des Fraunhofer Instituts gegründet und gehört seit 2014 zu Barco Audio Technologies. Iosono entwickelt vor allem 3D-Audio-Lösungen. Die Iosono-Technologie basiert auf dem Prinzip der Wellenfeldsynthese. Iosono verfolgt dabei den Ansatz, über den gesamten Wiedergaberaum einen idealen Klangeindruck zu erzeugen und diesen auch bei Verlassen des Sweet Spots aufrecht zu erhalten beziehungsweise den Sweet Spot zu minimieren. Über die Simulation einer Lautsprecherentfernung, die größer ist als es die realen Raummaße zulassen würden, wird der Sweet Spot (beim Iosono-System auch Sweet Area genannt) erweitert. Im Gegensatz zum kanalbasierten Auro-Format, bei welchem Lautsprecherkonfigurationen fest definiert sind, ist im objektbasierten Iosono-Format das Wiedergabesystem flexibel und kann aus nahezu beliebig vielen Lautsprechern bestehen. Bei entsprechender Konfiguration und einer diskreten Zuweisung von Audioobjekten zu einem Lautsprecherkanal können mit dem objektbasierten Iosono-Format auch kanalbasierte Multichannel-Formate bedient werden. Der Iosono Core kennt dabei die Maße der Hörfläche und die exakten Positionen jedes einzelnen Lautsprechers. Mit diesen Informationen und den Positionsdaten jedes einzelnen Audio-Objektes kann der Core das dreidimensionale Schallfeld für jeden im System vorgesehenen Lautsprecherkanal rendern und ausspielen.

### **3.2.1 Verteilung der Schallfelder**

Auch im Iosono-Format besteht die Aufgabe darin, Hörereignisse bestmöglich wiederzugeben. Eine gute Lokalisation und eine möglichst geringe Verfälschung der Klangfarbe der aufgenommenen Instrumente werden genau wie im Auro-Format angestrebt. Da es sich im Iosono-Format um eine Kombination unterschiedlicher Reproduktionsmethoden handelt, hängt die Verteilung des Schallfeldes stark vom gewählten Algorithmus ab. Wie im Auro-Format kommt es auch im Iosono-Format bei einfachen Reproduktionsmethoden über herkömmliche Panning-Algorithmen mit steigender Lautsprecheranzahl zur Verfärbung durch Phantomschallquellen. Auch hier sollte das Überschneiden zwischen den aufgenommenen Signalen möglichst gering sein um Verfärbungen während der Reproduktion zu vermeiden. In jedem Fall lässt sich die größtmögliche Immersion mit unkorrelierten Signalen erzeugen, da es hierbei zu keiner Klangverfärbung und Signalauslöschung kommt. Die Abbildung von Direktschallquellen in diskret zugewiesenen Lautsprechern hängt stark von deren Konfiguration ab. Eine Vielzahl an Lautsprechern ermöglicht beispielsweise das Zuweisen der Direktschallquellen zu einem jeweiligen diskreten Lautsprecher oder ermöglichen

über Isoono-Algorithmen eine hohe Ortungsstabilität. Untersuchungen haben gezeigt, dass sich bei Hörerpositionen innerhalb des Sweet Spots gute Ergebnisse in Bezug auf Lokalisierung mit dem integrierten VBAP Algorithmus erzielen lassen. Für nicht ideale Hörer-Positionen ermöglicht der Isoono-Algorithmus auch bei größerem Lautsprecherabstand vergleichbare Ergebnisse. In jedem Fall empfiehlt es sich je nach Anwendung unterschiedliche Reproduktionsmethoden auszuprobieren.<sup>81</sup>

### **3.2.2 Kanalübersprechen**

Kohärente Signale auf mehr als zwei Lautsprechern wirken sich auch im Isoono-Format zunehmend negativ auf die Lokalisationsschärfe und Klangfarbe aus. Die Abhängigkeit zur Hörposition verkleinert sich dabei wesentlich. Gute Hörergebnisse ohne eine gravierende Verfälschung in der Klangfarbe lassen sich auch bei einer Vielzahl an Audio-Objekten vor allem durch unkorrelierte Signale erzielen. Bei dreidimensionalen Aufnahme-Verfahren gelten auch bei der Reproduktion im Isoono dieselben Abhängigkeiten wie in *Kapitel 3.1.2* beschrieben.

### **3.2.3 Mischpult-Integration**

Ein Isoono kompatibles Mischpult muss in der Lage sein Mehrkanal-Busse mit der Anzahl der anzusteuernenden Lautsprecher verwalten zu können. Die Implementierung der Isoono-Algorithmen in das Mischpultsystem muss gegeben sein. Aktuell gibt es keine Live- beziehungsweise Hardwareconsole, die das Isoono-Format direkt unterstützt. In einer Live-Umgebung kann zum gegenwärtigen Zeitpunkt die Platzierung der Audio-Objekte im Schallfeld immer nur in Verbindung mit dem Isoono Core erfolgen. Dabei ist die Ansteuerung via iPad, PC oder anderen UDP-fähigen Programmoberflächen die das Isoono-Scene-Protocol implementiert haben, möglich.

---

<sup>81</sup> vgl. Melchior et al 2011

### 3.2.4 Iosono-Core

Das Iosono Core System besteht aus dem Iosono Core Processor, ein Linux-basierter Rechner, der in einem vier Höheneinheiten großen 19" Gehäuse integriert ist.



Abbildung 24: Iosono-Core<sup>82</sup>

Der Core verfügt wahlweise über Madi- oder AES-Audioschnittstellen. Eine Synchronisation ist über Wordclock oder Blackburst Video-Sync möglich. Über LTC-Timecode kann der Core mit anderen Medien und Systemen zeitlich synchronisiert werden. Ein Core kann maximal 128 Audio-Objekte gleichzeitig rendern und auf alle Ausgangskanäle ausgeben. Es besteht die Möglichkeit für eine größere Anzahl an Audio-Objekten und Ausgangskanälen Hardware-Erweiterungen an den Core anzubinden und zum Verbund zusammenzuschließen. Der Core hat Basisfunktionen, wie Lautstärkeregelung, Abrufen von Presets oder IP-Einstellungen, die über das integrierte Touchdisplay und den Lautstärkeregler an der Vorderseite des Gehäuses gesteuert werden können. Über die dazugehörige Remote Control Software können alle weiteren Funktionen, wie Lautsprecher-Setups, Processing-Einstellungen, I/O-Routings oder Filtereinstellungen, über eine TCP/IP Netzwerkverbindung konfiguriert werden.

---

<sup>82</sup> Iosono 2014: 14

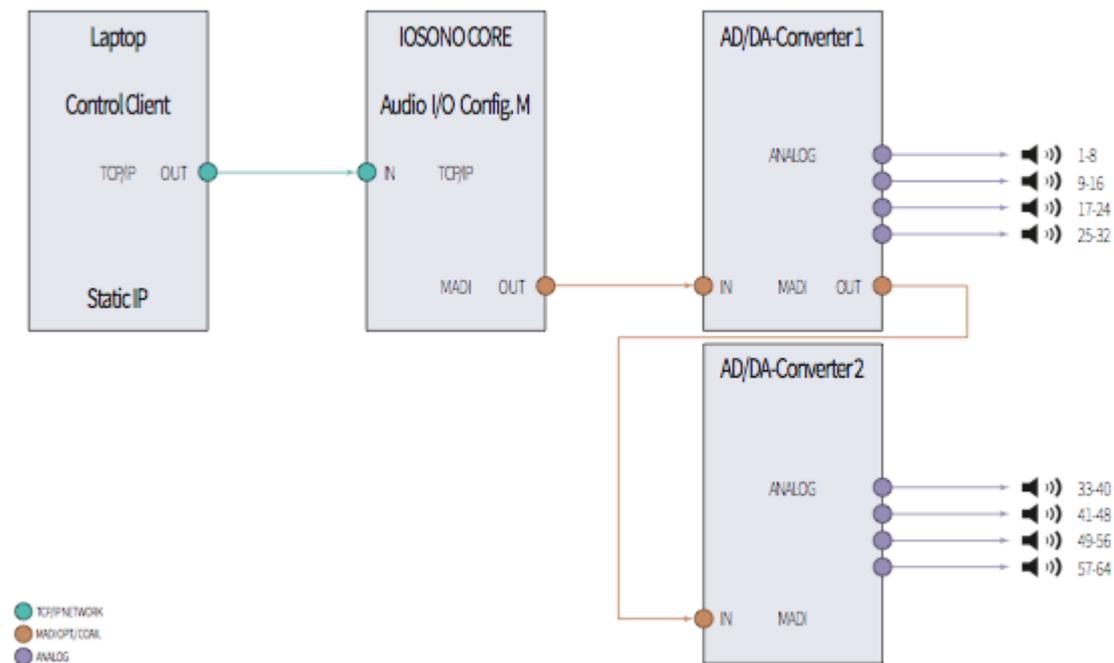


Abbildung 25: Standard-Verkabelung Iosono-Core für 64 Madi-Kanäle<sup>83</sup>

Das Erstellen und Steuern von Audioobjekten kann über den Core-eigenen Scene-Designer oder über die Spatial Audio Workstation als Plug-in in der Nuendo-DAW kontrolliert werden. Dabei werden alle szenischen Informationen der Audio-Objekte über eine UDP-Netzwerkverbindung übertragen.

### 3.2.5 Iosono Control Unit

Die Iosono Control Unit dient als zentrale Konfigurationsschnittstelle und Oberfläche für den Iosono Core. Nach dem Login können über die Oberfläche Lautsprecher-Setups, Processing-Einstellungen, I/O-Routings und Filterkonfigurationen vorgenommen werden. Alle festgelegten Konfigurationen können in Presets verwaltet werden. Um die Leistung zu optimieren, wird dabei im Preset immer nur das verwendete Softwaremodul gespeichert beziehungsweise geladen.

<sup>83</sup> Iosono 2014: 21

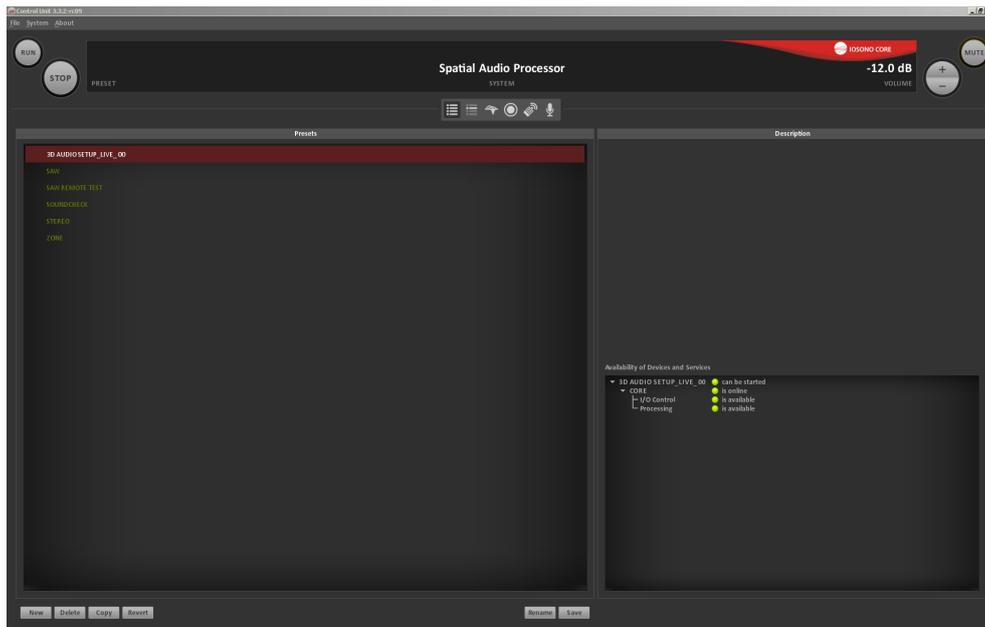


Abbildung 26: Iosono Control Unit<sup>84</sup>

Über die I/O Control Einstellungen wird Grundlegendes wie Synchronisation und die Art der Audio-Schnittstelle bestimmt. Der I/O Control Status bietet dabei eine hilfreiche Übersicht der ankommenden, prozessierten und ausgehenden Signale.

Abbildung 27: Iosono Control Unit - I/O Control<sup>85</sup>

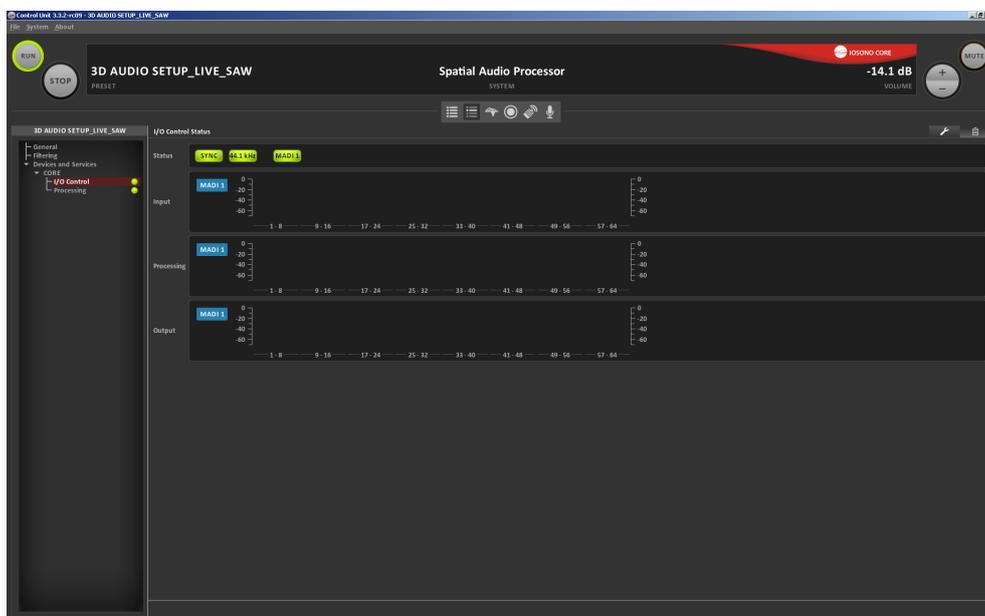


Abbildung 28: Iosono Control Unit - I/O Control<sup>86</sup>

<sup>84</sup> Screenshot

<sup>85</sup> Screenshot

Außerdem bietet der Core die Option, das System einzumessen und automatische Lautsprecherfilter zu erstellen und zu verwalten. Der Core hat einen Audio-Player und einen Testtongenerator integriert. Über den implementierten Scene-Designer können statische, objektbasierte Audioszenen erstellt werden. Ein schematischer Signalfluss der Core-Module soll diesen Punkt verdeutlichen.

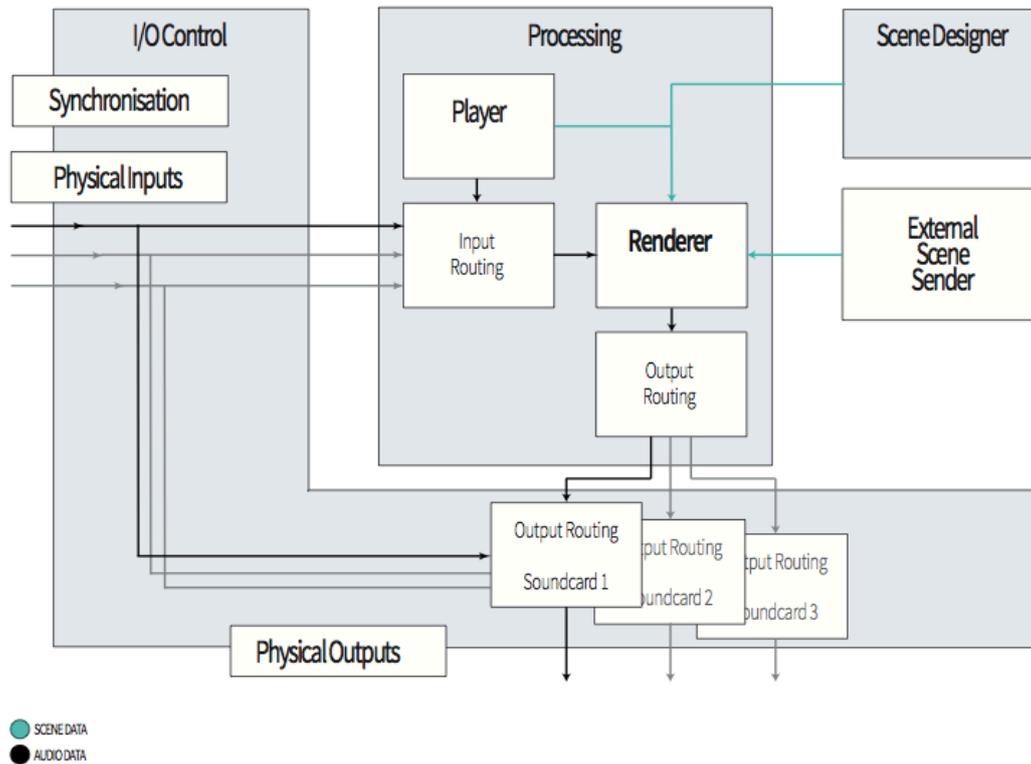


Abbildung 29: Iosono-Core - Signalfluss<sup>87</sup>

## Lautsprecher-Setup

Das im Lautsprecher-Setup erstellte Lautsprecher-File enthält die Anzahl, Art und exakte Position der im Setup verwendeten Lautsprecher. Zudem wird in dieser Oberfläche unter “Audience Area” die Größe und Höhe des Hörbereiches definiert. Jeder Layer wird als Lautsprechergruppe angelegt. Diese Gruppen können dabei quadratisch oder kreisförmig angelegt werden, zudem können Lautsprecher unabhängigen Layer zugewiesen oder als Subwoofer-Gruppe konfiguriert werden.

<sup>86</sup> Screenshot

<sup>87</sup> Iosono 2014: 44

Die genaue Positionierung erfolgt in zwei Schritten. Zunächst werden an jeder Seite (vorne, hinten, links, rechts) alle Lautsprecher gleichmäßig verteilt. Mit Platzhaltern (Left and Right Spacer) können die seitlichen Offsets bestimmt werden. Auch der Referenzpunkt, auf welchen alle Laufzeiten angepasst werden, kann hier festgelegt werden. Dieser liegt standardmäßig auf dem Nullpunkt des X,Y Koordinatensystems, besitzt aber die Höhe der Audience Area, was wiederum der Ohrhöhe entspricht.

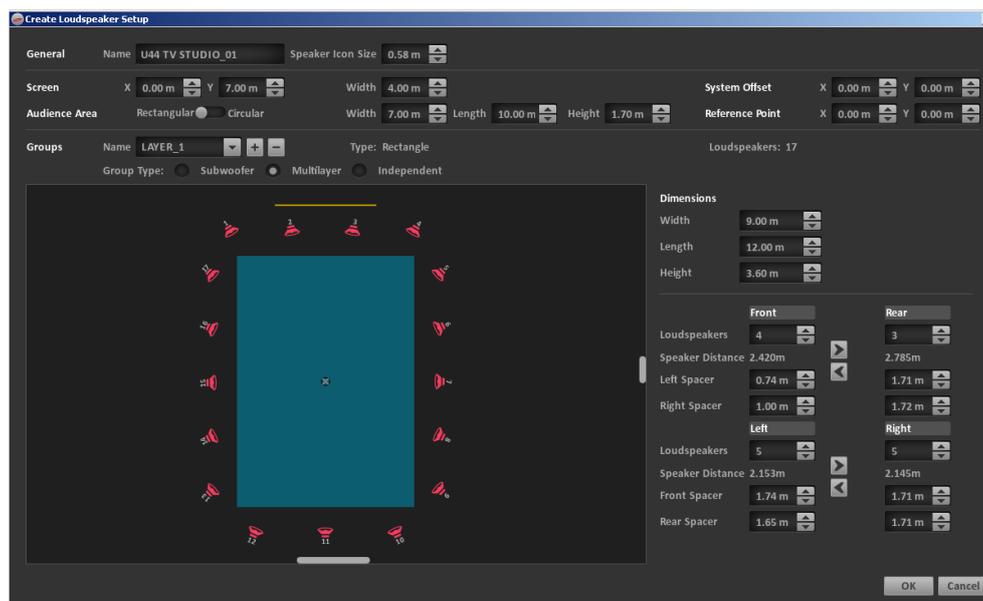


Abbildung 30: Iosono Loudspeaker-Setup<sup>88</sup>

In einem späteren Schritt kann über das Lautsprecher-Setup jeder einzelne Lautsprecher über die Eingabe von Polarkoordinaten oder kartesischer Koordinaten in seiner Position und Ausrichtung justiert werden.

<sup>88</sup> Screenshot

Je Lautsprecher-Gruppe werden die Lautsprecher im Uhrzeigersinn vorne links bei 1 startend automatisch durchgehend nummeriert. Für die Benennung der einzelnen Lautsprecher wählt das System den Namen der Gruppe mit dem Zusatz `_SPK001`, `_SPK002`, usw.

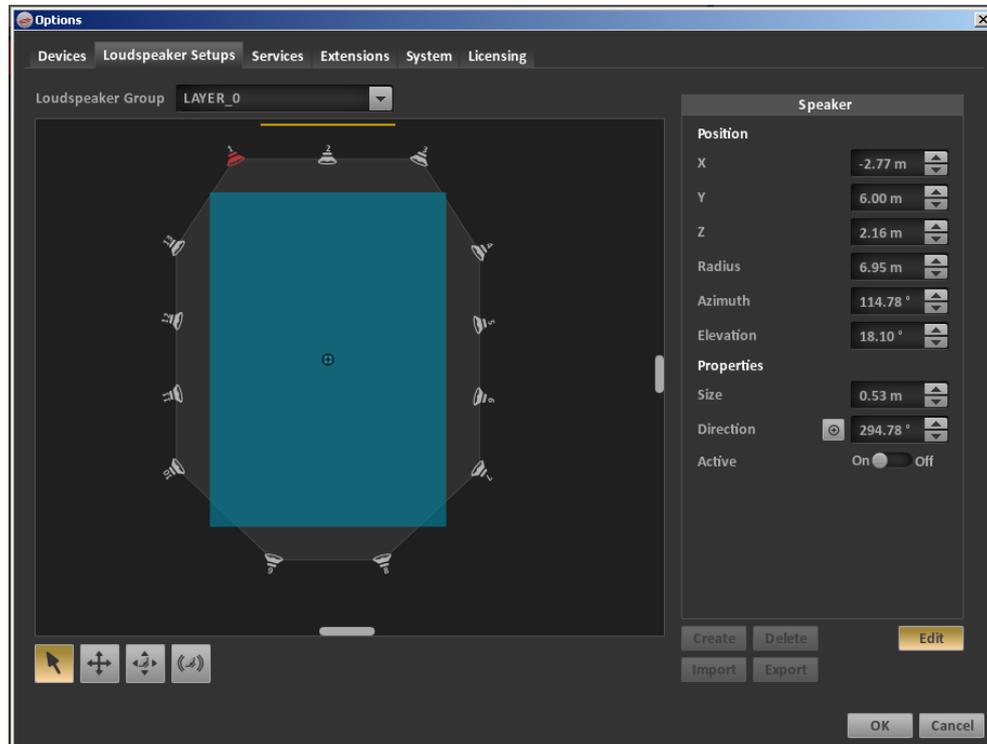


Abbildung 31: Iosono Loudspeaker-Setup<sup>89</sup>

## Processing Einstellungen

Nach dem Erstellen eines Lautsprecher-Setups gibt es die Möglichkeit über die Processing-Einstellungen eine Auto-Konfiguration vorzunehmen. Dabei erstellt der Prozessor pro Gruppe, in diesem Fall pro Lautsprecher-Layer, einen Render Slot mit entsprechender Layer-ID. Über die Layer-ID kann später das Panning in der Höhe adressiert werden. In dem jeweiligen Render-Slot wird je nach Lautsprecher-Anordnung und -Anzahl der passende Reproduktions-Algorithmus ausgewählt. Zur Auswahl steht der Iosono-Algorithmus, der mit einer ausreichenden Anzahl an Lautsprechern in einer Reihe, Wellenfeldsynthese zur Reproduktion verwendet. Bei einer geringen Anzahl an Lautsprechern, die eine Wellenfeldsynthese nicht zulässt kann man unter 2D- und 3D-Ceiling-Panning-Algorithmen wählen. Dabei entspricht 2D-Panning einem Intensitäts-Panning, wie

<sup>89</sup> Screenshot

man es aus Stereo-Setups kennt. Der 3D-Ceiling-Algorithmus arbeitet nach dem VBAP Prinzip und kann dabei mehr als zwei Lautsprecher gleichzeitig ansprechen.

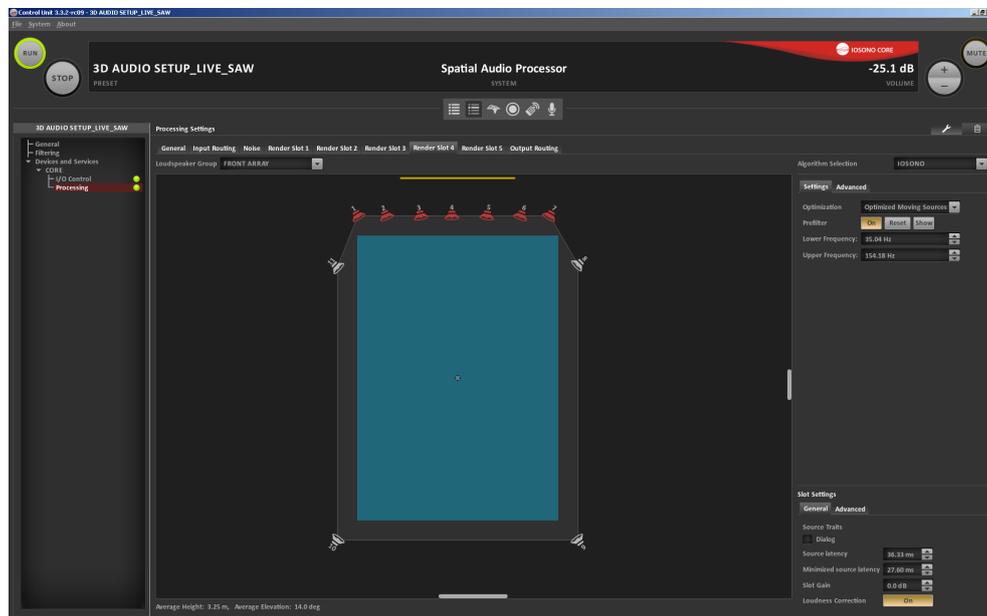


Abbildung 32: Iosono Control Unit - Processing Settings<sup>90</sup>

Da der Iosono-Algorithmus wesentlich rechenintensiver ist, kann man in den geringer bestückten Lautsprecher-Gruppen zur Schonung der Core-Ressourcen den Panning-Algorithmus wählen. Bei bewegten Audio-Objekten bietet der Iosono-Algorithmus eine Optimierungsfunktion, die für jedes Audio-Objekt eine Delay-Interpolation in den Reproduktions-Algorithmus einfließen lässt. Bei Line-Arrays entstehen durch dicht aneinander platzierte Lautsprecher Kopplungen im Bassbereich. Diese können über einen Prefilter im entsprechenden Render-Slot ausgeglichen werden. Im Render-Slot des Subwoofer können in den Einstellungen Lautsprecher-Extensions ausgewählt werden. Die nach den Layer-ID (Lautsprecher-Gruppen) benannten Check-Boxen ermöglichen bei einer Aktivierung, die Lautsprecher-Gruppe zu einem Fullrange-System zu erweitern, sofern keine Fullrange-Lautsprecher im System installiert sind. Das bedeutet, dass alle Bass-Anteile der im Layer platzierten Audio-Objekte ab einer Grenzfrequenz auf dem Subwoofer-Layer über einen 2D Panning-Algorithmus wiedergegeben werden. Die Zuweisung der physikalischen Audio-I/Os zu den entsprechenden Render-Slots erfolgt über die beiden Reiter „Input-Routing“ und „Output-Routing“.

<sup>90</sup> Screenshot

## Filter-Einstellungen

Filtereinstellungen für jeden einzelnen Lautsprecher können zum Einen ausschließlich für ein erstelltes Core-Preset, zum Anderen aber auch als globaler Filter im System für alle Presets erstellt und gesetzt werden. Dabei besteht die Wahl zwischen Filter mit unendlicher Impulsantwort (IIR) und Filter mit endlicher Impulsantwort (FIR), welche vor allem über die Einmess-Funktion des Systems automatisch generiert werden können. Die FIR-Filterliste kann nach der Einmessung des Systems als globale Filter-Konfiguration geladen werden und automatisch den entsprechenden Lautsprechern zugewiesen werden.

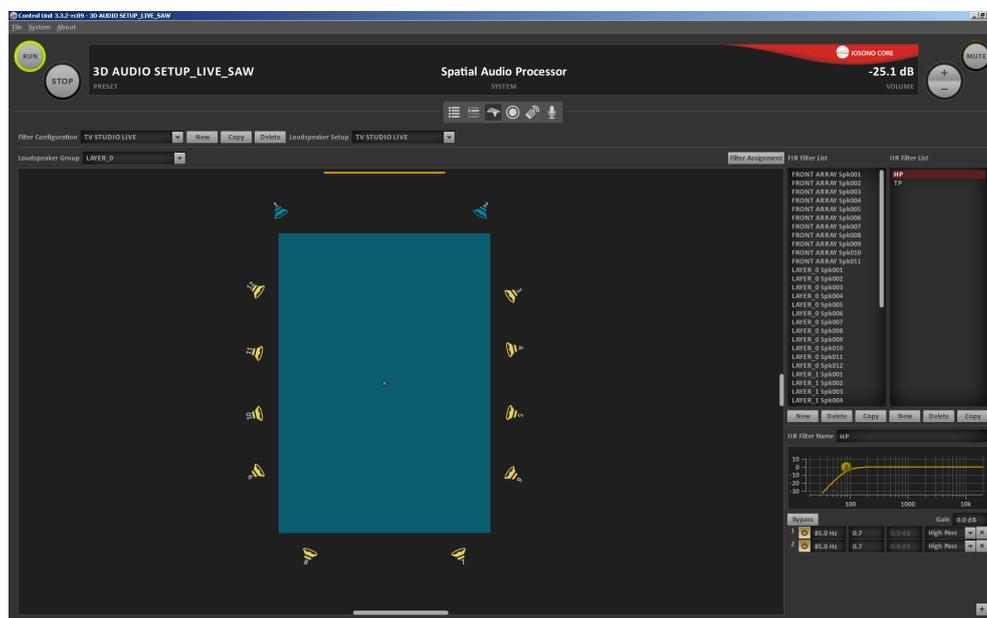


Abbildung 33: Iosono Control Unit – Filter Configuration<sup>91</sup>  
(Globale Filter-Einstellung: blau = aktiver IIR-Filter gelb = aktiver FIR-Filter)

## System Latenz und Auslastung

In den General-Settings des Processing-Moduls wird die Buffer-Size bestimmt, die Zeit, in der das Audio-Processing des Core durchgeführt werden soll. Dabei ist zu beachten, dass die eigentliche Gesamtlatenz aus der doppelten eingestellten Buffer-Size plus 32 Samples für die Berechnung der Rendering-Algorithmen besteht. Die Verdopplung der Buffer-Size ergibt sich durch das blockweise einlesen der eingestellten Samples, der 32 Sample Vorlaufzeit zur Berechnung des Render-Algorithmus und der blockweisen Ausgabe der gleichen Anzahl wie eingehender Samples.

<sup>91</sup> Screenshot

Ein Anhaltspunkt über die niedrigste Buffer-Size gibt die Processing-Status-Oberfläche. Sie gibt eine Auskunft über die prozentuale Auslastung der vier Processing-Kerne des Iosono Cores. Eine zu hohe Auslastung mit Peaks führt zu hörbaren Aussetzern beziehungsweise Stottern im Audio. Die Auslastung des Cores steht in direktem Zusammenhang mit der Anzahl an Lautsprechern im System und der Anzahl an Input-Quellen. Außerdem ist von Bedeutung welcher Algorithmus auf den jeweiligen Lautsprecher-Gruppen gewählt wurde. Ein Iosono-Algorithmus mit Delay-Interpolation nimmt dabei wesentlich mehr Rechenleistung in Anspruch als herkömmliche Panning-Algorithmen. Auch FIR-Filter, die für jeden Lautsprecher gefaltet werden, sind wesentlich aufwändiger zu berechnen als IIR Filter. Eine Pauschalisierung der Systemauslastung ist allerdings auf Grund der vielen Variablen im System und im Algorithmus nicht möglich. Es hilft nur der Blick auf den Processing-Status und ein gesundes Maß an Puffer in der Buffer-Size.

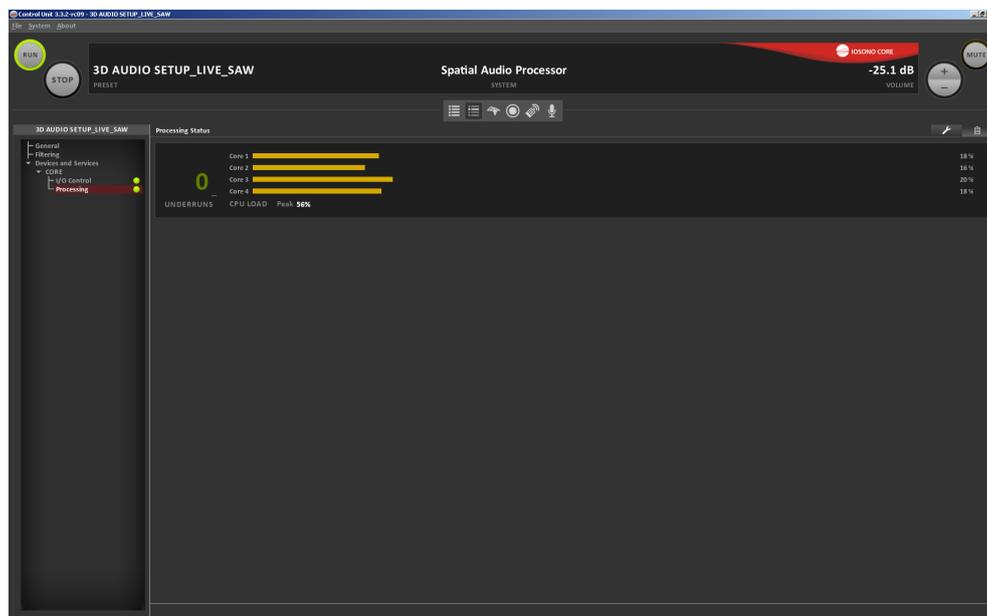


Abbildung 34: Iosono Control Unit - Processing Status<sup>92</sup>

Die Buffer-Size hat direkte Auswirkung auf die Quellenverzögerung. Das Iosono-System berechnet ein zeitgleiches Eintreffen aller Quellen in Bezug auf den in Kapitel "Lautsprecher-Setup" definierten Referenzpunkt. Neben der Buffer-Size hat bei der Berechnung der Delay-Werte vor allem die Position der Lautsprecher Einfluss auf den Wert.

<sup>92</sup> Screenshot

Im Detail setzt sich die Source-Latency aus folgenden Verzögerungen zusammen:

$$\text{Milliseconds}_{\text{Buffer}} + \text{Milliseconds}_{\text{Predelay}} + \text{Milliseconds}_{\text{Propagation}} + \text{Milliseconds}_{\text{Delay-Offset}}$$

Das Predelay dient zur Erzeugung fokussierter Quellen, die dafür Sorge tragen, dass ein Signal an einem bestimmten Lautsprecher früher als an einem anderen wiedergegeben wird, um die Wellenform entsprechend zu krümmen und den Schall auf einen Punkt zu fokussieren. Dieses Delay ergibt sich aus der Entfernung des am weitesten entfernten Lautsprechers und des nächstgelegenen. Das Predelay wird bei gleichmäßig im Kreis angeordneten Setups relativ klein, bei unförmig, beispielsweise langgezogenen Setups hingegen erhöht sich das Predelay. Hinzukommt im Weiteren das Propagation-Delay, das der reinen Schall-Laufzeit vom Lautsprecher bis hin zum Referenzpunkt entspricht. Um die spätere System-Latenz zu berechnen, müsste diese Laufzeit genau genommen abgezogen werden. Das Delay-Offset gleicht den Latenz-Unterschied in den einzelnen Render-Slots aus, damit die Signale mit schnellerem Panning-Algorithmus weiterhin zur gleichen Zeit im Referenzpunkt eintreffen, wie die Signale die mit langsameren Isono-Algorithmen berechnet werden. Das bedeutet, das System optimiert immer auf den langsamsten Render-Slot mit den am weitesten entfernten Lautsprechern.

### 3.2.6 Spatial Audio Workstation

Das Spatial Audio Workstation (SAW) Plug-in kann als Steuerschnittstelle für den Isono-Prozessor genutzt werden. Sie arbeitet mit der Software Nuendo als Basis-Host unter Windows. Um eine schnelle und effiziente Nutzung des Plug-ins zu ermöglichen, wird es tiefer in die DAW integriert als beispielsweise die bereits beschriebenen Auro 3D Authoring Tools. Der Unterschied liegt darin, dass die SAW auf die Busstruktur von Steinberg's Nuendo zugreifen darf und somit die eigentlichen Funktionen des Programms, wie beispielsweise Solo und Mute, erhalten bleiben. Mit Hilfe des Spatial Audio Workstation Plug-ins kann der Isono-Prozessor und damit das dreidimensionale Panning für alle dem Isono-Prozessor zugeführten Audiosignale ferngesteuert werden. Die SAW erstellt bei Aktivierung in Nuendo automatisch Ausgangskanäle, die daraufhin dem Isono-Prozessor zugeführt werden. Für jeden dieser Audiokanäle übermittelt das Plug-in dabei gleichzeitig Steuerdaten.

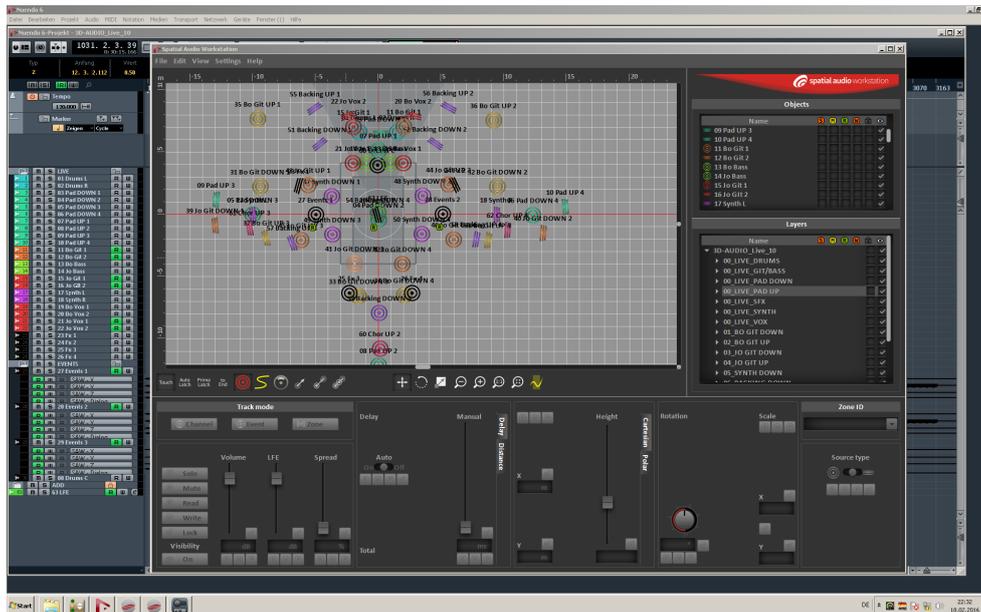


Abbildung 35: Spatial Audio Workstation Plug-in (SAW)<sup>93</sup>

Das SAW Plug-in kann in Nuendo geöffnet werden. Mit dem erstmaligen Öffnen des Plug-ins sind noch keine Objekte, also Soundquellen, erstellt. Die SAW zeigt lediglich das Lautsprecher-Setup. In der SAW gibt es zwei unterschiedliche Objektentypen:

- **Kanal-Objekte:** Quellen, die eine oder mehrere zusammengemischte Audioevents enthalten, aber aus einer Richtung des Hörbereichs kommen. Kanalobjekte sind virtuelle Lautsprecher.
- **Event-Objekte:** ein einzelnes Audioevent, das nur zu einer definierten Zeit existiert. Event-Objekte können gruppiert und somit gemeinsam bearbeitet werden.

Jedes angelegte Audio-Objekt kann entweder eine Punktschallquelle oder eine ebene Welle wiedergeben.

- **Punktschallquelle:** Eine Punktschallquelle wird sphärisch und ausgehend von einem Punkt im Raum abgestrahlt. Als Beispiel dient ein vorbeifahrendes Auto, welches als Punktschallquelle an jeder Position im Raum während des Vorbeifahrens mit einer

<sup>93</sup> Screenshot

Veränderung im Klang wahrgenommen wird. Die Punktschallquelle wird zudem von Zuhörern je nach Abhängigkeit ihrer Position im Raum unterschiedlich geortet.

- **Ebene Welle:** Im Gegensatz zur Punktschallquelle werden ebene Wellen für alle Zuhörer aus einer allgemeinen Richtung wahrgenommen. Eine ebene Welle kann als Punktschallquelle in unendlicher Entfernung vom Hörer betrachtet werden. Ebene Wellen sind nützlich um jeden Zuhörer das gleiche Klangbild eines Audio-Objektes zu vermitteln.

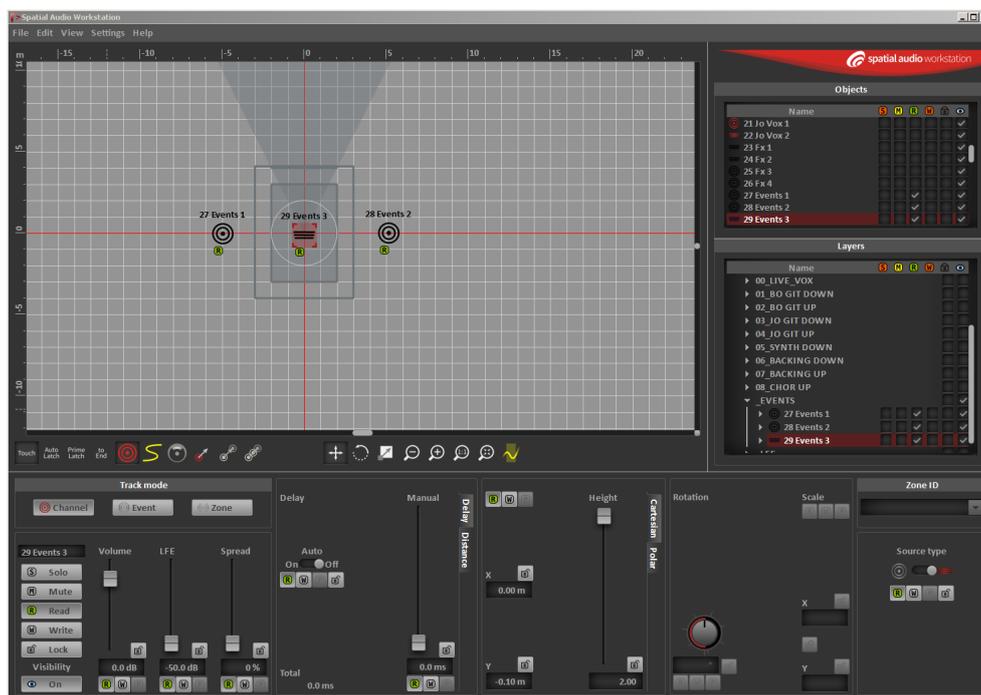


Abbildung 36: Spatial Audio Wokrstation - Punktschallquelle und Ebene Welle<sup>94</sup>

Über die SAW Oberfläche kann im rechten unteren Bereich schnell zwischen beiden Arten von Schallquellen umgeschaltet werden. Im rechten oberen Bereich kann im Objects-Fenster das gewünschte zu bearbeitende Objekt ausgewählt werden. Das Layer-Fenster dient der besseren Strukturierung einer Session. Im linken oberen Bereich der Oberfläche wird ein Überblick der sichtbaren Audio-Objekte und der aktuellen Szene angezeigt. Eine Darstellung der Automationsdaten über gelbe Verlaufslinien kann aktiviert werden, benötigt aber große Systemressourcen auf Grund der grafischen Implementierung. Die Session wird dadurch sehr träge und benötigt viel Zeit beim Kopieren von Automationsdaten und Bearbeitungen in der SAW.

<sup>94</sup> Screenshot

Im unteren, linken Bereich der Oberfläche wird die Lautstärke, der LFE-Anteil und die Verteilung auf umliegende Lautsprecher gesteuert. Im Fenster daneben besteht die Möglichkeit, die automatische Verzögerung in Abhängigkeit zur Audio-Objekt-Position zu deaktivieren und manuell zu steuern. Die mit Abstand wichtigste Funktion der SAW Oberfläche befindet sich im unteren mittleren Bereich. In diesem Fenster können die Audio-Objekte entweder über Polarkoordinaten oder über das kartesische Koordinatensystem in der Lautsprecher-Konfiguration bewegt werden. Im späteren Verlauf wird vor allem die Positionierung über kartesische Koordinaten genutzt. Dabei kann die Zuweisung eines Audio-Objektes auf einen bestimmten Höhen-Layer, wesentlich vereinfacht, über den Höhenfader, dessen Position der Layer-IDs entspricht, erfolgen.

Die SAW kann Mono- und Stereokanäle aus Nuendo verarbeiten. Um ein Objekt in der SAW anzulegen, muss in Nuendo eine Audiospur erzeugt werden. Der Ausgang jeder erzeugten Audiospur wird auf den vorher erstellten Iosono-Internal-Mix Bus geroutet. Ebenso wird der Mixer jeder erstellten Audiospur auf den Iosono-Spatial-Panner umgestellt. Mit diesen beiden Einstellungen wird ein Objekt in der SAW erzeugt. Die Audiokanäle werden als Kanalobjekte angelegt, können aber beliebig in Eventobjekte umgewandelt werden. Die Objekte können innerhalb und außerhalb des Lautsprecherarrays positioniert werden. Trotz der erzeugten Audiospuren, liegt an Nuendo kein Audiomaterial an. Das Audiomaterial wird während der Liveproduktion direkt von Pro Tools an den Iosono-Prozessor weitergeleitet. In Nuendo und der SAW werden lediglich die Steuerdaten des Iosono-Prozessors verwaltet. Damit kann unnötiges Routing vermieden werden, sowie Prozessorleistung des Nuendosystems, welches in Zusammenhang mit der SAW zum gegenwärtigen Zeitpunkt nur als 32 Bit System aufgesetzt werden kann und somit bei einer Vielzahl an Audio-Objekten und Automationsdaten schnell an seine Leistungsgrenzen kommen würde.

## 4 Der Selbstversuch: Vorbereitung des 3D-Konzerts

3D-Audio ist in Bezug auf Mischung und Produktion noch nicht allzu sehr vertreten und man findet wenig Literatur über die Handhabung und das Mischverfahren in einer Produktion populärer Musik. Für die Recherche der Mischung werden neben Beispielen von gewöhnlichem 2 Kanal-Stereo Mischungsverfahren, Surround-Produktionen herangezogen. Sie bilden eine eventuelle Grundlage für den Umgang mit einer dreidimensionalen Mischung.

Für das Systemdesign einer dreidimensionalen Beschallungsanlage und Produktionsumgebung gibt es eine Vielzahl an unterschiedlichen Ansätzen und Möglichkeiten. In der Praxis gibt es jedoch erst wenige Umsetzungen. Für die Recherche bezüglich der Gestaltung einer Produktionsumgebung wurden folgende Produktionen und Installationen be- und untersucht.

Die deutschen Pioniere elektronischer Musik, Kraftwerk, setzten für ihre "Catalogue 12345678"-Tour ein Isoono-System ein. Mit 24 Ring-Lautsprechern, 7 Front-Lautsprechern, 4 Front-Fills und 9 Subwoofern deckten sie bei der Gestaltung der Beschallungsanlage jedoch nur einen Layer ab. Ein Besuch des Konzertes am 30. November 2015 in der Stuttgarter Liederhalle war sehr hilfreich, um die Grenzen eines Systems dieser Größenordnung besser beurteilen und einordnen zu können. Bezüglich der Mischung resultiert die Erkenntnis darin, dass der Umgang mit einzelnen Quellen sehr ausgewogen sein sollte, damit das Schallfeld nicht kippt und der Zuhörer aus seiner Umhüllung herausgezogen wird. Auch unterschiedliche Positionen im Zuschauerraum sorgten für sehr unterschiedliche Hörerlebnisse, in der Sweet Area sehr ausgewogen, an den Randbereichen aber in Abhängigkeit der Mischung doch etwas seitenlastig. Das Konzert stellte eine Erfahrung dar, die im Umgang mit der Sweet Area in Abhängigkeit zur Mischung weitergeholfen hat.

Ein Blick auf das technische Design der Produktion lieferte folgende Erkenntnis für den Entwurf einer Produktionsumgebung:

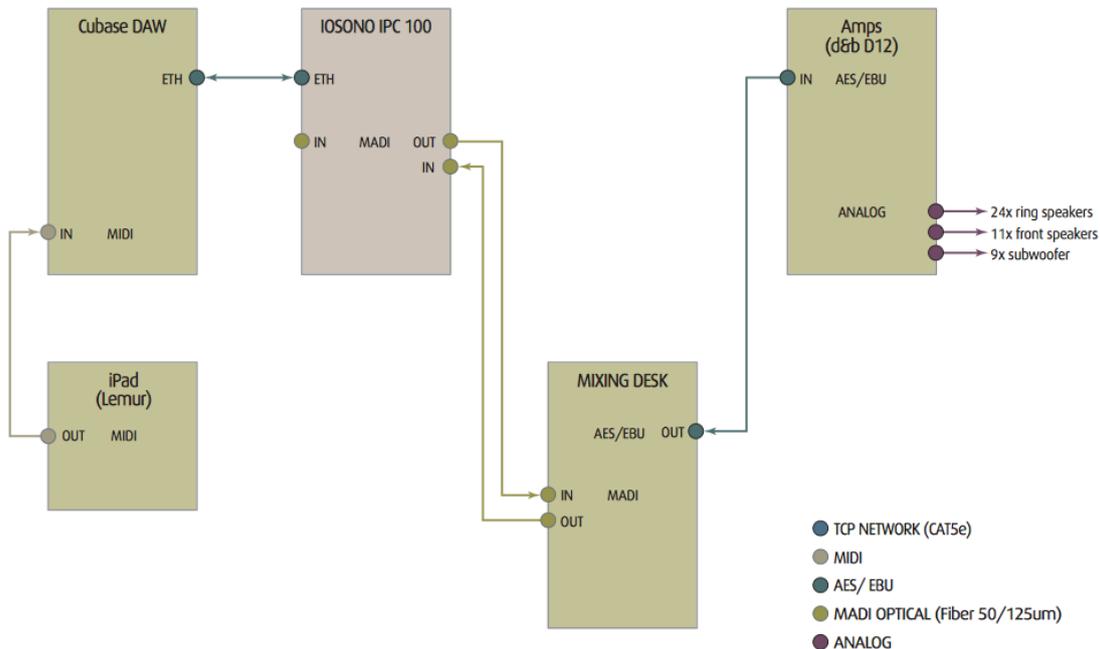


Abbildung 37: Signalfluss - Kraftwerk Livekonzert<sup>95</sup>

In jedem Fall wird eine Mischpult-Oberfläche benötigt um live-kritische Signale, wie Vocals, Gitarre oder Drums, während des Live-Konzertes im direkten Zugriff zu haben. Im Weiteren wird ein intuitiver Zugriff auf die Positionen und Echtzeitbearbeitung der Live-Signale und Audio-Objekte benötigt. Das Systemdesign muss skalierbar sein um eine höhere Anzahl an Lautsprecherkanälen zu verwalten.

Im Rahmen der Vorbereitungen wurde auch eine Iosono-Installation für das Fohn Soundlab in Reutlingen besucht und im Detail betrachtet. Das Soundlab, welches vor allem als Messraum eigener Produkte aber auch für dreidimensionale Audiokunst dient, besteht neben dem Iosono-Prozessor aus einer Installation von 24 Ring-Lautsprechern und weiteren 18 Decken-Lautsprechern. Es handelt sich damit um ein Lautsprecher-Setup, welches der gewünschten Intention von einem dreidimensionalen Live-Setup etwas näher kam, da es sich bereits auf zwei Höhen-Layer ausweitet.

<sup>95</sup> Iosono 2015b:

Ein Blick auf das Flussdiagramm und der Besuch unterschiedlicher Demonstrationen sowie Klanginstallationen ergab folgende Erkenntnis:

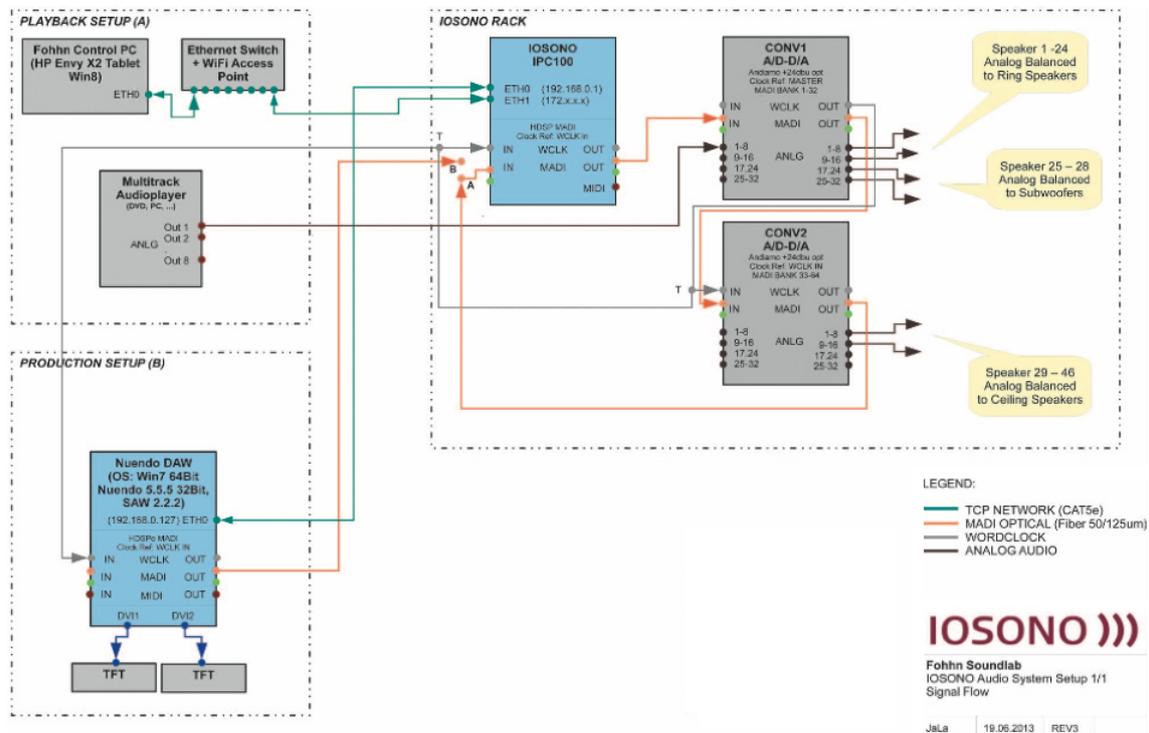


Abbildung 38: Signalfluss - Fohhn Soundlab<sup>96</sup>

Die Umhüllung über weitere Höhen-Layer ist ein wichtiges Element für die Klanggestaltung. Komplikationen mit der Nuendo DAW auf Grund der Instabilität des 32Bit-Systems führten zu dem Entschluss das Nuendo-System als reine Remote-DAW ohne anliegendes Audio einzusetzen. Die Verteilung der Audiosignale via Madi schafft eine hohe Flexibilität. Nach der beispielhaften Betrachtung der beschriebenen 3D-Audio-Systeme gilt es in Bezug auf das technische System-Design der Produktionsumgebung folgende Punkte zu erarbeiten. (siehe Kapitel 4.3.2.)

- Digitale Live-Mixing-Console vs. Echtzeitbearbeitung in Pro Tools
- Latenzfreies Monitoring für die Band
- Audio-Matrix für Signal-Distribution
- Vorverstärkung und Wandlung der Bühnensignale
- Wandlung der diskreten Lautsprecher-Signale

<sup>96</sup> Professional System 2013: 4

## 4.1 Mischung

### 4.1.1 Allgemeines

Im Duden werden folgende Synonyme zum Begriff 'mischen' aufgeführt: zusammenschütten; (umgangssprachlich) vermischen, verpanschen, durcheinanderbringen, zusammenwürfeln.

Das Mischen von Musik ist ein Prozess, der kein 'unkontrolliertes' Durcheinander schaffen, sondern der 'Geräusche' zu einem einheitlichen Klangbild 'vereinigen'<sup>97</sup> soll.

*„Ich setze alles in meiner Macht stehende daran, nicht unbedingt um ein musikalisches Event unverändert zu reproduzieren, aber um entweder mein Konzept der Musik oder das des Künstlers in der besten Art zu verwirklichen, die mir möglich ist.“ (Bruce Swedien)<sup>98</sup>*

Die Mischung in der Popmusik spiegelt nicht zwangsweise möglichst realistisch und unverfälscht ein musikalisches Ereignis wider. Sie soll das musikalische Ereignis unterstützen und dabei die emotionale Wirkung der Musik verstärken. Nichtsdestotrotz steht im Mittelpunkt der Mischung die Musik: die Mischung dient ihr.

Mischen ist ein kreativer Prozess. In der Mischung populärer Musik gibt es keine direkten Vorgaben und Regeln, die einzuhalten sind. Es werden auf Grund der Abhörsysteme bestimmte Grenzen gesetzt, die einen entsprechenden Rahmen bilden (beispielsweise die Kanalanzahl und die damit verbundenen Möglichkeiten der Platzierung des Signals im möglichen Panorama). Ein weiterer Rahmen wird durch Hörgewohnheiten gesetzt. Die Zuhörer werden durch die von ihnen gehörte Musik geprägt und geschult. Ein Einhalten der entsprechenden Vorgaben und Einbeziehen von Hörgewohnheiten kann somit von Vorteil sein. Alles Weitere liegt im Ermessen des Produzierenden.

Mischungen zu erstellen ist ein technischer sowie kreativer Prozess. Es werden grundsätzlich ästhetische und technisch funktionale Entscheidungen getroffen. Die Wirkung der Audibearbeitung kann grundsätzlich in physikalischer und psychologischer Ebene unterschieden werden. Durch eine Audibearbeitung wird das Audiomaterial physikalisch verändert, beispielsweise ändern sich Amplituden- und Zeitwerte. Der dadurch neu entstehende Klang des Audiomaterials wird als reiner Wahrnehmungsinhalt in den grundlegenden Merkmalen wie

---

<sup>97</sup> Duden 2016

<sup>98</sup> Owsinski 2007: 311

beispielsweise Lautstärke, Klangfarbe und Raumeindruck psychologisch repräsentiert. Darüber hinaus beeinflussen zahlreiche weitere psychischen Verarbeitungsstufen die Wahrnehmung wie beispielsweise die Aufmerksamkeit, Bedeutung, Bewertung, Erinnerungen und Emotionen.<sup>99</sup>

Es stehen folgende Mittel zur Bearbeitung der technisch kreativen Klanggestaltung zur Verfügung:

- Lautstärkeverhältnisse
- Panorama
- dynamische Klangbearbeitung
- frequentielle Klangbearbeitung
- räumliche Klangbearbeitung
- harmonische Klangbearbeitung

In der Mischung werden die für nötig befundenen Bearbeitungen getätigt, um die für die Musik aus Sicht der Produzierenden passende Ästhetik und Wirkung zu erreichen. Wie die Musik wirken soll wird durch die Musiker und die Produktion festgelegt sowie durch die Mischung erweitert und verstärkt. Eine Musikmischung ist interpretatorisch, es fließen künstlerische, ästhetische und technische Entscheidungen Dritter ein.<sup>100</sup>

*„The mixes should be a good sonic representation of the musical style and work well for the intended medium [...]. They should be well balanced, clear, and sound good in a variety of settings and on different speakers.“<sup>101</sup>*

Musik ist Emotionsträger. Unterschiedliche Gestaltungsmittel können den Höreindruck verändern und damit Einfluss auf die Stimmung des Zuhörers nehmen. Ob eine gestalterische Mischung die beabsichtigte Wirkung erzeugt, lässt sich nicht planen. Es kann nur mit Hilfe von durch Wahrnehmung geprägter Gestaltungsmittel die Wahrscheinlichkeit dafür erhöht werden.<sup>102</sup>

---

<sup>99</sup> vgl. Weinzierl 2008: 776

<sup>100</sup> vgl. Weinzierl 2008: 775

<sup>101</sup> Mixerman 2010: 143

<sup>102</sup> vgl. Mechthild 2003: 76

Bei der technischen beziehungsweise medialen Übertragung von akustischen Darbietungen kommt es nach Weinzierl zur mehrfachen Transformation, die das Erleben von Musik beeinflusst:

- *„akustisch (Original- und reproduziertes Schallfeld sind nicht identisch)*
- *klanglich (das auditive Perzept ist ein anderes)*
- *interpretatorisch (in die elektroakustische Übertragung fließen künstlerische, ästhetische und technische Entscheidungen Dritter ein)*
- *räumlich (die Wiedergabe erfolgt in einem anderen Raum)*
- *zeitlich (meist werden Aufzeichnungen gehört)*
- *kontextuell (akustische, optische, motivationale, emotionale, kognitive und soziale Bedingungen der Rezeption sind verändert)<sup>103</sup>*

Weinzierl sagt, dass die Bearbeitung des Audiomaterials einen wesentlichen Einfluss auf das Erleben von Musik nimmt, da sie sich „auf den technisch manifesten Audioinhalt auswirkt und daher für seine weitere Übertragung weitgehend Bestand hat, während hingegen Übertragungs- und Rezeptionsbedingungen [!] variieren können.“<sup>104</sup> Nach Weinzierl kann Einfluss auf die substantiellen Eigenschaften bei Musikaufnahmen durch klangliche Bearbeitungsmaßnahmen genommen werden und damit eine klangbeschreibende, ästhetische, emotionale und strukturbeschreibende Änderung erzeugt werden.<sup>105</sup>

Ein Bewusstsein für die räumliche Transformation und der damit verbundene Einfluss auf das Erleben der Musik in der Mischung ist von Bedeutung. Das Endprodukt soll auf möglichst vielen gleichartigen Systemen die vorgesehene Ästhetik wiedergeben. Ein System wird durch die Kanalanzahl, sowie deren Anordnung und Wiedergabeart beschrieben. Ändert sich das Abhörsystem, so stimmen der gestalterische Ausgangspunkt und das Abhörsystem nicht überein und stehen in einem Konflikt. Upmixalgorithmen werden nicht mit einbezogen, da sie keine gestalterische Kontrolle bieten.

---

<sup>103</sup> Weinzierl 2008: 775

<sup>104</sup> Weinzierl 2008: 775

<sup>105</sup> vgl. Weinzierl 2008: 775

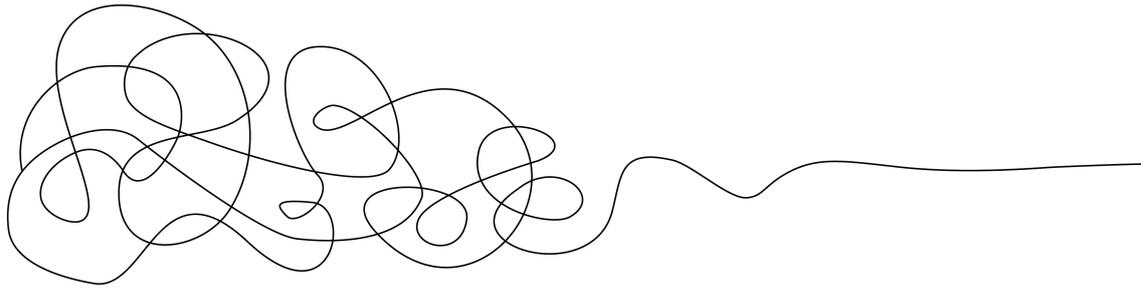


Abbildung 39: Fuzzy-Front-End - Designprozess <sup>106</sup>

Diese Darstellung des Designprozesses kann auf den Mischprozess übertragen werden und soll diesen graphisch verdeutlichen. In der Mischung kann man den technischen und den ästhetischen Teil nicht voneinander abgrenzen. Das Mischen ist ein stetig heterogener und iterativer Prozess, in dem sich beide Teile der Mischung gegenseitig beeinflussen.

*„When one part comes up, other parts need to go down. Everything has a relationship in mixing. That's what makes it such an art.“<sup>107</sup>*

Ein Beispiel hierfür ist ein Lowcut Filter, der die für das Signal irrelevanten Tiefenanteile dämpft, die technisch gesehen unerwünscht sind (bspw. Trittschall). Sie verändern den Klang des Instrumentes nicht offensichtlich, haben aber Einfluss auf die Wahrnehmung des Signals im Kontext der Mischung und führen zu einem ästhetisch hörbaren Unterschied.

*„The mixing and recording of popular music for threedimensional presentation is in its infancy.“<sup>108</sup>*

Es gibt aktuell nur wenig dreidimensionale Mischungen im Bereich der Popmusik und damit wenig Erfahrung in der Postproduktion. Die vorliegende Arbeit setzt sich auch mit den Möglichkeiten der Postproduktion einer dreidimensionalen Mischung mit dreidimensionalen Panning Tools sowie gewöhnlichen Werkzeugen einer Stereo-Produktion auseinander. Das im Rahmen des Versuchs genutzte Processing und die eingesetzten Plug-ins sind herkömmliche Werkzeuge, entwickelt zur Bearbeitung von mono- und stereophonen Quellen. In der Mischung wurde im Handel erhältlicher Hall, Delay und weitere Processing-Plug-ins verwendet. Die einzigen zum Einsatz kommenden Mehrkanalwerkzeuge sind Panning Tools der Firmen Auro Technologies und IoSono.

---

<sup>106</sup> modifiziert nach der Fuzzy-Front-End-Darstellung in Sanders 2013: 68

<sup>107</sup> Mixerman 2010: 149

<sup>108</sup> Martin 2015: 6

Die Arbeit versucht unter anderem zu beleuchten, in wie weit eine dreidimensionale Mischung mit konventionellen Stereo-Werkzeugen, ausgenommen dem Panner, erstellt werden kann, die dem Endformat auf eine positive Weise dient. Dabei stehen Immersion und Dreidimensionalität sowie Arbeitsfluss im Mittelpunkt.

#### 4.1.2 Grundlagen Wahrnehmung

*"Unter Wahrnehmung versteht man jene Prozesse, die eintreffende Informationen von den Sinnesorganen auswählen, ordnen und interpretieren."<sup>109</sup>*

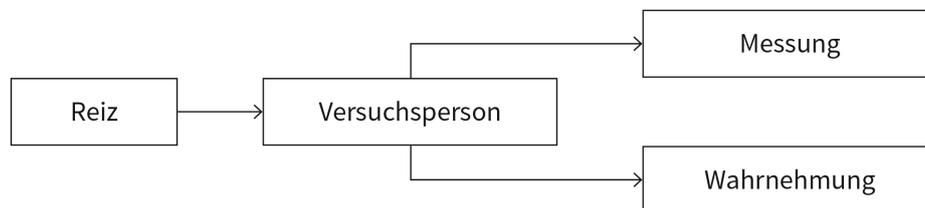


Abbildung 40: Messung und Wahrnehmung<sup>110</sup>

Messung und Wahrnehmung sind zwei unterschiedliche Bereiche des Hörens. Das Hören kann nicht auf eine physikalische Größe reduziert werden. Die psychische Größen spielen hierbei eine nicht zu vernachlässigende Rolle. Dabei steht das Hören immer im Zusammenhang mit dem Individuum.<sup>111</sup>

*„The aim should rather be to achieve a high accordance of perceptually relevant attributes between original and virtual sound fields. This can be fundamentally different from a high accordance of physical parameters. There may, for instance, be physical differences between original and virtual sound fields that do not have any perceptual consequences.“<sup>112</sup>*

Die Entscheidungen in der Mischung basieren auf der Wahrnehmung und Empfindung der Protagonisten. Die Wahrnehmung steht im Fokus der Mischung. Wahrnehmung ist subjektiv, da

---

<sup>109</sup> Mietzel 1998: 128

<sup>110</sup> Eigene Darstellung, basierend auf: Zymara 2002

<sup>111</sup> vgl. Zymara 2002: 6

<sup>112</sup> vgl. Wittek 2007: 6

jegliche wahrgenommenen Daten von einem individuellen Hintergrund verarbeitet und interpretiert werden.<sup>113</sup>

*„Die Wahrnehmung vermittelt keine objektive Wirklichkeit, sondern eine subjektive Welt; wir nehmen das wahr, was unseren Bedürfnissen, Erfahrungen, Erwartungen entspricht und nicht die objektiv gegebenen Reize.“<sup>114</sup>*

Das Wiedererkennen von Eigenschaften eines Objektes kann für ein schnelles Zurechtfinden hilfreich sein. Ist dem Rezipienten ein Objekt als 'bekannt' abgespeichert, so kann er das Objekt mit seinen bereits bekannten Daten abgleichen. Jeder Mensch besitzt eine individuell sowie kulturell geprägte Wahrnehmung.<sup>115</sup> Zu eben diesem Wiedererkennen zählen auch Hörgewohnheiten. Hörmuster werden implizit oder explizit aus der dem Rezipienten gebotenen Umwelt gelernt. Neben den individuellen Hörmustern bauen sich kulturelle Hörerfahrungen auf.<sup>116</sup> Auf eben diese kulturell gebildeten Hörerfahrungen soll in der Mischung Rücksicht genommen werden.

#### **4.1.3 Signal-Bearbeitung als stilistisches Mittel**

Die Aufnahme und Audibearbeitung von Musik ist nicht nur ein technischer und ästhetischer, sondern vor allem auch ein musikalischer Prozess.<sup>117</sup> Man kann grundsätzlich zwischen drei unterschiedlichen Transformationsprozessen unterscheiden.

- positivistisches Ideal: die physikalisch korrekte Nachbildung und damit eine möglichst geringe Manipulation
- illusionistisches Ideal: Wiederherstellung des psychischen Erlebens und damit einer Illusion mit Hilfe von wohl dosierter Überhöhung der klanglichen Merkmale
- medial-autonomes Ideal: eine freie Interpretation der Darbietung

Eine medial-autonome Darstellung entsteht nicht vor dem Medium, sondern im Medium. Der klanggestalterischen Manipulation sind keine Grenzen gesetzt.<sup>118</sup>

---

<sup>113</sup> vgl. Plassmann

<sup>114</sup> Hobmair 1997: 91

<sup>115</sup> vgl. Plassmann

<sup>116</sup> vgl. Mechthild 2003: 85

<sup>117</sup> vgl. Schlemm 1997

<sup>118</sup> vgl. Dickreiter 2014: 415

Klangbearbeitungsprozessoren, wie beispielsweise Equalizer und Kompressoren können für technische Bearbeitungen und klanggestalterischen Formung genutzt werden. Die technischen, sowie die daraus resultierenden klanglichen Eigenschaften des jeweiligen Effektes müssen dem Mischenden für einen bewussten Umgang mit diesen geläufig sein. Neben der grundlegenden Bearbeitungs-idee des jeweiligen Gerätes sind die durch das Gerät produzierbaren Modulationsartefakte für Klangfarbenänderungen am Signal interessant.

*„Ausnahmen sind nicht immer Bestätigung der alten Regel; sie können auch Vorboten einer neuen Regel sein.“<sup>119</sup>*

Modulationsartefakte können erzeugt werden indem beispielsweise ein Regelverstärker in einer Weise arbeitet, bei der der Zuhörer die Bewegung des Regelverstärkers in und aus der Pegelreduktion bewusst wahrnimmt. Dabei wird das technische zum gestalterischen Mittel, zum Instrument und zur Klangfarbenbearbeitung eingesetzt.

*„Es gibt technische und gestalterische Gründe für die Veränderung des Frequenzganges beziehungsweise der Klangfarbe einer Aufnahme.“<sup>120</sup>*

---

<sup>119</sup> Ebner-Eschenbach 1893: 43

<sup>120</sup> vgl. Dickreiter 1997: 351

## Lautstärkeverhältnisse

*„The level faders are the most basic tool for balancing mixes. The process of adjusting the level faders to achieve a satisfactory balance is called leveling.“<sup>121</sup>*

Mit der Änderung des Pegels wird die Amplitude eines Signals frequenzneutral variiert. Die Amplitude eines Signals bestimmt die Lautstärke des Schallereignisses. Steht das Signal in Abhängigkeit zu anderen Signalen so beeinflusst die Signalamplitude auch wahrgenommene Entfernung und Größe der Schallquelle sowie auf Grund der Kurven gleicher Lautstärke und der Verdeckungseffekte ihre Klangfarbe. Mit Hilfe des Pegelstellers kann eine korrekte Aussteuerung sowie eine „künstlerisch-technische Klangregie“<sup>122</sup> erzeugt werden.

Durch die Anpassung der Lautstärkeverhältnisse in der Mischung kann in der einfachsten Weise gestalterisch manipulativ Einfluss auf die Musik genommen werden. Das lauteste Signal in der Mischung erhält auf Grund seiner Gewichtung eine für den Rezipienten wahrnehmbare besondere Rolle.<sup>123</sup>

## Panorama

Das Wort Panorama leitet sich aus dem altgriechischen Wort  $\pi\acute{\alpha}\nu$  *pan* „alles“ und  $\acute{\omicron}\rho\acute{\alpha}\omega$  *horao* „sehen“ ab. In der Musikmischung beschreibt das Panorama alle im System möglich einzunehmenden Positionen. Ein Tonsignal kann in jede beliebige Abbildungsrichtung im dem von den Lautsprechern aufgespannten Panorama positioniert und bewegt werden.<sup>124</sup> Nach welchem Prinzip das Signal platziert wird, steht in Abhängigkeit zum verwendeten System.

*„Den größten Freiraum bei der Positionierung von Klangquellen bieten Hörspiel, Klangkunst und elektroakustische Musik.“<sup>125</sup>*

Bei bildbezogenen Audioinhalten orientiert sich die Position des Schallereignisses meist an den tatsächlichen, beziehungsweise den vom Rezipienten gesehenen Positionen.

---

<sup>121</sup> Thomas 2009: 20

<sup>122</sup> Weinzierl 2008: 725

<sup>123</sup> vgl. Thomas 2009: 20

<sup>124</sup> vgl. Dickreiter 1997: 373

<sup>125</sup> Weinzierl 2008: 729

## Dynamische Klangbearbeitung

Die dynamische Klangbearbeitung bearbeitet und manipuliert das dynamische Verhalten eines Signals in Abhängigkeit zur Zeit. Ein Regelverstärker ändert in Abhängigkeit des Eingangssignals seinen Verstärkungsfaktor. Neben der Einstellbarkeit der Ansprechschwelle und dem Verhältnis der Dynamikveränderung kann die Reaktionszeit sowie die Abklingzeit kontrolliert werden.<sup>126</sup> Ein Regelverstärker kann beispielsweise den dynamischen Umfang eines Signals soweit einschränken, dass der Pegel des Signals über die Dauer des Songs stetig bleibt. Damit kann der Pegel für den Zuhörer auf einer songdienlichen Lautstärke gehalten werden.

Dynamische Klangbearbeitung kann das Signal auf kreative Art und Weise verändern. Dabei spielt die Bauform des Dynamikprozessors und die jeweilige Arbeitsweise eine Rolle. Schaltungstechnisch unterschiedliche Bauformen des Kompressors können effektivere, artefaktreichere, artefaktärmere, kreativere oder zuverlässigere Dynamikbearbeitung ermöglichen. Eine wahrnehmbare Klangverdichtung des Kompressors sowie eine vom Zuhörer wahrnehmbare zeitliche Veränderung können gestalterisch genutzt werden.<sup>127</sup> Diese Modulationsartefakte können als gestalterisches Mittel zum Einsatz kommen. Die angewendete Kompression manipuliert den Charakter des Signals und wird dabei selbst zum Instrument. Folgende Tabelle stellt vereinfacht einen möglichen Zusammenhang zwischen physikalischer und perceptiver Wirkung des Dynamikprozessors dar:<sup>128</sup>

	Physikalische Wirkungen*	Perzeptive Wirkungen*	Nutzeffekt (+) Störeffekt (-)
Attack lang ( $\Delta t_{\text{att}} \geq 3 \text{ ms}$ )	Langsames Ausregeln.  Kurze Spitzen werden nicht abgeregelt.	Unauffällige Dynamikverminderung; Lautheitserhöhung bei Nutzung des Kompressionshubs nicht maximal möglich. Verzerrungen durch Übersteuerung der nachfolgenden Übertragungskette möglich.	+  -
Attack kurz ( $\Delta t_{\text{att}} < 3 \text{ ms}$ )	Schnelles Ausregeln.  Kurze Spitzen werden abgeregelt; geringere Impulshaftigkeit.  Signaldeformation im tieffrequenten Bereich; nichtlineare Verzerrung; erhöhter Klirrfaktor.	Auffällige Dynamikverminderung; Signalspitzen bestimmen die Lautstärke; Klangverdichtung; Lautheitserhöhung bei Nutzung des Kompressionshubs maximal möglich. Unter Umständen größeres Empfinden von Weichheit (z. B. bei perkussiven Klängen). Verzerrungen bis hin zu knackähnlichen Störungen.	- +(Gestaltungsmittel)  +  -

Abbildung 41: Physikalische und perzeptive Wirkung von Regelzeiten - Attack<sup>129</sup>

<sup>126</sup> vgl. Weinzierl 2008: 730

<sup>127</sup> vgl. Weinzierl 2008: 736

<sup>128</sup> vgl. Weinzierl 2008: 737

<sup>129</sup> Weinzierl 2008: 737

	Physikalische Wirkungen*	Perzeptive Wirkungen*	Nutzeffekt (+) Störeffekt (-)
Release lang ( $\Delta t_{rel} \geq 1$ s)	Seltenes Regeln.  Geringer mittlerer Pegel; bei kurzem Attack regeln kurze Signalspitzen nachhaltig den Pegel.	Unauffällige Dynamikverminderung; Lautheitserhöhung bei Nutzung des Kompressionshubs nicht maximal möglich. Geringe mittlere Lautstärke; bei kurzem Attack sind plötzliche nachhaltige Lautstärkeeinbrüche möglich.	+  -
Release kurz ( $\Delta t_{rel} < 1$ s)	Durch häufiges Regeln Modulation von langsam veränderlichen Signalanteilen.	Erhöhung von Rauigkeit oder Pumpen; Klangverdichtung; Lautheitserhöhung bei Nutzung des Kompressionshubs maximal möglich.	- + (Gestaltungsmittel)

\* Stärke und Häufigkeit der Wirkungen hängen außerdem von den Eigenschaften des Nutzsignals und dem gewählten Threshold-Wert ab.

Abbildung 42: Physikalische und perzeptive Wirkung von Regelzeiten - Release<sup>130</sup>

## Frequentielle Klangbearbeitung

Frequentielle Klangbearbeitung bedeutet die Bearbeitung der Frequenz eines Signals mit Hilfe eines frequenzverändernden Gerätes zur Entzerrung, Verzerrung, Filterung oder Tongestaltung. Ähnlich wie bei der dynamischen Bearbeitung kann auch hier unterschieden werden zwischen einer Bearbeitung zur Entfernung technischer Unsauberkeiten, und einer Bearbeitung als gestalterisches Mittel zur Erzeugung einer gewünschten Klangfarbe. Eine Klangfarbenänderung durch frequenzgangverändernde Geräte ist in weiten Grenzen möglich.<sup>131</sup>

*„Daneben eignet sich der grafische Entzerrer [...] für die Gestaltung der Klangfarbe ebenso wie für eine verbesserte Trennung der Mikrofonsignale bei Einzel- oder Stützmikrofonverfahren.“<sup>132</sup>*

Technische Aspekte einer Frequenzkorrektur können tieffrequenter Störschall wie beispielsweise Trittschall oder Windgeräusche sein, Resonanzen oder Rauschen.<sup>133</sup> Verschiedene Baumformen und Bedienkonzepte ermöglichen unterschiedliche Eingriffs- und Formungsmöglichkeiten. Mit frequentieller Klangbearbeitung können beispielsweise Störsignale entfernt, störende Formanten kompensiert, das Signal aufgehellt oder abgedunkelt, bis hin zu stark verfremdet werden.<sup>134</sup>

<sup>130</sup> Weinzierl 2008: 737

<sup>131</sup> vgl. Dickreiter 1997: 351

<sup>132</sup> Dickreiter 1997: 362

<sup>133</sup> vgl. Dickreiter 1997: 351

<sup>134</sup> vgl. Weinzierl 2008: 747

## **Verzerrer**

In der Elektrotechnik beschreibt die Verzerrung eine Veränderung der Form einer zeitveränderlichen Größe. Es wird unterschieden zwischen linearer und nichtlinearer Verzerrung. Die nichtlineare Verzerrung eines Signals fügt dem Ausgangssignal Frequenzen in Form von Obertönen hinzu, die im Eingangssignal nicht existent sind. Verzerrer sind Audiotbearbeitungsmittel mit denen nichtlineare Verzerrungen hervorgerufen und damit deutlich hörbare Artefakte generiert werden können.<sup>135</sup>

Es gibt verschiedene Arten von Verzerrung durch unterschiedliche Bauformen von Effekten. Das Signal kann mit Hilfe der Verzerrer mehr oder weniger stark deformiert oder abgeschnitten werden. Folgen sind eine Pegelbegrenzung sowie die Addition von Partialtönen.<sup>136</sup> Harmonische Klangbearbeitung kann von subtil bis offensichtlich zur Gestaltung des Audiosignals genutzt werden.

### **4.1.4 Räumliche Klangbearbeitung**

*„Der Raumeindruck ist ein wesentlicher Bestandteil einer Aufnahme; seine Bedeutung für das Hörereignis ist deshalb so groß, weil er den optischen Eindruck ersetzen muss.“<sup>137</sup>*

Durch das Einzelmikrofonverfahren und die dabei meist dichte Mikrofonierung in der Populärmusik gibt es einen nur sehr geringen räumlichen Anteil sowie Tiefe im Signal. Die räumliche Perspektive wird in der späteren Produktion durch künstliche räumliche Klangbearbeitung erzeugt.<sup>138</sup> Für die Bearbeitung des Raumeindrucks kommen Hallgeräte sowie Delays zum Einsatz. Diese sorgen für einen passenden räumlichen Höreindruck. Der Raumeindruck muss nicht zwingend einer Natürlichkeit entsprechen, um einen optischen Höreindruck zu ersetzen. Die räumliche Klangbearbeitung kann dazu genutzt werden, eine für das Signal passende ästhetische Gestaltung vorzunehmen.

---

<sup>135</sup> vgl. Weinzierl 2008: 747

<sup>136</sup> vgl. Weinzierl 2008: 747

<sup>137</sup> Dickreiter 1997: 382

<sup>138</sup> vgl. Weinzierl 2008: 751

#### 4.1.5 Einfluss der Dreidimensionalität auf die Bearbeitung

Es stellen sich bezogen auf die Mischung zwei Forschungsfragen:

- (1) In wie weit lassen sich Techniken der Stereo-Mischung auf eine dreidimensionale Mischung übertragen?
- (2) In wie weit sind die im Stereo angewendeten Techniken für das dreidimensionale System auf Grund seiner größeren Dimension sinnvoll?

In einer gewöhnlichen zweikanaligen Stereo-Mischung kann das Signal lediglich in einer Dimension angeordnet werden. Die Möglichkeitsspanne wird durch das Wiedergabesystem eingeschränkt. Eine dreidimensionale Mischung bietet die Möglichkeit der Platzierung eines Signals in x-, y-, sowie z-Richtung. Daraus resultiert ein größerer technischer sowie kreativer Entscheidungsspielraum.

In einer Zweikanal-Stereo-Mischung werden verschiedene Techniken angewendet, um die Vielzahl der in der Popmusik vorkommenden Signale im begrenzten System hörbar zu machen. Eine Trennung der Signale kann durch die Positionierung im Panorama erreicht werden. Eine weitere Trennung kann durch Frequentierung, Lautstärkeverhältnisse und Dynamikbearbeitung erreicht werden. Jedem Signal wird im Mix mit Hilfe technischer Mittel die passende Position und der passende Stellenwert zugeordnet. Dabei ist es wichtig, dass die Mischung auf Grund der Vielzahl an Signalen nicht unverständlich wird. Dieser Sachverhalt ändert sich in einer dreidimensionalen Mischung auf Grund der Dimensionierung des Systems. Ein Beispiel: Die Bearbeitung mit Equalizern wird im Stereo auch dafür genutzt, um Platz für andere Instrumente zu schaffen, beziehungsweise um einen homogenen und auf einem 2-Kanal-Stereo-Abhörsystem durchsichtigen Mix zu schaffen. Dabei spielen die Kurven gleicher Lautstärke sowie Verdeckungseffekte eine wichtige Rolle. Bei einem vollen Arrangement werden die verschiedenen Instrumente eines Songs im Frequenzbereich klar strukturiert und eventuelle vorherrschende Frequenzbereiche zugeordnet.<sup>139</sup> Die Frequenzwiedergabe des Systems ändert sich bei einem dreidimensionalen System nicht, jedoch können die Signale nicht nur in y-Richtung, sondern auch in x- und z-Richtung positioniert werden. Die Möglichkeiten der Platzierung von Signalen erweitert die Bandbreite des Systems und eine erweiterte örtliche Trennung steht in Korrelation mit der Trennung durch Frequentierung, Trennung durch Lautstärkeverhältnisse und Dynamik-

---

<sup>139</sup> vgl. Owsinski 2007: 52

bearbeitung sowie der Signalanzahl im Song. Bei gleichbleibender Signalanzahl in einer Stereo-Mischung und einer dreidimensionalen Mischung ist eine stärkere Trennung durch die Platzierung möglich. Frequentielle und dynamische Trennung muss nicht so stark genutzt werden, um Durchsetzungsfähigkeit und Struktur im Mix zu schaffen. Dieser Sachverhalt trifft nur bei gleichbleibender Anzahl von Audioquellen zu. Wie eben angesprochen sind die Trennung, das System und die Signalanzahl kohärent. Werden in einer dreidimensionalen Mischung auf Grund der Möglichkeiten entsprechend mehr Audiosignale verwendet, wie beispielsweise Dopplungen, so wird auch hier eine Trennung durch alle genannten Möglichkeiten stattfinden müssen.

In der E-Musik wird meist darauf verwiesen, dass im Surround eine „überzeugende Darstellung mit musikalisch befriedigenden Eigenschaften“ des musikalischen Ereignisses gestaltet werden soll. Gernemann verweist darauf, dass es sich vor allem um Musikgattungen handelt, „für deren Aufnahme in der Regel sog. 'Haupt'- und 'Stütz'-Mikrofone bevorzugt zum Einsatz kommen. Zumeist betrifft diese Ensemblesmusik mit überwiegend akustischen Instrumenten beziehungsweise Gesangsstimmen. Rein (mehrkanalige) elektroakustische Musik wird hier nicht betrachtet.“<sup>140</sup>

*„Technicians want to make it as natural as it can be. We need to put the Technology into the hands of artists to let them see what they can do with it, [...] and get something out of it that you never anticipated that's when it's exciting.“ (Carlos Barth, Manager Mando Diao)*

Betrachtet man die dreidimensionale Mischung populärer Musik, steht hierbei nicht die überzeugende natürliche Darstellung des akustischen Ereignisses im Vordergrund, sondern der gestalterische Einsatz der gegebenen Mittel.

*„You really have to make [...] drums larger than life, so they cut through. So when they're soloed, they're just ridiculous.“ (Bob Horn)<sup>141</sup>*

Eine unnatürliche Darstellung des akustischen Ereignisses, das dem musikalischen Ereignis zum Vorteil verhilft, steht in keinem Konflikt.

---

<sup>140</sup> Gernemann: 3

<sup>141</sup> Signature Sound Studio 2015

#### 4.1.6 Das System und seine gestalterischen Möglichkeiten

Das Abhörssystem setzt auf Grund seiner technischen Struktur einen gestalterischen Rahmen. Beispielweise werden die Positionierungsmöglichkeiten der Signale im System durch eben jenes begrenzt. Das Wiedergabesystem nimmt somit passiv Einfluss auf den gestalterischen Prozess der Mischung.

*„Das Erweitern eines fokussierten Monotons in die Stereobreite und weiter in einen umhüllenden Klang ist ein wesentliches Gestaltungsmittel von Surround, ebenso wie die Verengung der Raumperspektive nach einer Surroundsequenz. Diese Gestaltung kann durchaus mit einer Orchesterinstrumentierung verglichen werden, wo ja auch nicht permanent nur Soli oder Tutti gespielt (komponiert) werden, sondern eine Solostimme in eine Tutti mündet oder sich aus einem Tutti über die Reduktion auf Instrumentengruppe zu zwei bis drei Instrumente hin zum Solo entwickelt.“<sup>142</sup>*

Lensing zeigt, dass das Wiedergabesystem ein wesentliches Gestaltungsmittel der Mischung ist. Im Fall der Produktion soll das Wiedergabesystem als solches genutzt werden. Eine medial-autonome Darstellung von Musik beschreibt den Schaffungsprozess in und mit dem gegebenen Medium. Die Eigenschaften des Systems und die damit verbundenen Vor- und Nachteile müssen dem Mischenden bewusst sein, damit ein gezielter Umgang dessen als gestalterisches Mittel möglich ist.

#### 4.1.7 Hörgewohnheiten

Hörgewohnheit beschreibt die Bildung von Gedächtnisinhalten durch das Wahrnehmen über das Hören. Für das Verständnis akustischer Information benötigt es einer Interpretation und Bewertung der akustischen Daten, der Schallereignisse. Musik als Schallereignis wird vom Rezipienten interpretiert und mit bereits erlernten Hörerfahrungen abgeglichen. Hörgewohnheiten werden durch individuelle und kulturelle Erfahrungen gebildet. Die Prägung der Hörmuster durch die Erfahrung dient der Identifikation akustischer Reize. Eine differenzierte Wahrnehmung akustischer Reize wird durch Hörgewohnheiten gefördert und nimmt somit Einfluss auf die Hörerfahrung.<sup>143</sup>

---

<sup>142</sup> Lensing 2009:182

<sup>143</sup> vgl. Mechthild 2006: 85

In den letzten Jahrzehnten waren zweikanal-stereophone Produktionen prägend. Eventuelle 'Vorgaben' durch Hörgewohnheiten der Zweikanal-Mischung sollen in der Arbeit und der dreidimensionalen Mischung berücksichtigt und mit einbezogen werden. Das durch verschiedene Stereo-Produktionen etablierte Panorama der Instrumente soll in der Basis beibehalten werden.

#### **4.1.8 Dreidimensionales Mischen als gestalterisches Mittel**

Das System und die Bearbeitungsmittel werden dazu genutzt, gestalterisch auf die Musik einzuwirken, um deren Spannungsbogen zu unterstützen sowie die Aussage und emotionale Wirkung des Songs zu verstärken oder zu prägen.

Aktuell gibt es nur sehr wenig Ansätze für das dreidimensionale Mischen. Es gibt verschiedene Ansätze zum Mischen in 5.1 Surround. Diese sollen im Folgenden aufgegriffen werden, da sich eventuelle Mischmethoden extrahieren und auf ein dreidimensionales Setup anpassen lassen.

Es wird zwischen zwei Hauptansätzen in der 5.1-Surround-Mischung unterschieden:

- Das Mischen aus Sicht des Publikums (Publikumsansicht)
- Das Mischen aus Sicht der Musiker (Bühnenansicht)

Die *Publikumsansicht* beschreibt die Mischung entsprechend eines Zuhörers als Zuschauer. Der virtuelle Zuhörer befindet sich frontal zur Band. Die *Bühnenansicht* platziert den Zuhörer zwischen die Band. Er wird zum Teil der Band und wird von ihrem Klang umhüllt. Aus den beiden unterschiedlichen Herangehensweisen ergibt sich die jeweilige grundlegende Positionierung der Instrumente im Panorama und die dabei entstehende Grundästhetik.<sup>144</sup> Die Bühnenansicht hat aufgrund ihres starken Pannings eine dramatischere Wirkung auf den Zuhörer, kann jedoch unpassend sein: beispielsweise wenn dieser im Live-Konzert auch gleichzeitig Zuschauer ist und ein visueller Impuls von der frontal befindlichen Bühne gegeben wird, der nicht zum gehörten Klangbild passt. Die Publikumsansicht wirkt für den Zuhörer gewöhnlicher und aus diesem Grund natürlicher, aber auch wesentlich unspektakulärer. Die Publikumsansicht verfolgt einen illusionistischen Ansatz, indem das emotionale Erlebnis des Live-Konzertes aus Sicht des Zuschauers wiederhergestellt werden soll.

---

<sup>144</sup> vgl. Owsinski 2007: 155

Die dieser Arbeit zu Grunde liegende praktische Umsetzung bedient sich beider Konzepte. Das Panning, die dreidimensionale Präsentation der Musik, orientiert sich an einer frontalen Anordnung aus Sicht eines Zuhörers. In der praktischen Arbeit sollen auditive und visuelle Ereignisse im Einklang stehen. Im Panning, sollen bestimmte Hörgewohnheiten aufgegriffen werden, um dem Zuhörer ein teilweise gewohntes Hörbild zu präsentieren. Er soll in seiner Rolle Zuhörer beziehungsweise Zuschauer bleiben und von einer dreidimensionalen Klangwelt umhüllt werden. Diese Dreidimensionalität soll losgelöst von einer klaren strukturellen Wahrnehmung sein. Sie soll mit Hilfe von künstlich erstellten Dopplungen, Zuspielern und Räumlichkeit erzeugt werden. Dabei wird darauf geachtet, dass trotz Dopplungen das Schallereignis an der gewünschten Position geortet wird. Mit Hilfe von frequentieller und zeitlicher Bearbeitung kann die Position eindeutig festgelegt werden. Neben der klaren durch den visuellen Inhalt vorgegebenen Positionierung der Singale im Panorama werden vereinzelt Signale kontrovers zum visuellen Ereignis positioniert und bewusst durch den Raum bewegt, um dem Zuhörer das System und dessen Möglichkeiten zu verdeutlichen.

*„Making a choreography for 10 Loudsoeakers“ (Ronald Prent, Chief Sound Engineer Wisseloord Studios)<sup>145</sup>*

Ronald Prent verwendet in einem Werbevideo über den Mischprozess des Mando Diao Albums 'Aelita' den Begriff 'Choreography' für die Erstellung einer 3D Audio Mischung. Dieser Begriff scheint passend und wird im Folgenden weiter verwendet. Es soll eine für die Musik passende Mischung und damit eine Choreographie für das Wiedergabesystem hergestellt werden, die das dreidimensionale Wiedergabesetup in seiner Gänze ausnutzt und dabei dem Zuschauer den Mehrwert bezogen auf die Musik und das Konzert vermittelt. Dabei soll jedoch der Grundsatz, dass das System der Musik dient, nicht vernachlässigt werden.

---

<sup>145</sup>Lichtmond 2014

#### **4.1.9 Immersion und Dreidimensionalität**

Immersion beschreibt das Eintauchen oder gar Auflösen von einem Objekt in einem anderen. Der Begriff wird in verschiedensten Disziplinen und Bereichen in unterschiedlicher Bedeutung für einen solchen Vorgang des Eintauchens oder den Zustand des Versunken seins verwendet. Curtis spricht davon, dass in der heutigen Medienwissenschaft das Wort Immersion meist als Synonym einer Illudierungserfahrung eingesetzt wird. Diese Erfahrung kann mit Hilfe von technischen Mitteln realisiert werden. Immersion beschreibt in diesem Fall die Illusion des Eintauchens in eine virtuelle Welt.<sup>146</sup> Allison Griffiths sagt in ihrem 2008 erschienenen Buch 'Shivers Down Your Spine: Cinema, Museums and the Immersive View', dass Realismus zentraler Bestandteil der immersiven Erfahrung sein muss.<sup>147</sup>

Mit Hilfe von dreidimensionaler Beschallung kann eine illusionistische Klangwelt erschaffen werden, die dem Zuhörer ein realistisches Klangbild vortäuscht. In der Produktion des dreidimensionalen Live-Konzertes kann der Zuhörer auf Grund der dreidimensionalen Beschallung in eine durch die Band und Produktion geschaffene Klangwelt eintauchen. Die Präsentation stellt jedoch keine Audioinhalte nach, wie sie in der Realität vorzufinden sind. Es wird mit Hilfe eines dreidimensionalen Wiedergabesystems eine speziell für die Musik geschaffene Klangästhetik, beziehungsweise Klangwelt erzeugt. Ein hoher Immersionsgrad kann und soll hierbei nicht angestrebt werden, da das Ziel der Produktion nicht die Bildung einer möglichst glaubhaften und realen virtuellen Welt darstellt, sondern die musikalisch kreative Einwirkung des Systems auf die Musik und dabei den Zuhörer.

#### **4.1.10 Räumlichkeit & Plausibilität<sup>148</sup>**

Ist das Hörereignis für den Zuhörer annehmbar, nachvollziehbar beziehungsweise glaubhaft, so wird der Mischung für das Live-Konzert eine hohe Plausibilität zugesprochen. Kernfrage der Plausibilität bezogen auf eine dreidimensionale Mischung und Beschallung eines Live-Konzertes ist, in wie weit die Mischung für den Rezipienten unplausibel klingen darf beziehungsweise, wie plausibel sie minimal sein sollte, sodass der Zuhörer das Hörerlebnis nicht aktiv hinterfragt. Eine Mischung wird umso spektakulärer und fantastischer, je effektvoller sie bezogen auf die Dreidimensionalität gestaltet wird. Das bedeutet aber auch, dass sie weniger mit der vom Zuschauer gewohnten Präsentation eines gewöhnlichen Live-Konzertes zu tun hat. Bildinhalt, die

---

<sup>146</sup> vgl. Curtis 2008: 1

<sup>147</sup> vgl. Griffiths 2008

<sup>148</sup> Glaubhaftigkeit, Nachvollziehbarkeit, Verständlichkeit.

Band steht frontal zum Publikum, und Toninhalt, die Wiedergabe der Band, gehen bei einer spektakulären Mischung weiter auseinander. Die Mischung soll in der grundlegenden Basis 'gewöhnlich' gehalten werden. Die Ausweitung des kreativen Spielraums auf Grund der erweiterten Panningmöglichkeiten durch das System soll nicht kontraproduktiv zur grundlegenden Ästhetik aktueller Mischungen wirken. Das bewusste Brechen der Plausibilität soll den Rezipienten für das Wiedergabesystem und die dadurch entstehenden Möglichkeiten sensibilisieren. Im Mittelpunkt der Mischung steht das Musikstück. Das System dient der Darstellung und der musikalischen Unterstützung.

#### **4.1.11 Wahrnehmung im Kontext der Liveproduktion**

Wahrnehmung ist der Prozess der Gewinnung und Verarbeitung von Informationen aus Sinnesreizen. Im Folgenden soll die Wahrnehmung angesprochen durch den auditiven und visuellen Inhalt in Bezug auf die Mischung und Beschallung beleuchtet werden.

*„The goal is to match the sound to what you are seeing; thus, in a concert video, if a musician is featured prominently onscreen, the tendency is to raise the level of that instrument.“<sup>149</sup>*

Es ist wichtig den auditiven Inhalt mit dem visuellen Ereignis abzustimmen. Das bedeutet nicht, dass der auditive Inhalt sich nicht anders verhalten darf als der visuelle. Ein bewusstes Brechen der Wahrnehmungen von Bild und Ton wird benötigt, um dreidimensionale Effekte zu generieren, die offensichtlich nichts mit dem visuellen Inhalt der Band gemein haben. Die Sensibilisierung für das Wiedergabeformat und die technischen Neuerungen des Konzertes wird gefördert. Der Zusammenhang zwischen Bild und Ton ist auch hier durch eine Plausibilitätskontrolle zu prüfen. In wie weit darf sich der auditive Inhalt vom visuellen Ereignis unterscheiden, ohne dass es für den Rezipienten unnatürlich beziehungsweise unglaublich wirkt.

*„Die optischen Theorien gehen von der Annahme aus, dass der Hörereignisort auch davon abhängt, was die Versuchsperson während der Schalldarbietung sieht und wo sie es sieht.“<sup>150</sup>*

---

<sup>149</sup> Ainlay 2004: 4-1

<sup>150</sup> Blauert 2014: 171

Blauert sagt, dass die optischen Theorien anhand verschiedener Beobachtungen davon ausgehen, dass ein Zusammenhang zwischen Schalldarbietung und Hörereignisort besteht. Das bedeutet, dass ein Hörereignis eines Hörers nicht gänzlich losgelöst vom visuellen Input betrachtet werden kann. In Bezug auf die dreidimensionale Mischung bedeutet das, dass ein Schallereignis unpassend zur visuellen Erwartungshaltung einen ungewollten Effekt hervorrufen kann.

In einem Versuchsaufbau zeigte Klemm 1918, dass die visuelle Wahrnehmung das Schallereignis aus einer anderen Richtung erscheinen lassen kann, als aus der das Signal eigentlich kommt. In seinem Versuchsaufbau sind zwei Mikrofone links und rechts vor einer Kopfhörer tragenden Versuchsperson aufgebaut. Die durch das linke Mikrofon aufgenommenen Signale wurden dem rechten Ohr zugefügt und umgekehrt. Vor beiden Mikrofonen war ein Schallhammer aufgebaut. Hat der Proband die Augen geöffnet und den Schallhammer beim Schlag beobachtet, so ordnete er das Signal der visuellen Darbietung unter und nahm es aus der jeweiligen visuellen Richtung wahr. Mit geschlossenen Augen nahmen die Versuchspersonen das Signal aus der eigentlich schallrichtigen Seite wahr, da kein optischer Reiz die Wahrnehmung 'beeinflussen' konnte.<sup>151</sup> Im Konzert wird darauf geachtet, dass der für den Zuschauer klar ersichtliche visuelle Inhalt deckungsgleich zum auditiven ist. Dadurch können ungewollte Effekte, wie beispielsweise falsche Lokalisierung verhindert werden.

Owsinski beschreibt den 'Ausgangsschild-Effekt' als ungewollten Effekt, der den Blick des Zuschauers auf Grund eines aufmerksamkeitsziehenden Hörereignisses in den Surroundkanälen von der Leinwand auf die Ausgangshinweise zieht.<sup>152</sup> In der dreidimensionalen Mischung ist darauf zu achten, dass eine homogene Mischung passend zum visuell erwarteten Ereignis entsteht. Dabei soll sich der Zuhörer und Zuschauer zwar auf Grund eines Hörereignisses im Raum drehen können, jedoch nicht ungewollt die Orientierung und das optische Ziel bezogen auf die Band als Zentrum des Ereignisses verlieren. Flückiger bezeichnet das Phänomen als eine 'Problematik des Surround-Kanals', wenn eine Schallquelle von hinten die Fiktion der Leinwand durchbricht und dabei die auditive Wahrnehmung verzerrt.<sup>153</sup> Flückiger bezieht sich in ihrem Buch 'Sound Design' primär auf Soundeffekte, wobei sich das Verhalten auf das Konzert übertragen lässt, das gleichermaßen visuellen und auditiven Inhalt präsentiert. Im Veranstaltungsort des Konzertes wird die Band frontal zu den Zuschauern stehen. Visuelle Komponenten werden in der Mischung mit akustischen

---

<sup>151</sup> vgl. Blauert 2014: 171

<sup>152</sup> vgl. Owsinski 2007: 154

<sup>153</sup> vgl. Flückiger 2001: 56ff

verknüpft. Dabei werden akustische Signale, die für den Zuschauer eindeutig von einem der Musiker kommen, erkennbar in die jeweilige Richtung positioniert. Hörereignisort und Schalldarbietung sollen möglichst sinnvoll verknüpft werden. Weitere visuelle Inhalte sollen die dreidimensionale Beschallung stimmig verstärken beziehungsweise ergänzen.

## 4.2 Die Musik

### 4.2.1 Musikalische Auswahlkriterien

Die Arbeit beschreibt die Umsetzung eines dreidimensional beschallten Live-Konzertes populärer Musik. Bei der Auswahl der Musik und damit der Band wurde die Frage gestellt, welche Elemente der Musik in einem dreidimensionalen Kontext in welcher Funktion kreativ verarbeitet werden können. Hierzu sollen folgende Faktoren bei der Auswahl beachtet werden:

- Instrumentierung, die eine synthetische Mischung ästhetisch zulässt
- Möglichkeit der ungewohnten Panoramapositionen auf Grund der Instrumentierung
- Vielzahl an Dopplungen, die eine homogene Raumverteilung zulassen

*„Its mostly electronic music like very many synthesizers, effects, delays and stuff. This is something that really explores the possibilities of a 3D Mix, making it even more spectacular.“ (Patrick Lemmens, Senior Sound Engineer, Galaxy Studios about Lichtmond 3 – Mixing in Auro 3D)<sup>154</sup>*

Synthetische Instrumente und Elemente geben die Möglichkeit ein unrealistisches Panorama zu gestalten, ohne dass es für den Rezipienten unnatürlich erscheint.

### 4.2.2 Die Band: „Eau Rouge“

Eau Rouge ist eine dreiköpfige Elektro-Rock Band, die Popmusik spielen. Sie beschreiben sich selbst mit folgenden Worten:

*„EAU ROUGE machen Musik für Nachtschwärmer, machen Sounds für Somnambulisten. Ihre Klangteppiche zwischen Indie, Electro, Shoegaze, Post Punk und Trip Hop liebäugeln subtil mit der großen Pop-Geste.“<sup>155</sup>*

Eau Rouge's Musik eignet sich auf Grund ihres halb synthetisch elektronischen Musikstils und ihrer akustisch rockigen Instrumentierung für eine dreidimensional angeordnete Präsentation für die der Arbeit zu Grunde liegende praktische Anwendung. Die akustischen Elemente können passend zum visuellen Inhalt des Live-Konzertes plausibel angeordnet werden. Synthetische Dopplungen verschiedener Signale erweitern die Dimension ohne direkten Einfluss auf die Plausibilitäts-

---

<sup>154</sup> Lichtmond 2014

<sup>155</sup> Eau Rouge 2013

kontrolle zu nehmen. Elektronische Elemente der Songs eignen sich für eine dem Zuhörer ungewohnte Platzierung im Panorama. Damit kann das dreidimensionale System in seiner Gänze ausgenutzt werden.

### **4.2.3 Instrumentierung**

Die Band ist wie folgt instrumentiert:

- Schlagzeug und Samplepad
- E-Gitarre/ Bass und Gesang
- E-Gitarre/ Bass, Synthesizer und Gesang

Die beiden Gitarristen haben ein Setup, mit dem sie das Eingangssignal der Gitarren splitten und jeweils in eine Gitarreneffektkette und parallel in eine Bassamp-Simulation schicken können. Es können zwei unterschiedliche Signale mit einem Instrument erzeugt werden und, wenn es so gewollt ist, auch parallel.

Die Instrumentenstilistik lässt sich in fünf unterschiedliche Kategorien unterteilen. Als rhythmische Basis dient das akustische wie auch das elektronische Schlagzeug. Für die Tiefanteile in der Musik sorgt jeweils der Bass, gesplittet aus den Gitarrensensoren, oder ein Synthesizer. Als frequenzielle breite Basis dient eine schwere, meist verzerrte und leicht synthetisierte Rhythmusgitarre oder Synthesizer. Diese Instrumentengruppe schafft ein breites harmonisches Bett der Musik. Für die helle melodische Grundstruktur sorgt die Leadgitarre, die für die Melodieführung und den 'Glitzer und Glanz' zuständig ist. Als Leadinstrumente dienen die Gesänge und die Sologitarre. Die grundästhetischen Effekte der Einzelsignale werden von den Instrumentalisten selbst generiert. Dabei verleihen neben der Spielweise verschiedene Bodeneffektegeräte, Delays, Arpeggiator, Pitchshifter und Kompressoren den Signalen ihren Klangcharakter.

## 4.3 Die Vorproduktion

### 4.3.1 Vorbereitung

Das Kapitel beschreibt die Produktion des Inhaltes für das Live-Konzert. Dabei handelt es sich um kein gewöhnliches Live-Konzert, sondern wird definiert als Live-Produktion. Sie strebt nicht an, möglichst alle Elemente des Konzertes live zu performen, sondern legt den Fokus auf ein dreidimensionales Klangereignis verbunden mit einer Live-Performance der Künstler. Dabei kommen Zuspielder sowie Automationen vieler Signale zum Einsatz, die bereits vor Stattfinden des Konzertes vorproduziert, gemischt und eingerichtet werden.

In der Produktion selbst wird zwischen Live-Inhalt und Zuspielder unterschieden. Des Weiteren wird in der dreidimensionalen Mischung zwischen folgenden Objekten differenziert:

- Basisobjekte: Objekte die zum visuellen Inhalt kohärent sind
- umhüllende Objekte: Objekte, die subtil das Originalsignal stützen und 'verbreitern', um die Musik in alle Panoramrichtungen akustisch anzureichern
- plakative Objekte: Objekte, die offensiv das 3D Setup betonen und sich dabei beispielsweise im Raum bewegen (Sensibilisierung der Zuschauer für das System)

Ziel ist es, eine dreidimensionale Beschallung für die Live-Produktion zu realisieren. Der Raum wird mit akustischen Dopplungen und Hall gefüllt, um einen umhüllenden Raumklang zu schaffen, damit der Zuhörer in eine dreidimensionalen Klangwelt eintauchen kann. Dabei wird die Dreidimensionalität der Mischung nicht offensichtlich herausgestellt. Aus diesem Grund gibt es Teile der Mischung, die als plakative Anteile dienen, um dem Zuhörer aufzuzeigen, dass dreidimensional beschallt wird. Mithilfe der plakativen Objekte soll der Zuschauer für das dreidimensionale System sensibilisiert werden.

Die Abstufungen zwischen verschiedenen Wiedergabesystemen in der Mischung (2 Kanal Stereo, 5.1 Surround und 3D) sollen neben der Unterstützung des Spannungsbogens des Songs dem Zuhörer den Mehrwert des Setups während des Konzertes verdeutlichen. Wie Lensing in seinem Buch über die Gestaltung von Filmtönen, angesprochen in (*Kapitel 4.1.6*), aufgezeigt hat, wird das System maßgeblich als gestalterisches Mittel zur Unterstützung der Komposition und des Spannungsbogens genutzt.

*"But simply, for something to sound up-front in a mix, something else needs to sound further away, and one of the biggest mistakes people make at the mixing stage, is to try to make each sound as big and up-front as possible." (Paul White)<sup>156</sup>*

Für den Zuhörer werden akustische Zusammenhänge vor allem dann klar ersichtlich, wenn sie deutlich vorgeführt werden. Die Mischung wird so aufgebaut, dass ein Verständnis für das dreidimensionale Setup während des Konzertes entwickelt werden kann. Nichtsdestotrotz steht im Fokus der Produktion der Interpret und dessen Musik. Es wird darauf geachtet, dass trotz der vielseitigen Möglichkeiten der dreidimensionalen Effekte durch das System, die Aufmerksamkeit der Rezipienten beim Interpretieren bleibt.

### **Inhalt der musikalischen Vorproduktion**

Die musikalische Vorproduktion ist in drei Abschnitte unterteilt. Im ersten Arbeitsschritt werden die Zuspielder in Pro Tools vorproduziert. Ausgesuchte bereits für eine CD produziert vorliegende Audiospuren der Musiker werden im dreidimensionalen Raum angeordnet. Im zweiten Schritt findet in der gleichen Pro-Tools-Session die Vorbereitung der Mischung des Live-Materials statt. Während des Live-Konzertes werden die Audiosignale der Protagonisten technisch und ästhetisch in Echtzeit bearbeitet. Im dritten Schritt werden die in Pro Tools vorgemischten Zuspielder, sowie die live bearbeiteten Signale durch den Isono-Prozessor, gesteuert über die Spatial Audio Workstation (SAW), im dreidimensionalen Panorama angeordnet und ausgewählte Ereignisse vorautomatisiert.

Alle Systeme laufen timecodebasiert. An vorbestimmten Timecodepositionen liegt jeweils ein Song an. Die Band spielt auf Click. Mit dem Start des Songs an der festgelegten Position startet der Song im passenden Tempo, sodass tempobezogene Vorgänge synchronisiert ablaufen. Pro Tools gibt dabei ein Timecodesignal der entsprechenden Position aus. Alle weiteren zeitbezogenen Vorgänge können mit Hilfe des Timecodes synchronisiert werden. Es wird darauf geachtet, dass die in der CD-Produktion grundlegend hergestellte Klangästhetik beibehalten wird.

---

<sup>156</sup> SOS 1994

### 4.3.2 Technisches Setup der Vorproduktion

Auf Grund der angestrebten Flexibilität im Rahmen der Vorproduktion und der Ungewissheit bezüglich des finalen Produktionsortes sollte das technische Setup eine gewisse Mobilität gewährleisten. Mit dem im Folgenden beschriebenen Setup-Entwurf wäre zu jedem Zeitpunkt ein Ortswechsel möglich gewesen. Zentraler Bestandteil des Setups ist ein MacBook Pro: Mid 2012, 2,6 Ghz i7 Prozessor und 16GB Ram. Der Laptop erwies sich als stabiler Produktionsrechner, auf welchem sowohl die Auro-Vorproduktion, als auch die spätere Live-Session bearbeitet werden konnte. Pro-Tools-Sessions konnten damit, ohne auf gewisse Räumlichkeiten angewiesen zu sein, vorbereitet werden. Dadurch wurde ein Arbeitsablauf ermöglicht, der die eigentlichen Mischungen im HdM-Tonstudio beschleunigte.

#### Lautsprecherkonfiguration: Auro 3D

Für die Vorproduktion in Auro wurde die Lautsprecher-Anordnung des HdM-Tonstudios genutzt. Ein 5.1 Surround bestehend aus Adam S2-A Aktivmonitoren und einem Adam Subwoofer wurden mit vier Genelec 1029 als Höhen-Lautsprecher ergänzt. Daraus ergab sich ein Auro 9.1 Wiedergabe-System der Abbildung zu entnehmenden Kanal-Zuweisung.

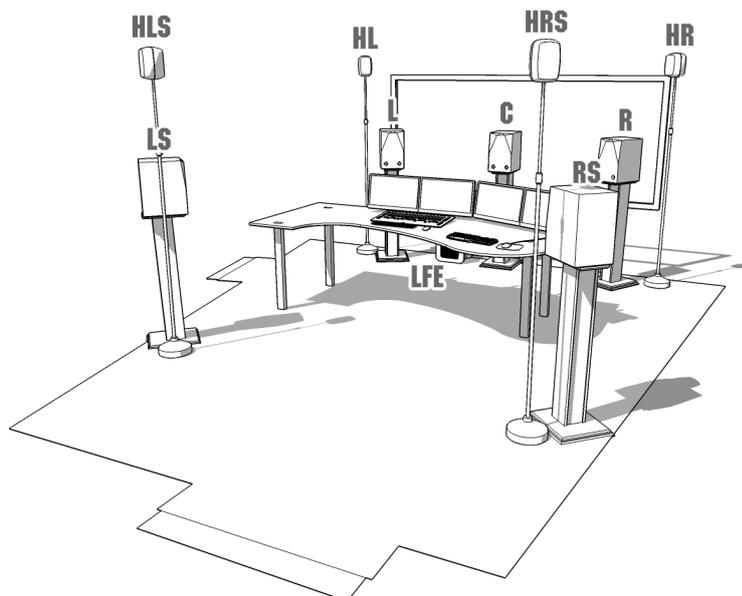


Abbildung 43: Auro 3D (9.1) Lautsprecherkonfiguration - HdM Tonstudio<sup>157</sup>

<sup>157</sup> Eigene Darstellung

## Lautsprecherkonfiguration: Iosono

Um sich während der Vorproduktion mit dem Iosono-System vertraut zu machen und ein Gefühl für zusätzliche Lautsprecher seitlich und hinter dem Zuhörer zu bekommen, wurde die Auro 9.1 Lautsprecher-Anordnung im HdM-Tonstudio um sechs weitere aktive Studiomonitore erweitert. Dabei blieb der Höhenlayer unverändert, jedoch wurde der Surround-Ring um jeweils zwei Lautsprecher an der Seite und hinten erweitert.

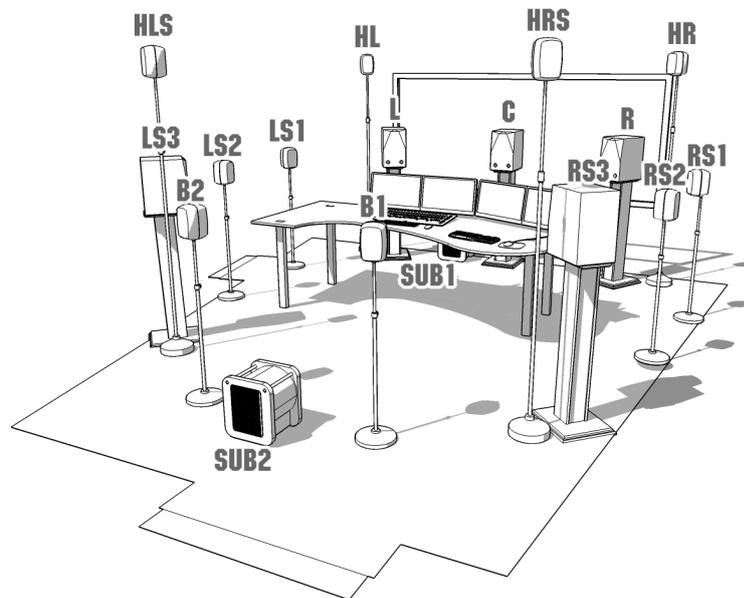


Abbildung 44: Iosono (15.2) Lautsprecherkonfiguration - HdM Tonstudio<sup>158</sup>

Für die Benennung der Lautsprecher im Iosono-Setup hat sich eine Bezeichnung ähnlich der des Auro-Setups bewährt. Sofern es auf einer Seite eine größere Anzahl von Surround-Lautsprechern gibt, werden diese von 1 beginnend durchnummeriert. In der Lautsprecheranordnung der Vorproduktion kommen mit der Bezeichnung B für Back Lautsprecher hinter dem Zuhörer hinzu. Die Bezeichnung des LFE wird zu SUB, weil der Subwoofer im Iosono-Ansatz als Erweiterung der Ringlautsprecher zu ‚Fullrange‘-Lautsprechern dienen soll und im System selbst, auch als Sub benannt wird. In diesem Setup kommt ein weiterer Subwoofer im hinteren Bereich der Anordnung hinzu, um als Extension der Audio-Objekte im hinteren Bereich für ausreichend Bassreserven zu sorgen.

<sup>158</sup> Eigene Darstellung

Erste Hörversuche mit unterschiedlichen Reproduktions-Algorithmen ergaben in dieser Aufstellung keinen Unterschied. Eine Reproduktion erfolgt ausschließlich über den Panning-Algorithmus, da der Abstand zwischen den Lautsprechern im unteren Ring-Layer zu groß ist, um mit Wellenfeldsynthese wiederzugeben. Dennoch war dieses Setup ein wichtiger Teil der Vorproduktion, um den Signalfluss und die Sessions vorzubereiten.

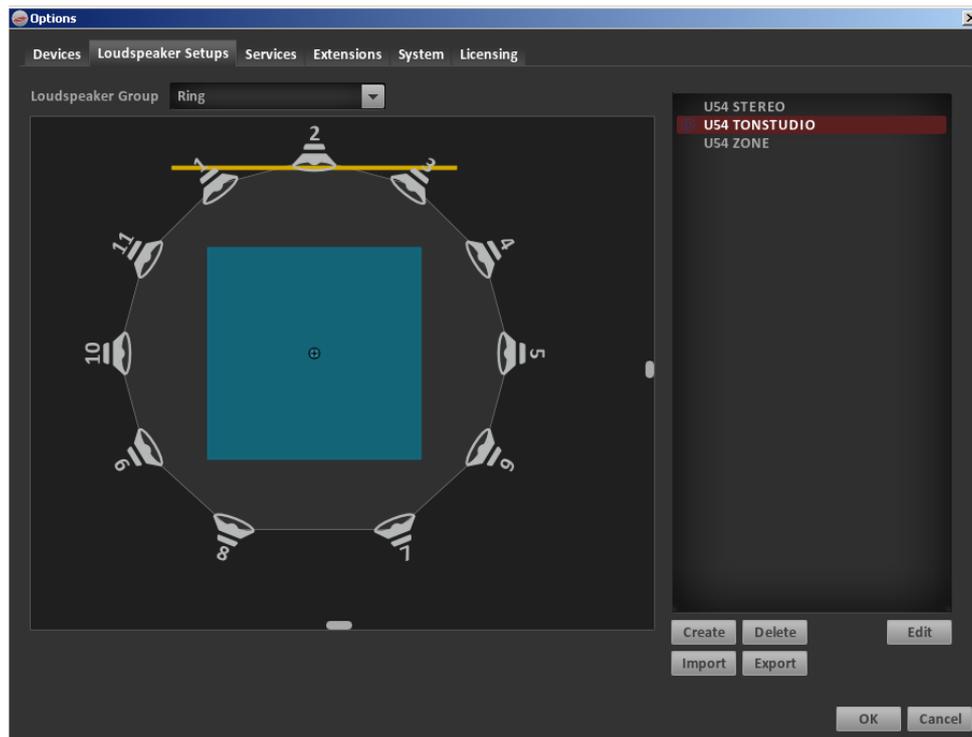


Abbildung 45: Iosono Loudspeaker Setup - HdM Tonstudio<sup>159</sup>

---

<sup>159</sup> Screenshot

## Signalkette: Auro 3D

Um den regulären Betrieb im Studio nicht einzuschränken und den Aufwand für die Einrichtung vor jeder 3D-Mischung zu minimieren, wurde das Studio um ein Motu 828 Audio Interface und ein XLR-Anschlussfeld mit Leitungen zu jedem einzelnen Lautsprecher erweitert. Somit war vor jeder Mischung lediglich ein Umstecken der XLR-Leitung direkt am Lautsprecher nötig. Eine Firewire-Leitung wurde direkt zum Arbeitsplatz verlegt, sodass nach Anschluss des Produktions-Laptops an das Audio-Interface sofort mit der Mischung begonnen werden konnte.

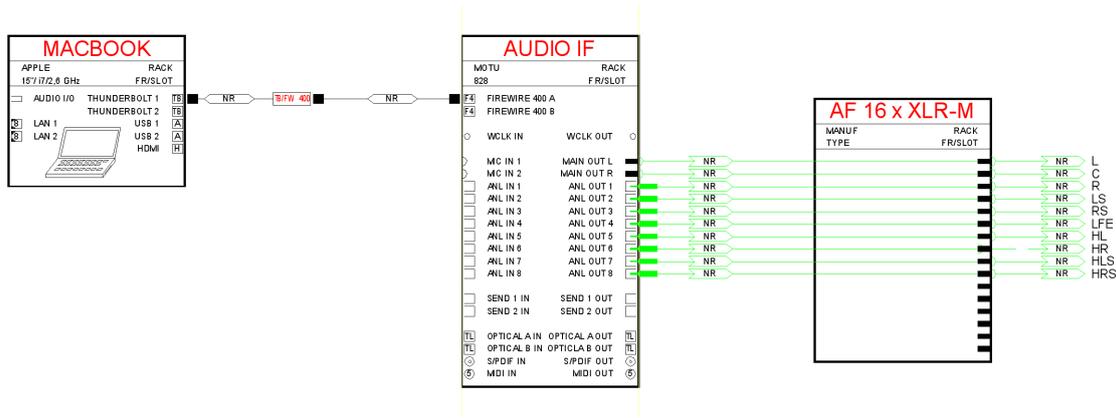


Abbildung 46: Blockschaltbild Auro 9.1 Setup - HdM Tonstudio<sup>160</sup>

## Signalkette: Iosono

Die folgende Erklärung bezieht sich auf *Anhang I*, das Blockschaltbild des Iosono System-Aufbaus. Im System-Aufbau bleibt das Mac Book als Produktionsrechner weiterhin von zentraler Bedeutung. Über Thunderbolt ist dieses mit einem PCIe-Gehäuse verbunden, in dem zwei Pro Tools HDX-Karten installiert sind. Eine davon dient als Backup-Karte. Über zwei Digilink-Verbindungen stehen dem Pro Tools System 64 Kanäle des modularen Penta 720 Pro-Tools-Interfaces der Firma NTP zur Verfügung. Dieses bekommt über Multimode-Duplex-Lichtwellenleiter 64 Kanäle aus dem im Blockschaltbild als „MADI-RTR“ bezeichneten RME Madiface XT, welches im Setup als zentrale Audio-Kreuzschiene dient. An dem zweiten der drei Madi-Ports des MADI-RTR liegen die AD/DA-Wandler an. Als kompakte Möglichkeit für die Wandlung von 32 Mikrofonsignalen und 64 Ausgangssignalen sind im Setup Wandler der Firma Direct.Out integriert. Eine Andiamo.MC wandelt bis zu 32 Mikrofonsignale zu Madi und kann von Madi bis zu 32 Signale D/A wandeln. Im Madi-Loop mit einer Andiamo.2 stehen D/A-Wandler für 32 zusätzliche Ausgangssignale zur

<sup>160</sup> Eigene Darstellung

Verfügung. Über SUB-D-25-Pin-Verbindungen sind die Andiamos an 19“ Breakout-Boxen angeschlossen, die als 1HE XLR -Streifen in entsprechender Bestückung mit XLR-F und XLR-M Anschlüssen einen sauberen Übergabepunkt bilden. Der dritte Madiport des Madi-RTR ist über Coaxial-Kabel mit dem Iosono-Core verbunden. Somit stehen auch diesem 64 Audio I/Os zur Verfügung. Als Steuerrechner für die Control Unit des Iosono Cores und DAW für die Spatial Audio Workstation steht eine HP Workstation zur Verfügung. Eine direkte Netzwerkverbindung zwischen Nuendo PC und Iosono Core sorgt für einen sicheren Austausch der TCP/IP-Daten für die Control Unit und der UDP-Daten für die Steuerung der Audio-Objekte. Zusätzlich läuft auf dem Nuendo PC die RME-Software TotalMix für das Monitoring der Band. Das über USB mit dem Nuendo PC verbundene Madiface XT wird über die TotalMix Software dank integrierter DSP zu einer Art digitalem Mischpult. Über die TotalMix Oberfläche werden alle Signal-Routings im System und die Monitor-Mixe der Band verwaltet.

Alle Geräte mit AD/DA-Wandlung werden im System über die Wordclock des Avid Sync HD synchronisiert. Ein Sonifex Wordclock-Distributionsverstärker sorgt für die Wordclock-Verteilung im System. Für die zeitliche Synchronisierung der DAWs und die Möglichkeit, die DAWs über Timecodesignale der Band zu starten, ist ein von der Bühne kommender Kanal für LTC auf den analogen Ausgang des Madiface XT geroutet. Dieser ist mit dem LTC-Eingang des Sync HDs verbunden. Das Sync HD ist in der Lage LTC in MTC umzuwandeln. Für die Synchronisierung der Nuendo DAW ist darum das Sync HD via MIDI mit dem Madiface XT verbunden. Somit gibt es folgende drei mögliche Timecode-Szenarien:

- LTC wird von der Band gestartet und synchronisiert sowohl Pro Tools als auch Nuendo
- LTC wird in Pro Tools gestartet und synchronisiert über den im Sync HD gewandelten MTC die Nuendo DAW
- MTC wird in Nuendo gestartet und synchronisiert über den im Sync HD gewandelten LTC die Pro Tools DAW

Für die Steuerungs- und Konfigurations-Software des Pro-Tools-Interface und des Iosono-Core sowie für einen schnellen Datenaustausch zwischen MacBook und Nuendo PCs sind alle Geräte via 100/1000 Mbit Netzwerk über einen Router verbunden. Für die in *Kapitel 4.2* angesprochene Abwägung zwischen digitaler Mixing-Console und Pro Tools als alleiniges Mischpult-System wurde der Entschluss gefasst, dass, sofern wichtige Signale über eine Faderoberfläche im direkten Zugriff

liegen, ein Mischpult nicht zwingend nötig ist. Eine AVID S3 Console mit 16 Fadern und der Möglichkeit, Mixer-Layouts zu speichern, bietet im Zusammenspiel mit Pro Tools eine intuitive Mix-Möglichkeit. Dies gilt auch für das Live-Konzert. Angebunden via Netzwerk kann der Pro-Tools-Laptop über das Avid eigene Eucon-Protokoll sicher und schnell die entsprechenden Daten austauschen.

Zur Konfiguration der Andiamo.MC gibt es eine weitere USB-Verbindung zwischen Nuendo PC und Andiamo.MC. Die Andiamo.MC kann in ihren Ein- und Ausgängen über eine Software-Oberfläche gemonitort werden. Über diese Oberfläche werden auch Pegel und Phantomspeisungen der Mikrofoneingänge gesteuert.

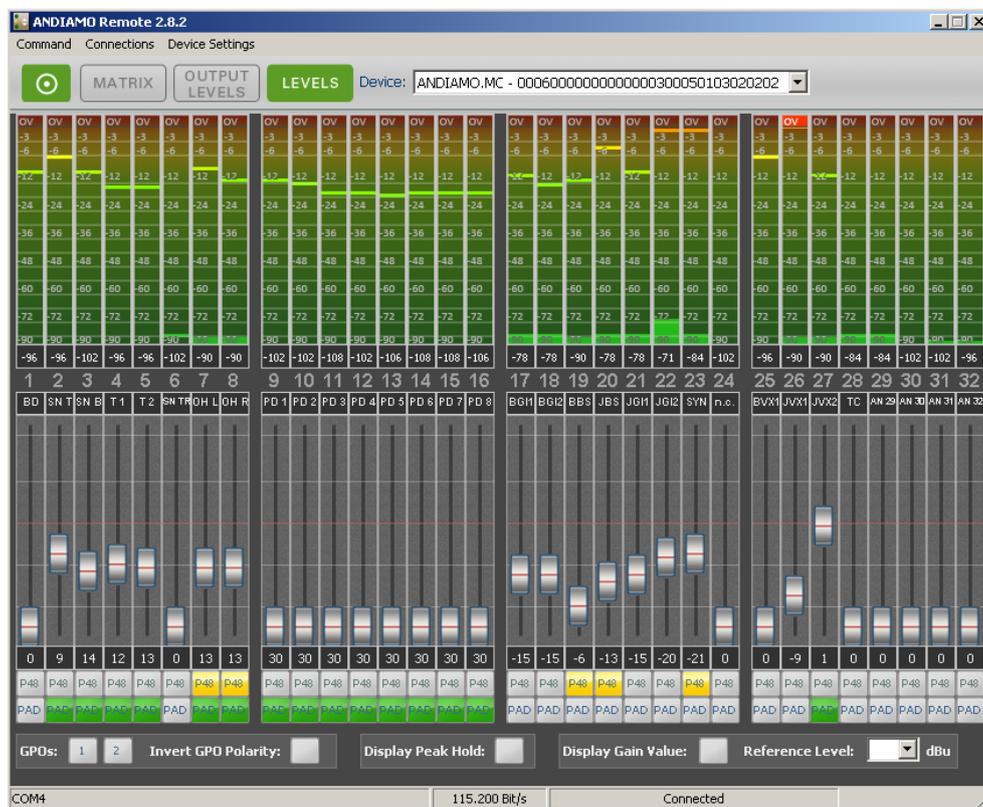


Abbildung 47: Andiamo Remote Software<sup>161</sup>

<sup>161</sup> Screenshot

Eine detaillierte Vorbereitung ist bei dieser Vielzahl an Kanälen und Routings unabdingbar. Die Signalliste in *Anhang III* verdeutlicht den Signalfluss sowie interne Routings im Verlauf der Vorproduktion. Zur Verdeutlichung wird der Signalfluss des Basedrum-Signals (BD) sowie des fertigen Schlagzeug-Mixes an Hand dieser Signalliste beispielhaft erklärt:

SIGNAL	MIC	BREAKOUT	ANDIAMO MC/2 (LWL)	RME MADI 2 (LWL)	RME MADI 1 (LWL)	PENTA In 1 (LWL)	PROTOOLS
BD	BETA 52	BO 1-1	MIC IN 1	O 1	I 1	O 1	I 1

Basedrum, mikrofoniert mit Shure Beta 52, geht über den ersten Anschluss der Breakout-Box 1 auf die Andiamo MC auf den Anschluss MIC IN 1. In beiden Andiamos wurde zunächst ein 1 zu 1 Routing vorgenommen, wobei auf Grund des Madi-Loops zwischen den Wandlern die Ausgangskanäle der Andiamo.MC auf die MADI-Kanäle 1-32 geroutet sind und die der Andiamo.2 auf die MADI-Kanäle 33-64. Mit diesem Routing kommt über Madi-Out 1 der AD-Wandler das Signal auf Kanal 1 am Madi-Port 2 des RME Madiface an. Hier sind alle ankommenden Signale des zweiten Madi-Ports auf die Ausgänge des ersten geroutet. So liegt am Pro-Tools-Interface Penta 720 auf Kanal 1 des ankommenden Madi-Signals das der Basedrum an.

PROTOOLS	PENTA Out 1 (LWL)	RME MADI 1 (LWL)	RME MADI 3 (COAX)	IOSONO CORE
DRUMS 1	O 1	I 1	Iosono Obj 1	O 1
DRUMS 2	O 2	I 2	Iosono Obj 2	O 2

Nach der Mischung in Pro Tools wird eine abgemischte Drum-Stereo-Spur über die Pro-Tools-Interface Madi-Ausgangskanäle 1 und 2 zurück an die Madiface-Eingängs-Kanäle 1 und 2 des Madi-Port 1 geroutet. Im Madiface werden alle ankommenden Signale des Madi-Port 1 direkt an die Ausgangskanäle des Madi-Port 3 geroutet. So gelangen die in Pro Tools abgemischten Audio-Objekte über Coaxial-Madi auf Kanal 1 und 2 in den Iosono-Core.

IOSONO CORE	RME MADI 3 (COAX)	RME MADI 2 (LWL)	ANDIAMO MC/2 (LWL)	BREAKOUT	SIGNAL
O 1	I 1 Iosono Render Slot 1-1	L	O 1	I 1	L (ADAM)
O 2	I 2 Iosono Render Slot 1-2	C	O 2	I 2	C (ADAM)
O 3	I 3 Iosono Render Slot 1-3	R	O 3	I 3	R (ADAM)
O 4	I 4 Iosono Render Slot 1-4	RS 1	O 4	I 4	RS 1 (GENELEC)
O 5	I 5 Iosono Render Slot 1-5	RS 2	O 5	I 5	RS 2 (GENELEC)
O 6	I 6 Iosono Render Slot 1-6	RS 3	O 6	I 6	RS 3 (GENELEC)
O 7	I 7 Iosono Render Slot 1-7	B 1	O 7	I 7	B 1 (ADAM)
O 8	I 8 Iosono Render Slot 1-8	B 2	O 8	I 8	B 2 (ADAM)
O 9	I 9 Iosono Render Slot 1-9	LS 1	O 9	I 9	LS 3 (GENELEC)
O 10	I 10 Iosono Render Slot 1-10	LS 2	O 10	I 10	LS 2 (GENELEC)
O 11	I 11 Iosono Render Slot 1-11	LS 3	O 11	I 11	LS 1 (GENELEC)
O 12	I 12 Iosono Render Slot 2-1	HL	O 12	I 12	HL (GENELEC)
O 13	I 13 Iosono Render Slot 2-2	HR	O 13	I 13	HR (GENELEC)
O 14	I 14 Iosono Render Slot 2-3	HRS	O 14	I 14	HRS (GENELEC)
O 15	I 15 Iosono Render Slot 2-4	HLS	O 15	I 15	HLS (GENELEC)
O 16	I 16 Iosono Render Slot 3-1	SUB 1	O 16	I 16	SUB 1 (ADAM)
O 17	I 17 Iosono Render Slot 3-2	SUB 2	O 17	I 17	SUB 2 (ADAM)

Im Iosono-Core werden die Audio-Objekte entsprechend dem Lautsprecher-Layout und der Daten aus der Spatial Audio Workstation gerendert. Die Ausgangskanäle der jeweiligen Render-Slots gelangen über die Coaxial-Madi-Verbindung zurück zum Madiface-Port 3. Dort werden die

gerenderten Signale für die Lautsprecher eins zu eins auf die Madi-Ausgangskanäle des Port 2 geroutet und gelangen so zurück zu den DA-Wandlern. In diesem Fall werden die gerenderten Signale von der Andiamo.MC gewandelt und gehen über die Anschlüsse 1-16 der Breakout-Box 3 und dem ersten Anschluss der Break-Out-Box 6 zu den aktiven Monitoren in der Regie des HdM Tonstudios.

O 56	56		O 56	→	156		ANL OUT 24		
O 57	57		O 57	→	157		ANL OUT 25		
O 58	58		O 58	→	158		ANL OUT 26		
O 59	59		O 59	→	159	→	ANL OUT 27	BO 6-11	→ IN-EAR Jonas L
O 60	60	TOTAL MIX / JO L	O 60	→	160	→	ANL OUT 28	BO 6-12	→ IN-EAR Jonas R
O 61	61	TOTAL MIX / JO R	O 61	→	161	→	ANL OUT 29	BO 6-13	→ IN-EAR Bo L
O 62	62	TOTAL MIX / BO L	O 62	→	162	→	ANL OUT 30	BO 6-14	→ IN-EAR Bo R
O 63	63	TOTAL MIX / BO R	O 63	→	163	→	ANL OUT 31	BO 6-15	→ IN-EAR Magnus L
O 64	64	TOTAL MIX / MAGN L	O 64	→	164	→	ANL OUT 32	BO 6-16	→ IN-EAR Magnus R
		TOTAL MIX / MAGN R							

Der Signalliste kann man auch entnehmen, dass auf die Madiface-Ausgangskanäle des Madi-Ports 2 keine diskreten Signale geroutet sind, da auf diesen Kanälen intern TotalMix-Monitormischungen für die Musiker erstellt wurden. Diese werden über die Andiamo.2 gewandelt und gehen über die Anschlüsse 11-16 der Breakout-Box 6 zu den In-Ear-Kopfhörer-Vorverstärkern der Musiker.

### 4.3.3 Kontroll- & Interaktionsschnittstellen

Für die Steuerung der Audio-Objekte im dreidimensionalen Raum gibt es die Möglichkeit, Audio-Objekte über die SAW im Nuendo oder dem Core-eigenen Scene Designer zu positionieren und zu steuern. In einem Versuchsaufbau wurden Ansätze entwickelt, Interaktionsmöglichkeiten für den Mischer selbst und für die Musiker zu schaffen. Mögliche Szenarien könnten folgendermaßen aussehen:

- Während des Live-Konzertes steht dem Mischer ein iPad zur Verfügung, auf dem in einer App alle Audio-Objekte mit ihrer aktuellen Position angezeigt werden. Nach Auswahl des Audio-Objektes kann es über einfache Wisch-Bewegung auf dem Touchscreen im Raum bewegt werden. Vorprogrammierte Automationsabläufe, wie Kreisbewegungen, Auf- und Ab-Bewegungen sowie Vorne-Hinten-Verläufe können aktiviert und über eine TAP-Funktion dem Tempo des Songs angepasst werden.
- Während des Live-Konzertes hat der Keyboarder an seinem Instrument beispielsweise einen lichtempfindlichen Sensor. Während der Musiker den Song aufführt, hat er die Möglichkeit über einfache Handbewegungen im Sensor einen Helligkeitsverlauf zu erzeugen. Die Informationen des Sensors steuern dabei beispielsweise den Auf- und Ab-Verlauf des entsprechenden Audio-Objektes in der dreidimensionalen Lautsprecherkonfiguration.
- Während des Live-Konzertes hat der Sänger die Möglichkeit über Körpergestik Audio-Objekte zu bewegen. Mit Hilfe eines Infrarot-Tiefensensors, vergleichbar mit dem eines Microsofts Kinect Systems<sup>162</sup>, können beispielsweise Informationen aus Hand- oder Körperbewegungen den Positionsverlauf von Audio-Objekten für die Gesangsspuren steuern.

Die Problematik im aktuellen Steuerkonzept ist, dass sowohl der Scene-Designer im Core als auch die Spatial Audio Workstation keine Logik implementiert haben, die prüft, ob externe Sender existieren und im Moment Daten für ein Audio-Objekt senden. Die einzige Möglichkeit wäre eine konkrete Aufteilung und Zuweisung der Audio-Objekte zur SAW oder einem externen Sender.

---

<sup>162</sup> Hardware zur Steuerung der Videospielekonsole Xbox 360

Ein Ansatz verfolgt die Möglichkeit, das grundlegende Panning und Positionieren der Audio-Objekte weiterhin über die SAW zu konfigurieren. Die Objekt-Daten sollen nur bei Betätigen der Interaktionsschnittstelle überschrieben werden. Nach der Interaktion sollen die Objekte automatisch auf die Steuerdaten der SAW zurück gesetzt werden. Der Testaufbau sah dabei folgendermaßen aus:

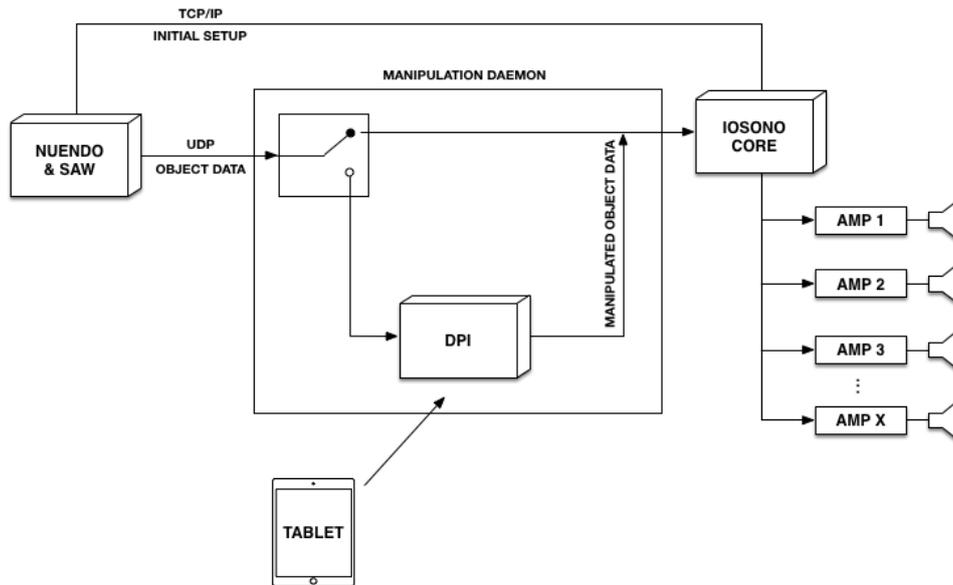


Abbildung 48: Iosono Interaktionsschnittstelle<sup>163</sup>

Im Testaufbau wurden die Positionsdaten der Audio-Objekte im UDP Datenstrom-der SAW über einen Manipulation Daemon verändert. Die Deep Packet Inspection (DPI) des Manipulation Daemon bestand dabei aus einer Matlab-Programmierung, die in der Lage war, die UDP-Pakete auszulesen. Bei der Deep Packet Inspection besteht der Unterschied zu klassischen Paketfiltern darin, dass sowohl der Headerteil als auch der Datenteil, also der Payload eines UDP-Paketes, überprüft werden und die Regulierung der Datenströme ermöglicht wird. An Hand der folgenden Iosono-„Vertical Pan“-Protocol-Definition war möglich, die 104 Bytes pro UDP-Paket in die einzelnen Bestandteile zu unterteilen und anhand programmierter Methoden den gewünschten Datentyp, beispielsweise einen ‚float32‘ für die Steuerung des Height Layers, im UDP-Paket zu überschreiben.

<sup>163</sup> Eigene Darstellung

No	Type	Bytes	Data Content / Example	Field Name	Description
1	char[]	35	/iosono/renderer/verticalpan/v1/src	OSC address	Protocol header
2	char[]	1	0		Filling zero bytes
3	char[]	15	,iiffffiifiii	OSC data types	
4	char[]	1	0		Filling zero bytes
5	int32	4	7	Channel/source number	Counting starts with zero
6	int32	4	1 (plane wave, normal latency)	Algorithm hints	Bit 0 is used as source type bit.. Bit 0 = 0: point source Bit 0 = 1: plane wave Bit 1..15: reserved  Bit 16 is used to indicate minimal latency mode which reduces the latency of the iosono algorithm at the cost of disabling focused sources Bit16 = 0: focused source enabled - normal latency Bit16 = 1: focused source disabled - minimal latency Bit 17..31: reserved
7	float32	4	0.5	Source position, azimuth	radiant or relative angle(see scaling chapter)
8	float32	4	0.5	Height Layer	Relative panning value
9	float32	4	4	Source position, radius	In meters or in relative distance (see scaling chapter)
10	float32	4	0.5	Volume	Linear scaling factor
11	float32	4	0.75	LFE Volume	Linear scaling factor
12	float32	4	1.2	Source delay	In seconds
13	int32	4	0	Scaling	0: off; 1: on
14	int32	4	1	On-screen-flag	0: off; 1: on
15	float32	4	0.2	Spreading	In percent [0 .. 100]
16	int32	4	0	Source trait	Bit field Bit 0: dialog Bit 1: reserved Bit 2: Aux1 Bit 3: Aux2 Bit 4-31: reserved
17	int32	4	-1	Bottom Layer ID	ID <= 0: The layer id of the bottom layer in the mixing setup
18	int32	4	2	Top Layer ID	ID >=0: The layer id of the top layer in the mixing setup

Abbildung 49: iosono Vertical Pan Protocol<sup>164</sup>

Folgende vereinfachte Methoden in Matlab ermöglichten das Aus-, und Einlesen sowie Überschreiben der Datenstruktur:

% Erstellen der Datentypen entsprechend des iosono Protocols

```

iosonoPacketPrototype.OSCAddress      = '/iosono/renderer/verticalpan/v1/src';
iosonoPacketPrototype.paddingByte1    = '0';
iosonoPacketPrototype.OSCdataTypes    = ',iiffffiifiii';
iosonoPacketPrototype.paddingByte2    = '0';
iosonoPacketPrototype.channelNumber    = int32(1);
iosonoPacketPrototype.algorithmHints   = int32(0);
iosonoPacketPrototype.sourcePosition1_Azimuth = single(0);
iosonoPacketPrototype.heightLayer      = single(0);
iosonoPacketPrototype.sourcePosition2_Radius = single(0);
iosonoPacketPrototype.volume           = single(1);
iosonoPacketPrototype.LFEVolume        = single(0);
iosonoPacketPrototype.sourceDelay      = single(0);
iosonoPacketPrototype.scaling          = int32(1);
iosonoPacketPrototype.onScreenFlag     = int32(0);
iosonoPacketPrototype.spreading        = single(0);
iosonoPacketPrototype.sourceTrait      = int32(0);
iosonoPacketPrototype.bottomLayerID    = int32(-1);
iosonoPacketPrototype.topLayerID       = int32(2);

```

<sup>164</sup> Barco Audio Technologies 2015a: 10

```

% Ausgabe der erstellten Daten-Struktur
%      Datatype      Size      Variable Name      Nr. of Elements/ByteOffset/isChar

'uint8' [1x2 double] 'OSCaddress' [35] [ 0] [1]
'uint8' [1x2 double] 'paddingByte1' [1] [35] [1]
'uint8' [1x2 double] 'OSCdataTypes' [15] [36] [1]
'uint8' [1x2 double] 'paddingByte2' [1] [51] [1]
'int32' [1x2 double] 'channelNumber' [4] [52] [0]
'int32' [1x2 double] 'algorithmHints' [4] [56] [0]
'single' [1x2 double] 'sourcePosition1_A?' [4] [60] [0]
'single' [1x2 double] 'heightLayer' [4] [64] [0]
'single' [1x2 double] 'sourcePosition2_R?' [4] [68] [0]
'single' [1x2 double] 'volume' [4] [72] [0]
'single' [1x2 double] 'LFEVolume' [4] [76] [0]
'single' [1x2 double] 'sourceDelay' [4] [80] [0]
'int32' [1x2 double] 'scaling' [4] [84] [0]
'int32' [1x2 double] 'onScreenFlag' [4] [88] [0]
'single' [1x2 double] 'spreading' [4] [92] [0]
'int32' [1x2 double] 'sourceTrait' [4] [96] [0]
'int32' [1x2 double] 'bottomLayerID' [4] [100] [0]
'int32' [1x2 double] 'topLayerID' [4] [104] [0]

% Abbilden der Struktur in einem Daten-Array
data = mapStructToData(iosonoPacketPrototype);

% Zum Test wandeln des Daten-Array zurück zur Struktur
myStruct = mapDataToStruct(data, iosonoPacketDescription);

% Sende an folgender Adresse und Port / Adresse des Iosono-Core
remoteAddress = '192.168.1.200';
remotePort = 4243;

% Höre auf folgenden Port / UDP-Port der SAW
localPort = 4244;

% Erstelle UDP Kommunikationsobjekt.
u1 = udp(remoteAddress, remotePort, 'InputBufferSize', 1024*8);

% While-Schleife zu Manipulation Daten
while(true)

% Adressierung des Audio-Objektes
iosonoPacketPrototype.channelNumber = int32(Audio-Objekt Nr. X);

% Manipulation der Audio-Objekt-Position Radius
iosonoPacketPrototype.sourcePosition2_Radius = single(f(x));

% Manipulation der Audio-Objekt-Position Azimuth
iosonoPacketPrototype.sourcePosition1_Azimuth = single(g(x));

% Manipulation der Audio-Objekt-Position Azimuth
data = mapStructToData(iosonoPacketPrototype);
fwrite(u1, data);
end

```

Mit Hilfe der hier beschriebenen Methoden war es möglich ein Audio-Objekt in der SAW zu positionieren und anschließend über Matlab in seiner Position über die Funktionen  $f(x)$  und  $g(x)$ , beispielsweise über Sinus-Funktionen in einer Kreisbewegung zu automatisieren. Der nächste Schritt wäre nun eine saubere und professionelle Implementierung eines Manipulation Daemons als Software-Applikation. Die Software müsste entsprechende Steuerschnittstellen bieten um die gewünschten Information der Interaktions-Interfaces wie iPad oder Sensor auszuwerten. Die Fortsetzung der Implementierung einer finalen Applikation für mobile Endgeräte hätte den Rahmen dieser Masterarbeit überschritten und wird an dieser Stelle nicht weiter ausgeführt.

#### 4.3.4 Zuspieler

##### Ausgangsmaterial

Die Arbeit beschreibt die Herstellung einer dreidimensionalen Mischung populärer Musik für ein Live-Konzert. Betrachtet man populäre Musik im Ausgangsmaterial, so handelt es sich meist um monophone Signale.

*„...it is possible to create believable three-dimensional immersion using conventional stereo spatial tools in a 22.2 multi-channel playback environment from monophonic multi-track source material.“<sup>165</sup>*

Von der Produktion des Albums des Interpreten werden sechs Titel ausgewählt die im dreidimensionalen Konzert präsentiert werden sollen. Dabei wird darauf geachtet, dass die ausgewählten Songs die bereits beschriebenen Kriterien erfüllen und eine entsprechende Produktion ermöglichen. Benötigte Zuspieler werden aus dem vorhandenen Audiomaterial des Albums angefertigt. Für die Mischung wird fast ausschließlich monophones Ausgangsmaterial, sowie zur Nachbearbeitung gewöhnliche Stereo-Effektgeräte genutzt. In der Mischung kommen dreidimensionale Panner zum Einsatz. Es wird auf Upmix-Algorithmen verzichtet.

##### Idee & Ziel

Wie bereits angesprochen wird in drei unterschiedliche Objektaufgaben der dreidimensionalen Mischung kategorisiert: Basisobjekte, umhüllende beziehungsweise füllende Objekte und plakative Objekte. Die Zuspieler, die im ersten Arbeitsschritt in Pro Tools erstellt werden, haben die Aufgabe, das dreidimensionale Panorama mit Dopplungen und Erweiterungen anzureichern und aufzufüllen. Es sind umhüllende Objekte. Die Zuspieler sollen das Gefühl vermitteln, dass der Zuhörer sich in einer Klangwelt befindet, die sich um ihn herum aufbaut, wobei hierbei der Fokus zur Bühne nicht verloren gehen soll. In der Mischung der Zuspieler in Pro Tools werden Panoramapositionen, Lautstärkeverhältnisse sowie Hall und Delay dazu genutzt, ein stimmiges Klangbild zu erschaffen, dass ein dezentes Klangbett bildet. Die Zuspieler finden im Hintergrund statt und sollen die Band auf passive Weise im dreidimensionalen Raum stützen. Der Zuhörer soll Zuspieler und Live-Material nicht differenzieren können.

*„The surround sound mixer should avoid mixing for one 'sweet spot'“<sup>166</sup>*

---

<sup>165</sup> Martin 2015: 6

<sup>166</sup> PEW 2004: 2

Das Basisklangbild soll an jeder zugelassenen Position im Raum ein im Rahmen stimmiges Klangbild für den Zuhörer ergeben. Die Zuspeler werden meist so angeordnet, dass ein komplementäres Klangbild entsteht, sodass man an jeder Position die für den Zuhörer wichtige Informationen hören kann. Das bedeutet, wenn an der entsprechenden Raumposition keine Dopplung platziert wird, ist an dieser Stelle ein anderes Objekt platziert, das eine ähnliche Funktion übernimmt und im Panorama bezogen auf den Sweetspot ein Gleichgewicht herstellt. Somit wird versucht sicherzustellen, dass jeder Zuhörer zu jeder Zeit an jeder zugelassenen Position im Raum ein 'hörbares' und ästhetisch hörrichtiges Hörerlebnis geboten bekommt.

Pro Tools unterstützt in der aktuellen Version kein natives dreidimensionales Panning. Nativ unterstützt wird von Avid ein System bis zu einer Ausbaustufe von maximal 7.1 Surround auf einer Ebene. Das dreidimensionale Panning kann mit Hilfe von Plug-ins von Drittanbietern, wie beispielsweise dem Dolby Atmos Panner, den Auro 3D Authoring Tools, oder auch dem Spatial Audio Designer realisiert werden. Die Plug-ins übernehmen das Routing und die Busstruktur im Programm und ermöglichen somit ein dreidimensionales Panning in der DAW. Die Grundidee ist es, die Zuspeler im Vorherein im dreidimensionalen Raum anzuordnen und sie während des Live-Konzertes ergänzend einzusetzen.

Als Panning Tool sollten ursprünglich die Auro 3D Authoring Tools zum Einsatz kommen. Der Auro 3D Panner ermöglicht ein freies Panning im dreidimensionalen Raum. Das Plug-in ist nicht echtzeitfähig, somit müssen die Signale vor dem Konzert gerendert vorliegen. Um das Panning im Plug-in realisieren zu können, wird das komplette Routing wie bereits in *Kapitel 3.1.4* beschrieben in der DAW von den Auro 3D Authoring Tools abgewickelt.

### **Systemschwächen**

Die Auro 3D Authoring Tools ermöglichen ein intuitives Panning auf jede dem System mögliche Position und fördern somit den kreativen Umgang mit der Positionierung der Signale im dreidimensionalen Raum. Sie verlangsamen jedoch viele grundlegende Arbeitsschritte auf Grund der aktuellen Architektur der Plug-ins, denn verschiedene Parameter der DAW werden nicht an das Plug-in übergeben. Einfache Funktionen wie Solo, Mute oder PegelEinstellung können nur im Plug-in vorgenommen werden. Ein schnelles und übersichtliches Arbeiten sowie ein schneller Zugriff auf die Positionierung der Signale ist nicht gegeben. Auf einfache Befehle sowie ein visuelles Feedback der Einstellungen kann nur in zweiter Ebene, im geöffneten Plug-in, zugegriffen werden. Der Arbeitsablauf wird verlangsamt und verliert an Komfort sowie Agilität, und damit an Effizienz.

Zu einem späteren Zeitpunkt der Produktion wurde außerdem festgestellt, dass ein Eingriff in die Anordnung der einzelnen Signale der Zuspierer im dreidimensionalen Panorama zu jedem Zeitpunkt wichtig ist. Aus diesen Gründen wurde in der Vorproduktion der Zuspierer auf dreidimensionale Panning-Tools verzichtet. Das Panning wird mit Hilfe von zwei 4.0 Surround Pannern in der unteren und oberen Ebene des Auro 3D Abhörsystems realisiert. Dabei wird auf eine variable z-Positionierung verzichtet.

### **Umgebung: Vorproduktion der Zuspierer in Pro Tools**

Das Setup für den ersten Schritt der Vorproduktion besteht aus einem Pro Tools System und einem Auro 3D Abhörsystem. Benutzt wird ein Pro Tools 11 HDX System, sodass ausreichend Audiospuren und Auxwege, sowie Ein- und Ausgänge zur Verfügung stehen und Surroundpanning möglich ist.

### **Sessionaufbau**

Die Signale werden auf Grund des auf 64 Eingangskanäle begrenzten Iosono-Prozessors mit Hilfe des Pro Tools eigenen Panners in x- und y-Richtung räumlich angeordnet und damit auf eine geringere Kanalanzahl reduziert. Aus den ca. 80 Audiosignalen werden Quads erstellt, um die Signale auf sinnvolle Subgruppen zu reduzieren und sie dann zu einem späteren Zeitpunkt der Produktion dem Iosono-Prozessor zuzuführen. Folgende Subgruppen liegen vor:

- Bo Gitarre (Down)
- Bo Gitarre (Up)
- Jonas Gitarre(Down)
- Jonas Gitarre (Up)
- Synthesizer
- Backings (Down)
- Backings (Up)
- Chor

Die Subgruppen werden instrumentenspezifisch kategorisiert, um die Kontrolle und Übersicht zu steigern. Das Auro 3D Abhörsystem wird in zwei Quads unterteilt. Ein Quad befindet sich im oberen Layer, ein anderer Quad im unteren Layer des Auro 3D Systems. Ein Panning zwischen den beiden Ebenen wird in diesem Arbeitsschritt ausgeschlossen. Mit Hilfe der SAW und des Iosono-Prozessors werden die Quads zu einem späteren Zeitpunkt im dreidimensionalen Raum angeordnet. Der

Centerkanal sowie der LFE werden in der Vormischung der Zuspieler nicht genutzt. In den Centerkanälen sollen sich während des Live-Konzertes auf Grund von besserer Differenzierbarkeit ausschließlich Live-Signale befinden. Das Bassmanagement wird durch das Isoono-System realisiert.

Die Grundstruktur der Zuspieler ist bei allen Songs gleich und statisch. Zu den Dopplungen und Ergänzungen des Live-Materials gehören Rhythmusgitarren, Synthesizer, E-Bass, Chöre und Gesangsdopplungen. Die Signale werden hinsichtlich Räumlichkeit und Pegel so angeordnet und mit Hall und Delay angereichert, dass ein homogenes Klangbild entsteht, das das Live-Material unauffällig erweitert.

Mit Hilfe von drei extra dafür angelegten Monoauxwegen, auf die das Signal parallel geroutet werden kann, lassen sich gezielt einzelne Elemente im Song mit Hilfe der Spatial Audio Workstation bewegen. Damit können dynamische Objekte erzeugt werden. In der SAW kann der jeweilige Auxweg im Panorama automatisiert werden. Beispielsweise gibt es Synthesizer, die sich passend zu ihrem Spiel von vorne nach hinten durch den Raum am Zuhörer vorbei bewegen. Die Bewegung wird im Studio vorproduziert und während des Live-Konzertes per Timecode synchron abgespielt.

#### **4.3.5 Live-Inhalt**

##### Idee & Ziel

Die technische sowie ästhetische Bearbeitung der Live-Signale soll im Vorfeld vorbereitet werden. Für die Vorbereitung werden im ersten Schritt die Rohsignale der CD Produktion verwendet, die im späteren Konzert live performt werden. Diese Rohsignale werden in die passenden Audiospuren der Pro Tools Session eingesetzt durch die während des Konzertes die Signale der Band in Echtzeit prozessiert werden. Somit können grundlegende Entscheidungen im Vorfeld getroffen werden.

Durch die Arbeit mit den Rohsignalen der CD kann eine erste Vorbereitung sowie eine Einarbeitung in das System ohne Band ermöglicht werden. Dadurch ist die Vorproduktion an keinen festen zeitlichen Rahmen gebunden. Nachteil dabei ist, dass die Rohsignale der CD nicht identisch zu den später live gespielten Signalen der Musiker sind. Es wird nachkorrigiert werden müssen.

### **Umgebung: Vorproduktion des Live-Inhaltes in Pro Tools**

Für die technische und ästhetische Bearbeitung der Live-Signale in Echtzeit wird das Pro Tools HDX System genutzt. Bei der Vorbereitung des Live-Inhaltes wird ein Auro 3D Abhörsystem genutzt. Mit Hilfe des DSP gestützten Systems können Audibearbeitungen mit sehr geringer Latenz realisiert werden. Die meisten in der DAW zum Einsatz kommenden Plug-ins sind DSP gestützte Versionen, nur wenige Ausnahmen sind nativ<sup>167</sup> basiert. Im Fall, dass kein vergleichbares DSP gestütztes Plug-in existiert, muss auf ein natives Plug-in zurückgegriffen werden. Pro Tools HDX verbietet den Einsatz von nativen Plug-ins im Insert einer Audiospur bei gleichzeitigem Inputmonitoring. Das bedeutet, dass bei den in Echtzeit zu bearbeitenden Inputkanälen nur DSP gestützte Plug-ins zum Einsatz kommen können. Soll ein natives Plug-in in Echtzeit genutzt werden, so wird in der Session ein Auxweg mit dem jeweiligen Effekt erstellt.

Das Pro Tools System arbeitet für die Produktion mit einer Systemlatenz von 256 Samples. Der Latenzausgleich im Programm ist aktiviert. Das bedeutet, dass das Programm alle Signale und deren Latenz durch die Bearbeitung durch verschiedene Plug-ins bis zu einem maximalen Versatz von 256 Samples aneinander angleicht.

### **Mikrofonierung**

Es wird das Einzelmikrofonverfahren angewendet. Ziel dieses Verfahrens ist die weitgehende akustische Trennung der Einzelsignale. Grund dafür ist, dass sich die einzelnen Schallquellen in ihrer natürlichen Lautstärke stark unterscheiden und sich somit kein ausbalanciertes Klangbild einstellt.

*„Es fordert von der Tonregie in einem hohen Maße die Gestaltung des Klangbildes, das in der akustischen Realität des Aufnahmerraums nicht vorhanden ist. Tonmeister und Toningenieure werden gerade bei dieser Aufnahmetechnik in hohem Maße künstlerische Mitgestalter“<sup>168</sup>*

---

<sup>167</sup> Native Plug-Ins werden auf dem Prozessor des Computers prozessiert.

<sup>168</sup> Dickreiter 1997: 308

Des Weiteren sollen die mikrofonierten Signale teils stark nachbearbeitet werden. Eine akustische Trennung kann mit Hilfe von distanzierter Anordnung der Schallquellen im Raum, sowie gerichteten Mikrofonen und Trennwänden erzielt werden. Durch einen geringen Mikrofonabstand spielt die Akustik des Aufnahmeraums eine zu vernachlässigende Rolle.<sup>169</sup>

### **Sessionaufbau**

Der live gespielte Inhalt hat in der Mischung eine statische Grundstruktur. Der Zuhörer hat ein in der Basis gleichbleibendes Hörbild der live gespielten Schallereignisse passend zur visuellen Präsentation. Das Hörbild des Live-Inhalts wird nur in gezielt ausgewählten Momenten bewusst gebrochen. Es sind Basisobjekte.

Folgende Signale werden durch die Band live performt:

- akustisches Schlagzeug
- elektronisches Drumpad (Doppel Quadro)
- stereo Gitarre 1
- stereo Gitarre 2
- Bass 1
- Bass 2
- Synthesizer
- Vocals 1
- Vocals 1 Effekt
- Vocals 2

Die Signale werden in der selben Pro-Tools-Session wie die Zuspeler technisch und ästhetisch in Echtzeit mit Hilfe von unterschiedlichen Plug-ins bearbeitet. Zu diesem Zeitpunkt wird die räumliche Anordnung der Signale ohne Panning gelöst. Die Signale werden hart auf die gewünschten Ausgangskanäle des Auro 3D Setups geroutet. In Pro Tools wird für jedes Eingangssignal eine Audiospur erstellt. Folgende, der aus *Anhang III* und *Anhang IV* zu entnehmenden Signale, liegen an Pro Tools an.

---

<sup>169</sup> vgl. Dickreiter 1997: 307

## Panorama

Die akustischen Signale der Protagonisten werden passend zu ihrer visuellen Position angeordnet. Mittig sitzt das Schlagzeug. Links und rechts daneben stehen die beiden Sänger, sowie deren Gitarren. Der Synthesizer sowie das elektronische Drumpad bilden eine Ausnahme. Sie werden meist als plakative Objekte genutzt. Das Panning dieser Objekte entspricht nicht der visuellen Position. Wie bereits angesprochen ermöglichen elektronische Tonerzeuger ein unrealistisches Panning, ohne dass es für den Rezipienten unnatürlich erscheint. Die Positionierung der akustischen Schallquellen entspricht und verdeutlicht den Platz der Musiker im Raum. Der Gesang wird während der Vorbereitungen im Studio auf den Centerkanal geroutet und zu einem späteren Zeitpunkt der Produktion mit Hilfe des Isoono-Systems auf die beiden ihnen zugeordneten Centerkanäle gelegt. Die Lokalisation wird dadurch positionsunabhängig. Die Gitarren der Musiker speisen jeweils nach ihren stereophonen Gitarreneffektgeräten zwei Verstärker. Das erlaubt es, die Gitarren jeweils als Stereoquelle im Hörbild anzuordnen. Mit Hilfe von Laufzeit- sowie Frequenzanpassung werden die Stereosignale der Gitarren in die jeweilige visuell passende Richtung positioniert. Die beiden Bilder zeigen beispielhaft die frequentielle und zeitliche Bearbeitung eines der beiden Stereosignale für eine für den Zuhörer klare Lokalisation.

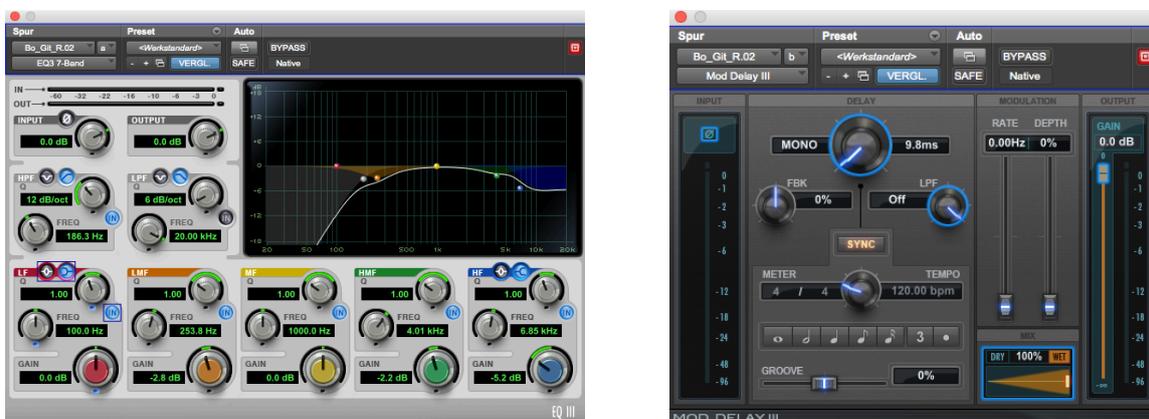


Abbildung 50: Lokalisationsbearbeitung<sup>170</sup>

<sup>170</sup> Screenshot

In der Pro-Tools-Session werden folgende Auxwege für das Livematerial angelegt. Sie werden als Hall, Delay, Subgruppen oder Trigger genutzt. Subgruppen werden dann gebildet, wenn Einzelsignale von einer gemeinsamen Bearbeitung profitieren. Angelegt werden folgende Auxwege mit den in der Tabelle aufgelisteten Effekten:

Auxweg	Effekt	Auxweg	Effekt	Auxweg	Effekt
Bo Gitarre FX 1	Delay	Jonas Vox FX 1	Reverb: Plate	Bo Vox FX 1	Reverb: Plate
Bo Gitarre FX 2	Delay	Jonas Vox FX 2	Reverb: Plate	Bo Vox FX 2	Reverb: Plate
Sologitarre FX 3	Delay	Jonas Vox FX 3	Delay	Bo Vox FX 3	Delay
Jonas Gitarre FX 1	Delay	Jonas Vox FX 4	Delay	Bo Vox FX 4	Delay
Jonas Gitarre FX 2	Delay	Special FX 1	Pitched Feedback Reverb	SN Trigger 1/2/3	Trigger
Drum Bus	Subgruppe	Special FX 2	Pitched Feedback Reverb		

Die jeweiligen Effekte werden im Panorama komplementär zum Direktsignal gelagert, um eine künstliche Tiefe beziehungsweise Breite zu erzeugen. Gleiche Effekte haben geringe Unterschiede in den eingestellten Parametern und sind auf unterschiedliche Raumpositionen im Panorama platziert. Wenn ein Instrument beispielsweise mehrere Delays zugeordnet bekommt, so sind diese im Timing unterschiedlich angepasst. Sie sind außerdem auf unterschiedliche Panoramapositionen im Raum platziert. Damit überlagern sie sich nicht, sondern es entstehen unterschiedliche Delays in unterschiedlichen Panoramarichtungen, die als Einheit agieren.



Abbildung 51: Valhalla Shimmer Plug-in<sup>171</sup>

Eine Besonderheit der in der Session verwendeten Sendeffekte ist das Valhalla-Shimmer-Plug-in. Es ist ein algorithmischer Hall. Die Besonderheit ist die Pitch-, sowie Feedbackfunktion. Das Plug-in kann das Inputsignal nach Wunsch gleichzeitig nach oben und unten pitchen und dabei eine Feedbackschleife erzeugen. Das Ergebnis sind lange, je nach Einstellung, harmonisch 'glitzernde'

<sup>171</sup> Screenshot

Hallfahnen. Das Shimmer-Hall-Plug-in wird in der Session dazu genutzt, eine lang nachklingende 'glitzernde' Fläche zu erzeugen, die die Signale harmonisch zum Song passend weich einbettet.

## Instrumentenbearbeitung

Die Einzelsignale der Instrumente werden mit Hilfe von Kompressoren, Filtern und Verzerrern technisch und gestalterisch bearbeitet. Die technische Bearbeitung mit Hilfe von Filtern reduziert störende und unwichtige Frequenzanteile. Mit Hilfe von Kompressoren kann eine technische dynamische Anpassung auf ein für den Zuhörer songdienliches Level erreicht werden. Die gestalterische Klangformung der Signale steht in Abhängigkeit mit der technischen Bearbeitung.

Eine Besonderheit in der Bearbeitung stellt das Schlagzeug dar, dass folgend genauer beschrieben werden soll. Das Schlagzeug besteht aus mehreren Einzelsignalen. Neben denen, die mikrofoniert wurden, gibt es Trigger-Signale, die bei Bedarf dem Originalsignal beigemischt werden können. Jede Samplespur triggert ein eigenes für den Song dienliches Sample, das je nach Gebrauch dazu



Abbildung 52: Steven Slate Trigger Plug-in<sup>172</sup>

gemischt werden kann. Somit können verschiedene Effekte erzeugt werden. Das zusätzliche Snaresignal wird mit Hilfe des Slate-Digital-Trigger-2-AAX-Plug-in ausgelöst. Dabei liegt das Mikrofonsignal der Snare in Pro Tools an drei Auxkanälen an, die jeweils im Insert einen Slate Digital 2 Trigger haben. Die Trigger laufen mit einem Sampleversatz von ca. 170 Samples. Das sind bei einer Abtastrate von 44,1 kHz ca. 3,8 Millisekunden. Die einzelnen Schlagzeugsignale laufen auf einer gemeinsamen Schlagzeug-Subgruppe zusammen. Die hauptsächliche Klanggestaltung des Schlagzeugs findet in dieser Subgruppe statt, die als Stereogruppe angelegt wird. Diese Vorgehensweise hat den Vorteil, dass die Auswahl der Plug-ins zur Bearbeitung wesentlich größer

<sup>172</sup> Screenshot

ist, als wenn eine LCR-Gruppe anlegt würde. Der dabei entstehende Nachteil ist die Abbildung der Mitte des Schlagzeuges als fiktive Schallquelle. Solche sogenannten Phantomschallquellen sind in ihrer Lokalisation störanfällig. Findet eine Bewegung des Zuhörers statt, so bleibt die Phantomschallquelle nicht ortungsstabil.

*„Audio engineers mixing in stereo have long relied on the ability to create a "phantom" center by placing equal amounts of signal in the left and right channels. This creates the illusion of the sound appearing mid-way between the two speakers, but it is a fragile image that shifts closer to one speaker or the other as the listener moves out of the sweet spot.“<sup>173</sup>*

Das Schlagzeug wird wie gewohnt mittig vor dem Zuhörer platziert. Um ein für den Zuhörer stabiles Hörbild zu schaffen, wird ein Centerspeaker dazu genutzt, das Signal zu stützen. Die akustische Positionierung soll auf jede im Raum bezogene Position ähnlich wirken. Das Snare- und Basedrumsignal sind die für die Ästhetik wichtigsten Signale und werden deshalb durch den Centerkanal gestützt.

Das Schlagzeug wird als geschlossener Klangkörper bearbeitet. Hierbei kann Kompression, Verzerrung und Filterung auf das ganze Set eine andere Wirkung haben, als im Einzelsignal, da sie sich untereinander beeinflussen. Bearbeitet man die Einzelsignale des Instruments, so bearbeitet man eine Trommel. Soll das Schlagzeug als geschossenes Instrument gestaltet werden, so muss die Subgruppe bearbeitet werden. Es wird in zwei Schritten bearbeitet: zuerst werden die Einzelsignale gesäubert und ästhetisch aufbereitet, sodass sie die gewünschte Klangfarbe nach Bearbeitung der Gruppe als geschlossenes Instrument haben. Die Bearbeitung der Einzelsignale und der Schlagzeuggruppe ist ein komplementärer und iterativer Arbeitsvorgang.

### **Latenzzeiten und Livekompatibilität**

Digitale Bearbeitung von Audiosignalen benötigt Zeit, um den Plug-in Algorithmus zu berechnen. Es kommt zu Latenz-, soll heißen Verzögerungszeiten. Es können zweierlei Probleme auftreten. Einerseits verschiebt sich ohne ein Ausgleichen der Latenzzeiten das musikalische Timing, andererseits können Phasenprobleme bei korrelierten Signalen hervorgerufen werden. Die automatische Delay-Kompensation von Pro Tools kann je nach Einstellung der Größe des Sample-Buffers automatisch die Latenzzeiten der einzelnen Arbeitsschritte aufeinander abstimmen. Alle Signale bleiben zeit- und phasenkorrekt zueinander.

---

<sup>173</sup> PEW 2004: 4-3

Wenn das Programm zur Echtzeitbearbeitung genutzt werden soll, ist darauf zu achten, dass ein gewisser Wert an Verzögerung nicht überschritten wird. Denn in einem solchen Fall wären damit verschiedene Nachteile verbunden:

- Das über das System gehörte Monitorsignal ist zum eigentlich vom Musiker gespielten Signal derart verzögert, dass ein reibungsloses Spiel eingeschränkt wird.
- Das vom Zuschauer gehörte Signal ist asynchron zum visuellen Inhalt.
- Das vom Zuschauer über die PA gehörte Signal ist zum akustischen Signal des Instrumentes derart verzögert, dass für den Zuhörer nachteilige Artefakte entstehen.

Es stellt sich die Frage, welche maximale Latenzzeit der Zuhörer und Zuschauer tolerieren kann: Schall breitet sich im vereinfachten Beispiel mit einer Geschwindigkeit von 343 m/s aus. Das bedeutet, dass der Schall nach ca. 2,92 ms einen Weg von einem Meter zurückgelegt hat. Ein Schallereignis mit einer Distanz zum Hörer von einem Meter, wird um 2,92ms verzögert wahrgenommen.

*„In einem musikalischen Kontext betrachtet bedeutet es, dass zwei 1/32teln bei einem Tempo „Allegretto“ von 120bpm 62,5 ms voneinander entfernt sind.“<sup>174</sup>*

Befindet sich beispielsweise der Schlagzeuger in 50 cm Entfernung zum Instrument, so hört er den Klang seines Schlagzeuges um ca. 1,46 ms verzögert.

Die Produktion wird mit einer Abtastrate von 44,1 kHz ausgeführt. Das bedeutet, dass das System mit einer Latenz von einem Sample umgerechnet ca. 0,02268 ms Versatz erzeugt. Der Latenzausgleich in Pro Tools ist aktiviert. Die Session wird mit einem Samplebuffer von 256 Samples bei einer Abtastrate von 44,1 kHz um maximal ca. 4,8 ms verzögert. Das würde sich auf den Musiker so verhalten, als würde er sein Instrument in ca. zwei Meter Entfernung vor sich wahrnehmen und sollte sich somit im spielbaren Rahmen für den Musiker befinden. Um Letzterem jedoch ein möglichst latenzfreies Monitoring zu ermöglichen, wird das für den Musiker hörbare Signal direkt am RME Madiface abgegriffen und damit der Monitormix der Musiker erstellt. Eine genauere Beschreibung des Setups folgt im späteren Verlauf der Arbeit.

---

<sup>174</sup> San Segundo 2008

### 4.3.6 Liveproduktion in Iosono

Im Iosono-System werden die Positionsdaten aller Audiosignale für das Konzert verwaltet. Iosono ist für die Steuerung des Pannings aller für den Zuhörer hörbaren Audiosignale im dreidimensional beschallten Raum zuständig.

#### Idee

Die einzelnen Audiosignale sind zu diesem Zeitpunkt technisch und ästhetisch aufeinander abgestimmt. Änderungen werden auf Grund der ständigen Weiterentwicklung der Mischung nötig sein. Die bereits mit Pro Tools und dem Auro-3D-Abhörsystem vorbereitete Liveproduktion wird in folgendem Arbeitsschritt mit Hilfe eines Iosono-Prozessors räumlich dreidimensional für das Konzert aufbereitet. Der Iosono-Prozessor übernimmt das dreidimensionale Panning. Für diesen Arbeitsschritt wird ein Testaufbau in kleinerer Ausführung gemacht. Es wird ein 15.2 Abhörsystem für die weitere Mischung der Pannings eingerichtet. Zu diesem Zeitpunkt wird die Band eingeladen, um das Setup für den effektiven Liveeinsatz zu testen.

### 4.3.7 Systemvorstellung: Vorproduktion der Liveproduktion in Iosono

Die in Pro Tools gemischten Audiosignale werden über 64 Kanäle direkt an den Iosono-Prozessor geschickt. In Nuendo wird für jeden Eingangskanal eine Audiospur angelegt. Mit jedem in Nuendo angelegten Audiotrack wird ein statisches Kanalobjekt, als virtueller Lautsprecher in der SAW erzeugt. Mit Hilfe der Spatial Audio Workstation kann der Iosono-Prozessor ferngesteuert werden.

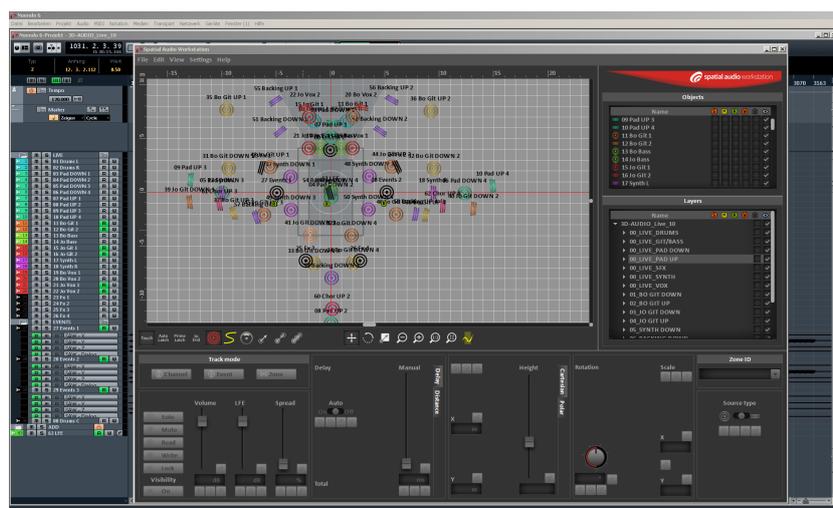


Abbildung 53: Iosono SAW - Audio-Objekte<sup>175</sup>

<sup>175</sup> Screenshot

Die in Pro Tools gemischten Audiosignale werden über 64 Kanäle direkt an den Iosono-Prozessor geschickt. In Nuendo wird für jeden Eingangskanal eine Audiospur angelegt. Mit jedem Audiotrack wird ein statisches Kanalobjekt als virtueller Lautsprecher in der SAW erzeugt. Die in Pro Tools bereits quadrofonisch angeordneten Events werden mit Hilfe der SAW im dreidimensionalen Raum mit beliebiger Skalierung und Positionierung in x-, y- und z-Richtung angelegt. Die live gespielten Signale werden ebenfalls mit Hilfe der SAW und des Iosono-Prozessors im Raum passend zu den Protagonisten positioniert.

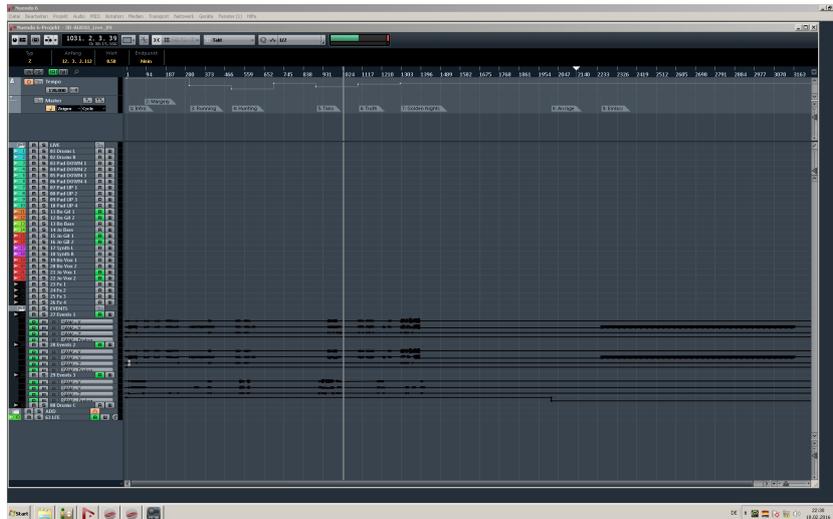


Abbildung 54: Nuendo SAW - Automationsdaten<sup>176</sup>

Für die plakativen Objekte der Mischung, die dem Zuhörer den dreidimensional beschallten Raum bewusst aufzeigen sollen, wird in der SAW ebenfalls je ein Kanalobjekt angelegt. Diese drei Kanalobjekte werden nach Bedarf durch vorgeschriebene Automationen bewegt.

## LFE

Für jedes im SAW angelegte Objekt kann im Objektparameter-Fenster, wie in gewöhnlichen Systemen mit LFE-Controllern der LFE-Anteil festgelegt werden. Dadurch kann der Bassanteil im Projekt für einzelne Objekte gesteuert und automatisiert werden. Während des Konzertes kann der Bassanteil verschiedener Signale und der damit verbundene LFE gestalterisch genutzt werden. Neben der direkten Kontrolle des LFE Kanals, kann das System die Subwoofer zur verbesserten Frequenzwiedergabe nutzen.

<sup>176</sup> Screenshot

## **Latenzzeiten und Livekompatibilität**

Wenn in Nuendo ein Stereo Audio- oder Gruppenkanal angelegt wird, so entspricht diesem in der SAW ein Upmix-Objekt. Ein solches Objekt ist kein virtueller Lautsprecher, sondern es transformiert das Stereosignal mit Hilfe des Upmixalgorithmus der Iosono-Software in ein dreidimensionales Audioevent. Upmix-Objekte können im Live-Setup nicht genutzt werden, da die dabei verursachte Latenz für den Konzertkontext zu groß ist.

In der Veranstaltungsstätte hat der am nächsten zur Band stehende Zuschauer einen Abstand von ca. 2 Metern und der am weitesten entfernte Zuschauer einen Abstand von ca. 10 Metern. Die durchschnittliche Distanz zur Band beträgt somit 6 Meter. Dieser gemittelte Abstand wird als Referenzpunkt genutzt, der angibt, auf welche Distanz das System maximal verzögern darf. Bei einer angenommenen Schallgeschwindigkeit von 343 m/s dauert es ca. 17,49 ms bis der abgestrahlte Schall 6 Meter zurückgelegt hat. Sehr einfach betrachtet bedeutet es, dass die auditive Wahrnehmung zur visuellen ca. 11,66 ms verzögert ist. Hierbei wird die Zeit der menschlichen Verarbeitung von auditivem und visuellem Inhalt vernachlässigt. Die Latenzzeit des Nuendo Systems und der Iosono-Beschallung darf eine maximale Latenzzeit von ca. 17,49 ms nicht überschreiten. 17,49 ms entsprechen bei einer Abtastrate von 44,1 kHz ca. 771 Samples. Rechnet man die Latenz des Pro-Tools-Systems mit ein, so kann das Nuendosystem mit einer Latenz von ca. 515 Samples arbeiten, ohne dass es dem Zuschauer im Mittel negativ auffallen wird.

## 5 Der Selbstversuch: Umsetzung des 3D-Konzerts

### 5.1 Die Mischung: exemplarisch beschrieben am Song „Circles“

Im Folgenden wird die Mischung und das zuvor beschriebene Setup am praktischen Beispiel des Songs 'Running' im Detail erläutert. Das Beispiel soll ein besseres Verständnis des Setups und der Mischung ermöglichen, sowie einen detaillierten Einblick in die Produktion geben.

Der Song 'Running' hat elektronische und akustische Elemente. Die Struktur der Produktion und des Songwriting ermöglichen eine gestalterisch aufbauende, sowie übernatürliche Größe in der Mischung.

#### 5.1.1 Panorama

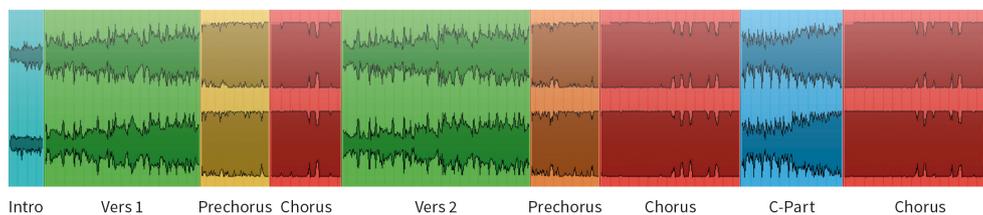


Abbildung 55: Circles - Songstruktur<sup>177</sup>

Die Abbildung zeigt den Songablauf. Die Farben unterteilen den Song in Intro, Vers, Prechorus, Chorus und C-Part. Wie bereits angesprochen, werden die einzelnen Songteile durch Nutzung des Panoramas voneinander abgetrennt. Durch die abschnittsbezogene Nutzung des dreidimensional beschallten Raums kann der ausgedrückte Spannungsbogen des Songs verstärkt werden. Das dreidimensionale Panorama wird wie folgt genutzt, um die Songteile voneinander abzugrenzen. Dabei sind Hall- und Delayanteile, die den Signalen eine räumliche Definition geben, ausgeschlossen. Diese Effekte haben eine statische Position im dreidimensionalen Panorama. Ausgenommen sind ebenfalls Effektsignale, also plakative Objekte.

Das Intro wird fast ausschließlich mono gehalten. Der Vers wird stereolastig gemischt. Im Übergang zum Prechorus füllt für einen Moment ein halliges Sample den dreidimensional beschallten Raum. Der PreChorus selbst soll sich hauptsächlich in einer Ebene des Isonos abspielen, die Höhenkanäle werden bis auf Hall- und Delayeffekte nicht genutzt. Im Chorus soll die

---

<sup>177</sup> Eigene Darstellung

komplette Dimension des Setups genutzt werden und es werden bewusst Direktsignale in z-Richtung positioniert. Alle Übergänge sind möglichst hart und auffällig gestaltet.

### 5.1.2 Arrangement

Das Schaubild zeigt neben dem Songablauf die Instrumentierung der einzelnen Songteile. Jede Instrumentengruppe hat eine Aufgabe in der Mischung. Mit Hilfe der Panoramaeinstellungen, Hall und Delay, sowie Dopplungen soll diese Aufgabe unterstützt werden.

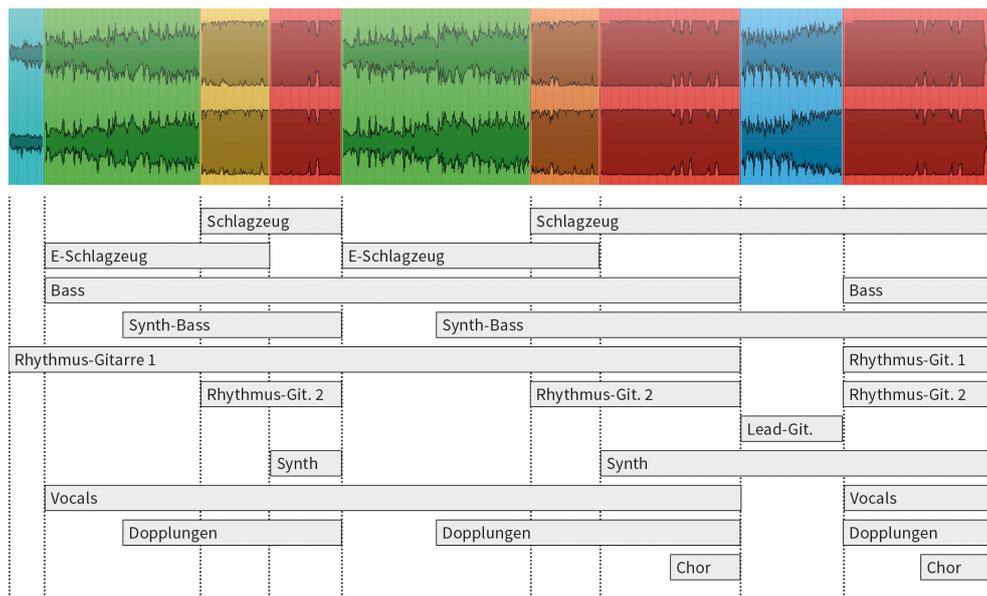


Abbildung 56: Circles – Instrumentierung<sup>178</sup>

Im Schaubild wird zwischen live und zugespiltem Material unterschieden. Zuspielder und Live-Material, die gestalterisch zueinander gehören, wie beispielsweise Rhythmusgitarre 1 und deren Dopplungen, sind in der gleichen Farbe dargestellt.

#### Schlagzeug

Die rhythmische Führung übernimmt das Schlagzeug. Es ist im Song in drei Bereiche unterteilt. Die Verse werden durch das E-Drum geführt, der Prechorus wird mit E-Drum und echtem Schlagzeug gespielt, der Chorus und der Cpart werden ausschließlich durch das akustische Schlagzeug begleitet. Die Intensität und Dichte des Schlagzeuges wird passend zum Song aufgebaut. Im Chorus ist es am intensivsten und dichtesten, um die Größe und Schwere des Refrains zu

<sup>178</sup> Eigene Darstellung

unterstreichen. Es liegt passend zu seiner visuellen Position vorne im Panorama. Um dem Schlagzeug eine gefühlte räumliche Tiefe zu verschaffen, speist es parallel einen Hallraum, der nach hinten oben gepanned ist. Der genutzte Hallraum ist ein mittelgroßer, mittig klingender Raum. Die Reflexionen durch den Hall von hinten sollen das Schlagzeug auf eine natürliche Art und Weise größer erscheinen lassen.

Die Einzelsignale des Schlagzeugs werden mit Hilfe von Filter und Gate gesäubert, um einen besseren Zugriff in der gestalterischen Einzelsignal sowie der Gruppenbearbeitung zu ermöglichen. Das Noise-Gate ist ein Regelverstärker, der den Übertragungsweg unterhalb eines einstellbaren Pegels sperrt. Damit kann das Nutzsignal vom Störsignal weitgehend isoliert werden.<sup>179</sup> In den Einzelsignalen werden die Trommeln gated. Außerdem werden mit Hilfe von Filtern ungewollte Resonanzen sowie störendes Low- und Highend reduziert. Darauf folgend werden Basedrum sowie Snaresdrum ästhetisch geformt. Um die Basedrum größer wirken zu lassen, hebt ein Filter bei 60 Hz 18 db mit einer Güte von 1,4 an. Die Basedrum erhält ein voluminöses Lowend, bleibt aber auf Grund der geringen Güte von 1,4, die in der Bandbreite etwa einer Oktave entspricht, natürlich im Klang. Um den Nebeneffekt dieser breiten Anhebung, die Anhebung der Frequenzen über 60 Hz, zu kontrollieren, wird bei 160 Hz breitbandig abgesenkt. Damit wird der pappig klingende Anteil der Basedrum reduziert.



Abbildung 57: Circles - Basedrum Bearbeitung<sup>180</sup>

Die Snaresdrum wird nach der Säuberung des Nutzsignals mit Hilfe eines Kompressors dynamisch und harmonisch bearbeitet. Der Kompressor deformiert das Signal mit Hilfe der extrem kurz eingestellten Regelzeiten. Auf Grund der Attackzeit im Mikrosekundenbereich entwickelt der

<sup>179</sup> Dickreiter 1997: 400

<sup>180</sup> Screenshot

Kompressor mit jedem Ansprechen ein dezentes Knackgeräusch, das den eingehenden Transienten der Snare prägt. Ein nachgeschalteter Verzerrer bearbeitet das Signal weiter harmonisch und dynamisch. Der Verzerrer rundet den Transienten ab und lässt die Snares drum zwar weicher, aber auch voluminöser wirken. Im gleichen Moment werden mit Hilfe des Verzerrers harmonische Obertöne generiert, die dem Signal eine größere Durchsetzungskraft verleihen.



Abbildung 58: Circles - Snare Bearbeitung<sup>181</sup>

Das klangästhetisch geformte Snaresignal wird anschließend mit Hilfe eines kurzen Halls angereichert. Dieser Hall lässt das Snaresignal breiter wirken. Er wird durch einen nachgeschalteten Kompressor verdichtet und dynamisch passend zur Geschwindigkeit des Songs geregelt.

<sup>181</sup> Screenshot

In der Gruppe wird das Schlagzeug als geschlossenes System komprimiert und gefiltert sowie mit Hilfe von mehreren Kompressionsstufen dynamisch gestaltet. Auf Grund der Regelzeiten der Kompressoren wirkt das Schlagzeug nach der Bearbeitung druckvoller und massiver.



Abbildung 59: Circles - Drumgruppe Bearbeitung<sup>182</sup>

Die Kompressoren arbeiten mit schnellen Regelzeiten, um eine auffällige Dynamikverminderung hervorzurufen. Kurze Pegelspitzen werden reduziert. Das Originalsignal erfährt eine Verringerung der Impulshaftigkeit, jedoch wirkt das Schlagzeug auf Grund der starken Klangverdichtung und Lautheitserhöhung intensiver und massiver. Im tieffrequenten Bereich kann es zur Signaldeformation kommen und dabei nichtlineare Verzerrungen bis hin zu Knackgeräuschen erzeugen. Die Release-Zeit der Kompressoren ist kurz eingestellt, um das Pumpen und damit die Rauigkeit zu erhöhen. Die Modulationsartefakte werden gestalterisch genutzt, um dem Schlagzeug Größe und Schwere zu verleihen. Am Ende wird die Schlagzeuggruppe mit einem

<sup>182</sup> Screenshot

Equalizer in den Bässen und Höhen erweitert. Der Subbassbereich bei 20 Hz wird dezent reduziert, um das Schlagzeug in den tiefen Frequenzen weniger schwammig wirken zu lassen. Gleichzeitig wird der Bassbereich bei 40 Hz angehoben, um eine wuchtigere Wirkung zu erzielen.

Angereichert wird das Schlagzeug durch die live ausgelösten Samples. Das Snaresignal löst mit Hilfe des Slate Digital Triggers im Song drei unterschiedliche Samples aus, die das Schlagzeug in ausgewählten Teilen des Songs ästhetisch ergänzen und übernatürlicher wirken lassen.

## Bass

Die Bassstruktur des Songs wird von unterschiedlichen Elementen im Laufe des Songs übernommen und geprägt. Wie bereits angesprochen werden die Signale der Gitarristen gesplittet, um das komplette Signal oder einen Teil Octavern sowie Basssimulationen zuzuführen. Im Song sorgen die in folgender Grafik hervorgehobenen Instrumentengruppen für den Bassanteil.

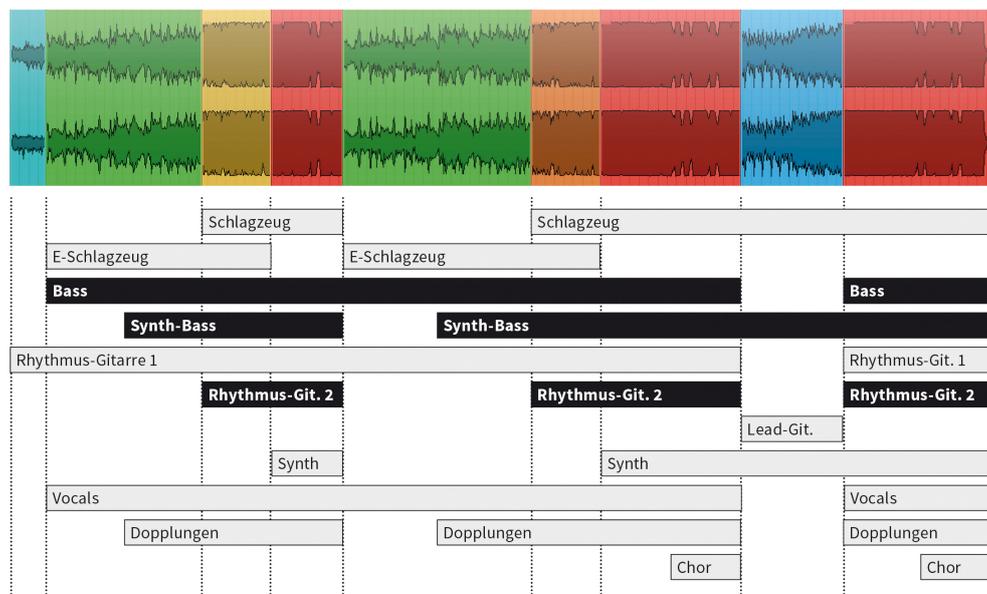


Abbildung 60: Circles – Bassinstrumentierung<sup>183</sup>

Um den dynamischen Umfang des live gespielten Basssignals auf einem für den Zuhörer stetigen Level zu halten, wird es mit Hilfe eines Kompressors dynamisch eingeschränkt. Im Refrain des Songs ergänzt ein synthetischer Bass den live gespielten Bass. Außerdem kommen für einen effektvollen Subbassbereich gespielte Subbass-Samples hinzu, die der Schlagzeuger über das Samplepad im Refrain triggert. Der synthetische Bass sowie der Subbass der Samples sollen die

<sup>183</sup> Screenshot

Frequentierung im Tiefenbereich des Refrains erweitern. Neben der dimensionalen Erweiterung durch ausschließlich offensichtliche Nutzung des dreidimensionalen Panoramas für Direktsignale im Refrain, soll die Erweiterung des Frequenzspektrums den Refrain größer wirken lassen.



Abbildung 61: Circles – Kompression<sup>184</sup>

## Gitarren

Die Rhythmusgitarren bilden im Song die harmonische Grundstruktur. Der störende Tiefenanteil der Live-Signale der Gitarren werden mit Hilfe eines Equalizers reduziert. Ein Deesser, ein frequenzabhängiger Regelverstärker, schränkt die unkontrollierten Höhen des Gitarrensings auf unauffällige Weise ein. Damit werden ungewollt stechende Höhen dynamisch kontrolliert. Neben den live gespielten Rhythmusgitarren erweitern Dopplungen dieser Gitarren das Panorama.

Um den Song im Refrain lauter und intensiver zu gestalten, kommen neben der Panoramaerweiterung, der frequentiellen Erweiterung im Bassbereich, sowie der größeren Dichte auf Grund der Menge an Instrumenten, Elemente zur stärkeren Betonung der Mittenfrequenzen hinzu. Die Rhythmusgitarren 2 füllen den Mittenbereich des Frequenzspektrums und geben dem Song eine größere Lautheit.

<sup>184</sup> Screenshot

Die Kurven gleicher Lautstärke verdeutlichen die Empfindlichkeit des menschlichen Gehörs im Mittenbereich.

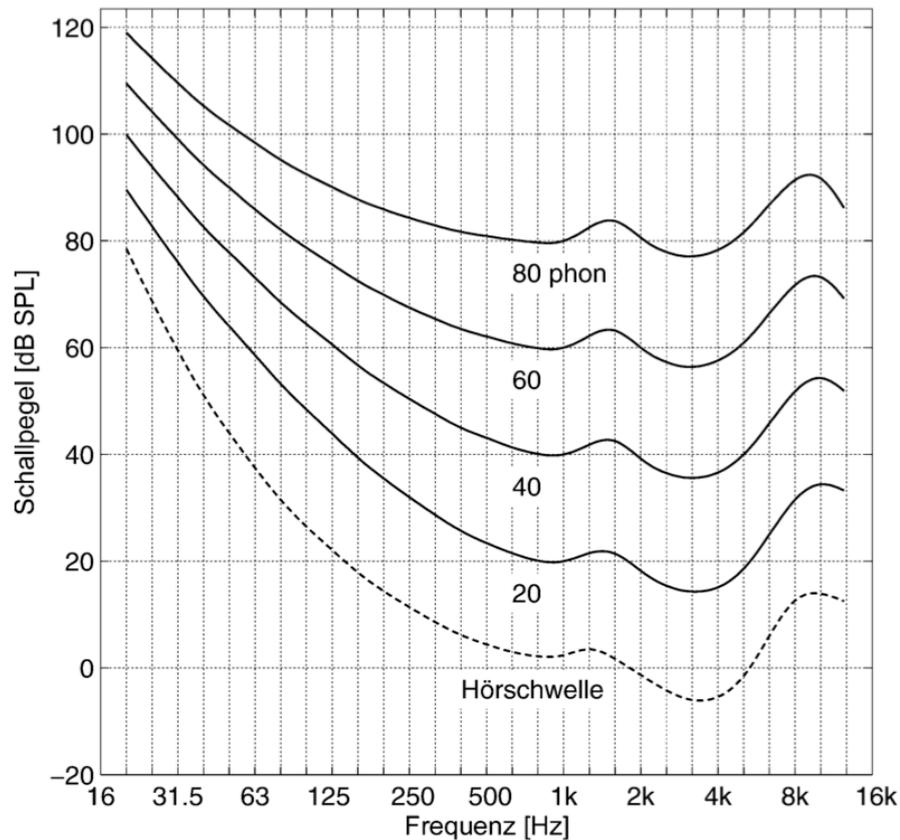


Abbildung 62: Kurven gleicher Lautstärke<sup>185</sup>

Durch dessen Ausprägung im Refrain wird eine größere Lautheit erzielt, die diesen weiter in den Aufmerksamkeitsmittelpunkt rücken.

Die räumliche Größe des Songs wird durch die Leadgitarren und deren Hallräume geprägt. Die helle Struktur der Gitarren sollen dem Song zur Offenheit und Weite verhelfen. Die Leadgitarren profitieren von den gepitchten Feedback-Reverbs, die räumlich nach hinten oben und hinten unten im dreidimensional beschallten Raum gepanned sind. Die Pitched-Feedback-Reverbs bilden einen stetigen Klangteppich, der sich der harmonischen Struktur der Leadgitarren anpasst und eine Weite im Song generiert.

<sup>185</sup> Weinzierl 2008: 54

## Vocals

Im Mittelpunkt des Songs steht der Gesang. Er wird mit Equalizern und Kompressoren bearbeitet. Es wird darauf geachtet, dass ungewollte tieffrequente Anteile des Signals vor der dynamischen Bearbeitung kontrolliert werden. Der Kompressor soll nicht durch ungewollte Signale angesprochen werden. Die Kompression beeinflusst nach der Filterung des Gesangs das gewünschte Signal und schränkt es in der Dynamik ein. Für die dynamische Bearbeitung der Vocals wird ein optischer Kompressor gewählt, dessen Regelverhalten eine charakteristische Trägheit aufweist und einen unauffälligeren und doch starken Eingriff zulässt.

Neben der Hauptstimme gibt es verschiedene Dopplungen, die zugespielt werden sowie spezielle Sendeffekte in Pro Tools, die den Gesang synthetisch erweitern. Letzterer wird passend zur visuellen Position platziert, sodass er eindeutig dem Sänger zugeordnet werden kann. Die Gesangsdopplungen erweitern die Stimme des Sängers in unterschiedliche Dimensionen im Raum passend zum Songablauf und zur Songklimax.

Im Vers dienen die vorkommenden Dopplungen dem ästhetischen Andicken der Stimme mit dem Ziel einer volleren Wirkung. Die Dopplungen sind meist auf eine ähnliche Position wie die Hauptstimme gepanned. Sie setzen sich aus verschiedenen Harmonien und gesungenen Intensitäten zusammen. Um die Harmonien unauffällig zu halten sind es oftmals hohe und tiefe Oktaven, die den Hauptgesang anreichern.

Im Prechorus und Chorus werden die Dopplungen der Stimme genutzt, um die Hauptstimme räumlich größer wirken zu lassen. Die Dopplungen sind so angelegt, dass sie die Dimension des gegebenen Raumes nutzen und den Sänger nicht nur durch Hall und Delay, sondern auch durch die Dopplungen übernatürlich größer wirken lassen. Der Gesang soll sich auf Grund der Dopplungen, sowie der Hall- und Delayeffekte von der gewöhnlich erwarteten Ästhetik des Raumes lösen.

Die Hauptstimmen der Sänger werden im Song mit jeweils zwei Delays, sowie zwei unterschiedlichen Hall-Effekten unterstützt. Die Delays unterscheiden sich in ihrer räumlichen Positionierung und in ihrem Timing. Die Effekte sollen dem Gesang als eine ästhetische, räumliche Einbettung in den Song dienen.

## 5.2 Veranstaltungsort

Die Intention war es, das 3D-Audio-Live-Konzert an der Hochschule der Medien statt finden zu lassen und dabei einen fachlichen Austausch unter Interessierten zu ermöglichen. Der einzige Raum, der für diese Art von Produktion eine ausreichende Größe und Ausstattung aufweist, ist das Fernsehstudio U44. Nach einer Begehung und der Aufnahme der limitierenden Faktoren konnte an Hand des Grundrisses und mit Hilfe der Software Sketchup eine erste Idee entstehen. Unter Beachtung sicherheitsrelevanter Vorschriften entstand in dem 240m<sup>2</sup> großen TV-Studio ein Grundrissplan-Entwurf, der eine Traversenkonstruktion von 9m x 12m x 4,50m vorsah. Details können dem Grundriss- und Bühnenplan in *Anhang II* entnommen werden.

Der limitierende Faktor Höhe verkleinerte die nutzbare Fläche enorm. Die Traversenkonstruktion muss, um eine Höhe von 4,50m zu ermöglichen, innerhalb der im Plan rot markierten, umlaufenden Vorhangschiene platziert werden, da diese eine Höhe von nur 4m bis zur Unterkante aufweist. Mit der Voraussicht, dass es trotz einer Vielzahl an Lautsprechern eine begrenzte Sweet Area geben wird, wurde der Zuschauerbereich innerhalb der Traversenkonstruktion um jeweils ein bis zwei Meter eingegrenzt um die finale Besucherkapazität zu bestimmen. Mittels einer großzügigen Besucherformel mit 1,5 Besuchern pro Quadratmeter (nach Versammlungstättenverordnung üblicherweise zwei Besucher pro Quadratmeter)<sup>186</sup> wurde die maximale Besucheranzahl von 90 Personen berechnet.

Das Fernsehstudio verfügt über eine sehr gute Ausstattung. Dennoch muss ein Großteil, vor allem der Audio-Technik aus externen Quellen zur Verfügung gestellt werden. Teile der Studiobeleuchtung und vor allem der elektrischen Versorgung wurden in die Produktion eingeplant. Erste technische Entwürfe wurden stets mit den technischen Mitarbeitern des TV-Studio diskutiert und auf Machbarkeit geprüft. Zur besseren Verdeutlichung und für die Sponsorsuche wurde folgendes Sketchup-Modell entworfen. Die Nützlichkeit einer dreidimensionalen Visualisierung des Konzeptes kam immer wieder zur Geltung, ganz egal ob bei Planungstreffen, bei Gesprächen mit Herstellern und Sponsoren oder während der Produktion als Orientierung und Planungsunterlagen für externe Beteiligte.

---

<sup>186</sup> Schliermann, 2015: 17

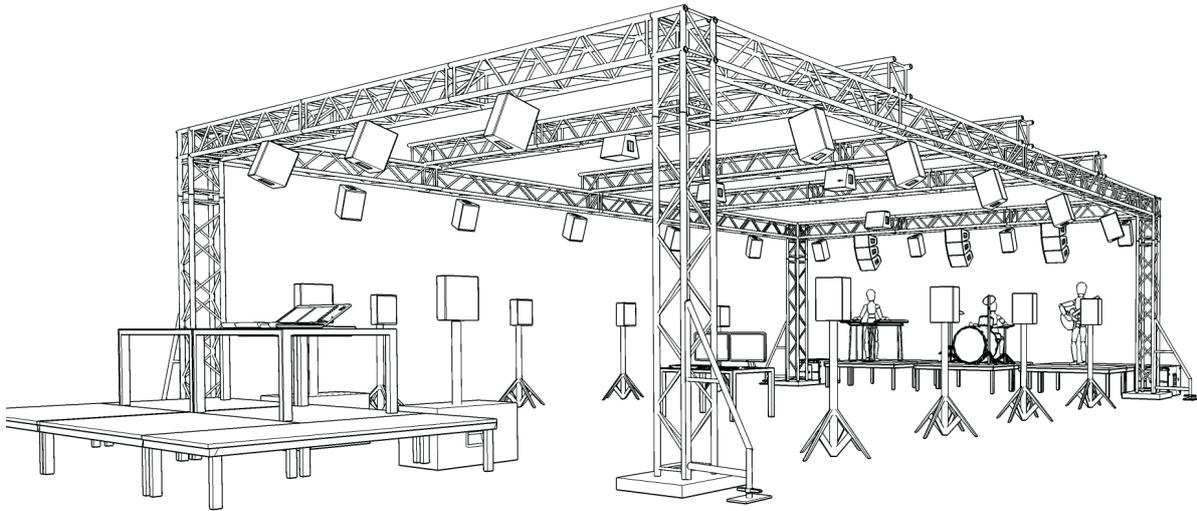


Abbildung 63: Entwurf: Lautsprecherkonfiguration (34.4) HdM U44 TV-Studio<sup>187</sup>

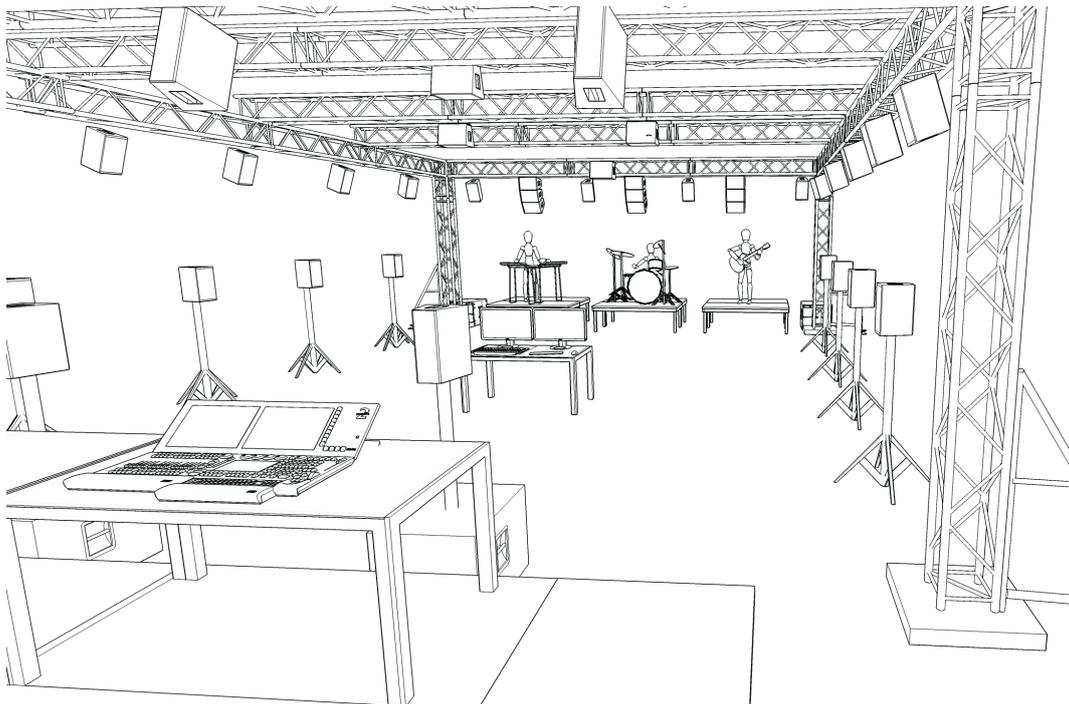


Abbildung 64: Entwurf: Lautsprecherkonfiguration (34.4) HdM U44 TV-Studio<sup>188</sup>

---

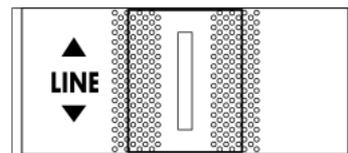
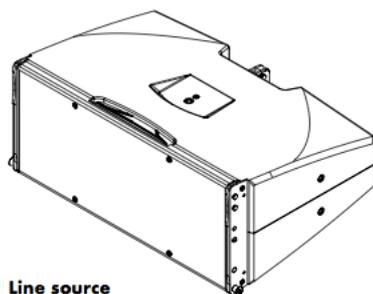
<sup>187</sup> Eigene Darstellung

<sup>188</sup> Eigene Darstellung

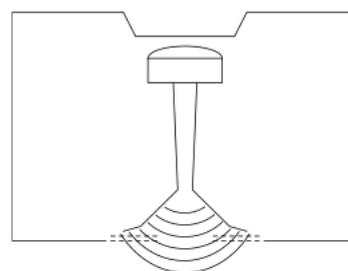
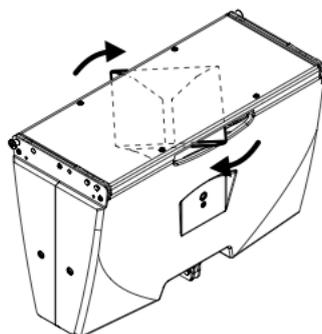
## 5.3 Technisches Setup der Produktion

### 5.3.1 Lautsprecherkonfiguration

Die Lautsprecherkonfiguration für das Live-Konzert besteht aus 40 T10 Lautsprechern und vier Q-Subwoofern der Firma d&b Audio. Der T10 kann sowohl als Line-Array als auch als Punktschallquellen-Lautsprecher mit hoher Direktivität eingesetzt werden. Es stehen zwei unterschiedliche Abstrahlcharakteristika zur Verfügung. Werkzeuglos kann der drehbare Wellenformer mit Horn und akustischer Linse gedreht werden. Das Horn erzeugt dabei einen horizontalen Abstrahlwinkel von  $90^\circ$ . Die Linse, integriert im Frontgitter des Lautsprechers, erweitert das Abstrahlverhalten im Line-Array-Modus im Hochtonbereich auf  $105^\circ$ . Bei Aufrechter Position des Lautsprechers als Punktquelle wird mit Hilfe der Linse die Wellenfront der Linienquelle gekrümmt. Daraus resultiert eine Abstrahlcharakteristik von  $90^\circ \times 35^\circ$ .



$105^\circ$



T10 horn and lens in line source setup

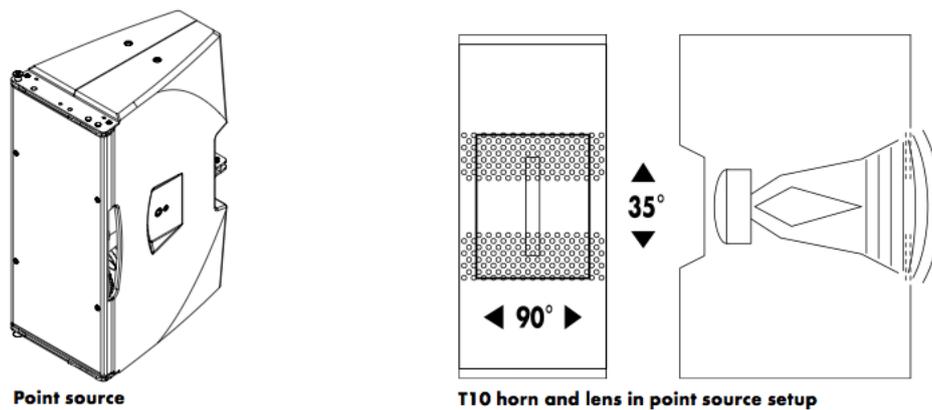


Abbildung 65: d&b Audio T10 Lautsprecher - Abstrahlcharakteristik<sup>189</sup>

Mit dieser Möglichkeit und den 40 Lautsprechern entstand folgende 34.4 Lautsprecherkonfiguration:

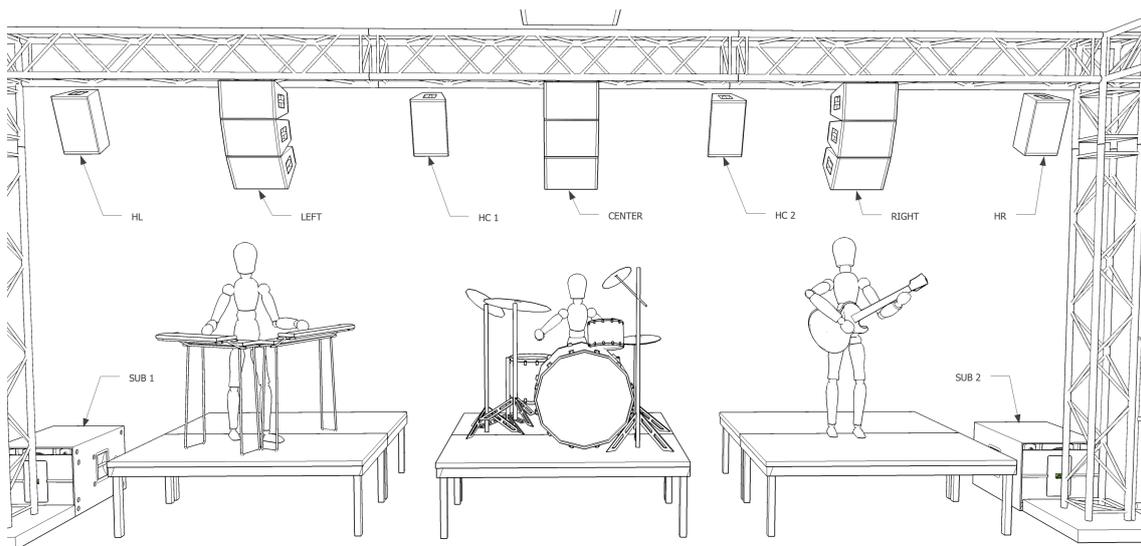


Abbildung 66: Lautsprecherkonfiguration HdM U44 TV-Studio - Front Array<sup>190</sup>

Aus dem grundlegenden mischtechnischen Ansatz einer klaren Ortung der Band von vorne, wie es im Live-Konzert-Kontext üblich ist, resultiert die größere Dimensionierung des Front-Arrays. Dieses besteht aus vier Punktquellen-Lautsprechern und drei Line-Array-Clustern aus jeweils drei Lautsprechern. Für eine gleichmäßige Lautstärken- und Leistungsverteilung im vorderen Bereich war eine Bestückung aus insgesamt sieben Line-Array-Clustern zu je drei Lautsprechern geplant.

<sup>189</sup> d&b Audio 2016: 11

<sup>190</sup> Eigene Darstellung

Diese war aber auf Grund der begrenzten Lautsprecheranzahl nicht realisierbar. Im Iosono-Core konnte der systembedingte Lautstärkenunterschied über die Control Unit angepasst werden.

Die Benennung der Lautsprecher in der Abbildung resultiert aus der ursprünglichen Überlegung, im vorderen Bereich der Lautsprecherkonfiguration zwei Layer zu ermöglichen. Darum gibt es L/C/R für den unteren Layer und HL/HC1/HC2/HR für den Höhen-Layer. Diese Idee wurde allerdings wieder verworfen. Eine Installation von Lautsprechern auf Höhe der Musiker beziehungsweise leicht über der Kopfhöhe der Besucher stört zum Einen das visuelle Erscheinungsbild und kommt zum Anderen lediglich den vorderen Zuschauer-Reihen zu Gute. Darum und um auch die Vorzüge des Iosono-Algorithmus im Front-Array zu nutzen, viel der Entschluss, alle sieben Positionen mit entsprechend weniger Abstand zwischen den Lautsprechern auf einen Layer zu setzen. Dadurch können im Mix die von vorne kommenden Signale auf Grundlage der im Iosono-Algorithmus integrierten Wellenfeldsynthese wesentlich ortungsstabiler platziert werden.

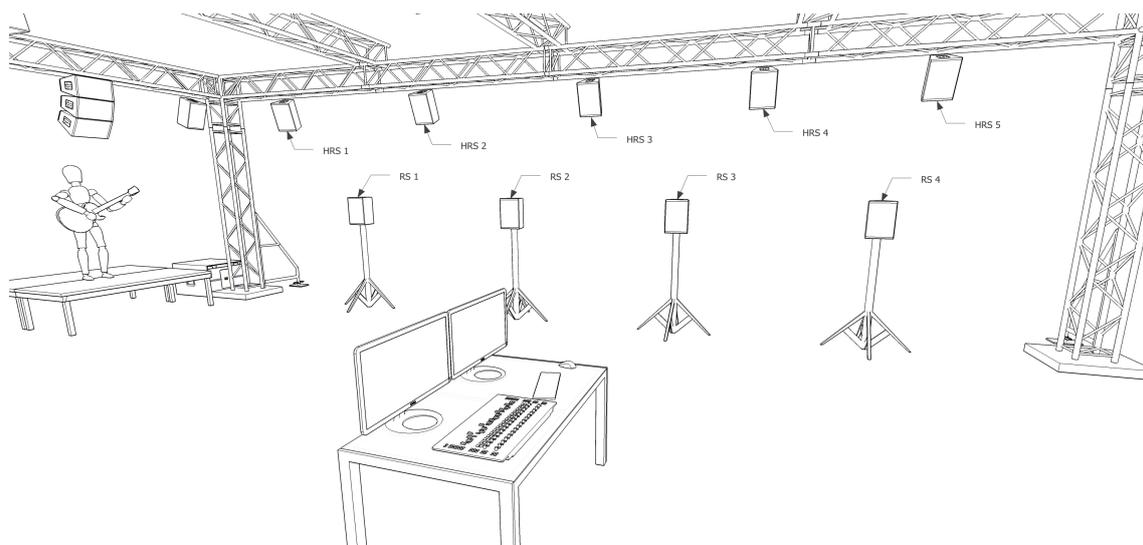


Abbildung 67: Lautsprecherkonfiguration HdM U44 TV-Studio - Right<sup>191</sup>

Die seitliche Lautsprecherkonfiguration sieht weiterhin zwei Layer vor. RS 1-4 als Right-Surround-Lautsprecher gehören zu dem unteren Layer. Diese sind auf ca. 2.20m installiert, damit der Schall nicht durch die seitlich stehenden Besucher verdeckt wird. Auf dem Höhen-Layer hängen fünf Height-Surround-Lautsprecher. Hinter den Zuhörern wird dieses Prinzip weitergeführt. Mit zwei Back-Surround-Lautsprechern auf dem unteren Layer und drei Height-Back-Surround-Lautsprechern auf dem oberen Layer.

---

<sup>191</sup> Eigene Darstellung

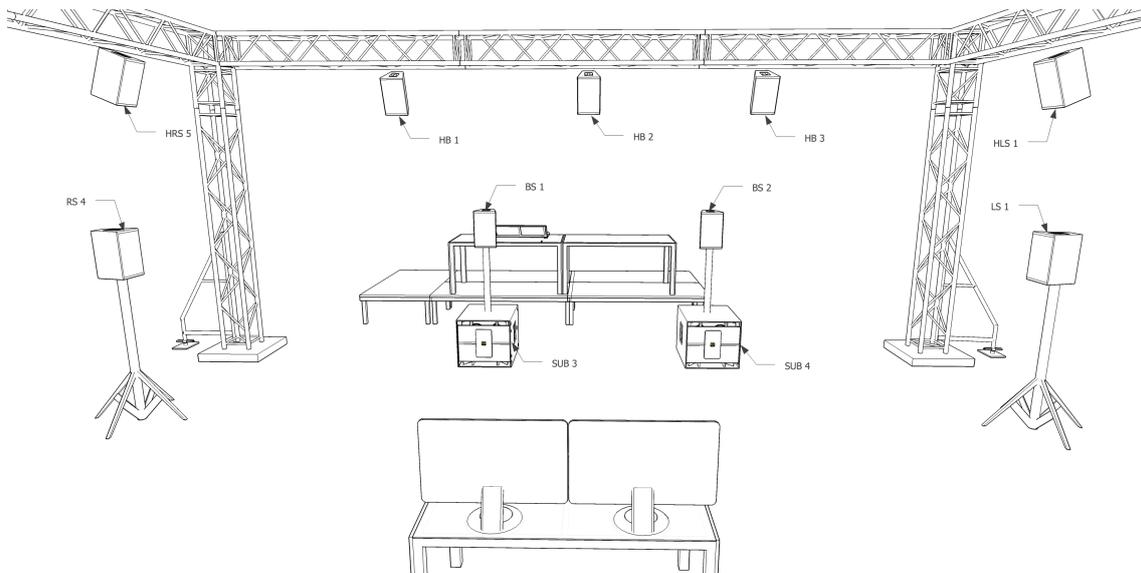


Abbildung 68: Lautsprecherkonfiguration HdM U44 TV-Studio - Back<sup>192</sup>

Die Lautsprecherkonfiguration auf der linken Seite ist identisch zur rechten Seite. Systembedingt ist die Nummerierung umgekehrt. Das resultiert aus der Lautsprecher-Konfiguration im Iso-sono-Core (siehe Kapitel 3.2.5 – Lautsprecher-Setup).

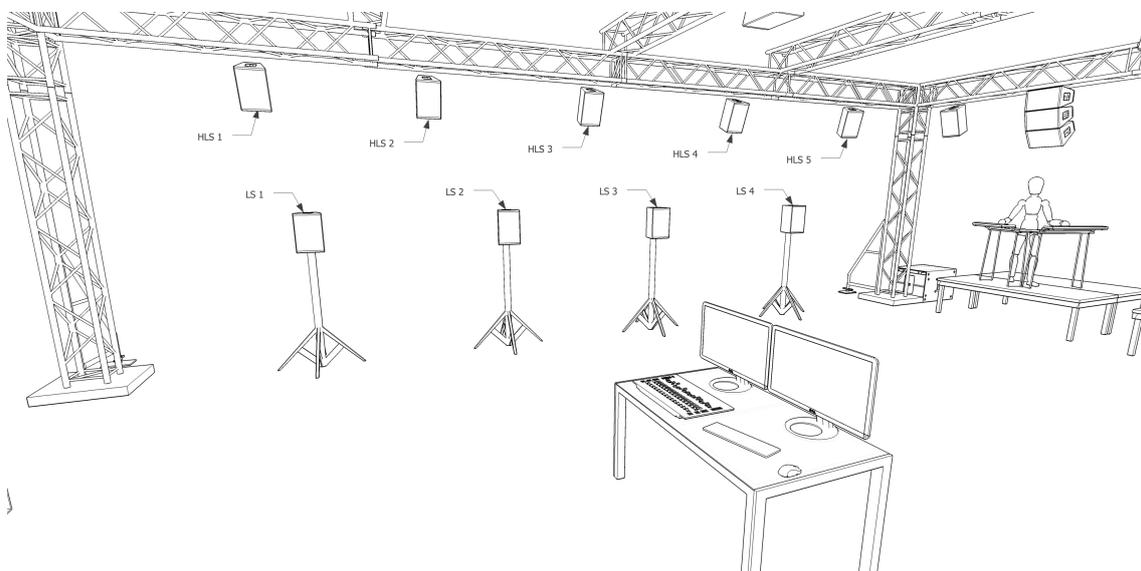


Abbildung 69: Lautsprecherkonfiguration HdM U44 TV-Studio - Left<sup>193</sup>

Die Konfiguration des Top-Layers bestand aus vier Punktquellen-Lautsprechern. Auch hier hat die ursprüngliche Planung eine höhere Anzahl an Lautsprechern vorgesehen, um den Top-Layer

<sup>192</sup> Eigene Darstellung

<sup>193</sup> Eigene Darstellung

gleichmäßiger abzudecken. Auf Grund der limitierten Höhe im Raum war eine höhere Anbringung der Top-Lautsprecher leider nicht möglich. Zudem stand für die ursprünglich angedachte Konfiguration mit drei Lautsprechern in der vorderen Reihe, vier Lautsprechern in der mittleren und wieder drei Lautsprechern in der hinteren Reihe leider nicht genügend Equipment zur Verfügung. Da der Isoono-Core den Top-Layer mit VBAP-Panning reproduziert, werden die Lautsprecher immer versetzt zueinander angebracht, damit Lautsprecher-Trippel für eine saubere VBAP-Reproduktion entstehen.

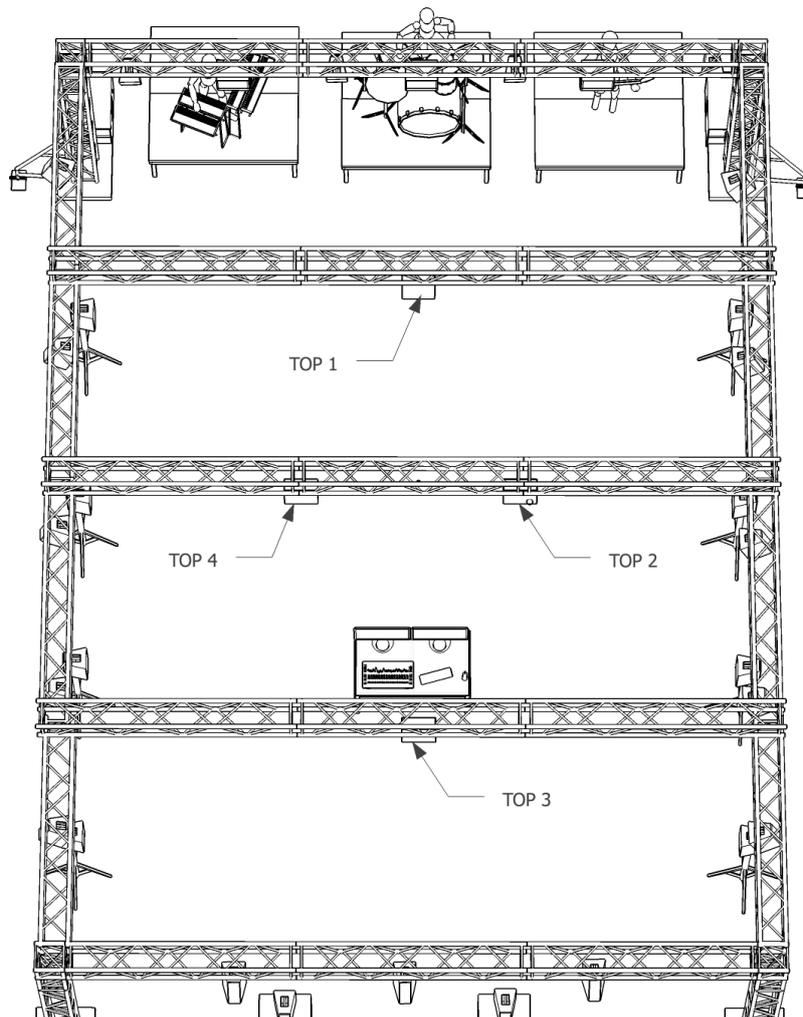


Abbildung 70: Lautsprecherkonfiguration HdM U44 TV-Studio - Top<sup>194</sup>

<sup>194</sup> Eigene Darstellung

Über die Lautsprecher Hersteller Software ArrayCalc wurden die horizontalen und vertikalen Neigungswinkel sowie die Splay-Winkel<sup>195</sup> innerhalb der Arrays berechnet.

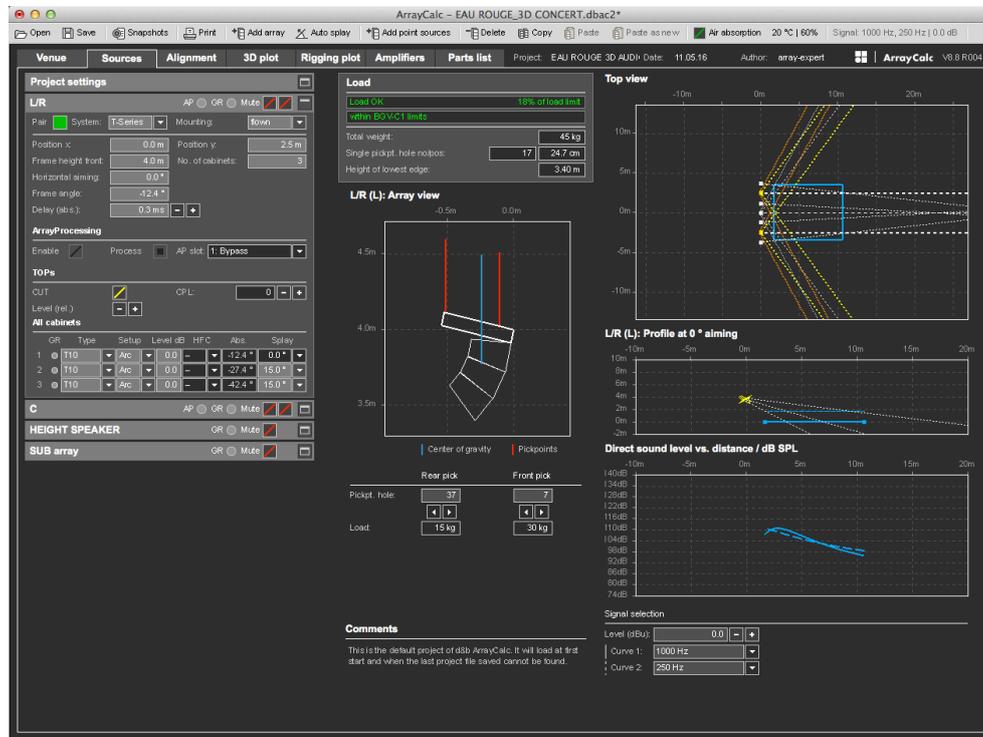


Abbildung 71: ArrayCalc Software – Winkelberechnung<sup>196</sup>

Für die Line-Arrays ergab sich daraus eine geneigte Aufhängung von ca. 13° und einen Splay-Winkel der Lautsprecher innerhalb des Arrays von jeweils 15°. Die Punktquellen-Lautsprecher sind auch um jeweils 15° geneigt. Für die gleichmäßige horizontale Abdeckung sind die L/R Line-Arrays um ca. 5°, die äußeren Punktquellen-Lautsprecher um ca. 10° und die inneren um ca. 5° zur Mitte eingedreht. Die „Top view“- Darstellung zeigt das horizontale Abstrahlverhalten der Lautsprecherquellen. In der Darstellung „Profile at 0° aiming“ ist die gleichmäßige Schallausbreitung über die Zuschauerfläche auf Ohrhöhe (blaue gestrichelte Linie) zu sehen. Aus der Darstellung darunter kann man entnehmen, dass der Schalldruckpegel eines Pink Noise Signales mit 0dBu Aussteuerung mit den im linken Bereich ausgewählten Lautsprechern im Vergleich zwischen vorderem und hinterem Zuhörerbereich um ca. 12dB SPL abnimmt.

<sup>195</sup> Vertikaler Neigungswinkel der Lautsprecher innerhalb eines Line-Array

<sup>196</sup> Screenshot

Über das SPL-Mapping der Software in der 3D Plotansicht konnten Lautstärkenverteilungen im System simuliert und berechnet werden.

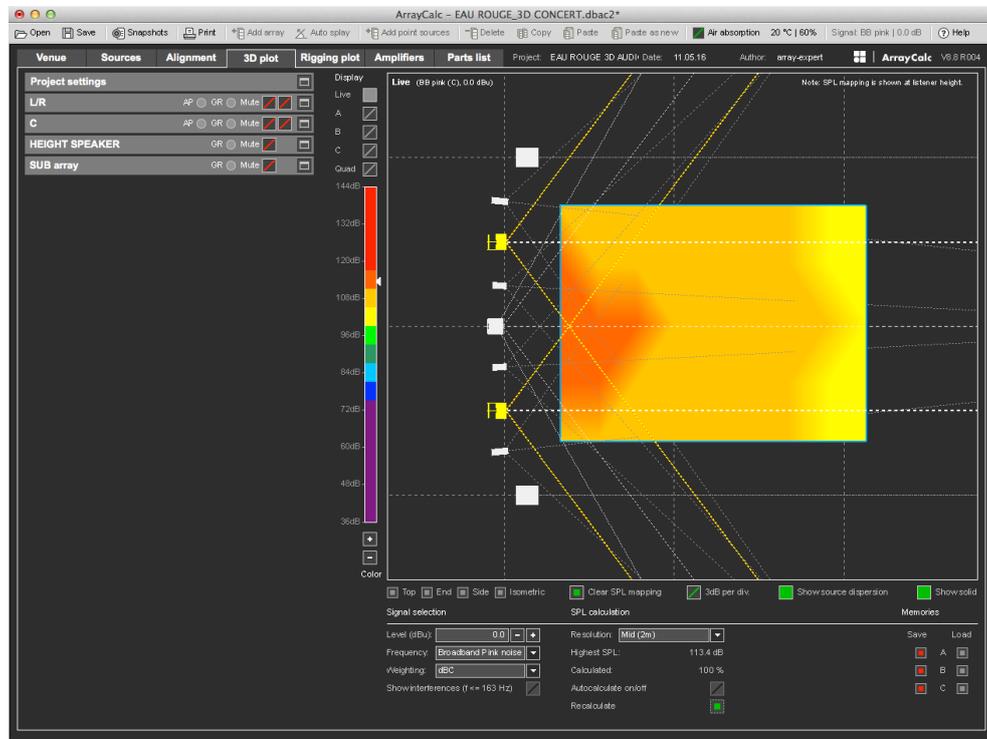


Abbildung 72: ArrayCalc - SPL Mapping Front-Array<sup>197</sup>

Dabei ergab sich für das Front-Array an Hand eines Pink Noise Signals mit 0 dB<sub>u</sub> Pegel und einer in lauten Live-Umgebungen üblichen C-gewichteten Messung, ein maximaler Schalldruckpegel von 113.4 dB<sub>SPL</sub> im vorderen Zuschauerbereich. Im mittleren Zuschauerbereich erzeugt das Front-Array ca. 108 dB<sub>SPL</sub> und im hinteren Bereich ca. 104 dB<sub>SPL</sub>. Diese Berechnung und Optimierung der Ausrichtung wurde dabei für jede Seite im der Lautsprecherkonfiguration vorgenommen.

<sup>197</sup> Screenshot

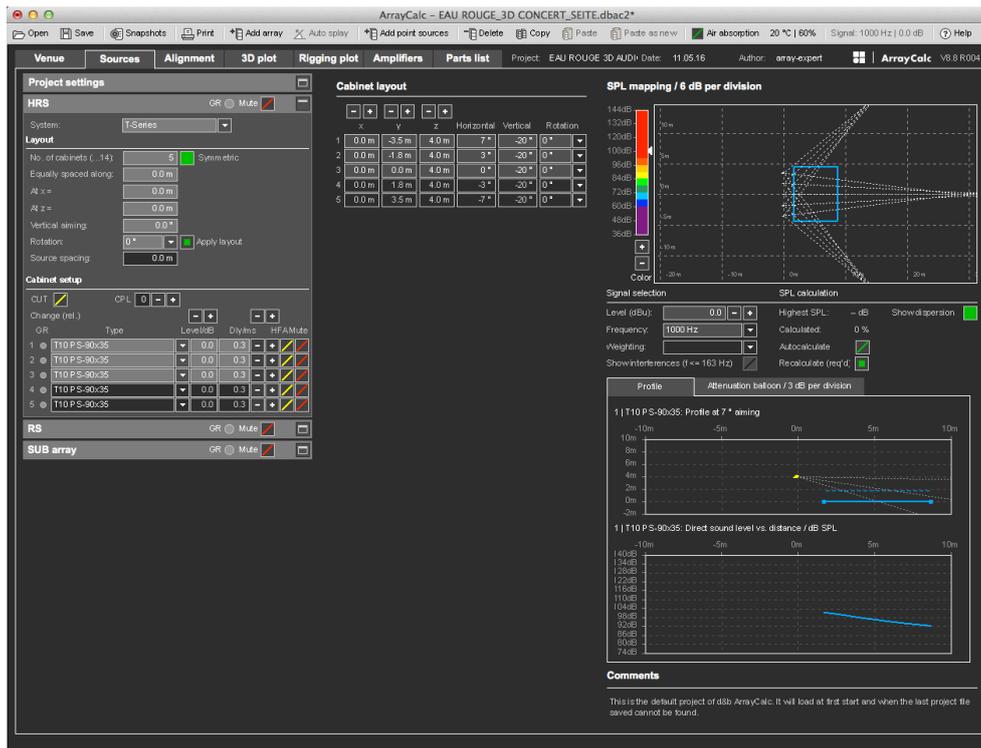


Abbildung 73: ArrayCalc – Cabinet Layout und Winkelberechnung für Seitenlautsprecher<sup>198</sup>

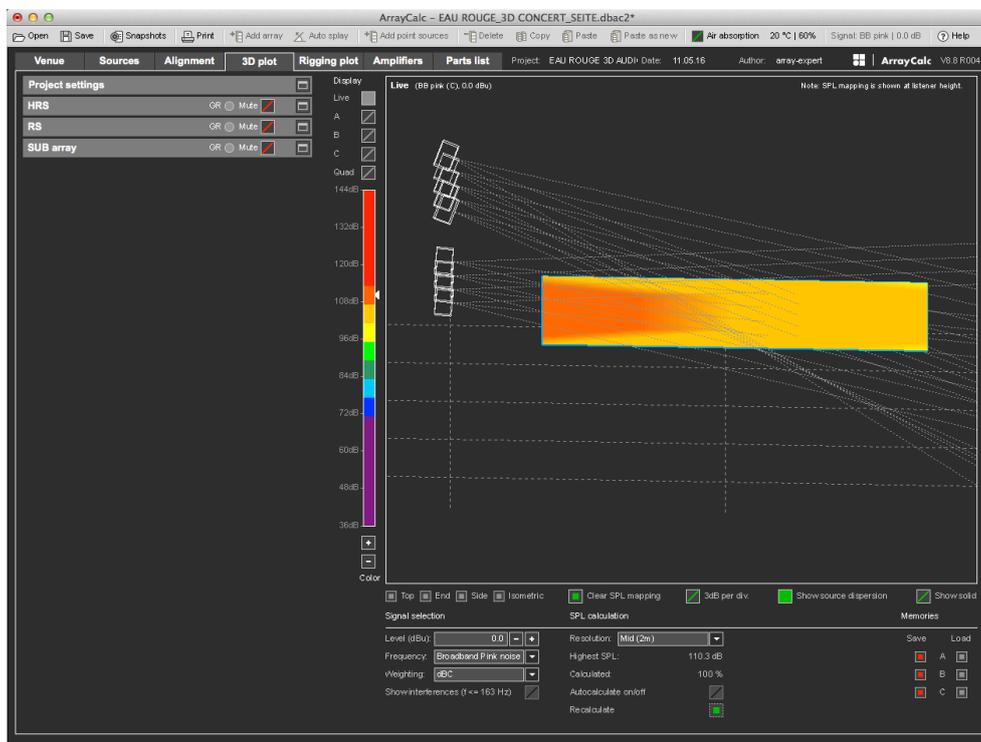


Abbildung 74: ArrayCalc – SPL Mapping Seitenlautsprecher<sup>199</sup>

<sup>198</sup> Screenshot

<sup>199</sup> Screenshot

Anhand der ArrayCalc-Berechnungen konnten die passenden Parameter für die Konfiguration der Verstärker ermittelt werden. Auch eine Stückliste aller benötigten d&b Lautsprecher-Komponenten konnte generiert werden und war dabei eine Hilfe für die Erstellung von Material- und Bestückungs-Listen.

Für die Verstärkung der Lautsprecher-Signale standen 24 d&b Audio D12 Verstärker zur Verfügung. Da für die Produktion das Beschallungs-Equipment von drei unterschiedlichen Firmen zur Verfügung gestellt wurde, war eine detaillierte Material- und Verkabelungsplanung nötig. Den T10 Lautsprecher und die D12 Verstärker gibt es mit unterschiedlichen Anschlussausführungen (EP5 oder NLT4). Das hat wiederum Auswirkung auf die weiterführende Verkabelung sowie die Multicore-Aufteilungen, die je nach Verleiher variieren. Bei der Planung der Lautsprecherkonfiguration für das Live-Setup wurde eine detaillierte Multicore-Liste erstellt (*siehe Anhang VII*). Dieser Liste kann, zu jedem Lautsprecher der zugewiesenen Verstärker, die Anschlussart, sowie die Aufteilung, Nummer und Länge des Multicores entnommen werden. In der rechten Spalte ist die Art der Multicore-Breakoutbox sowie die Restlänge bis zum Lautsprecher aufgeführt. Eine Bestückungs-Liste (*siehe Anhang VI*) gibt Informationen über die Art und Herkunft der Lautsprecher Aufhängung und Befestigung. Eine Material-Liste (*siehe Anhang V*) gibt einen Gesamtüberblick der Art und Herkunft des Equipments.



Abbildung 75: Lautsprecherkonfiguration U44 HdM TV-Studio<sup>200</sup>

---

<sup>200</sup> Eigenes Foto

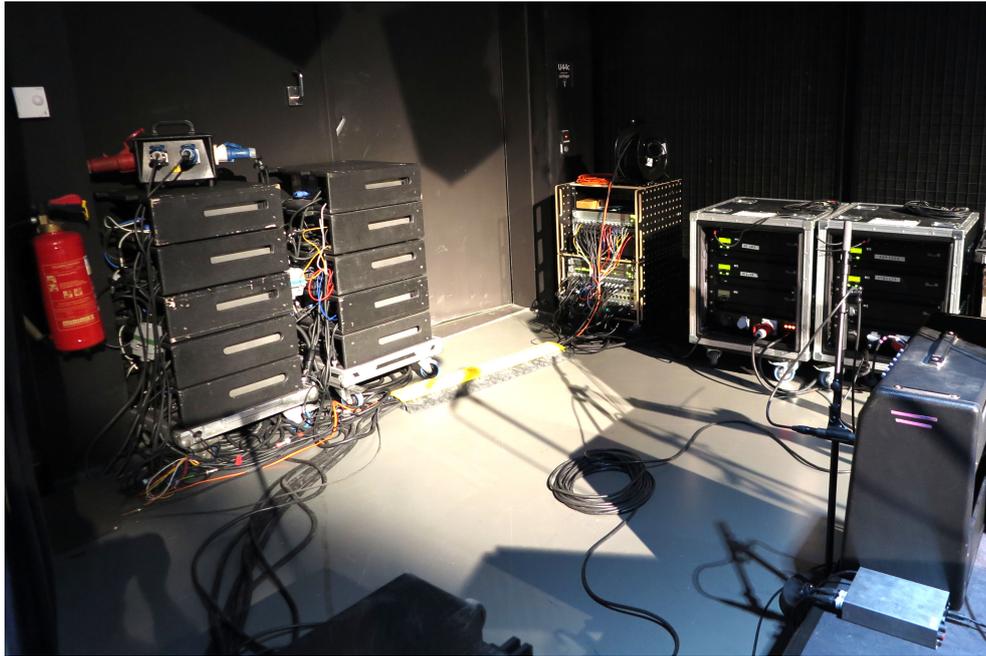


Abbildung 76: d&b Audio - D12 Verstärker<sup>201</sup>

### 5.3.2 Iosono Lautsprecher-Setup

Trotz einer guten Planung und Berechnung über die Software ArrayCalc empfiehlt es sich die exakten Positionen der Lautsprecher vor dem Erstellen des Iosono-Lautsprecher-File, erneut auszumessen. Für die einfache Eingabe in der Control-Unit-Oberfläche werden dabei, mit Hilfe eines Laser-Messgerätes, immer die Abstände zur Mitte der jeweiligen Seite des Layers gemessen. Bei der Konfiguration des Lautsprecher-File für den Iosono Core gibt es noch folgende Besonderheit: Die in Kapitel 5.3.1 beschriebene Herangehensweise in Bezug auf das Front-Array setzt im Lautsprecher-File einen unabhängigen Layer voraus, der im vorderen Bereich der Lautsprecherkonfiguration auf jede Layer-ID reagiert. Egal ob sich ein Audio-Objekt aus dem unteren Layer (Layer-ID 0) oder aus dem Höhen-Layer (Layer-ID 1) in das Front-Array bewegt, sollte der Übergang nicht hörbar sein. Für eine saubere Berechnung dieser Übergänge braucht der Iosono-Algorithmus zusätzliche Dummy-Lautsprecher. Diese werden im Output-Routing keinem physikalischen Ausgangskanal zugewiesen und dienen so der reinen Berechnung des Iosono-Algorithmus. Ein weiterer Mechanismus um Audio-Objekte auf das unabhängige Front-Array zu binden, bietet die Dialog-Checkbox unter „Source Traits“ in den Redner-Slot-Einstellungen. Im späteren Mix können damit über die SAW Audio-Objekten „Dialogue-Flags“ mitgegeben werden,

---

<sup>201</sup> Eigenes Foto

was soviel bedeutet, dass sie die Bildschirm-Fläche beziehungsweise das Array welches entsprechend definiert wurde, nicht verlassen. Eine Funktion die aus der Kino-Mischung resultiert, bei der Dialoge den Leinwandbereich nicht verlassen sollen.

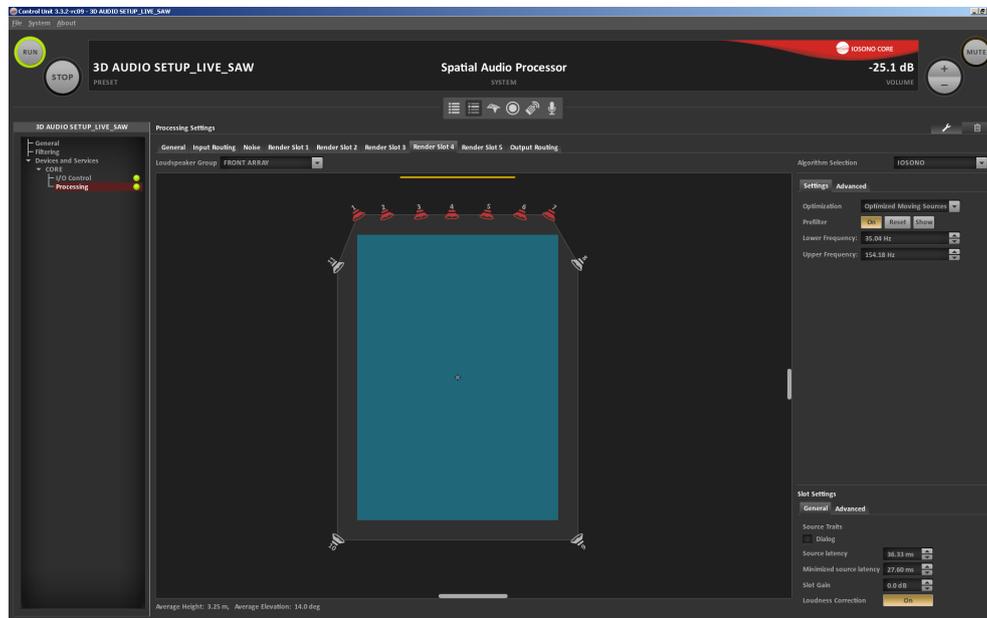


Abbildung 77: Iosono Control Unit - Processing-Settings - Front-Array<sup>202</sup>

In der Abbildung sind die aktiven, auf einen physikalischen Ausgang gerouteten Lautsprecher des Front-Arrays rot eingefärbt. Die vier grauen Symbole, sind die beschriebenen Dummy-Lautsprecher. Nach dem Erstellen von Lautsprecher-File und Preset der Produktion kann über den internen Rauschgenerator jeder Lautsprecher im System angesprochen werden. Für einen reinen Signaltest ist das ein hilfreicher Weg. Um die Übergänge zwischen den Layern, sowohl horizontal, aber auch vertikal zu prüfen, wurde über den internen Audio-Player in Verbindung mit dem internen Scene-Designer ein Testobjekt angelegt. So kann man das System auf Lautstärken-Unterschiede, Einbrüche im Panning, oder defekte Komponenten überprüfen. Es empfiehlt sich, erst dann mit der Systemeinstellung zu beginnen wenn alle groben Fehler im System ausgeschlossen sind.

<sup>202</sup> Screenshot

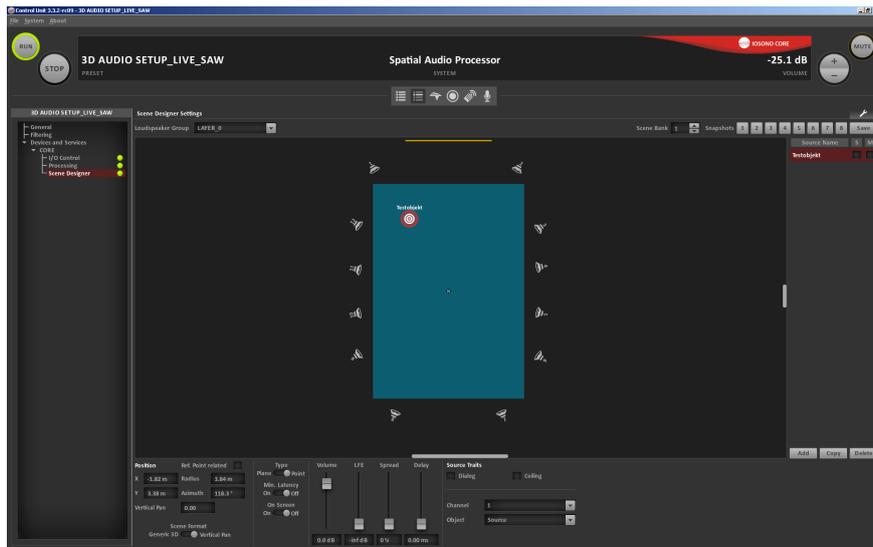


Abbildung 78: Iosono - Scene Designer Test-Objekt<sup>203</sup>

### 5.3.3 Iosono Systemeinmessung

Der Iosono-Core hat ein optionales Software-Modul zum Einmessen des Systems. Dabei wurde das System über vier Messpunkte mit Kugel-Messmikrofonen aufgenommen. Über jeden Lautsprecher im jeweiligen Layer wird dabei nacheinander ein ca. 2Sekunden langer Sinus-Sweep von 20Hz bis 20kHz abgespielt. Aus der aufgenommenen Information kann der Core pro Lautsprecher im System die Frequenzfilter-Kurve und die Impulsantwort anzeigen.

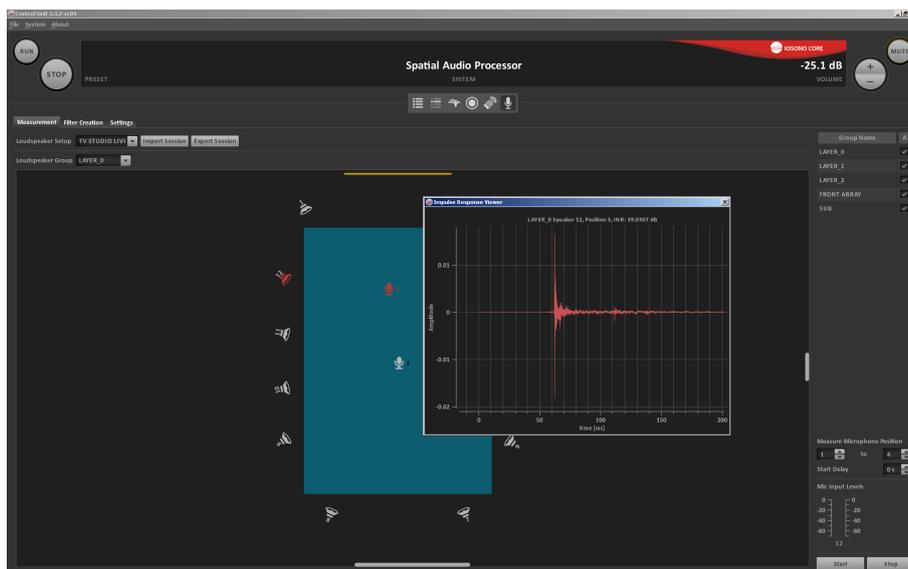


Abbildung 79: Iosono Measurement<sup>204</sup>

<sup>203</sup> Screenshot

<sup>204</sup> Screenshot

In den Measurement-Einstellungen können neben Audio-Routings, die Grenzwerte für Signal Noise-Ratio (SNR) und Minimum Impulse-Response-Noise-Ratio (INR) definiert werden. Bei entsprechenden Messungen lag der SNR bei 10.00 dB und der INR bei 45.00 dB.

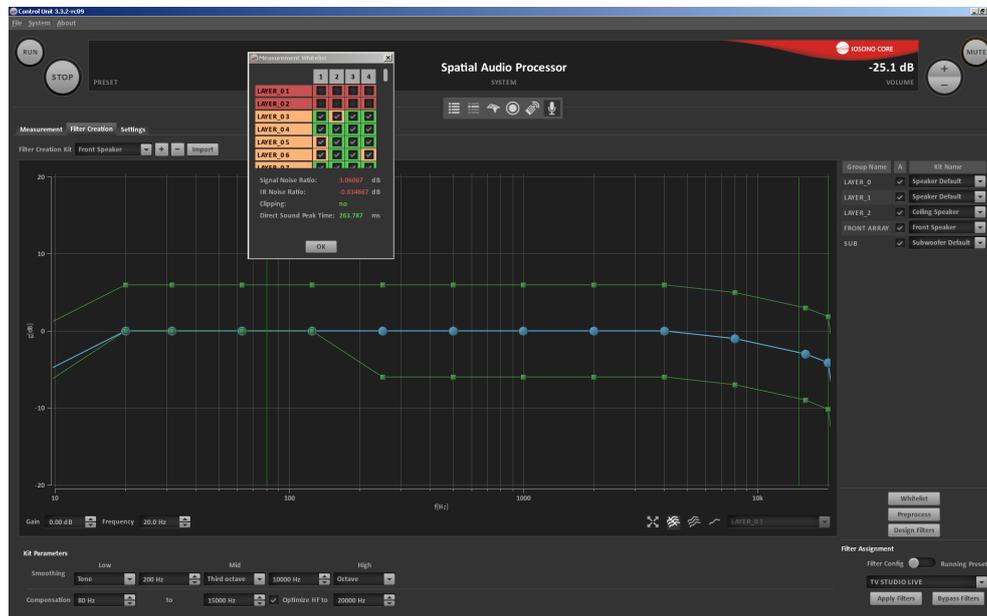


Abbildung 80: Iosono - Filter Creation - Whitelist<sup>205</sup>

An Hand der definierten SNR und INR Daten erstellt der Iosono-Core eine Whitelist mit allen Messpunkten und markiert dabei Messungen außerhalb der Grenzwerte rot, knapp unterhalb der Grenzwerte orange und Messungen innerhalb der Grenzwerte grün. Die in der Abbildung rot markierten Messpunkte ‚Layer\_01‘ und ‚Layer\_02‘ sind die Messpunkte der Dummy-Lautsprecher. Nach Anwahl der Messpunkte kann an Hand der Messinformationen eine FIR-Filter-Library zu jedem einzelnen Lautsprecher erstellt werden. Dabei stellt man über die blaue Linie die ideale angestrebte Frequenzkurve an. Über die grünen Linien kann man den Frequenzbereich einschränken, der zur Filterberechnung gewichtet werden soll. Die Filteranhebung von 20Hz bis ca. 120Hz wird bewusst nicht gewichtet, da die Messergebnisse auf Grund der Wellenlänge im Verhältnis zum Raum und der Abstände der Messpunkte zu nicht aussagekräftigen Ergebnissen führten. Zudem decken die T10 Lautsprecher einen tiefen Frequenzbereich bis maximal 85Hz ab. Die Anpassung des Bassbereiches erfolgt über einen besser kontrollierbaren und Ressourcen schonenderen, globalen IIR-Filter.

<sup>205</sup> Screenshot

Das aus den Messungen generierte Filter-Creation-Kit wurde im Anschluss über das globale Filter-Setup den jeweiligen Lautsprechern zugeordnet. Um das Wiedergabeverhalten im Bassbereich zu definieren, wurden jedem T10 Ring-Lautsprecher ein Hochpassfilter mit einer Grenzfrequenz von 85Hz und einer Flankensteilheit von 6dB pro Oktave zugewiesen. Um die Flankensteilheit zu erhöhen wurde der gleiche Filter zweimal aktiviert, daraus ergab sich eine Flankensteilheit von 12 dB pro Oktave.

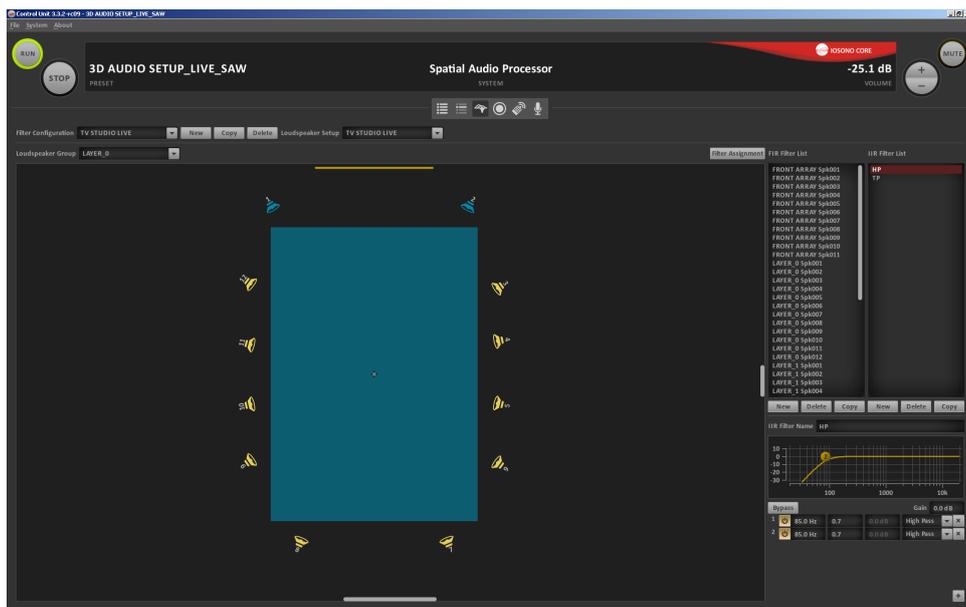


Abbildung 81: Iosono Filter Configuration<sup>206</sup>

Am Ende der Systemeinmessung sollte man das System in jedem Fall mit eigenen Referenzen anhören und sich auf das eigene Ohr und die Hörerfahrung verlassen. In diesem Fall wurde für die Lautsprecher im Front-Array, auf Grund der Überpräsenz in den mittleren Höhen, ein Filter bei 4000 Hz gesetzt und dieser dabei mit einer Güte von 0.7 um -3dB absenkt.

<sup>206</sup> Screenshot

### 5.3.4 Signalkette

Die Signalkette des Live-Konzertes entspricht der gleichen, die während der Vorproduktion erarbeitet wurde und unter Signalkette in *Kapitel 4.4.2* detailliert beschrieben ist. Die einzige Ergänzungen im System für das Live-Konzert besteht in einem Backup-Rechner als Redundanz für das MacBook und einem USB-Netzwerk-Extender zum Absetzen der USB-Verbindung für die Steuerung der Andiamo.MC. Der genaue Systemaufbau kann dem Blockschaltbild in *Anhang I* entnommen werden. Auch eine für das Live-Konzert ergänzte Signalliste befindet sich in *Anhang IV*. Hier besteht die Änderung zu dem in *Kapitel 4.3.2* beschriebenen Signalfluss darin, dass auf Grund des geänderten Lautsprecher-Setup im Iosono-Core eine andere Zusammensetzung von Render-Slots vorliegt. Auch die Anzahl der Lautsprecher im System ist gestiegen, sodass sich nach dem Processing im Core eine Erweiterung im Routing des Madiface XT ergibt. Auch die genaue Kanal-Zuweisung der neuen Lautsprecher-Konfiguration kann der Signalliste in *Anhang IV* entnommen werden.

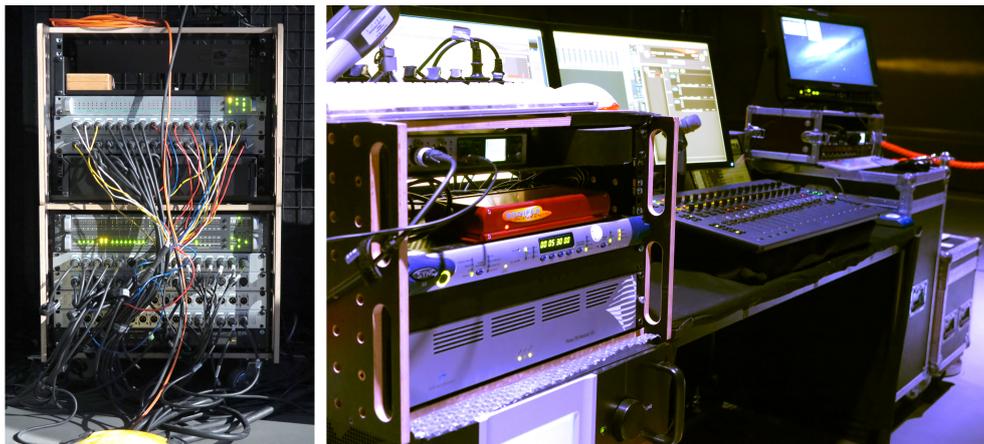


Abbildung 82: 3D Live Setup: AD/DA Wandler (links) Audio-Regie (rechts)<sup>207</sup>

---

<sup>207</sup> Eigenes Foto

### 5.3.5 Gefährdungsanalyse

Im Rahmen dieser Produktion gibt es eine Vielzahl an potentiellen Gefahren. In einer ausführlichen Gefährdungsanalyse wurde auf die jeweiligen potentiellen Gefahren hingewiesen und die entsprechende zu treffende Maßnahme beschrieben. Es wird die Gefährdung in den Bereichen Bühne, Arbeitsumgebung, Veranstaltungsraum, Beleuchtung, Tontechnik, elektrische Geräte, Kabel und Leitungen sowie Gefahren während des Aufbaus und der technischen Einrichtung analysiert (*siehe Anhang VIII*).

Unter Abschnitt 3.5 der Gefährdungsanalyse wird auf die Gefahr der Gehörschädigung für Darsteller und Publikum hingewiesen. Wie in *Kapitel 5.3.1* beschrieben, wurden die zu erwartenden Schalldruckpegel simuliert und berechnet. Während der Veranstaltung wurden die Schalldruckpegel kontinuierlich überwacht. Schalldruckpegelmessungen mit geeichtem dB<sub>SPL</sub> Messgerät der HdM Haustechnik ergaben während der Konzerte unter Vollaussteuerung C-bewertete Spitzenpegel von 115dB<sub>SPL</sub> und ein A-bewerteter gemittelter Schalldruckpegel über die gesamte Länge des Konzertes von 98dB. Nach DIN 15095-5 wird damit die Norm erfüllt, sodass keine Gefahr durch zu hohe Schalldruckpegel vorliegt.<sup>208</sup>

Unter Abschnitt 5.5. wird auf die Gefahr durch Stress, Überarbeitung und Konzentrationsschwäche hingewiesen. Ein Punkt der im Rahmen von studentischen Produktionen von großer Relevanz ist. Da die Entwicklung neuer Produktionen immer Unbekanntes mit sich bringt, hilft nur eine großzügige Zeitplanung und Personaleinteilung sowie eine gewisse Selbstdisziplin, um sich die nötigen Pausen zu gönnen. In *Anhang IX* befindet sich darum eine Zeitplanung der Produktionswoche in der darauf geachtet wurde, dass alle Beteiligten ihre Pausen einhalten können.

## 5.4 Positionierung der Band

Bei einer 3D-Audio-Produktion, in der die Zuschauer aus allen Richtungen mit Audio umhüllt werden, war zunächst die Überlegung, die Band auf einer Drehbühne in der Mitte des Zuschauerbereiches zu platzieren. Sowohl der visuelle Eindruck als auch die Immersion würde dadurch an Bedeutung gewinnen. Aus Platzgründen und aufgrund technischen Aufwandes war diese Idee leider nicht zu verwirklichen.

---

<sup>208</sup> vgl. Ebner 2009

Die Band wurde nach gewohnter Konzertaufstellung im vorderen Bereich der Bühne platziert. Damit die Ortung des Schlagzeugers mit der Mix-Ästhetik übereinstimmt wurde dieser in der Mitte unter dem Center-Lautsprecher platziert. Die beiden Sänger und Gitarristen standen jeweils links und rechts neben dem Schlagzeug. Um das Übersprechen des Schlagzeuges auf den Gesangsspuren zu minimieren, wurden die Gesangsmikrofone eingedreht, sodass nur die 90°- beziehungsweise 270°-Achse der Nieren-Richtcharakteristik in Richtung Schlagzeug zeigte. Zudem war das Schlagzeug mit einer Schallschutz-Wand aus Plexiglas umstellt. Dank der guten Raumakustik mit einem guten Absorptionsgrad der Wände und dem Vorhang hinter der Band gab es sehr wenig Signalübersprechen zwischen den einzelnen Signalen und wenige Reflexionen in den Raum. Dadurch und durch die Minimierung des Bühnen-Sounds konnte eine bestmögliche Verteilung des Schallfeldes in der vorhandenen Lautsprecher-Konfiguration umgesetzt werden. Klangliche Verfärbungen durch korrelierte Signale, räumliches Übersprechen und Rückkopplungen waren nahezu nicht wahrnehmbar.



Abbildung 83: Eau Rouge 3D Konzert - Positionierung der Band<sup>209</sup>

Die Abbildung zeigt die Positionierung der Band. Für die Ortungsstabilität der Sänger und Gitarren wurde mit Hilfe der SAW jeweils eine Punktschallquelle in das Line-Array über dem jeweiligen Musiker gelegt. Das gedoppelte Signal des gleichen Instrumentes beziehungsweise der Stimme

---

<sup>209</sup> Eigenes Foto

wurde als ebene Welle weit dahinter positioniert. Die Intention dabei ist eine Ortungsstabilität der Sänger und Gitarristen an ihrer jeweiligen Position. Bei einer Zuhörerposition im linken Bereich würde der Zuhörer jedoch den Sänger von rechts dabei fast nicht mehr wahrnehmen, da das Array auf seiner Seite nicht angesprochen werden würde. Darum sorgt die ebene Welle hinter der Punktschallquelle für ein gleichmäßiges Ansprechen des Front-Arrays. Dennoch bleibt die Ortung der Musiker an ihrer realen Position dank der fokussierten Quelle, direkt über Ihnen im Line-Array, erhalten.

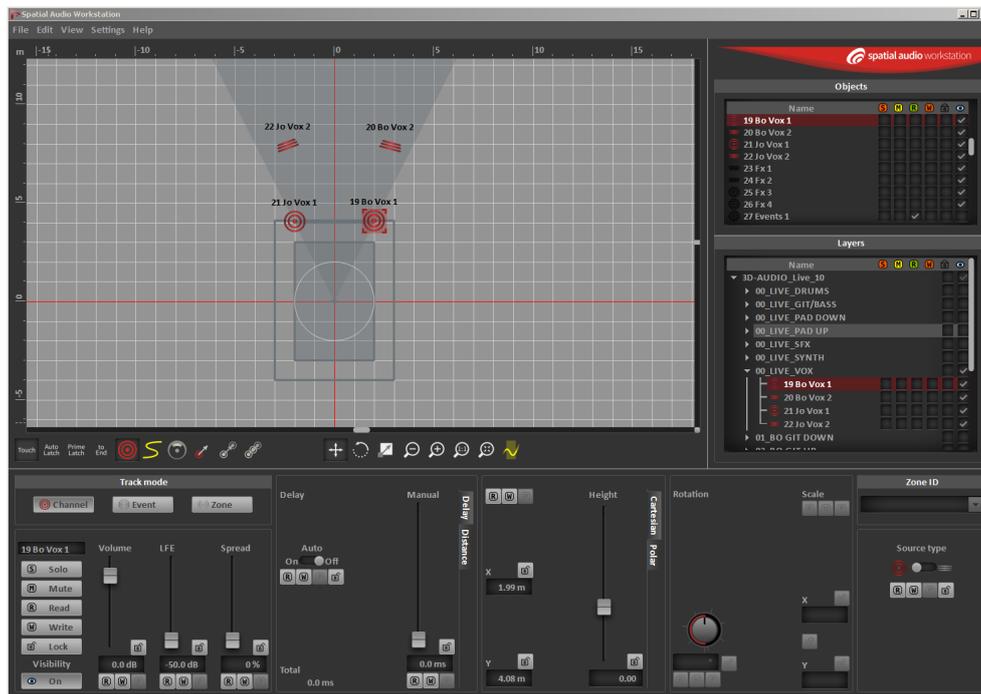


Abbildung 84: SAW - Positionierung der Vocal-Audio-Objekte<sup>210</sup>

<sup>210</sup> Screenshot

## 5.5 Monitoring

Das Monitoring der Band wird mit Hilfe des RME Totalmix FX des RME Madiface XT realisiert. TotalMix FX ist ein digitaler Echtzeitmischer und kann mit sehr geringen Latenzen Submixe für die Musiker erstellen. Das bedeutet, dass sich die für die Band relevanten Latenzen auf ein Minimum reduzieren. Direkten Einfluss auf die für die Band gehörten Latenzen haben folgende Systeme:

- A/D-Wandler
- Totalmix mit DSP
- D/A-Wandler

Wie in *Kapitel 4.3.2* beschrieben wird in Totalmix FX wird für jeden Musiker ein Ausgangskanal angelegt. Diese werden als Monitorkanal für die Musiker genutzt. Jedem Ausgangskanal kann jedes am RME Madiface XT anliegende Eingangssignal beliebig zugeführt werden. Das Madiface XT wird im Setup als Signalverteiler genutzt. Alle in der Liveproduktion genutzten Signale durchlaufen das Madiface XT. Mit Hilfe von DSPs können im Totalmix FX frequenzuelle und dynamische Eingriffe der Einzelsignale für das Monitorsignal getätigt, sowie Hall und Delay hinzugefügt werden. Damit kann ein für den Musiker angenehmerer Monitormix erstellt werden und ‚Snapshots‘ gespeichert und wieder abgerufen werden.

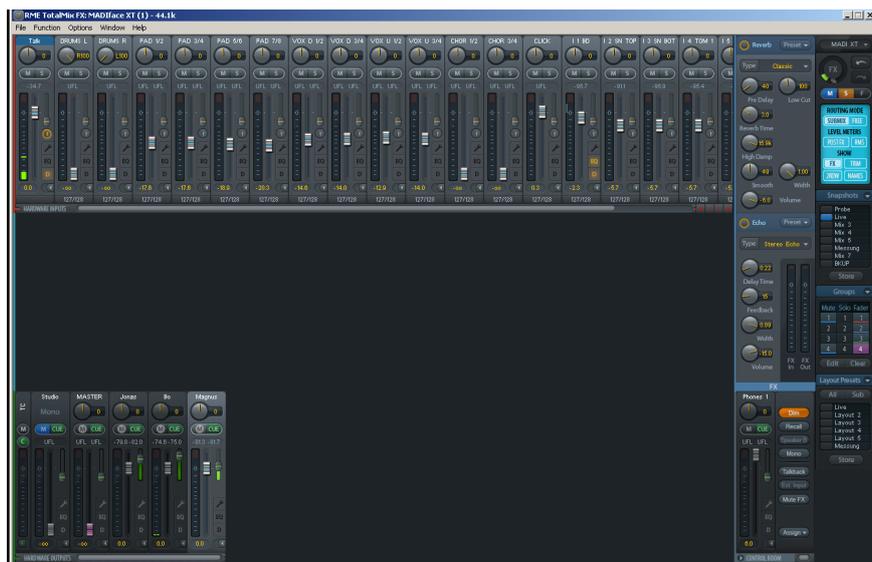


Abbildung 85: RME Totalmix FX<sup>211</sup>

<sup>211</sup> Screenshot

## 6 Aufzeichnung & Abspielen

Im Rahmen der Vorproduktion wurden Ausschnitte der dreidimensionalen Mischung im Studio U54 der Hochschule der Medien mit Kunstkopf-Stereofonie aufgezeichnet. Zum Einsatz kamen Soundman OKMII Elektret Mikrofone von geringer Qualität, die im Sweet Spot positioniert wurden. Die Ergebnisse waren unbefriedigend. Lokalisation, Distanz und Klangfärbung wurden verzerrt dargestellt und die wahrgenommene Größe der Aufnahme entsprach einer anderen als der, die während der Mischung im Sweet Spot angestrebt worden war. Auch die Bewegung der Audio-Objekte konnte nicht gut aufgelöst werden.

Die binaurale Aufnahme der Vorproduktion ermöglichte einen Einblick in die Probleme und Schwierigkeiten der binauralen Aufzeichnung populärer Musik. Es wurde festgestellt, dass eine solche Aufnahme des bereits gemischten Materials im Sweet Spot lediglich Teil einer wirkungsvollen Aufzeichnung sein kann. Übliche Inkopflokalisation der Kopfhörermischung können die Intensität und das Erlebnis der Musikwahrnehmung steigern. Für das Konzert soll mit Hilfe eines hochwertigen Kunstkopfes eine binaurale Aufnahme hergestellt werden, die in einer späteren Kopfhörermischung neben der gewöhnlichen Inkopflokalisation einen Raumeindruck vermitteln soll. Für die Aufnahmen der Live-Produktion wurde ein „HMS – AachenHead“ der Firma Head Acoustics eingesetzt, mit dem eine wesentlich höhere Klangtreue erzielt werden konnte. Die wahrgenommene Räumlichkeit der Aufnahme deckte sich allerdings nicht mit der Wahrnehmung im Raum. Auch bei dieser Aufnahme erzielten die bewegten Audio-Objekte nicht die gleiche Wirkung wie während der Wiedergabe auf der dreidimensionalen Lautsprecherkonfiguration.

Die nächsten Schritte in der Post-Produktion des Konzertes, welche auf Grund des Umfanges nicht mehr Teil dieser Arbeit sein können, werden zeigen, ob diese Aufnahmen im Rahmen einer binauralen Musik-Video-Mischung zu einem Mehrwert führen.

## 7 Schlussfolgerung und Ausblick

Mit dem aktuellen Stand der technischen Möglichkeiten im dreidimensionalen Audio können 3D-Audio-Live-Konzerte mit der entsprechenden Idee und Vorbereitungszeit umgesetzt werden. Diese Produktion hat gezeigt, dass es dabei viele Ansätze im Systemdesign und der dreidimensionalen Mischung geben kann. Während der Arbeit entstanden viele Ideen eines dreidimensionalen Beschallungs-Konzeptes. Die Möglichkeiten in der Mischung sind vielfältig und lassen den Audio-Ingenieur zum kreativen Teil der Performance werden. Der Einfluss von Technik und dreidimensionaler Mischung auf die Wirkung eines aufgeführten Songs darf nicht unterschätzt werden. Der Praxisversuch zeigt, dass eine dreidimensionale Mischung mit Hilfe von Techniken der gewöhnlichen Stereo-Arbeitsweise hergestellt werden kann.

Das finale Medium der Produktion, ob auf einem Live-Konzert oder auf der dreidimensionalen Hifi-Anlage zu Hause, orientiert sich stark an der Akzeptanz durch den Konsumenten. Zum gegenwärtigen Zeitpunkt schätzen wir die Bereitschaft, Mehrkosten für eine Live-Performance in einer dafür geeigneten Veranstaltungsstätte in Kauf zu nehmen, wesentlich höher ein, als diejenige, eine 11.1 Hifi-Anlage im eigenen Wohnzimmer zu betreiben. Zudem liegt die Einrichtung des technischen Setups bei einem Live-Konzert in den Händen professioneller Audio-Ingenieure. Somit kann die Richtigkeit der technischen Grundlagen des Systems kontrolliert und optimiert werden. Wohingegen eine 3D-Audio-Hifi-Anlage unter Umständen, nach für den Raum visuell ästhetischen Merkmalen positioniert und dabei die Wichtigkeit raumakustischer Merkmale bei der Aufstellung vernachlässigt wird. Solange ein solches Ungleichgewicht vorliegt und 3D-Audio-Standardisierungsversuche scheitern, gehen wir von einer stärker zunehmenden Entwicklung im Live-Bereich aus. Auch unter der Begründung, dass die Distribution von dreidimensionalem Inhalt bis zum Endkunden noch nicht final geklärt ist.

Das Verlangen nach visueller und auditiver Attraktion steigt kontinuierlich. Show- und Konzertkonzepte werden immer aufwändiger produziert und beeindrucken durch ein Verschmelzen von Künstler, Kunst und Technik. Dreidimensionales Audio wird sich im Event-, Unterhaltungs- und Live-Bereich weiterentwickeln und etablieren. Mit wachsenden Erwartungen der Zuschauer und Zuhörer steigt auch die Bereitschaft, für besondere audiovisuelle Erlebnisse einen höheren Preis zu bezahlen. Die auditive Präsentation wird dann am effektivsten, wenn es als Teil eines ganzheitlichen Erlebnisses betrachtet wird, das heißt im Verbund mit beispielsweise Licht- und Videogestaltung. Die technischen Gegebenheiten sollen die kreativen

Ausdrucksmöglichkeiten und Ausdrucksstärken der Künstler intensivieren und zu einem Gesamtkonzept werden lassen. In dieser Art von Gesamtkonzepten sehen wir die Zukunft mit dreidimensionalem Audio als festem Bestandteil.

Auch in unserer Produktion gab es ein Zusammenspiel zwischen Licht und Ton, welches das Konzert-Erlebnis durch weitere Wahrnehmungsebenen intensivierte. Dabei wurde anhand von Besucherreaktion und im Gespräch mit Konzertbesuchern ein von den Zuhörern empfundener Mehrwert der dreidimensionalen Beschallung bestätigt. Am Ende zählt der beeindruckende Moment, umhüllt von einem vielseitigen, schillernden und interdisziplinären Gesamtkunstwerk.

## 8 Referenzen

- Ainlay, Chuck, Chiccarelli, Joe, Clearmountain, Bob: 'The Recording Academy's Producers & Engineers Wing (2004): 'Recommendations For Surround Sound Production', unter:  
[http://www2.grammy.com/pdfs/recording\\_academy/producers\\_and\\_engineers/5\\_1\\_rec.pdf](http://www2.grammy.com/pdfs/recording_academy/producers_and_engineers/5_1_rec.pdf)  
[Zugriff am 17.05.2016]
- American Standard Association ASA (1960): American Standard Acoustical Terminology. New York
- Auro Technologies (2016): Auro-3D Authoring Tools - User Guide, Plug-in Version 2.2.0
- Barco Audio Technologies (2015a): Specification Iosono Scene Data Protocol V2.3.1
- Blauert, Jens (2014): 'Räumliches Hören', Hirzel Verlag, Stuttgart
- Baelen, W. (2011): 'Auro 3D - ein neuer Standard?', VDT Magazin , S. 13-18
- Baelen, W., Bert, T., Clapool, B., & Sinnaeve, T. (2012): White Paper - Auro 11.1, a new dimension in cinema sound. Barco, Inc.
- Bennett J.C., Barker K. and Edeko F.O. (1985): 'An approach to the assessment of stereophonic sound system performance', J. Audio Eng. Soc., Vol. 33, pp. 314–321
- Berkhout, A.J. (1988): 'A holographic approach to acoustic control'  
J. Audio Eng. Soc., Vol. 36, No. 12, pp. 977-995
- Berkhout, A.J., de Vries, D., Vogel, P. (1993): 'Acoustic control by wave field synthesis.', in:  
Journal of the Acoustical Society of America, Vol.93, No.5, May 1993, pp.2764-2778.
- Bitkom (2013): 'Kopfhörer-Trend durch Smartphone-Boom', unter:  
<https://www.bitkom.org/Presse/Presseinformation/Kopfhoerer-Trend-durch-Smartphone-Boom.html>  
[Zugriff am 16.05.2016].
- Bregman, A.S. (1994): 'Auditory Scene Analysis', Cambridge, MA. MIT Press
- Born, M., Wolf, E. (1975): 'Principles of Optics', New York: Pergamon Press, 5 1975, pp.370ff.
- Curtis, Robin (2008): 'Immersion und Einfühlung; Zwischen Repräsentationalität und Materialität bewegter Bilder', unter:  
[http://www.montage-av.de/pdf/172\\_2008/172\\_2008\\_Immersion\\_und\\_Einfuehlung.pdf](http://www.montage-av.de/pdf/172_2008/172_2008_Immersion_und_Einfuehlung.pdf)  
[Zugriff am 01.05.2016]
- Daele, B., & Baelen, W. (2012): Auro 3D White Paper - 'Productions in Auro 3D'
- d&b Audio (2016): 'Die T-Serie Broschüre', unter:  
<http://www.dbaudio.com/de/support/downloads/category/detail/dokumentation/broschueren.html#ssch-db-downloads-category-86>  
[Zugriff am 10.05.2016]

Dickreiter, Michael (1997 a): 'Handbuch der Tonstudioteknik', Band 1. 6. verbesserte Auflage, K.G. Saur, München

Dickreiter, Michael, Dittel, Volker, Hoeg, Wolfgang, Wöhr, Martin (2008): 'Handbuch der Tonstudioteknik'. K.G. Saur, München

Dickreiter, Michael; Dittel, Volker; Hoeg, Wolfgang; Wöhr, Martin (2014): 'Handbuch der Tonstudioteknik'. K.G. Saur, München

Duden: 'mischen', unter:

<http://www.duden.de/rechtschreibung/mischen#b2-Bedeutung-6>

[Zugriff am 29.12.2015]

Eben-Eschenbach, Marie (1893): 'Gesammelte Schriften', 1. Band. Aphorismen. Parabeln

Ebner M., Knoll J. (2009): 'Gehörgefährdung des Publikums bei Veranstaltungen' Kommentar zu DIN 15905-5, 1. Auflage, Beuth Verlag, Berlin

Eau Rouge (2013): 'Band', unter:

<http://eaurouge-music.com/?p=93>

[Zugriff am 07.12.2015]

Fraunhofer IDMT (2015): 'Spatial Sound Wave - Die neue Generation räumlicher Klangerlebnisse'

unter: <http://www.idmt.fraunhofer.de/content/dam/idmt/de/Dokumente/Publicationen/Produktinformationen/SpatialSound%20Wave/SpatialSoundWave.pdf>

[Zugriff am 03.01.2016].

Flückiger, Barbara (2001): 'Sound Desing. Die virtuelle Klangwelt des Films', Schüren Verlag, Marburg

Gerzon, M. (1973): 'With Height Sound Reproduction', AES Journal, Vol. 21, No. 1

Gerzon, M. (1975): 'The Design of Precisely Coincident Microphone Arrays for Stereo and Surround Sound', AES Convention 50, March 1975, Mathematical Institute, University of Oxford, England

Gernemann, Andreas: 'Surround-Sound Mikrophonierungen für Ensemblesmusik', unter:

<http://www.uni-koeln.de/phil-fak/muwi/ag/tec/sse.pdf>

[Zugriff am 07.01.2016]

Griffiths, Alison (2008): 'Shivers Down Your Spine: Cinema, Museums, and the Immersive View', Columbia University Press

Hammershøi D, Møller H (2005): 'Binaural technique – basic methods for recording, synthesis and reproduction', in: Blauert J (Hrsg.) 'Communication Acoustics', Springer, New York, S 223–254

Hagen, Mechthild (2006): 'Förderung des Hörens und Zuhörens in der Schule', Göttingen

Hoffman, Y.: 'Wellenfeldsynthese 3D: Ankunft akustischer Holografie in künstlerischer Praxis',

unter: <http://zkm.de/media/video/yannick-hofmann-wellenfeldsynthese-3d>

[Zugriff 15.05.2016]

Hobmair, Hermann (Hrsg.) (1997): 'Psychologie', Stam, Köln, München

Iosono GmbH (2014): 'Iosono Spatial Audio Workstation User Manual', unter:  
[http://www.iosono-sound.com/uploads/downloads/SpatialAudioWorkstation\\_Manual\\_2.2.pdf](http://www.iosono-sound.com/uploads/downloads/SpatialAudioWorkstation_Manual_2.2.pdf)  
[Zugriff am 15.12.2015]

Iosono GmbH (2014): 'Iosono Core User Manual', Erfurt

Iosono GmbH (2015a): 'References - Case Studies', unter:  
<http://www.iosono-sound.com/references/case-studies/>  
[Zugriff am 03.01.2016]

Iosono GmbH (2015b), unter:  
[http://www.iosonosound.com/fileadmin/user\\_upload/pdf/Case\\_Study\\_Live\\_Kraftwerk.pdf](http://www.iosonosound.com/fileadmin/user_upload/pdf/Case_Study_Live_Kraftwerk.pdf)  
[Zugriff am 15.05.2016]

Kronlacher, Matthias (2012): 'Ambisonics Binaural Dekoder Implementation als Audio Plug-in mit Headtracking zur Schallfeldrotation', Universität für Musik und darstellende Kunst Graz

Lawo AG (2016): 'Lawo präsentierte Neuheiten auf der NAB Show', unter:  
<https://www.lawo.com/de/aktuell/news/nab-showcase.html>  
[Zugriff am 17.05.2015]

Lensing, Jörg U. (2009): 'Sound-Design. Sound-Montage. Soundtrack-Komposition. Über die Gestaltung von Filmtönen', Fachverlag Schiele & Schön GmbH, Berlin

Lichtmond (2014): 'Lichtmond 3 – Mixing Auro 3D', unter:  
<https://www.youtube.com/watch?v=bZhFLzXmo9U>  
[Zugriff am 29.12.2015]

Martin, Bryan; King, Richard (2015): 'Three Dimensional Spatial Techniques in 22.2 Multi-channel Surround Sound for Popular Music Mixing', Convention Paper 9432 presented at the 139th Convention, New York. Audio Engineering Society

Melchior, Frank, Heusinger Udo, Liebetrau Judith (2011): 'Perceptual evaluation of a spatial audio algorithm based on wave field synthesis using a reduced number of loudspeakers', AES Convention Paper 8575, at the 131st Convention, New York USA

Mietzel, Gerd (1998): 'Wege in die Psychologie', Klett-Cotta, Stuttgart, Berlin

Mixerman (2010): 'Zen and the Art of Mixing', Milwaukee

Nachbar, C., F. Zotter, E. Deleflie und A. Sontacchi (2011): 'Ambix - A Suggested Ambisonics Format', in: Ambisonics Symposium 2011, Lexington

Owsinski, Bobby (2007): 'Mischen wie die Profis, Das Handbuch für Toningenieure', München

Plassmann, Ansgar A., Schmitt, Günter, Universität Duisburg-Essen: 'Soziale Aspekte der Wahrnehmung', in: Lern-Psychologie: Psychologie online lernen, unter:

<https://www.uni-due.de/edit/lp/common/sozialwahr.htm>  
[Zugriff am 28.04.16]

Pro Tools Expert (2014): 'Auro 3D: Using An AMS Neve DFC To Work In Auro 3D', unter:  
<https://www.youtube.com/watch?v=qoLErvB52c0>  
[Zugriff am 11.05.16]

Pulkki, V. (1997): 'Spatial Sound Generation and Perception By Amplitude Panning Techniques', Helsinki University of Technology Laboratory of Acoustics and Audio Signal Processing Espoo 2001 Report 62

Rumsey, F. (2002): 'Spatial Quality Evaluation for reproduced Sound: Terminology, Meaning, and a Scene-based paradigm', J. Audio Eng. Soc, Vol.50, No.9, September 2002, pp.651-666.

Rumsey et al. (2005): 'On the relative importance of spatial and timbral fidelities in judgments of degraded multichannel audio quality', JASA 118/2, August 2005, pp.968ff.

San Segundo, Carlos (2008): 'Die Latenz in der Musikproduktion', unter:  
<http://www.delamar.de/musikproduktion/die-latenz-in-der-musikproduktion-2838/>  
[Zugriff am 30.12.2015]

Sanders, Elizabeth B.-N. (2013): 'Perspectives on Participation in Design', in: Mareis, Claudia; Held, Matthias; Joost, Gesche (Hrsg.) (2013): 'Wer gestaltet die Gestaltung? Praxis, Theorie und Geschichte des partizipatorischen Designs', Bielefeld: Transcript S. 65 – 78

Schlemm, W.: 'Musikproduktion', in: Fischer, L. (Hrsg.) (1997): Die Musik in Geschichte und Gegenwart (MGG), Sachteil, Bd. 6, Sp. 1534ff, 2. Auflage, Bärenreiter, Metzler

Schliermann, D. (2015): 'Handbuch Bühnentechnik Bühnenbeleuchtung', Dieter Schliermann (Hrsg.), Nürnberg

Signature Sound Studio (2015): 'Bob Horn Quick Mix Walkthrough of the Kick and Snare', unter:  
<https://www.youtube.com/watch?v=uC06iQP2jAY>  
[Zugriff am 07.01.2016]

SMPTE (2014): 'Immersive Sound for Cinema, Standards Update Webcasts, Interoperable Immersive Sound Systems for Digital Cinema', unter:  
[https://www.smppte.org/sites/default/files/2014-11-04-ST-Immersive\\_Sound-Lude-v3-Handout.pdf](https://www.smppte.org/sites/default/files/2014-11-04-ST-Immersive_Sound-Lude-v3-Handout.pdf)  
[Zugriff am 03.01.2016]

SOS (1994): '3D Mixing: Giving Your Mixes More Space', in: SOS Ausgabe November 1994, unter:  
[http://www.soundonsound.com/sos/1994\\_articles/nov94/3dmixing.html](http://www.soundonsound.com/sos/1994_articles/nov94/3dmixing.html)  
[29.12.2015]

Sontacchi, A. (2003): 'Dreidimensionale Schallfeldreproduktion für Lautsprecher- und Kopfhöreranwendungen', Dissertation, Technische Universität Graz

Start, E. W. (1997): 'Direct Sound Enhancement by Wave Field Synthesis', PhD thesis, Technical University Delft

- Theile, G. (1981): 'Zur Theorie der optimalen Wiedergabe von stereofonen Signalen über Lautsprecher und Kopfhörer', Rundfunktechn. Mitt. 25, 1981, S. 155 - 170
- Theile, G. (1991): 'On the Naturalness of Two-Channel Stereo Sound', J. Audio Eng. Soc. 39, No. 10, 1991, S. 761 - 767
- Theile, G., Wittek, H., Reisinger, M. (2002): 'Wellenfeldsynthese-Verfahren: Ein Weg für neue Möglichkeiten der räumlichen Tongestaltung', Proceedings 21st Tonmeistertagung 2002, Hannover, Germany, November 2002.
- Theile, G., Wittek, H., Reisinger, M. (2003): 'Potential Wavefield Synthesis Applications in the Multichannel Stereophonic World', Proceedings 24th AES Int. Conference on Multichannel Sound, Banff, Canada, May 2003.
- Thomas, Nick (2009): 'Guide to Mixing v1.0', unter:  
<https://books.google.de/books?id=qzrRhccYnmgC&printsec=frontcover&hl=de#v=onepage&q&f>  
 [Zugriff am 31.04.2016]
- de Vries, D. and Berkhout, A.J. (1981): 'Wave theoretical approach to acoustical focusing', in: Journal of the Acoustical Society of America, Vol.70, No.3, September 1981, pp.740-748.
- de Vries, D. (1995): 'Sound Enhancement by Wave Field Synthesis: Adaptation of the Synthesis Operator to the Loudspeaker Directivity Characteristics'. Proceedings 98 vention, Paris, France, January 1995, Preprint No.3971.
- Weinzierl, S. (2008): 'Handbuch der Audiotechnik' Springer-Verlag, Berlin
- Weiss, P. (2015): 'Wenn Design die Materie verlässt: Sound. Das Design der Emotionen, der Imagination und der Lebendigkeit', Norderstedt, Deutschland: Books on Demand
- Wiggins, B. (2004): 'An Investigation into the Real-Time Manipulation and Control of Three-Dimensional Sound Fields', Dissertation, University of Derby
- Wittek, H. (2007): 'Perceptual differences between wavefield synthesis and stereophony', Department of Music and Sound Recording, University of Surrey
- Zahn, T. (2014): '3D-Sound im Fohhn SoundLab', Professional System 4/2013
- Zymara, Andreas (2002): 'Hören III: auditorische Wahrnehmung', in: Haider, F., Claessen, R.: Physik der Sinneswqahrnehmung, unter:  
[http://www.physik.uni-augsburg.de/~ferdi/sinnesorgane/hoeren\\_3/hoeren\\_3.pdf](http://www.physik.uni-augsburg.de/~ferdi/sinnesorgane/hoeren_3/hoeren_3.pdf)  
 [Zugriff am 08.05.16]

## 9 Abbildungsverzeichnis

Abbildung 1: Kanalbasiertes Audio.....	7
Abbildung 2: Objektbasiertes Audio .....	8
Abbildung 3: Kopfbezogenes Koordinatensystem .....	8
Abbildung 4: Die visuelle Perspektive kann mit der stereofonen Perspektive verglichen werden ...	12
Abbildung 5: Lokalisationswinkel der Phantomschallquelle $\theta$ und Versatzwinkel $\theta_0$ .....	14
Abbildung 6: Ambisonic Blockdiagramm .....	16
Abbildung 7: Erste Ordnung Ambisonics - W,X,Y und Z Kanal, dabei ist der In-Phase Anteil rot.....	17
Abbildung 8: Sphärische Harmonische .....	18
Abbildung 9: Huygen-Fresnel-Prinzip.....	19
Abbildung 10: Kirchoff-Helmholtz-Integral.....	20
Abbildung 11: Aliasing-Artefakte .....	22
Abbildung 12: Kunstkopfstereofonie.....	23
Abbildung 13: Binaurale Raumpulsantwort .....	24
Abbildung 14: Lautsprecheranordnung für Auro 3D (9.1) .....	26
Abbildung 15: Seitenansicht der Auro-3D Anordnung.....	28
Abbildung 16: Mögliche Verteilung der räumlichen Attribute bei Auro-3D (9.1) Wiedergabe; .....	29
Abbildung 17: AMS-Neve DFC Digital-Mixing-Konsole – 3D Panoramaregler .....	30
Abbildung 18: Auro 3D Blockdiagramm .....	31
Abbildung 19: Auro-Panner.....	32
Abbildung 20: Auro-Bus.....	32
Abbildung 21: Auro-Mixing-Engine .....	33
Abbildung 22: Auro-Return .....	33
Abbildung 23: Auro-Setitings - Fader Link.....	34
Abbildung 24: Iosono-Core.....	37
Abbildung 25: Standard-Verkabelung Iosono-Core für 64 Madi-Kanäle.....	38
Abbildung 26: Iosono Control Unit .....	39
Abbildung 27: Iosono Control Unit - I/O Control.....	39
Abbildung 28: Iosono Control Unit - I/O Control.....	39
Abbildung 29: Iosono-Core - Signalfluss.....	40
Abbildung 30: Iosono Loudspeaker-Setup.....	41

Abbildung 31: Iosono Loudspeaker-Setup .....	42
Abbildung 32: Iosono Control Unit - Processing Settings.....	43
Abbildung 33: Iosono Control Unit – Filter Configuration .....	44
Abbildung 34: Iosono Control Unit - Processing Status.....	45
Abbildung 35: Spatial Audio Workstation Plug-in (SAW) .....	47
Abbildung 36: Spatial Audio Wokrstation - Punktschallquelle und Ebene Welle.....	48
Abbildung 37: Signalfluss - Kraftwerk Livekonzert .....	51
Abbildung 38: Signalfluss - Fohhn Soundlab .....	52
Abbildung 39: The Squiggle - Designprozess .....	56
Abbildung 40: Messung und Wahrnehmung .....	57
Abbildung 41: Physikalische und perzeptive Wirkung von Regelzeiten - Attack.....	61
Abbildung 42: Physikalische und perzeptive Wirkung von Regelzeiten - Release .....	62
Abbildung 43: Auro 3D (9.1)Lautsprecherkonfiguration - HdM Tonstudio .....	77
Abbildung 44: Iosono (15.2) Lautsprecherkonfiguration - HdM Tonstudio .....	78
Abbildung 45: Iosono Loudspeaker Setup - HdM Tonstudio.....	79
Abbildung 46: Blockschaltbild Auro 9.1 Setup - HdM Tonstudio .....	80
Abbildung 47: Andiamo Remote Software .....	82
Abbildung 48: Iosono Interaktionsschnittstelle.....	86
Abbildung 49: Iosono Vertical Pan Protocol.....	87
Abbildung 50: Lokalisationsbearbeitung .....	95
Abbildung 51: Valhalla Shimmer Plug-in.....	96
Abbildung 52: Steven Slate Trigger Plug-in.....	97
Abbildung 53: Iosono SAW - Audio-Objekte .....	100
Abbildung 54: Nuendo SAW - Automationsdaten .....	101
Abbildung 55: Circles - Songstruktur .....	103
Abbildung 56: Circles – Instrumentierung.....	104
Abbildung 57: Circles - Basedrum Bearbeitung .....	105
Abbildung 58: Circles - Snare Bearbeitung.....	106
Abbildung 59: Circles - Drumgruppe Bearbeitung .....	107
Abbildung 60: Circles – Bassinstrumentierung .....	108
Abbildung 61: Circles – Kompression .....	109
Abbildung 62: Kurven gleicher Lautstärke .....	110

Abbildung 63: Entwurf: Lautsprecherkonfiguration (34.4) HdM U44 TV-Studio .....	113
Abbildung 64: Entwurf: Lautsprecherkonfiguration (34.4) HdM U44 TV-Studio .....	113
Abbildung 65: d&b Audio T10 Lautsprecher - Abstrahlcharakteristik .....	115
Abbildung 66: Lautsprecherkonfiguration HdM U44 TV-Studio - Front Array .....	115
Abbildung 67: Lautsprecherkonfiguration HdM U44 TV-Studio - Right .....	116
Abbildung 68: Lautsprecherkonfiguration HdM U44 TV-Studio - Back.....	117
Abbildung 69: Lautsprecherkonfiguration HdM U44 TV-Studio - Left .....	117
Abbildung 70: Lautsprecherkonfiguration HdM U44 TV-Studio - Top .....	118
Abbildung 71: ArrayCalc Software – Winkelberechnung.....	119
Abbildung 72: ArrayCalc - SPL Mapping Front-Array .....	120
Abbildung 73: ArrayCalc – Cabinet Layout und Winkelberechnung für Seitenlautsprecher.....	121
Abbildung 74: ArrayCalc – SPL Mapping Seitenlautsprecher .....	121
Abbildung 75: Lautsprecherkonfiguration U44 HdM TV-Studio.....	122
Abbildung 76: d&b Audio - D12 Verstärker .....	123
Abbildung 77: Iosono Control Unit - Processing-Settings - Front-Array .....	124
Abbildung 78: Iosono - Scene Designer Test-Objekt.....	125
Abbildung 79: Iosono Measurement.....	125
Abbildung 80: Iosono - Filter Creation - Whitelist .....	126
Abbildung 81: Iosono Filter Configuration .....	127
Abbildung 82: 3D Live Setup: AD/DA Wandler (links) Audio-Regie (rechts) .....	128
Abbildung 83: Eau Rouge 3D Konzert - Positionierung der Band .....	130
Abbildung 84: SAW - Positionierung der Vocal-Audio-Objekte .....	131
Abbildung 85: RME Totalmix FX .....	132

## 10 Anhang

Die Anhänge zur Masterarbeit können dem beiliegenden Datenträger entnommen werden.

Blockschaltbild Audio .....	I
Bühnenplan U44 HdM TV-Studio .....	II
Signalliste U54 HdM Tonstudio.....	III
Signalliste U44 HdM TV-Studio .....	IV
Materialliste .....	V
Bestückungsliste .....	VI
Multicore-Aufteilung.....	VII
Gefährdungsanalyse.....	VIII
Zeitplan .....	IX

Informationen und Medien über die Produktion können über [www.eaurouge3d.de](http://www.eaurouge3d.de) abgerufen werden



Fotograf & Urheber: Jan Potente und Ida Schneider



Fotograf & Urheber: Jan Potente und Ida Schneider



Fotograf & Urheber: Jan Potente und Ida Schneider