

---

# **BACHELORARBEIT**

---

**Herr Tobias Pannicke**

**Wodurch erreicht man  
auditive Immersion im System  
Auralite 3D und wo liegen die  
Unterschiede zu Dolby Atmos**

**2022**

# **BACHELORARBEIT**

---

## **Wodurch erreicht man auditive Immersion im System Auralite 3D und wo liegen die Unterschiede zu Dolby Atmos**

Verfasser:

**Herr Tobias Pannicke**

Studiengang:

**Media and Acoustical Engineering**

Seminargruppe:

**MG18wA-B**

Erstprüfer:

**Prof. Dr.-Ing. Michael Hösel**

Zweitprüfer:

**M. Sc. Sebastian Stingl**

Einreichung:

**Mittweida, 24.06.2022**

# BACHELOR THESIS

---

## **How auditive immersion is achieved in the system Auralite 3D and its differences to Dolby Atmos**

author:

**Mr. Tobias Pannicke**

course of studies:

**Media and Acoustical Engineering**

seminar group:

**MG18wA-B**

first examiner:

**Prof. Dr.-Ing. Michael Hösel**

second examiner:

**M. Sc. Sebastian Stingl**

submissionn:

**Mittweida, 24.06.2022**

## **Bibliografische Angaben**

Nachname, Vorname: Pannicke, Tobias

Thema der Bachelorarbeit: Wodurch erreicht man auditive Immersion im System Auralite 3D und wo liegen die Unterschiede zu Dolby Atmos

Topic of thesis: How auditive immersion is achieved in the system Auralite 3D and its differences to Dolby Atmos

52 Seiten, Hochschule Mittweida, University of Applied Sciences,

Fakultät Medien, Bachelorarbeit, 2022

## **Abstract**

Die vorliegende Arbeit hat zum Ziel dem Leser ein umfassendes Verständnis über immersive Audioproduktionen und damit einhergehende Wiedergabetechniken zu vermitteln. Dabei werden Grundlagen der Schallausbreitung, des menschlichen Gehörs sowie notwendige physikalische Gesetzmäßigkeiten erläutert. Ebenso werden etablierte Wiedergabetechniken für Audio behandelt und voneinander differenziert. Die gezielte Behandlung von immersivem dreidimensionalem Audio erfolgt anhand zweier Systembeispiele, welche in ihren Herangehensweisen unterschiedliche Philosophien verfolgen.

## **Danksagung**

Hiermit möchte ich mich bei all jenen bedanken, ohne deren Unterstützung die Anfertigung dieser Arbeit nicht möglich gewesen wäre. In erster Linie gilt mein Dank Prof. Dr.-Ing. Michael Hösel und M. Sc. Sebastian Stingl, welche durch ihre Betreuung diese Abschlussarbeit erst ermöglicht haben.

Des Weiteren bedanke ich mich bei der MediaApes GmbH, welche mir einen Einblick in ihr Studio und ihr immersives Audiosystem Auralite 3D gewährt haben. Darüber hinaus standen sie mir bei aufkommenden Detailfragen während meines Schreibzeitraums jederzeit zur Seite.

Ein besonderer Dank gilt meiner Familie und meinen Freunden für die enorme Unterstützung und den Rückhalt, den sie mir während der Anfertigung dieser Arbeit entgegenbrachten.

# Inhaltsverzeichnis

Inhaltsverzeichnis.....	I
Abbildungsverzeichnis.....	III
Abkürzungsverzeichnis .....	IV
1 Theoretische Grundlagen .....	1
1.1 Auditive Immersion.....	1
1.2 Schall und dessen Ausbreitung im Raum .....	2
1.3 Das menschliche Hören und die Lokalisation von Schallereignissen.....	4
1.4 Das Huygens-Prinzip.....	7
1.5 Das Kirchhoff-Helmholtz Integral.....	8
1.6 Die Wellenfeldsynthese.....	9
2 Wiedergabetechniken von Audiosignalen .....	11
2.1 Monophonie .....	11
2.2 Stereophonie.....	12
2.3 Surround-Sound.....	15
2.4 Kanal- und objektbasierte Audioproduktionen .....	16
2.4.1 Kanalbasiertes Audio .....	17
2.4.2 Objektbasiertes Audio .....	20
3 Dreidimensionales Audio .....	21
3.1 Auralite 3D .....	25
3.1.1 Technischer Aufbau .....	26
3.1.2 Funktionsweise .....	28
3.1.3 Ablauf einer Mischung.....	30
3.1.4 Einsatzmöglichkeiten .....	36
3.2 Dolby Atmos.....	37
3.2.1 Technischer Aufbau .....	38
3.2.2 Funktionsweise .....	41
3.2.3 Ablauf einer Mischung.....	43
3.2.4 Einsatzmöglichkeiten .....	48

3.3	Gegenüberstellung von Auralite 3D und Dolby Atmos .....	49
4	Fazit.....	51
	Literaturverzeichnis .....	V
	Anhang.....	IV
A.1	Ab Interview mit Nûjîn Kartal von der MediaApes GmbH .....	IV
	Eigenständigkeitserklärung .....	XIV

## Abbildungsverzeichnis

Abbildung 1 Überlagerung von zwei Schwingungen gleicher Frequenz.....	3
Abbildung 2 Beugung von Schallwellen um verschieden große Hindernisse .....	3
Abbildung 3 Querschnitt durch das menschliche Ohr .....	4
Abbildung 4 Koordinatensystem zur Lokalisation der Schalleinfallrichtung .....	6
Abbildung 5 Entscheidende Frequenzbänder für die Lokalisation in der Medianebene	7
Abbildung 6 Veranschaulichung des Huygensschen Prinzips.....	8
Abbildung 7 Bewegung einer Schallwelle durch eine perforierte Wand .....	9
Abbildung 8 Reproduktionen eines realen Schallfeldes .....	10
Abbildung 9 Optimale Abhörordnung bei Zweikanalstereophonie.....	13
Abbildung 10 Hörzonenbreiten bei Stereophonie mit 2, 3 und 4 Lautsprechern.....	14
Abbildung 11 Abhörkreis für 5.1 Surround-Sound .....	15
Abbildung 12 kanalbasierte Audiowiedergabe .....	17
Abbildung 13 Surround Panner einer 5.1 Mischung in Pro Tools.....	18
Abbildung 14 objektbasierte Audiowiedergabe .....	20
Abbildung 15 Dreidimensionales Amplituden-Panning .....	22
Abbildung 16 Aliasing Artefakte eines linearen Lautsprecherarrays .....	29
Abbildung 17 Output Routing in der DAW Nuendo .....	31
Abbildung 18 verschiedene Corpus-Konfigurationen .....	32
Abbildung 19 Übersicht aller aktiven Signalpegel während der Wiedergabe.....	33
Abbildung 20 Bedienoberfläche des Room Simulation Module - Configurator .....	34
Abbildung 21 Bedienoberfläche der Production App.....	35
Abbildung 22 Signalwege zwischen Pro Tools und der Dolby RMU .....	39
Abbildung 23 Lautsprecheranordnung von Dolby Atmos im Kinosaal.....	40
Abbildung 24 Funktionsprinzip eines Dolby Atmos Enabled Speaker .....	41
Abbildung 25 Dolby Atmos Panner Plug-In.....	45
Abbildung 26 Routing der Audiokanäle von Pro Tools zur Dolby RMU .....	46
Abbildung 27 Dolby Atmos Rendererer Hauptbildschirm .....	47



## Abkürzungsverzeichnis

<b>C</b>	Center
<b>DANTE</b>	Digital Audio Network through Ethernet
<b>DAW</b>	Digital Audio Workstation
<b>DBAP</b>	Distance Based Amplitude Panning
<b>HD</b>	High Definition
<b>HRTF</b>	Head Related Transfer Function
<b>L</b>	Left
<b>L OH</b>	Left Over Head
<b>LFE</b>	Low-Frequency Effects
<b>Lsr</b>	Left Surround Rear
<b>Lss</b>	Left Side Surround
<b>LTC</b>	Linear Timecode
<b>MADI</b>	Multi-Channel Audio Digital Interface
<b>R</b>	Right
<b>R OH</b>	Right Over Head
<b>RMU</b>	Rendering and Mastering Unit
<b>RSM</b>	Room Simulation Module
<b>Rsr</b>	Right Surround Rear
<b>Rss</b>	Right Side Surround
<b>VBAP</b>	Vector Base Amplitude Panning
<b>VR</b>	Virtual Reality
<b>WFS</b>	Wave Field Sythesis

# 1 Theoretische Grundlagen

Diese Arbeit beschäftigt sich mit immersiver Audiowiedergabe und zeigt, wie diese anhand zwei konkreter Systembeispiele mithilfe von Lautsprechern erzielt werden kann. Dieses Kapitel soll dem Leser einen Überblick über sämtliche Grundlagen bieten, welche für das Verständnis der restlichen Arbeit erforderlich sind.

## 1.1 Auditive Immersion

Zunächst gilt es zu definieren, was der Begriff Immersion aussagt und welche Anforderungen an die Audiosysteme dadurch resultieren. Immersiv gründet auf dem Lateinischen *immergere*, was eintauchen oder untertauchen bedeutet.<sup>1</sup> Somit hat immersives Audiomaterial zum Ziel, dem Zuhörer ein Eintauchen in die abgebildete Klangwelt zu ermöglichen. Alles Wahrgenommene soll sich möglichst real anfühlen und den Zuhörer im Zentrum des Geschehens positionieren. Dafür gibt es mehrere Möglichkeiten, welche sich anhand eines Beispiels anschaulich erläutern lassen. Man stelle sich ein Orchester in einem Konzertsaal vor. Sitzt man mittig und unmittelbar vor der Bühne, befinden sich die einzelnen Musikgruppen sowohl von links nach rechts als auch von vorne nach hinten verteilt. Sobald das Orchester zu spielen beginnt, sind die einzelnen Musiker oder Musikgruppen auch auditiv, also mithilfe der Ohren, lokalisierbar. Dies gilt es folglich auch in der akustisch immersiven Wiedergabe zu erzielen. Darüber hinaus sollten die einzelnen Schallquellen ortsstabil sein, wenn sich der Zuhörer im Schallfeld bewegt. Dreht der Rezipient beispielsweise seinen Kopf, sollten die Schallquellen stets an ihrer ursprünglichen Position wahrnehmbar sein und sich nicht der Kopfbewegung anpassen. Abweichend von der realen Positionierung des Rezipienten vor der Bühne könnte man ihn im virtuellen Klangraum auch mitten im Orchester platzieren. Die einzelnen Schallquellen wären dann um den Rezipienten herum verteilt. Dementsprechend bietet immersives Audio neben der möglichst realen Abbildung von Ereignissen auch die Möglichkeit den Zuhörer in völlig neue Erlebnisse zu integrieren. Die Einbettung des Adressaten in das abgebildete Schallfeld bleibt dabei die oberste Prämisse.

---

<sup>1</sup> Vgl. Josef M. Stowasser, Michael Petschenig und Franz Skutsch, *Stowasser: Lateinisch-deutsches Schulwörterbuch*, Auflage 1994 (Wien, München: Verlag Holder-Pichler-Tempsky R. Oldenbourg Verlag, 1994), 246.

Das Hören als Sinneswahrnehmung ist allein schon durch physische Individualität für jeden Menschen anders. Physikalisch definierbare Reize der Umwelt treffen als Schallereignisse auf das Gehör, werden darin weiterverarbeitet und schließlich an das Gehirn weitergeleitet. Erst durch die Analyse im Gehirn entstehen aus den physikalischen Schallereignissen emotionale Hörereignisse.<sup>2</sup> Trotz dieser individuellen und emotionalen Komponente des Hörens können verschiedene Mechanismen dazu genutzt werden, nahezu jeden Rezipienten in ein immersives Audioerlebnis eintauchen zu lassen. Entscheidend für einhüllendes Erlebnis ist, dass die technische Schallwiedergabe die menschliche Wahrnehmung in allen Aspekten anspricht.

## 1.2 Schall und dessen Ausbreitung im Raum

„Schall ist eine Form mechanischer Schwingungen und Wellen innerhalb eines elastischen Mediums“.<sup>3</sup> Diese werden durch ihre Frequenz, also ihre Schwingungsanzahl pro Sekunde, und ihre Auslenkung in Bezug auf ihre Ruhelage, auch Amplitude genannt, charakterisiert.

Innerhalb eines gasförmigen Mediums, beispielsweise Luft, breitet sich Schall in Form von Longitudinalwellen oder Kompressionswellen im Raum aus.<sup>4</sup> Das bedeutet, dass sich die einzelnen Luftmoleküle in Ausbreitungsrichtung der Schallwellen hin und her bewegen. Die Folge sind Über- und Unterdruckzonen, welche sich „periodisch im Abstand der Wellenlänge der Schwingung“<sup>5</sup> abwechseln. Diese periodischen Druckzonen können schließlich mithilfe der menschlichen Ohren wahrgenommen werden.

In der Realität interagieren Schwingungen, und somit auch Schallwellen, miteinander. Dem Umstand geschuldet, dass ein Molekül stets nur eine Schwingungsbewegung in einer Schwingungsrichtung ausführen kann<sup>6</sup>, kommt es zu verschiedenen Phänomenen, welche alle auf der Addition der einzelnen Schwingungsverläufe beruhen. Durch diese Überlagerung kann es in der resultierenden Schwingung zur Verstärkung (siehe Abbildung 1 links), Abschwächung oder Auslöschung (siehe Abbildung 1 rechts) der ursprünglichen Schwingungen kommen. Ebenso bestimmt die Phasendifferenz, also eine zeitliche Verschiebung der sich überlagernden Schwingungen, die Eigenschaften der resultierenden Schwingung.

---

<sup>2</sup> Vgl. Hubert Henle, *Das Tonstudio-Handbuch: Praktische Einführung in die professionelle Aufnahmetechnik*, 5., komplett überarbeitete Auflage (München: Carstensen, 2001), 50.

<sup>3</sup> Ebd., 27..

<sup>4</sup> Vgl. Andreas Friesecke, *Die Audio-Enzyklopädie: Ein Nachschlagewerk für Tontechniker*, 2. Auflage (Berlin: De Gruyter Saur, 2014), 21.

<sup>5</sup> Ebd.

<sup>6</sup> Vgl. Henle. 2001, 22.

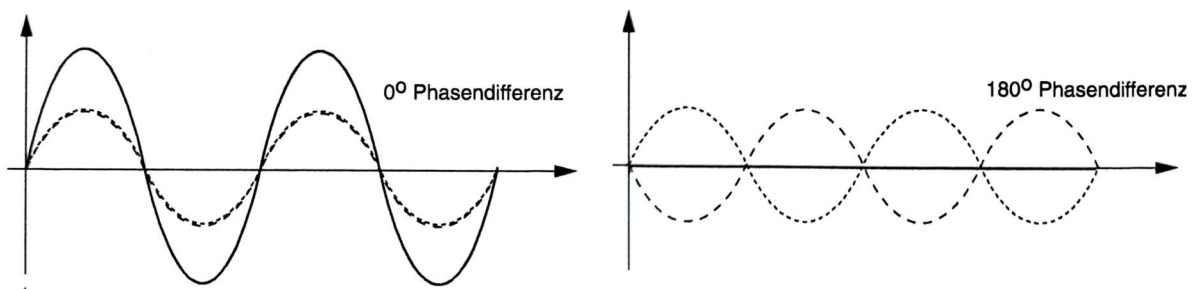


Abbildung 1 Überlagerung von zwei Schwingungen gleicher Frequenz<sup>7</sup>

Äquivalent zu den Gesetzen der Optik kann auch Schall reflektiert werden. An ebenen Flächen ist der Einfallswinkel der Schallwelle gleich ihrem Ausfallswinkel.<sup>8</sup> Voraussetzung dafür ist, dass die reflektierende Fläche in ihren Abmessungen groß gegenüber der Wellenlänge der reflektierten Schallwelle ist. Die Abmessungen sollten demnach mindestens einige Wellenlängen der reflektierten Schallwelle betragen.<sup>9</sup> Ist die Wellenlänge  $\lambda$  einer Schallwelle gleich oder größer ihrem Hindernis, so kommt es zur Beugung, also zur Umgehung des Hindernisses. Die Schallausbreitung wird infolgedessen nicht behindert (siehe Abbildung 2 oben).<sup>10</sup>

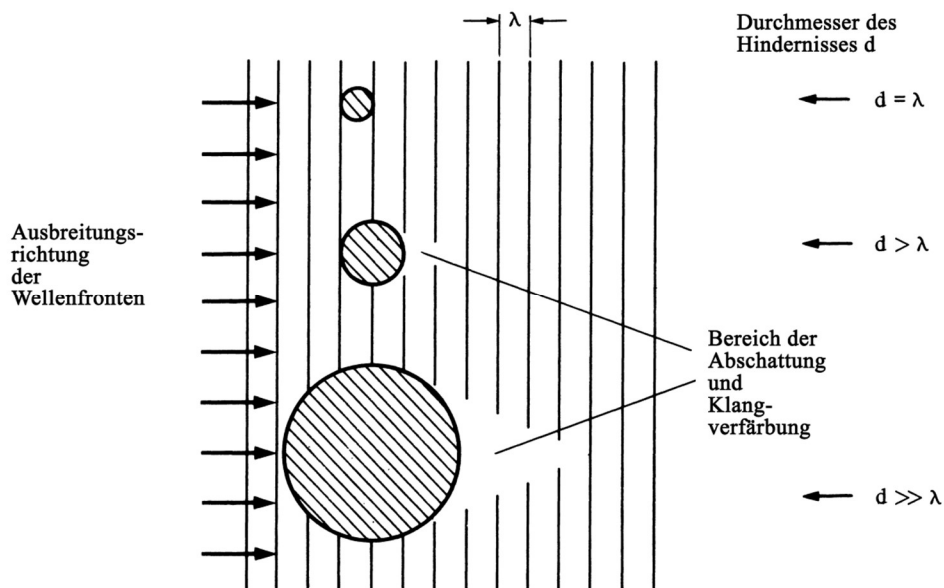


Abbildung 2 Beugung von Schallwellen um verschieden große Hindernisse<sup>11</sup>

<sup>7</sup> Henle. 2001, 22.

<sup>8</sup> Vgl. ebd., 38.

<sup>9</sup> Vgl. ebd.

<sup>10</sup> Vgl. ebd., 40.

<sup>11</sup> Michael Dickreiter et al., Hrsg., *Handbuch der Tonstudiotchnik*, 8. überarbeitete und erweiterte Auflage (München: Saur, 2014), 22.

Ein weiteres Phänomen in der Schallausbreitung zeigt sich in der Intensitätsabnahme des von einer Schallquelle abgegeben Schalls mit zunehmender Entfernung zu seinem Ursprung. Dies ist theoretisch auf zwei Faktoren zurückzuführen. Einerseits verteilt sich die Schallenergie im Normalfall auf eine immer größer werdende Fläche, wodurch die Energiedichte sinkt. Andererseits wird ein Teil der Schallenergie in andere Energieformen umgesetzt.<sup>12</sup> Bei einer kugelförmigen Schallausbreitung, wie sie bei den meisten Lautsprechern zu finden ist, nimmt die Schallintensität pro Abstandsverdopplung zur Schallquelle um sechs Dezibel ab.<sup>13</sup>

### 1.3 Das menschliche Hören und die Lokalisation von Schallereignissen

Das menschliche Gehör, welches im Idealfall Frequenzen von 16 bis 20 000 Hertz wahrnehmen kann<sup>14</sup>, setzt sich aus mehreren Komponenten zusammen. Seine physiologischen Bestandteile werden durch die Verarbeitung der aufgenommenen Informationen im Gehirn ergänzt.<sup>15</sup> Physiologisch besteht das Ohr aus dem Außenohr, Mittelohr und dem Innenohr, wobei jedem Abschnitt eine konkrete Aufgabe zuteilwird.

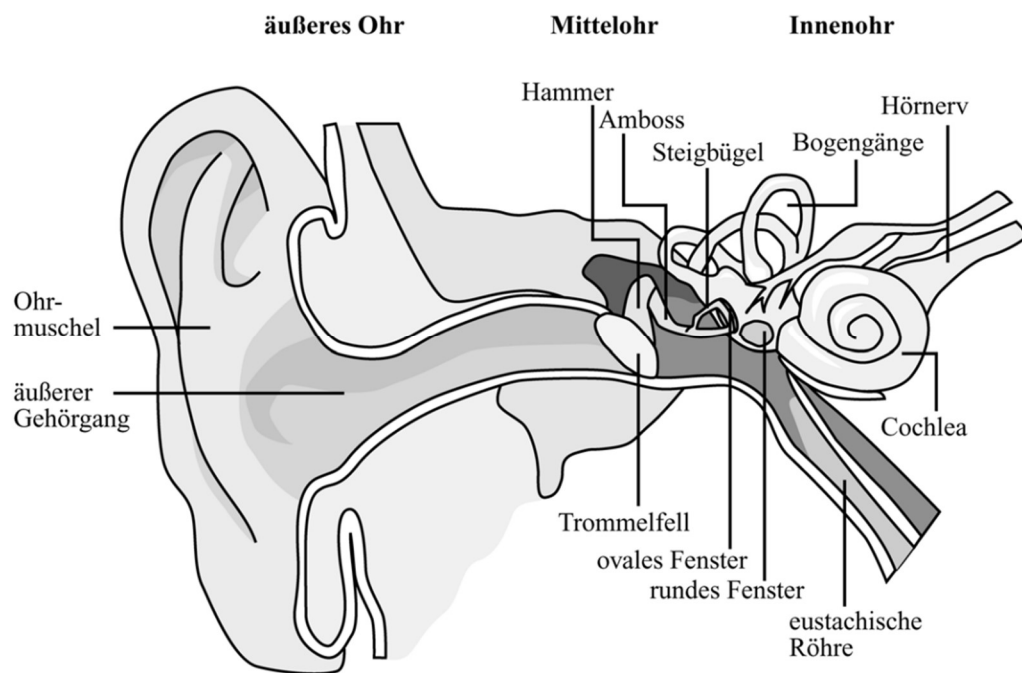


Abbildung 3 Querschnitt durch das menschliche Ohr<sup>16</sup>

<sup>12</sup> Vgl. E. Lübcke, *Schallausbreitung*. [https://link.springer.com/chapter/10.1007/978-3-642-47352-4\\_15](https://link.springer.com/chapter/10.1007/978-3-642-47352-4_15) (letzter Zugriff: 28. Mai 2022), 648. H. Backhaus et al., Hrsg., *Akustik*. Handbuch der Physik 8 (Berlin, Heidelberg: Springer, 1927).

<sup>13</sup> Vgl. Friesecke. 2014, 27.f.

<sup>14</sup> Vgl. ebd., 117.

<sup>15</sup> Vgl. Dickreiter et al. 2014, 115.

<sup>16</sup> Ebd.

*„Das Außenohr [, bestehend aus der Ohrmuschel und dem äußeren Gehörgang,] dient der Schallankopplung des Ohrs an das umgebende akustische Feld und der Umwandlung der Luftschwingungen in mechanische Schwingungen“.<sup>17</sup>*

Die individuelle Beschaffenheit der Ohrmuschel bedingt, je nach Einfallsrichtung des Schalls, unterschiedliche Reflexionen und somit Überlagerungen der aufgenommenen Schallwellen. Die dadurch veränderten akustischen Signale gelangen anschließend durch den äußeren Gehörgang zum Mittelohr.

Das durch die aufgenommenen Schallwellen in Schwingung versetzte Trommelfell bildet die Begrenzungsfläche des Außenohrs zum Mittelohr. An ihm ist ein Hebelsystem aus Gehörknöchelchen befestigt, welches die Vibrationen des Trommelfells durch das ovale Fenster an das Innenohr überträgt. „Der auf das ovale Fenster übertragene Druck ist bei kleiner Membranauslenkung [...] wesentlich größer als der Schalldruck, der auf das Trommelfell einwirkt; umgekehrt wird die Membranauslenkung entsprechend reduziert“.<sup>18</sup> Dadurch erfolgt eine Umwandlung vom Luftschall der Umgebung zum weiter verarbeitbaren Körperschall.

Im Innenohr wird der Körperschall schließlich in Nervenimpulse umgewandelt und zum Gehirn weitergeleitet. Im Gegensatz zum luftgefüllten Außen- und Mittelohr befindet sich im Inneren der Gehörschnecke des Innenohrs Flüssigkeit. Diese Flüssigkeit überträgt den ankommenden Körperschall und regt infolgedessen die sogenannten Haarzellen an. Aufgrund der Beschaffenheit der Gehörschnecke und der Haarzellen werden je nach Schallereignis nur bestimmte Nervenimpulse ausgelöst, wodurch das aufgenommene Schallereignis im Gehirn analysiert werden kann.

Für die Feststellung, aus welcher Richtung ein Schallereignis kommt, sind mehrere Faktoren entscheidend. Für eine eindeutige Zuordnung der Richtung bezieht man sich auf ein kopfbezogenes Koordinatensystem gemäß Abbildung 4.

---

<sup>17</sup> Dickreiter et al. 2014, 115.

<sup>18</sup> Ebd., 116.

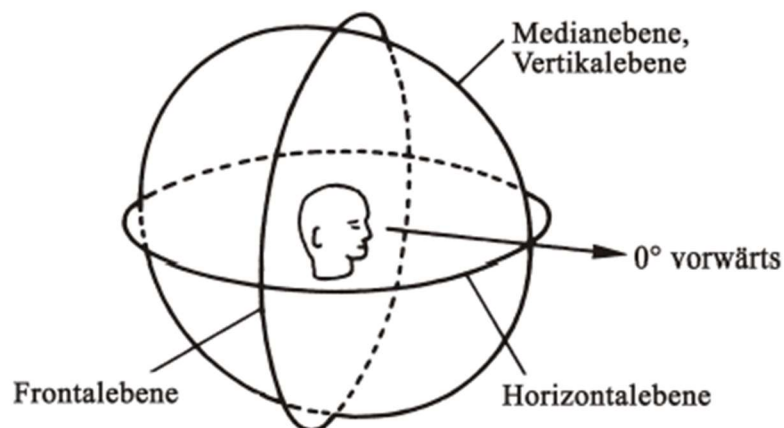


Abbildung 4 Koordinatensystem zur Lokalisation der Schalleinfallrichtung<sup>19</sup>

Laufzeit- und Pegelunterschiede zwischen den beiden Ohren bilden die fundamentale Grundlage der Lokalisation in der horizontalen Ebene<sup>20</sup>. Laufzeitunterschiede, also das zeitversetzte Eintreffen von Schallwellen an den Ohren, entstehen durch den Abstand der menschlichen Ohren zueinander. Zusätzlich treten dabei Pegelunterschiede auf, da der Schalldruckpegel einer Schallwelle mit zunehmendem Ausbreitungsweg abnimmt. Das Ohr, auf welches die Schallwellen zuerst treffen nimmt in der Regel dementsprechend auch mehr Schalldruckpegel wahr.

Ein weiterer Faktor der Lokalisation von Schallquellen sind Klangfarbenunterschiede, also unbewusst wahrgenommene Hervorhebungen bestimmter Frequenzbänder je nach Einfallrichtung des Schalls. Diese werden „durch die Form und Beschaffenheit von Kopf und Ohren verursacht“<sup>21</sup> und sind besonders in der vertikalen Medianebene ausschlaggebend, da hier keine Laufzeitunterschiede zwischen den Ohren auftreten<sup>22</sup>. Für die erfolgreiche Analyse der Klangfarbenunterschiede ist besonders der Bekanntheitsgrad einer Schallquelle relevant. Um auch unbekannte Schallquellen in der Medianebene möglichst genau lokalisieren zu können, drehen Menschen unbewusst ihren Kopf geringfügig. Somit werden zusätzlich Laufzeit- und Pegelunterschiede provoziert, welche eine exakte Lokalisation begünstigen.<sup>23</sup> Ausschließlich breitbandige Signale können in der Medianebene zuverlässig verortet werden.<sup>24</sup>

<sup>19</sup> Dickreiter et al. 2014, 128.

<sup>20</sup> Vgl. ebd., 129.

<sup>21</sup> Ebd., 131.

<sup>22</sup> Vgl. ebd.

<sup>23</sup> Vgl. Henle. 2001, 55.

<sup>24</sup> Vgl. Dickreiter et al. 2014, 132.

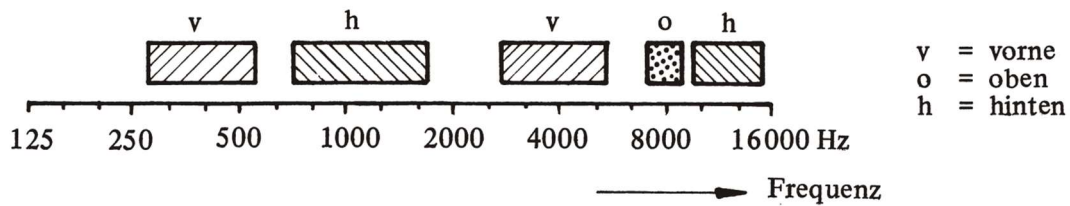


Abbildung 5 Entscheidende Frequenzbänder für die Lokalisation in der Medianebene<sup>25</sup>

Entfernungen werden neben Pegelwerten ebenfalls über Klangfärbung aber auch über das Zusammenspiel von Direkt- und Diffusschall, also dem direkten Signal und dessen Reflexionen im Raum, wahrgenommen.<sup>26</sup> Ein lautes und klares Direktsignal mit verhältnismäßig wenig Diffusschall wird als nah empfunden, wohingegen ein leises und dumpf klingendes Signal mit viel Raumanteil weit entfernt scheint.

Eine exakte Lokalisation tieffrequenter Schallereignisse ist physikalisch bedingt kaum bis gar nicht möglich. Je niedriger die Frequenz einer Schwingung ist, umso größer ist deren Wellenlänge. Demzufolge führen die kleinen Laufzeitunterschiede zwischen den Ohren lediglich zu einer minimalen Phasenverschiebung im Schwingungsverlauf der Ohrsignale. Zusätzlich können sich tieffrequente Schallwellen unterhalb von 300 Hertz um den Kopf beugen, wodurch Pegelunterschiede als Lokalisationsmechanismus ebenfalls entfallen.<sup>27</sup>

Lernprozesse spielen in jedem Aspekt der Lokalisation von Schallquellen eine entscheidende Rolle.<sup>28</sup> Je nach Bekanntheitsgrad und bereits gewonnenen Erfahrungswerten mit der Schallquelle, wird deren Lokalisation begünstigt beziehungsweise erschwert.<sup>29</sup>

## 1.4 Das Huygens-Prinzip

Die technische Erzeugung eines dreidimensionalen Schallfeldes beruht auf mehreren Gesetzmäßigkeiten. Das Huygens-Prinzip bildet dabei die erste Theorie. Es besagt, dass jeder Punkt einer Wellenfront Ausgangspunkt einer kugelförmigen Elementarwelle sei.<sup>30</sup> Die Einhüllende, also die Überlagerung aller Elementarwellen bildet „eine neue Wellenfront, die mit der ursprünglichen Wellenfront identisch ist“.<sup>31</sup> Somit ist es möglich „jede von einer Quelle

<sup>25</sup> Dickreiter et al. 2014, 131.

<sup>26</sup> Vgl. Friesecke. 2014, 144.

<sup>27</sup> Vgl. ebd., 135.

<sup>28</sup> Henle. 2001, 55.

<sup>29</sup> Vgl. Dickreiter et al. 2014, 132.

<sup>30</sup> Vgl. ebd., 353.

<sup>31</sup> Ebd.



[...] erzeugte Wellenfront - und damit auch sich ausbreitende[n] Schall - als Summe von unendlich vielen Einzelwellen<sup>32</sup> zu betrachten.

Folglich kann jedes beliebige Schallereignis von einer theoretisch unendlichen Anzahl an Punktschallquellen, beispielsweise Lautsprechern, realitätsgetreu wiedergegeben werden.

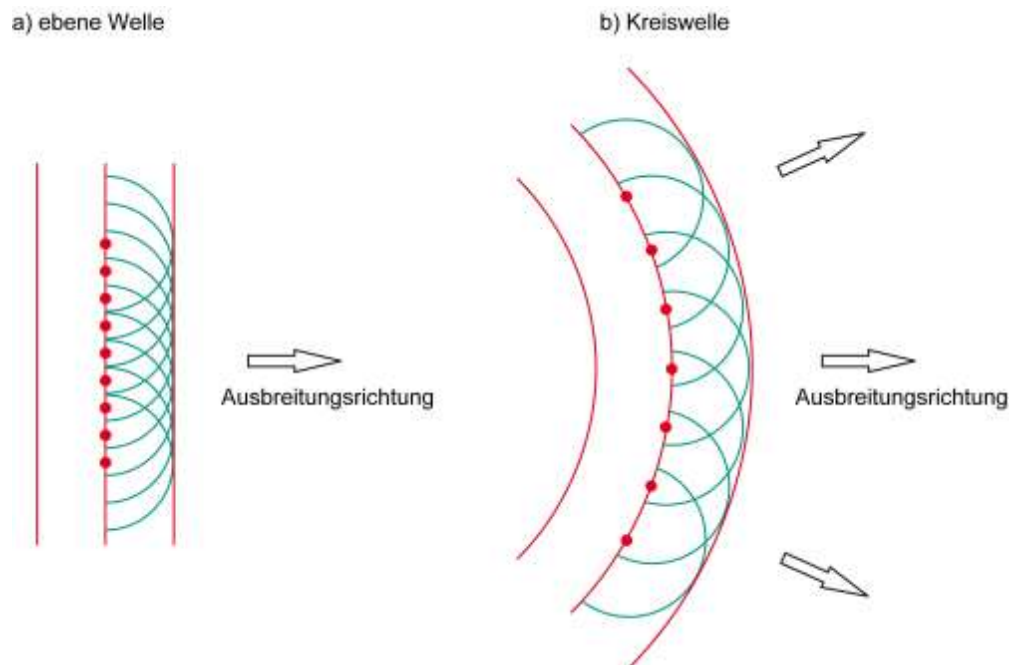


Abbildung 6 Veranschaulichung des Huygensschen Prinzips<sup>33</sup>

## 1.5 Das Kirchhoff-Helmholtz Integral

Die zweite theoretische Grundlage bildet das Kirchhoff-Helmholtz Integral.

*„Sind Schalldruck und Schallschnelle aller Punkte der Umgrenzung einer geschlossenen Fläche bekannt – oder eines Volumens im dreidimensionalen Modell –, kann der Schalldruck jedes Punktes in dieser Fläche bzw. in diesem Volumen berechnet werden.“<sup>34</sup>*

Demnach „ergibt sich der Schalldruck am Hörerplatz als Summe von Schallsignalen auf der Oberfläche“<sup>35</sup> des Raumes. Ein Zusammenspiel beider Theorien lässt sich im Modell des akustischen Vorhangs veranschaulichen. Dieser geht von einem Schallereignis und einer großflächig durchlöcherter Wand aus. Trifft die Wellenfront auf die perforierte Wand, so

<sup>32</sup> Dickreiter et al. 2014, 353.

<sup>33</sup> Hochschule der Medien, „Die Wellenfeldsynthese.“ <https://curdt.home.hdm-stuttgart.de/PDF/WFS%202.pdf> (letzter Zugriff: 28. August 2021), 4.

<sup>34</sup> Dickreiter et al. 2014, 353.

<sup>35</sup> Stefan Weinzierl, Hrsg., *Handbuch der Audiotechnik* (Berlin, Heidelberg: Springer, 2008), 666. f.

gelangen Teile der Wellenfront in die einzelnen Löcher. Diese Wellenfrontfragmente beinhalten gemäß des Huygensschen Prinzips in sich unendlich viele kugelförmige Elementarwellen, welche sich auf der anderen Wandseite räumlich ausbreiten. Dabei überlagern sie sich und bilden eine neue Wellenfront, welche identisch zur Ausgangswelle ist. Dadurch wird der Schall realitätsgetreu abgebildet.

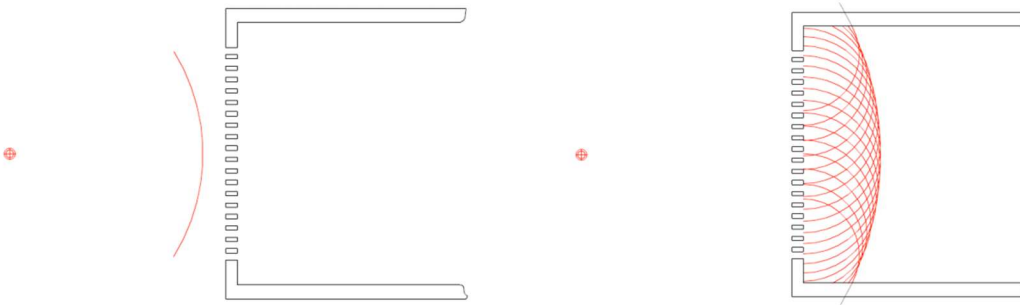


Abbildung 7 Bewegung einer Schallwelle durch eine perforierte Wand<sup>36</sup>

Ersetzt man jedes Loch jeweils mit einem Mikrofon und einem damit verbundenen Lautsprecher bleibt das Prinzip bestehen.<sup>37</sup> Das Mikrofon nimmt die Schallwelle auf der einen Wandseite auf und der Lautsprecher gibt sie auf der anderen wieder. Erneut entsteht eine Einhüllende, welche das Schallereignis auf der anderen Wandseite abbildet. Die Mikrofone und Lautsprecher stehen dabei sinngemäß für das Kirchhoff-Helmholtz Integral, indem sie die Schallinformationen an der Raumbegrenzungsfläche aufnehmen und anschließend für alle Punkte im Raum reproduzieren.

## 1.6 Die Wellenfeldsynthese

Beide Theorien finden auch im 1980 formulierten Konzept der Wellenfeldsynthese, auch Holophonie genannt, Anwendung.<sup>38</sup> Sie „ist ein Verfahren, mit dem ein Schallfeld eines Raums in einem anderen Raum realistisch nachgebildet werden kann“.<sup>39</sup> Dabei nimmt man jedoch nicht wie beim akustischen Vorhang die Schallwellen eines Ereignisses an möglichst vielen Punkten auf, sondern beschränkt sich auf zwei Faktoren. Man nimmt einerseits die von anderen Klängen isolierten Schallquellen und andererseits die akustischen Eigenschaften des darzustellenden Raumes auf. Die Raumeigenschaften werden über eine sogenannte Impulsantwort ermittelt, also den Nachhall des Raumes auf einen sehr kurzen Schallimpuls<sup>40</sup>.

<sup>36</sup> Hochschule der Medien, 5.

<sup>37</sup> Vgl. Dickreiter et al. 2014, 354.

<sup>38</sup> Vgl. Weinzierl. 2008, 667.

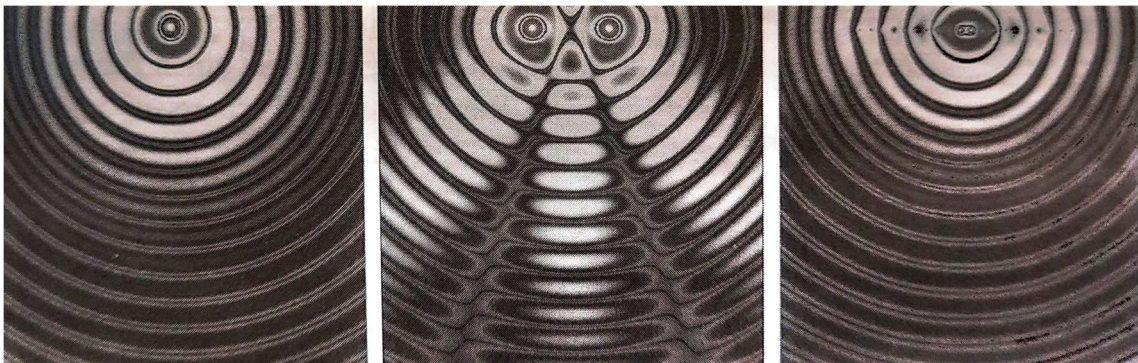
<sup>39</sup> Dickreiter et al. 2014, 352.

<sup>40</sup> Vgl. ebd., 354.

Mittels spezieller Software werden die Einzelschallquellen in einem virtuellen Raum positioniert und mit den aufgenommenen Raumeigenschaften verrechnet.

Wiedergabeseitig sind Lautsprecherarrays, also Aneinanderreihungen vieler einzelner Lautsprecher, erforderlich, welche an den Begrenzungsflächen des Beschallungsraumes verteilt angebracht sind. Diese geben Schalldruck sowie Schallschnelle an den Raumbegrenzungsflächen vor. Mithilfe des Kirchhoff-Helmholtz-Integrals können dadurch Schalldruck und -schnelle für jeden Punkt im geschlossenen Volumen des Beschallungsraumes berechnet und somit erzeugt werden. Die dadurch berechnete Wellenfront wird anschließend als Einhüllende der einzelnen Lautsprecherarrays wiedergegeben und bildet somit das ursprüngliche Schallfeld realitätsgetreu nach. Die einzelnen Lautsprecher fungieren dabei gemäß des Huygens-Prinzips als Elementarwellen, welche in Summe eine neue Wellenfront bilden.

Anders als bei der Stereophonie, welche eine akustische Differenzierung lediglich innerhalb einer im Verhältnis zum Beschallungsraum kleinen Hörzone, dem sogenannten Sweetspot, ermöglicht, wäre die Lokalisation einzelner Schallquellen innerhalb des Schallfeldes bei der Holophonie ortsunabhängig möglich. Der Zuhörer ist demnach nicht an einen Sweetspot gebunden und kann sich im Schallfeld frei bewegen.



*Abbildung 8 Reproduktionen eines realen Schallfeldes<sup>41</sup>*

Abbildung 8 zeigt den Unterschied der Wellenfeldstrukturen. Das reale Schallfeld einer punktförmigen Quelle (siehe Abbildung 8 links) kann mithilfe der Stereophonie zweier Lautsprecher im Abstand von 2 m zueinander (siehe Abbildung 8 Mitte) nur bedingt reproduziert werden. Die in der Mitte zwischen den Lautsprechern befindliche Phantomschallquelle kann nur in einem kleinen Bereich des Schallfeldes korrekt lokalisiert werden. Die Reproduktion des Schallfeldes mithilfe der Wellenfeldsynthese durch ein Lautsprecherarray aus 20 Lautsprechern im Abstand von 30 cm (siehe Abbildung 8 rechts) ist

---

<sup>41</sup> Weinzierl. 2008, 665.

wesentlich näher am natürlichen Original. Die korrekte Lokalisation der Phantomschallquelle ist fast im gesamten Schallfeld möglich.<sup>42</sup>

Die Theorie der Holofonie ist durch mehrere Faktoren limitiert und dementsprechend in der Praxis nicht unmittelbar umsetzbar. Beispielsweise lassen sich Lautsprecher aufgrund ihrer Baugröße nicht unendlich dicht nebeneinander anordnen, um das tatsächliche Verhalten von Elementarwellen zu reproduzieren. Dadurch entstehen ungewollte Interferenzen der einzelnen Lautsprechersignale. Zusätzlich verfälschen Reflexionen des Wiedergaberaums das synthetisierte Schallfeld. Selbst die erforderlichen Übertragungskanäle stellen ein Problem dar. Diese steigen für eine exakte Schallfeldsynthese in etwa proportional zur Anzahl der zu übertragenden Schallquellen.<sup>43</sup>

Trotz dieser Limitierungen zeigen die Systembeispiele, welche in Kapitel 3 explizit behandelt werden, dass praktische Umsetzungen diesem Ideal der realistischen Schallwiedergabe nahekommen. Auch andere Funktionsprinzipien können eine von der Abhörposition unabhängige Lokalisation von Schallquellen bewerkstelligen.

## **2 Wiedergabetechniken von Audiosignalen**

Die Funktionsweisen der in dieser Arbeit thematisierten Audiosysteme sind Weiterentwicklungen aller bisherigen Wiedergabemöglichkeiten. Der breiten Masse der Bevölkerung sind die verschiedenen Wiedergabetechniken für Audio zumindest begrifflich bekannt, jedoch in ihren Funktionsprinzipien nicht klar voneinander abgegrenzt. Zur konkreten Differenzierung der einzelnen Evolutionsstufen werden im Folgenden die primären Merkmale der einzelnen lautsprecherbasierten Techniken behandelt. Daraus ergeben sich die spezifischen Unterschiede sowie ein Verständnis darüber, was technisch zu leisten ist, um von einer punktuellen zu einer immersiven Audiowiedergabe zu gelangen. Mit zunehmender Komplexität der Wiedergabesysteme steigt grundsätzlich auch das immersive Abhörerlebnis aber auch gleichzeitig der Produktionsaufwand.

### **2.1 Monophonie**

Bei einer monophonen Audiowiedergabe steht nur ein Audiokanal für die Signalübertragung zur Verfügung. Somit kommen sämtliche Informationen punktuell aus ein und demselben Lautsprecher, wodurch eine räumliche Differenzierung der einzelnen Geräusche kaum möglich ist. Es kann lediglich ein grober Entfernungseindruck durch unterschiedliche

---

<sup>42</sup> Vgl. Weinzierl. 2008, 665.

<sup>43</sup> Dickreiter et al. 2014, 356.ff.

Hallanteile einzelner Geräusche erzeugt werden. Demzufolge bietet eine Audiowiedergabe über Mono kein immersives Abhörerbis. Trotzdem hat die Monowiedergabe auch in der heutigen Zeit noch ihre Daseinsberechtigung „für Anwendungen, wo der reine Informationsgehalt einer Wiedergabe oder Übertragung im Vordergrund steht, also beispielsweise bei Nachrichten und Dokumentationen“.<sup>44</sup> Ein weiterer Vorteil der Monotechnik liegt in der Lokalisationsstabilität der Schallquelle, welche unabhängig von der Abhörposition im Lautsprecher wahrgenommen wird.<sup>45</sup>

## 2.2 Stereophonie

*„Die Abbildung eines akustischen Geschehens durch Lautsprecher gewinnt mit der Stereotechnik an Durchsichtigkeit, Räumlichkeit und Klangfülle gegenüber der Abbildung mit Monotechnik“.<sup>46</sup>*

Technisch arbeitet Stereophonie mit der Überlagerung der Schallfelder zweier Lautsprecher, wodurch eine eindimensionale räumliche Verteilung von Schallquellen auf der Verbindungsachse zwischen benachbarten Lautsprechern möglich wird. Für das Gelingen dieser räumlichen Verteilung muss sich der Rezipient in einem bestimmten Verhältnis zum Lautsprecherpaar befinden. Ein gleichseitiges Dreieck gemäß Abbildung 9 gilt dabei als Standard, wobei sich die Abhörposition H und die Lautsprecher L1 sowie L2 in den Eckpunkten des Dreiecks befinden. Dabei ergeben sich in allen Eckpunkten des Dreiecks Innenwinkel von 60 Grad.

---

<sup>44</sup> Henle. 2001, 56.

<sup>45</sup> Vgl. Friesecke. 2014, 785.

<sup>46</sup> Dickreiter et al. 2014, 218.

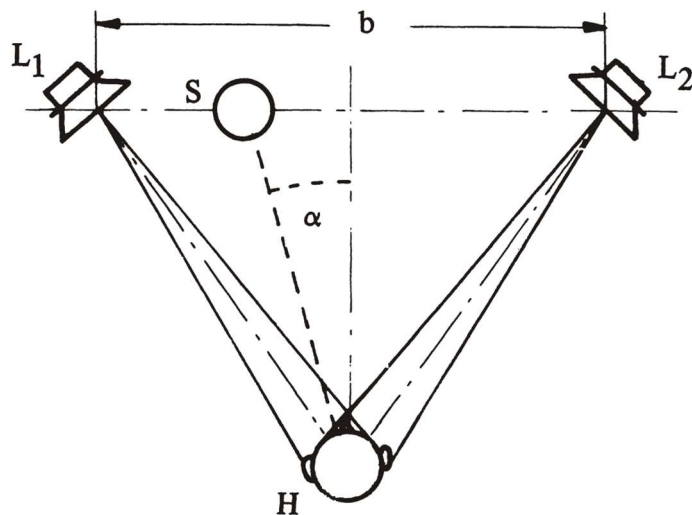


Abbildung 9 Optimale Abhörordnung bei Zweikanalstereophonie<sup>47</sup>

Geben beide Lautsprecher das gleiche Signal mit gleichem Pegel wieder, so nimmt der Zuhörer eine Phantomschallquelle S in der Mitte zwischen beiden Lautsprechern wahr. Diese „ist fiktiv [...], weil an diesem Punkt keine wirkliche Schallquelle existiert, der Ort der Schallquelle also nur in der auditiven Verarbeitung der Ohrsignale durch den Hörer gebildet wird“.<sup>48</sup> Durch Laufzeit- und Pegeldifferenzen zwischen dem Linken und dem Rechten Audiokanal können einzelne Geräusche beliebig zwischen dem Lautsprecherpaar im Mischvorgang positioniert und beim Abhören lokalisiert werden. Für die Lokalisation von Phantomschallquellen ist es zwar ausreichend, dass ausschließlich Laufzeit- oder Pegeldifferenzen zwischen den Lautsprechern auftreten, jedoch ist eine Kombination beider Mechanismen effektiver.<sup>49</sup> Besonders im Bereich großer Signaldifferenzen und somit großer Auslenkung der Phantomschallquelle aus der Mitte der Lautsprecher tragen wachsende Pegeldifferenzen positiv zur Lokalisationsschärfe bei, wohingegen eine Erhöhung der Laufzeitunterschiede der Lokalisationsschärfe konträr gegenübersteht.<sup>50</sup>

*„Verlässt der Hörer die Symmetrieachse der Lautsprecheranordnung, entstehen Verzerrungen der Richtungsdarstellung. Die sog. Abbildungsverzerrungen [...] ergeben sich, weil die von den Lautsprechern abgestrahlten Signale mit zusätzlichen Laufzeitdifferenzen beim Hörer eintreffen und damit ein Auswandern der Phantomschallquellen in Richtung des näheren Lautsprechers verursachen“.<sup>51</sup>*

<sup>47</sup> Dickreiter et al. 2014, 221.

<sup>48</sup> Ebd.

<sup>49</sup> Vgl. ebd., 222–225.

<sup>50</sup> Vgl. ebd., 225.f.

<sup>51</sup> Ebd., 228.

Zur Verbreiterung der Hörzone, welche bei der klassischen Zweikanal-Stereophonie in etwa 20 cm breit ist, werden bei der Mehrkanal-Stereophonie mehr Lautsprecher zwischen links und rechts verteilt, wodurch der Gesamtabstand in mehrere stereophone Teilbereiche gegliedert wird. Dies bewirkt neben der Verbreiterung der Hörzone eine Verbesserung der Lokalisationsschärfe.<sup>52</sup> Schon allein die gängige Aufteilung der Stereobasis in einen linken, einen Center und einen rechten Lautsprecher bewirkt ein realeres Hörerlebnis, da mittig positionierte Geräusche nicht aus zwei sich überlagernden Schallfeldern wahrgenommen werden müssen, sondern tatsächlich vom mittig positionierten Center Lautsprecher ausgehen.

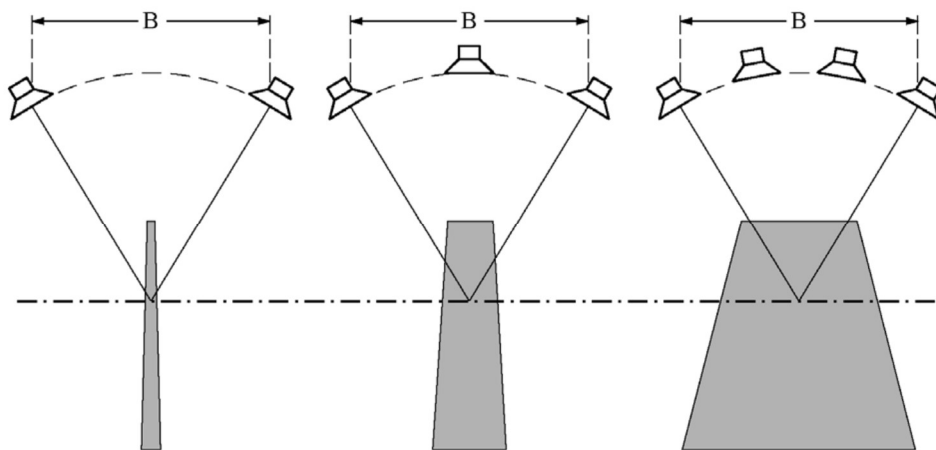


Abbildung 10 Hörzonenbreiten bei Stereophonie mit 2, 3 und 4 Lautsprechern<sup>53</sup>

Häufig findet sich bei Lautsprechersystemen zusätzlich zu den Lautsprechern, welche für die räumliche Signalverteilung verantwortlich sind, ein Subwoofer, welcher durch einen sogenannten LFE Kanal der Wiedergabe tieffrequenter Signale dient. Ein 2.1 Lautsprechersystem beinhaltet beispielsweise zwei Stereokanäle und einen LFE Kanal. Der Ursprung der dadurch produzierten tieffrequenten Schallwellen ist für das menschliche Gehör kaum bis gar nicht lokalisierbar, weshalb der Subwoofer in keinem festgelegten Verhältnis zu den anderen Lautsprechern positioniert werden muss.

<sup>52</sup> Vgl. Dickreiter et al. 2014, 228. f.

<sup>53</sup> Ebd., 229.

## 2.3 Surround-Sound

„Von Surround-Sound spricht man in der Regel dann, wenn der von den Lautsprechern einer Anlage produzierte Klang den Raum derart erfüllt, dass beim Hörer der Eindruck entsteht, er sei von Schallquellen umgeben“.<sup>54</sup>

Demnach gelingt Surround-Sound der Schritt in die zweite Klangdimension. Bei der Weiterentwicklung der Stereophonie wird deren Links-Rechts Verteilung durch zusätzliche Lautsprecherkanäle um eine Vorn-Hinten Komponente ergänzt. In der kleinstmöglichen Ausbaustufe wird eine Mehrkanal-Stereoanordnung bestehend aus Links, Center und Rechts um zwei Surroundlautsprecher links und rechts hinter dem Hörer erweitert. Eine zweckmäßige Lautsprecheranordnung in Verbindung mit einer korrekten Abhörposition ermöglicht dem Rezipienten nun Schallquellen in der ihn umgebenden horizontalen Ebene zu lokalisieren. Die korrekte Positionierung der Lautsprecher beruht auf bestimmten Winkelbeziehungen der Lautsprecher zueinander.

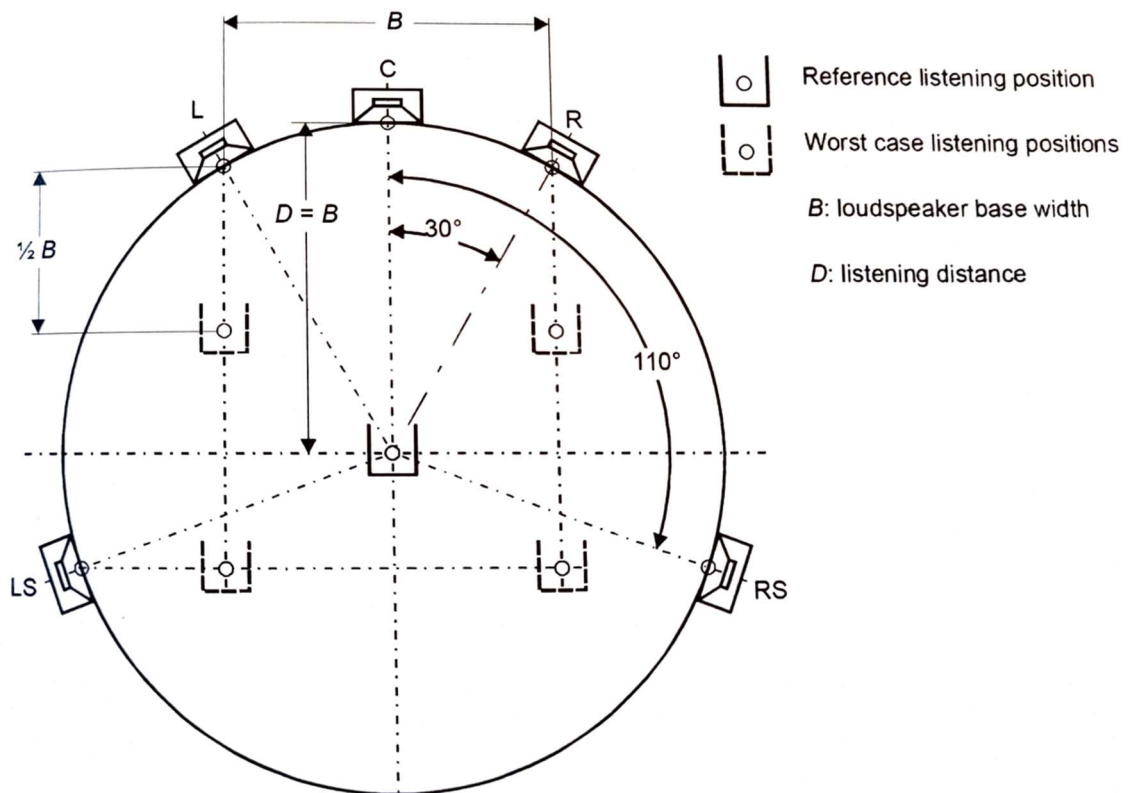


Abbildung 11 Abhörkreis für 5.1 Surround-Sound<sup>55</sup>

<sup>54</sup> Rainer Stahlmann, *Die verschiedenen Dolby Surround-Sound-Verfahren: Überblick, Analyse und Funktionsweise* (München: GRIN Verlag GmbH, 2002). <https://www.grin.com/document/5732> (letzter Zugriff: 3. Mai 2022), 2.

<sup>55</sup> Weinzierl. 2008, 643.



Für ein optimales Abhörerelebnis sollten bei Surround-Soundsystemen die Lautsprecher stets symmetrisch um den Rezipienten herum verteilt werden. Die Anordnung auf einer Kreisbahn mit der Abhörposition als Mittelpunkt sichert einen gleichen Abstand zu allen Lautsprechern (siehe Abbildung 11). Die Funktionsweise basiert erneut auf den sich überlagernden Schallfeldern der einzelnen Lautsprecher, ähnlich wie bei der Stereophonie. Ebenso können Phantomschallquellen zwischen benachbarten Lautsprechern erzeugt und mithilfe von Laufzeit- sowie Pegeldifferenzen beliebig positioniert werden.

Analog hat eine Erhöhung der Anzahl individuell angesteuerter Lautsprecherkanäle positive Auswirkungen auf die Größe der Hörzone sowie die Lokalisationsschärfe.<sup>56</sup> Demzufolge ermöglichen 7.1 Lautsprechersysteme ein differenzierteres Klangbild als 5.1 Systeme. Die zwei zusätzlichen betriebenen Surroundlautsprecher sollten dabei ebenfalls auf der Kreisbahn des Abhörkreises platziert werden.

## **2.4 Kanal- und objektbasierte Audioproduktionen**

Die Mischung von Audioproduktionen separiert sich in zwei verschiedene Teilbereiche, welche sowohl Einfluss auf den Workflow während der Mischung als auch auf die spätere Wiedergabe beim Endnutzer haben. Im Folgenden werden beide Herangehensweisen vorgestellt und deren Vor- sowie Nachteile erläutert.

---

<sup>56</sup> Friesecke. 2014, 789. f.

## 2.4.1 Kanalbasiertes Audio



Abbildung 12 kanalbasierte Audiowiedergabe<sup>57</sup>

Die kanalbasierte Herangehensweise ist die konventionelle Art Audiomaterial für ein festgelegtes Produktionsformat abzumischen. Sowohl Mono als auch Stereo und Surround-Sound basieren auf diesem Prinzip. Dabei werden zunächst die einzelnen Tonsignale in eine festgelegte Anzahl an Audiokanälen aufgeteilt, welche anschließend über eine entsprechende Lautsprecheranordnung wiedergegeben werden. Dies geschieht mithilfe sogenannter Panner. Während ein Stereo Panner durch einen Drehregler lediglich eine Signalverteilung zwischen einem linken und rechten Kanal bewerkstelligen muss, ist bei einem 5.1 Surround Panner bereits eine zweidimensionale Fläche für die richtungsbezogene Signalverteilung erforderlich. Der Ort der Phantomschallquelle wird dabei innerhalb dieser Fläche als Punkt platziert und dementsprechend auf die Wiedergabekanäle verteilt (siehe Abbildung 13).

---

<sup>57</sup> Fraunhofer IIS, „MPEG-H AUDIO: Das Audiosystem der nächsten Generation für UHDTV, Streaming und VR.“. <https://www.youtube.com/watch?v=9fU90uhHBm8> (letzter Zugriff: 9. Juni 2022).



Abbildung 13 Surround Panner einer 5.1 Mischung in Pro Tools<sup>58</sup>

Mithilfe von Laufzeit- sowie Pegelunterschieden bilden die sich überlagernden Schallfelder Phantomschallquellen aus, welche die akustischen Lücken zwischen benachbarten Lautsprechern schließen. Die jeweilige Produktionsart bestimmt die Anzahl der verfügbaren Audiokanäle. Im Vergleich zu einer 2.1 Stereoproduktion, welche zwei Audiokanäle und einen LFE Kanal besitzt, stehen bei einer 7.1 Surroundproduktion fünf weitere Audiokanäle zur Verfügung. Grundsätzlich steigt mit der Anzahl an individuell ansteuerbaren Lautsprechern das immersive Abhörerbis aber gleichzeitig erhöht sich auch der Produktionsaufwand. Mit steigender Kanalzahl müssen hardwareseitig entsprechend mehr Lautsprecher betrieben werden. Darüber hinaus steigt der Aufwand während der Mischung, da die einzelnen Signale differenzierter positioniert werden müssen.

Die exakte Lokalisation einzelner Schallquellen in der wiedergegebenen Mischung beschränkt sich, außer bei Mono, auf den Sweetspot. Bewegt sich der Hörer aus diesem hinaus, verschiebt sich die Lokalisation in der Regel zum nächstgelegenen Lautsprecher und verfälscht folglich das Hörerbis. Die Hörzone vergrößert sich zwar mit zunehmender Anzahl an individuell angesteuerten Lautsprechern, wird aber nie den kompletten Beschallungsraum ausfüllen, da für die Funktionsweise der Lokalisation ein möglichst gleichmäßiger Abstand zu allen Lautsprechern von Vorteil ist.

<sup>58</sup> Paul Maunder, „An introduction to surround mixing in Pro Tools.“  
<https://www.youtube.com/watch?v=tZeKaHyZ3ig> (letzter Zugriff: 12. Juni 2022).

*„Für eine befriedigende Mehrkanalwiedergabe [ist es] unumgänglich, dass bei Produktion und Wiedergabe zumindest annähernd ähnliche Bedingungen herrschen.“<sup>59</sup>*

Dies bezieht sich besonders auf die Anordnung der Lautsprecher. Eine 5.1 Surroundmischung erfordert, dass der Toningenieur während der Mischung und der Rezipient bei der Wiedergabe eine vergleichbare 5.1 Lautsprecherkonfiguration nutzen, damit der durch den Toningenieur konzipierte Klangeindruck entstehen kann. Die angefertigte Mischung ist demnach an eine Lautsprecheranordnung mit konkreten Abstands- und Winkelverhältnissen gebunden.

Innerhalb dieser Infrastruktur haben sich kanalbasierte Formate jedoch durch standardisierte Kanalbelegungen gut etabliert, weshalb der mehrkanalige Signalaustausch beispielsweise zwischen Studios und Sendeanstalten problemlos funktioniert<sup>60</sup>. Die Anzahl der Audiokanäle gibt dabei die minimale Anzahl an Lautsprechern und auch deren Position vor. Eine Anpassung auf andere Lautsprecher setups ist zwar möglich, jedoch mit viel Aufwand verbunden. Je nachdem, ob man die bestehende Mischung auf mehr oder weniger Audiokanäle anpassen muss, ist ein entsprechender Up- beziehungsweise Downmix erforderlich. Die Anfertigung einer völlig neuen dezidierten Mischung für das jeweilige Wiedergabeformat wäre meist qualitativ hochwertiger, da die Unterbringung der einzelnen Signale auf das Zielformat abgestimmt werden könnte. Beim Sprung von 5.1 auf 7.1 könnten im Mischprozess dadurch Richtungsinformationen gezielter vermittelt werden. Dies würde jedoch den Zugriff auf sämtliches Quellmaterial erfordern, da die Extraktion einzelner Geräusche innerhalb einer kanalbasierten Mischung kaum möglich ist.<sup>61</sup> Ein weiterer Nachteil ergibt sich im Hinblick auf die durch das Produktionsformat festgelegten Lautsprecheranordnungen. Der „Klangeindruck [leidet] deutlich, wenn die Lautsprecher bei der Wiedergabe nicht optimal positioniert sind.“<sup>62</sup>

---

<sup>59</sup> Weinzierl. 2008, 641.

<sup>60</sup> Vgl. ebd., 618.

<sup>61</sup> Vgl. Max Utke, „Objekt- und kanalbasierte Audioformate: Vor- und Nachteile.“ (Tonseminar, Hochschule der Medien, 2018). [https://curdt.home.hdm-stuttgart.de/PDF/Objekt-\\_u\\_kanalbasierte\\_Audioformate.pdf](https://curdt.home.hdm-stuttgart.de/PDF/Objekt-_u_kanalbasierte_Audioformate.pdf) (letzter Zugriff: 13. Mai 2022), 9.

<sup>62</sup> Fraunhofer - Institut für Digitale Medientechnologie 1, „Wellenfeldsynthese.“ [https://www.idmt.fraunhofer.de/content/dam/idmt/documents/IL/wave\\_field\\_synthesis\\_de.pdf](https://www.idmt.fraunhofer.de/content/dam/idmt/documents/IL/wave_field_synthesis_de.pdf) (letzter Zugriff: 6. April 2022), 1.

## 2.4.2 Objektbasiertes Audio



Abbildung 14 objektbasierte Audiowiedergabe<sup>63</sup>

Objektbasiertes Audio hat eine völlig andere Herangehensweise. Anstatt Audiosignale festen Lautsprecherkanälen zuzuordnen, erschafft man Audioszenen. Darin werden Audiosignale in einem virtuellen Raum als virtuelle Schallquelle positioniert. Jede virtuelle Schallquelle wird während der Mischung mit individuellen Metadaten versehen. „Metadaten sind beschreibende Daten, die zusätzlich zur eigentlichen Audioinformation übertragen werden.“<sup>64</sup> Sie können neben den Koordinaten des Audiosignals im virtuellen Raum beispielsweise auch Informationen über Bewegungsabläufe oder die räumliche Ausdehnung der virtuellen Schallquelle beinhalten. Dadurch, dass einzelne Audiosignale nicht wie bei einer kanalbasierten Mischung an feste Audiokanäle und somit Lautsprecher gebunden sind, ist eine objektbasierte Mischung wiedergabeseitig wesentlich flexibler.

*„Damit wird es möglich, das Audiosignal im jeweiligen Endgerät und direkt beim Kunden an die Bedingungen der Abhörsituation anzupassen und die Wiedergabe zu optimieren. Für die Produktion bedeutet dies, dass nicht verschiedene Mischungen und Bearbeitungen für unterschiedliche Zielmedien und Wiedergabebedingungen hergestellt werden müssen, sondern mit einem Bitstream alle Kunden bedient werden können.“<sup>65</sup>*

---

<sup>63</sup> Fraunhofer IIS. 2019.

<sup>64</sup> Weinzierl. 2008, 632.

<sup>65</sup> Ebd.

Für die korrekte Berechnung und Anpassung der virtuellen Raummischung an die jeweilige Abhörsituation ist neben der eigentlichen Audio-Produktionssoftware ein Renderer erforderlich.<sup>66</sup> Dieser benötigt Informationen über die Dimensionen des Beschallungsraumes und die Konfiguration des Abhörmediums, beispielsweise die Anzahl und Position aller verfügbaren Lautsprecher. Auf Grundlage dessen und der Informationen der wiederzugebenden Audioszene berechnet der Renderer Ausgangskanäle, welche hardwareseitig über den dafür vorgesehenen Lautsprecher abgespielt werden. Demnach ist die Anzahl der Lautsprecher gewissermaßen irrelevant, wodurch diese Wiedergabeweise enorm flexibel ist. Jedoch kann Audio selbstverständlich nur aus Richtungen kommen, in denen auch Lautsprecher positioniert sind. Je nach Produktion kann also erst ab einer bestimmten Anzahl an Lautsprecherkanälen in einer bestimmten Konfiguration der in der Mischung vorgesehene Klangeindruck entstehen. Beim Versuch eine Surroundmischung über lediglich zwei Lautsprecher wiederzugeben, würden technisch gesehen keine Audioinformationen verloren gehen jedoch könnten Geräusche, welche hinter dem Rezipienten abgespielt werden sollten, nicht realisiert werden.

### **3 Dreidimensionales Audio**

Dreidimensionale Audioproduktionen stellen die nächste Stufe klanglicher Immersion dar. Sie haben zum Ziel den Rezipienten noch intensiver in produzierte Klangwelten zu integrieren, als es Surroundformate könnten. Dabei orientiert man sich an der natürlichen Hörweise des Menschen. Die Realität erzeugt kontinuierlich ein dreidimensionales Schallfeld um jeden Menschen. Geräusche kommen dabei nicht ausschließlich frontal, seitlich oder von hinten, sondern aus allen Richtungen um den Rezipienten. In dreidimensionalen Audiosystemen wird ebendies versucht zu reproduzieren und das menschliche Gehör in sämtlichen Aspekten anzusprechen. Das ausschlaggebende Kriterium gegenüber Surround-Sound ist die Möglichkeit, den Hörer Schallquellen über sich wahrnehmen zu lassen, also den Schallquellen eine Elevation zu geben. Die logische Konsequenz sind Wiedergabesysteme mit Höhenlautsprechern, welche über der Ebene des Hörers angebracht werden. Dabei gibt es verschiedene Verfahren, die den Sprung aus der Ebene in den dreidimensionalen Raum ermöglichen. Es existieren sowohl kanal- als auch objektbasierte Wiedergabetechniken, welche meist unterschiedliche Prinzipien nutzen aber alle ein immersives Audioerlebnis zum Ziel haben.

---

<sup>66</sup> Vgl. Fraunhofer - Institut für Digitale Medientechnologie 2, „Object-based Spatial Audio Mastering.“. <https://www.youtube.com/watch?v=cqC1rWSxT4c> (letzter Zugriff: 6. April 2022).

Ein 1997 entwickeltes Verfahren namens Vector Base Amplitude Panning, kurz VBAP, beruht dabei auf der Lokalisation von Phantomschallquellen, welche mithilfe von sich überlagernden Lautsprechersignalen erzeugt werden.<sup>67</sup> Anders als bei einer stereophonen Wiedergabe kommen Höhenlautsprecher zum Einsatz, wodurch Phantomschallquellen nicht nur auf Verbindungslinien zwischen benachbarten Lautsprechern erzeugt werden können, sondern sich gemäß Abbildung 15 auf der eingeschlossenen Fläche eines Lautsprechertripels bilden. Dafür geben alle Lautsprecher dasselbe Signal in unterschiedlicher Intensität wieder. Vom Ortsvektor des Rezipienten, also dessen Raumposition aus, ergeben sich drei Richtungsvektoren, welche auf jeweils einen Lautsprecher deuten. Deren Länge wird durch Verstärkungsfaktoren bestimmt, welche sich aus der Wiedergabelautstärke der Lautsprecher ergeben. Der Richtungsvektor und damit der Hörereignisort der Phantomschallquelle kann auf Grundlage der linearen Addition aller Richtungsvektoren vorhergesagt und demzufolge ermittelt werden.<sup>68</sup> Somit geschieht die konkrete Signalberechnung im VBAP in Bezug auf eine feste Abhörposition.

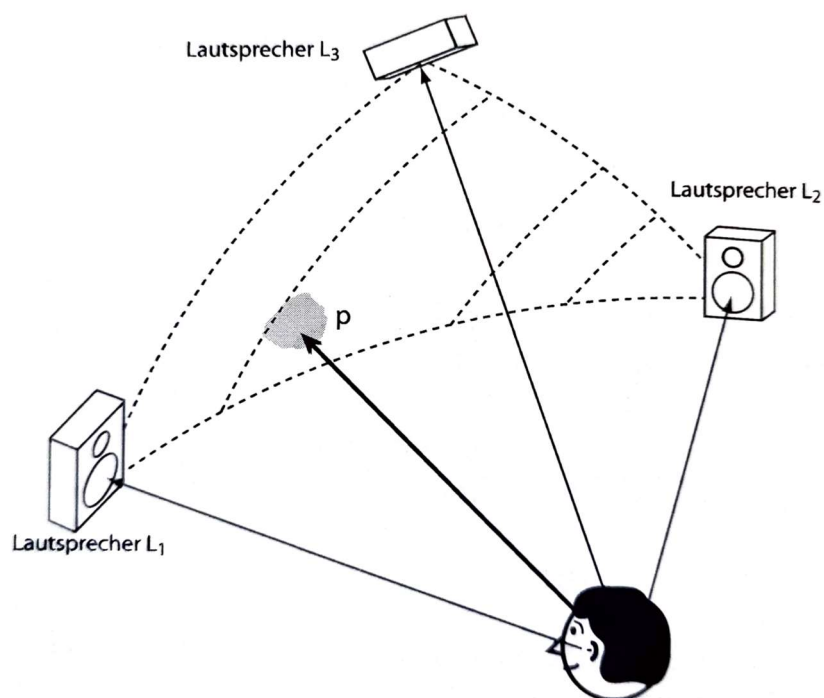


Abbildung 15 Dreidimensionales Amplituden-Panning<sup>69</sup>

<sup>67</sup> Vgl. Weinzierl. 2008, 657.

<sup>68</sup> Vgl. Ville Pulkki, „Virtual Sound Source Positioning Using Vector Base Amplitude Panning.“ <http://lib.tkk.fi/Diss/2001/isbn9512255324/article1.pdf> (letzter Zugriff: 17. Juni 2022), 458.

<sup>69</sup> Weinzierl. 2008, 660.

Auch andere Methoden des Amplitude Pannings können zur Erzeugung von Phantomschallquellen in dreidimensionalen Lautsprecheranordnungen genutzt werden, so auch das Distance Based Amplitude Panning, kurz DBAP.

*„Die Spatialisationstechnik arbeitet [ebenso] nach dem Amplitude-Panning-Prinzip, die Positionierung der virtuellen Schallquellen wird also über die Amplituden der einzelnen Lautsprecher gesteuert. [...] Die Amplituden der Lautsprecher werden dabei wenig rechenintensiv auf Basis der Abstände der virtuellen Schallquellen zu den Lautsprechern ermittelt.“<sup>70</sup>*

Ein enormer Vorteil gegenüber VBAP liegt in der Variabilität der Abhörposition. Für eine exakte Lokalisation bei einer VBAP-Audiowiedergabe muss der Rezipient sich ähnlich wie bei herkömmlichen stereophonen Wiedergabetechniken in einem konkreten Sweetspot befinden, welcher im Optimalfall äquidistant zu allen Lautsprechern ist. Andernfalls würde sich der Hörereignisort von Phantomschallquellen in Richtung des nächstgelegenen Lautsprechers verschieben, da dieser durch die geringere Distanz als lauter wahrgenommen werden würde. Bei DBAP hingegen werden die Wiedergabesignale ungeachtet einer Abhörposition berechnet, wodurch diese keinen Einfluss auf die Funktionsweise nimmt.<sup>71</sup>

Dem gegenüber stehen Ambisonics und die Wellenfeldsynthese, welche eine physikalische Synthese realer oder virtueller Schallfelder anstreben. Ambisonics nutzt zur Definition der synthetisierten Schallfelder und somit zu deren Entwicklung die Fourier-Bessel-Reihe. Dabei werden sogenannte sphärische Harmonische zunehmender Ordnung berechnet, welche jeweils den Schalldruck mit zunehmender Entfernung vom Ursprung beschreiben. Je mehr Harmonische berechnet werden, umso genauer ist die Schallfeldsynthese. Ebenso steigt die Genauigkeit mit der Anzahl der zur Verfügung stehenden Übertragungskanäle sowie der Anzahl an wiedergebenden Lautsprechern.<sup>72</sup>

---

<sup>70</sup> Markus Maier, „Distance based amplitude Panning.“ (Bachelorarbeit, Universität für Musik und darstellende Kunst, 2012). <https://phaidra.kug.ac.at/download/o:11015> (letzter Zugriff: 9. Juni 2022), 4.

<sup>71</sup> Vgl. Trond Lossius, Pascal Baltazar und Théo de La Hogue, „DBAP - DISTANCE-BASED AMPLITUDE PANNING.“. [https://d1wqtxts1xzle7.cloudfront.net/50048560/DBAP\\_-\\_DISTANCE-BASED\\_AMPLITUDE\\_PANNING20161101-1709-6hyhe4-with-cover-page-v2.pdf?Expires=1655483932&Signature=BlhYIrl08olplrainxA50dZsZvUiBokC84sDI7Cc4BfpTFMSntDNOdLzSSSVRSZlgCnj2wxgW5sS390~rTZvPIr9iUwbV0VkyYnYcRc1n2Ygyjc7FrROQtWVocNVluz243TzZxl7SA8s~dF2esUXthvHd51NcLqScVWhGi6nsTPoQz7G-5J~atRGFUiBKko19aD-5CkFa8AJBWP-H8gfp84uDuRrAohIFzBXO4EuoprAIAUJfU84w~w3ocSld14OrYD5kg1YZy3XBNr3ujQL16lrNHhxK2iyUI4C2Zxl5ntYimHEBld~f1eD814O48S8nsK0FyCOKMNCjh7ns5QxwHA\\_\\_&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA](https://d1wqtxts1xzle7.cloudfront.net/50048560/DBAP_-_DISTANCE-BASED_AMPLITUDE_PANNING20161101-1709-6hyhe4-with-cover-page-v2.pdf?Expires=1655483932&Signature=BlhYIrl08olplrainxA50dZsZvUiBokC84sDI7Cc4BfpTFMSntDNOdLzSSSVRSZlgCnj2wxgW5sS390~rTZvPIr9iUwbV0VkyYnYcRc1n2Ygyjc7FrROQtWVocNVluz243TzZxl7SA8s~dF2esUXthvHd51NcLqScVWhGi6nsTPoQz7G-5J~atRGFUiBKko19aD-5CkFa8AJBWP-H8gfp84uDuRrAohIFzBXO4EuoprAIAUJfU84w~w3ocSld14OrYD5kg1YZy3XBNr3ujQL16lrNHhxK2iyUI4C2Zxl5ntYimHEBld~f1eD814O48S8nsK0FyCOKMNCjh7ns5QxwHA__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA) (letzter Zugriff: 17. Juni 2022), 2.

<sup>72</sup> Vgl. Weinzierl. 2008, 659. ff.



Die in Kapitel 1.6 bereits vorgestellte Wellenfeldsynthese beruht auf der Simulation Huygensscher Elementarwellen durch einzelne, dicht zueinander positionierte Lautsprecher, welche in Summe eine virtuelle Wellenfront synthetisieren können. Die Schallinformationen an den Raumbegrenzungsflächen ermöglichen auf Grundlage des Kirchhoff-Helmholtz Integrals die exakte Reproduktion eines Schallfeldes in einem geschlossenen Volumen.

Dreidimensionales Audiomaterial kann jedoch nicht ausschließlich durch Lautsprecher reproduziert werden. Kopfhörer bieten andere Möglichkeiten den Hörer in immersive Audiowelten abtauchen zu lassen.

Bei einer immersiven Kopfhörerwiedergabe werden zahlreiche Richtungsinformationen versucht über manipulierte Klangfärbungen zu vermitteln, die die natürliche Wahrnehmungsweise von Schall simulieren sollen.<sup>73</sup> Diese Reproduktion der Ohrsignale über einen rechten und linken Audiokanal wird als binaurales Audio bezeichnet. Da jeder Mensch Schall anders wahrnimmt, ist die Effektivität einer universellen Binaural-Mischung limitiert. Um diese Funktionsweise auf jeden Menschen exakt abstimmen zu können und somit das Hörerlebnis zu maximieren ist eine sogenannte HRTF, eine Head Related Transfer Function, erforderlich. Diese Übertragungsfunktion berücksichtigt je nach physischer Beschaffenheit der Ohren die individuellen Reflexionserscheinungen und die daraus resultierenden Klangfärbungen für jeden einzelnen Rezipienten. Die körpereigene Schallbearbeitung wird dafür gemessen und in Form der HRTF auf das über Kopfhörer abgespielte Audiomaterial angewandt.<sup>74</sup> Dadurch können wesentlich umfangreichere Richtungseindrücke vermittelt werden als bei der Wiedergabe einer herkömmlichen Stereo-Mischung über Kopfhörer. Mehrkanalige Lautsprechersysteme bieten hingegen die Möglichkeit, dass sie Schall tatsächlich aus verschiedenen Richtungen auf den Rezipienten treffen lassen können. Dadurch entstehen natürliche Klangfärbungen als Folge der Überlagerung des wahrgenommenen Lautsprecherschalls mit Reflexionserscheinungen am menschlichen Körper. Diese sind nicht künstlich erzeugt und gemäß dessen physischer Erscheinung bereits auf den jeweiligen Menschen abgestimmt, wodurch sich das Hörerlebnis realer anfühlt als eine Kopfhörerwiedergabe.

---

<sup>73</sup> Vgl. Michael Wagner 1, „Q&A: Ambisonics vs. Atmos, Stereo vs. Binaural, Waves NX in Fairlight.“. <https://www.youtube.com/watch?v=R2QON1y5BCo> (letzter Zugriff: 24. Mai 2022).

<sup>74</sup> Vgl. Ville Pulkki. 1997, 456.

### 3.1 Auralite 3D

*„Not only in matters of club and event irradiation, but also in planetariums and theme parks, [an] Auralite 3D system can be reliably deployed for public addressing, as installed system, for promotional purposes or even three-dimensional audio-visual simulations“.*<sup>75</sup>

Auralite 3D ist ein immersives Audiosystem der MediaApes GmbH. Es erschien 2018 am Markt und basiert auf „der Technologie der Wellenfeldsynthese, die an der TU Delft erfunden und vom Fraunhofer IDMT zur Marktreife entwickelt wurde“.<sup>76</sup> Die Mischung erfolgt ausschließlich objektbasiert und ermöglicht dadurch die freie Platzierung sowie Bewegung von bis zu 64 Audioobjekten in einem virtuellen Raum. Die Signale der über den gesamten Wiedergaberaum frei verteilten Einzellausprecher werden in Echtzeit für jeden Lautsprecher individuell berechnet. Die Überlagerung der Lautsprechersignale reproduzieren das Schallfeld der einzelnen Audioobjekte bei der Wiedergabe. Anders als bei einer kanalbasierten Mischung, wo die konzipierte räumliche Lokalisation nur in einem bestimmten Sweetspot gewährleistet ist, kann bei dieser Technik „über nahezu den gesamten Wiedergaberaum ein natürlicher Raumeindruck geschaffen und eine akustisch realistische Einhüllung des Zuhörers erreicht“<sup>77</sup> werden.

Zusätzlich können mithilfe der zur Berechnung genutzten Algorithmen virtuelle Schallquellen im Wiedergaberaum zwischen den Lautsprechern und dem Zuhörer positioniert werden. Diese fokussierten Schallquellen sind aufgrund der im Vergleich zur Theorie der Wellenfeldsynthese geringen Lautsprecheranzahl mit Auralite 3D jedoch physisch nicht realisierbar.<sup>78</sup>

Das Anwendungsspektrum ist durch variable Lautsprecherpositionen und Live-Kompatibilität kaum limitiert. In Wiedergaberäumen können mit bis zu 128 Ausgangskanälen<sup>79</sup> immersive Audiowelten kreiert werden. Virtuelle Schallquellen sind bei der Wiedergabe nicht auf ihre ursprüngliche Form als trockenes Signal ohne Raumanteil begrenzt.

---

<sup>75</sup> Jakob Kirner, „Auralite3D Manual: Version 4.4.“, 6.

<sup>76</sup> Fraunhofer - Institut für Digitale Medientechnologie 1, 1.

<sup>77</sup> Ebd.

<sup>78</sup> Vgl. A.1.19.

<sup>79</sup> Vgl. Kirner, 16.

Mithilfe der sogenannten Raumhallsimulation ist es möglich Schallquellen in die komplexe Akustik von im Vorfeld gemessenen oder simulierten Räumen einzubinden. Somit kann ein regulärer Abhörraum akustisch beispielsweise in eine Kathedrale oder einen Konzertsaal verwandelt werden.<sup>80</sup>

Selbst eine Wiedergabe diskret abgemischter Audioproduktionen ist ohne den immensen Aufwand eines Up- oder Downmixes möglich. Somit ist das System nicht auf die Wiedergabe eigener Produktionen limitiert, sondern auch für die Wiedergabe kanalbasierter Audioproduktionen geeignet.<sup>81</sup>

Im Folgenden wird das System in seinen Bestandteilen vorgestellt. Neben den erforderlichen Hard- und Softwarekomponenten werden allgemeingültige Workflowschritte während einer Produktion analysiert. Daraus lässt sich auf bereits bestehende sowie potentielle Einsatzmöglichkeiten schließen.

### **3.1.1 Technischer Aufbau**

Auralite 3D setzt ein Zusammenspiel mehrerer Komponenten voraus. Das hohe Maß an Variabilität hat zur Folge, dass es kein Regalartikel ist, welcher in einer Konfiguration alle Eventualitäten abdeckt.<sup>82</sup> Bei vielen Aspekten bedarf es einer individuellen Anpassung an den angestrebten Anwendungsbereich sowie die Eigenschaften des Wiedergaberaumes. Hinzu kommt, dass es für spezifische Anforderungen unterschiedliche Ausbaustufen gibt. Ein Auralite 3D Player-System kann beispielsweise ausschließlich zur Wiedergabe von bereits angefertigten Mischungen verwendet werden. Die nächste Stufe ist die Production-Suite, mit welcher immersive Mischungen angefertigt werden können. Live-Kompatibilität wird in der Live-Suite realisiert. Komplexe Simulationen von dreidimensionalen Nachhallsignalen können ausschließlich in der endgültigen Ausbaustufe mithilfe der Raumhallsimulation durchgeführt werden. Die Anforderungen an die jeweilige Hardware steigen mit zunehmender Komplexität des Systems, weshalb nicht eine konkrete Form von Auralite 3D existiert. Nichtsdestotrotz können bereichsübergreifende Komponenten als essenziell benannt werden.

Als Quelle des Audiomaterials dient die Player Unit, ein beliebiger PC beziehungsweise Mac. Sämtliche Audiodateien werden darauf in einer beliebigen Digital Audio Workstation, kurz DAW, verwaltet. Ebenso erfolgt darin das grundlegende Editing und die Effektbearbeitung der einzelnen Mono- beziehungsweise Stereoaufnahmen. Äquivalent kann ein beliebiges digitales Mischpult als Player Unit agieren. Die Verbindung zum Auralite 3D-Renderer kann über drei Wege realisiert werden. Entweder werden die einzelnen Kanäle mithilfe des Multi-Channel

---

<sup>80</sup> Vgl. Fraunhofer - Institut für Digitale Medientechnologie 1, 1.

<sup>81</sup> Vgl. ebd., 2.

<sup>82</sup> Vgl. A.1.6.

Audio Digital Interface, kurz MADI, oder per Digital Audio Network through Ethernet, kurz DANTE, übertragen. Zusätzlich können Audiosignale mithilfe von Analog-Digitalwandlern übermittelt werden.<sup>83</sup>

Das Herzstück jeder Ausbaustufe von Auralite 3D bildet ein dezidiertes Renderer, ein Audio-Server mit dem Betriebssystem Linux CentOS.<sup>84</sup> Dieser berechnet basierend auf den an ihn übermittelten Audiosignalen und zugehörigen Metadaten die Wiedergabesignale für die verfügbaren Lautsprecher. Die mit einer minimalen Latenz berechneten Signale jedes individuellen Lautsprechers werden dabei von allen Audioobjekten der Mischung und deren Metadaten beeinflusst. Je nach Ausbaustufe ergibt sich die maximale Anzahl möglicher Audioobjekte. Bei der Vorproduktion von Inhalten stehen 64 Objektkanäle zur Verfügung, wohingegen bei Liveanwendungen und Raumhallsimulationen nicht mehr als 32 Objektkanäle verwendet werden können.<sup>85</sup>

Die Bedienung aller Parameter im Bereich des dreidimensionalen Audios, wie beispielsweise die Bewegung einzelner Audioobjekte, erfolgt im browserbasierten Graphical User Interface von Auralite 3D. „Auralite 3D is designed to exclusively synergize with Google Chrome“.<sup>86</sup> Einzelne Anwendungsbereiche sind dabei als separate Apps strukturiert. Der zugehörige Server ist Teil der Renderer-Unit. Mithilfe von Netzwerkverbindungen können Endgeräte wie PCs, Macs, Tablets oder Smartphones auf die Apps zugreifen und darüber Auralite 3D steuern.<sup>87</sup> Demnach ist das System abseits des Renderers an keine spezifische Hardware gebunden.

Ausgangsseitig wird Auralite 3D direkt an Audioverstärker angeschlossen. Damit können anschließend bis zu 128 Lautsprecher betrieben werden. Sowohl die Anzahl als auch die Positionierung der Lautsprecher ist dabei variabel. Es sind keine spezifischen Winkelbeziehungen der Lautsprecher zueinander notwendig, was den Wiedergabeaufbau enorm flexibel und an den jeweiligen Beschallungsraum anpassungsfähig macht. Das immersive Abhörererlebnis wächst jedoch mit der Anzahl an verwendeten Lautsprechern. Dies ist vordergründig auf zwei Faktoren zurückzuführen. Zunächst kann Schall selbstverständlich nur aus den Richtungen wiedergegeben und wahrgenommen werden, in denen sich Lautsprecher befinden. Bei einer ungeschlossenen Lautsprecheranordnung würden demnach Schallinformationen von Audioobjekten verloren gehen, an deren Position sich keine Lautsprecher für ihre Wiedergabe befänden. Darüber hinaus wird die Synthese des virtuellen Schallfeldes umso exakter, je mehr individuell angesteuerte Lautsprecher verwendet

---

<sup>83</sup> Vgl. Kirner, 7.

<sup>84</sup> Vgl. ebd.

<sup>85</sup> Vgl. 1.A.1.21.

<sup>86</sup> Kirner, 13.

<sup>87</sup> Vgl. ebd., 7.

werden.<sup>88</sup> Dies bedeutet jedoch nicht, dass die gesamte Raumbegrenzungsfläche mit Lautsprechern eingedeckt werden muss. „Even without a fully closed driver setup, Auralite 3D is a guarantor for acoustic immersion“.<sup>89</sup> Es wird empfohlen eine gleichmäßige Verteilung von mindesten 20 Lautsprechern um den Beschallungsraum herum zu positionieren, um ein näherungsweise geschlossenes Schallfeld erzeugen zu können. Möchte man Schallquellen von oben wahrnehmen können müssen dementsprechend Lautsprecher über der Hörfläche angebracht sein.

### 3.1.2 Funktionsweise

Die Funktionsweise von Auralite 3D basiert auf der Theorie der Wellenfeldsynthese. Audiosignale werden während der Mischung keinen Lautsprechern zugeordnet, sondern als virtuelle Schallquellen in einem virtuellen Raum positioniert. Während der Wiedergabe erfolgt die Berechnung der ersten Wellenfront dieser virtuellen Schallquellen und daraus resultierend die Berechnung der Lautsprechersignale. Demnach unterscheidet sich die Signalberechnung sowie -wiedergabe wesentlich von herkömmlichen Lautsprechersystemen.

*„[Die] Wellenfeldsynthese strebt keine Abbildung von Schallquellen als Phantomschallquellen zwischen benachbarten Lautsprechern an, sondern [...] eine physikalisch reale Resynthese eines räumlich ausgedehnten Schallfelds durch die Interferenz von zahlreichen dicht zueinander positionierten Lautsprechersignalen.“<sup>90</sup>*

In der Realität gibt es Abweichungen vom mathematisch exakten Konzept. So würde die unmittelbare Umsetzung der Theorie eine „unendlich große, [...] unendlich dicht angeordnete Anzahl von [...] Lautsprechern erfordern“<sup>91</sup>, welche auf einer geschlossenen Fläche um den Rezipienten angebracht werden müssten. In der Praxis können Lautsprecher ihrer Bauform geschuldet lediglich in einem diskreten Abstand zueinander angeordnet werden. Die Folge ist das Auftreten von Spatial Aliasing bei der Überlagerung der einzelnen Lautsprechersignale. Es entstehen also oberhalb einer Grenzfrequenz neben der korrekt synthetisierten Wellenfront ungewollte Schallartefakte „in Form von falsch orientierten Wellenfronten“.<sup>92</sup> Ist der Lautsprecherabstand geringer als die Wellenlänge der synthetisierten Schallwelle wird die Schallwelle exakt reproduziert. Versucht man bei gleichem Lautsprecherabstand eine Schallwelle höherer Frequenz zu synthetisieren, welche eine kleinere Wellenlänge gegenüber

---

<sup>88</sup> Vgl. 1.A.1.20.

<sup>89</sup> Vgl. Kirner, 6.

<sup>90</sup> Weinzierl, 2008, 664.

<sup>91</sup> Ebd., 667.

<sup>92</sup> Ebd., 669.

dem Lautsprecherabstand besitzt, so treten Artefakte auf. Abbildung 16 zeigt besagte Problematik. In beiden Grafiken werden synthetisierte Schallfelder dargestellt, welche durch dasselbe lineare Lautsprecherarray mit Lautsprecherabständen von 30 cm wiedergegeben werden. Das Schallfeld einer ebenen Welle mit der Frequenz 1000 Hertz, also einer Wellenlänge von 34 cm, wird fehlerfrei synthetisiert (siehe Abbildung 16 links). Bei der Wiedergabe einer ebenen Welle mit der Frequenz 1600 Hertz, also einer Wellenlänge von 21 cm, treten im synthetisierten Schallfeld Artefakte in Form von Alias-Wellen auf (siehe Abbildung 16 rechts).<sup>93</sup>

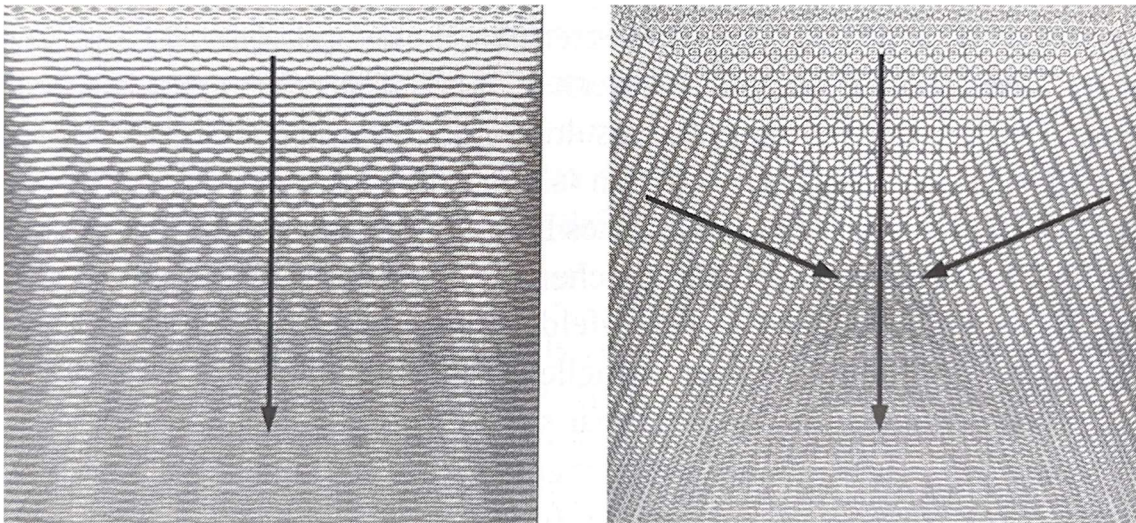


Abbildung 16 Aliasing Artefakte eines linearen Lautsprecherarrays<sup>94</sup>

Auch die Größe der Hörfläche reicht mit zunehmender Lautsprecheranzahl lediglich an das Ideal heran, den gesamten Beschallungsraum abzudecken. Dennoch wird eine größere Hörfläche geschaffen als bei herkömmlichen kanalbasierten Mischverfahren. Um trotz der im Vergleich zur Theorie geringen Anzahl an Lautsprechern ein geschlossenes Schallfeld generieren zu können, wurden zusätzlich Aspekte aus den Bereichen des Vector Base Amplitude Panning und dem Distance Based Amplitude Panning in Auralite 3D implementiert.

---

<sup>93</sup> Vgl. Weinzierl. 2008, 670.

<sup>94</sup> Ebd.

### 3.1.3 Ablauf einer Mischung

Wie auch bei einer herkömmlichen Soundmischung liegen sämtliche Entscheidungen, wie eine Produktion ablaufen und letztendlich klingen soll, beim jeweiligen Toningenieur, der sie anfertigt. Ebenso variiert die Herangehensweise je nach Produktionsinhalt. Deshalb wird im Folgenden auf eine logische Abfolge einzelner Arbeitsschritte bezüglich einer immersiven Audioproduktion mit Auralite 3D eingegangen. Der individuelle Workflow kann je nach persönlicher Präferenz davon abweichen. Auch die Übertragung der Signale ist variabel. Im Sinne der Übersichtlichkeit wird bei der Beschreibung von einer DAW als Player Unit und MADI als Methode zur Signalübertragung ausgegangen.

Zu Beginn steht die grundlegende Signalbearbeitung des im Vorfeld aufgenommenen Audio-Quellmaterials. Dieser Schritt erfolgt äquivalent zu herkömmlichen Audioproduktionen innerhalb einer DAW. Sämtliche Mono- beziehungsweise Stereosignale werden in einzelnen Kanälen getrennt voneinander strukturiert. Die Selektion produktionsrelevanter Inhalte, deren zeitliche Anordnung und die zugehörige Effektbearbeitung, beispielsweise durch Equalizer oder Kompressoren, stehen dabei im Vordergrund.<sup>95</sup> Auch externe Effektgeräte können auf herkömmliche Weise in die Mischung eingebunden werden. Da lediglich Monosignale zum Renderer übermittelt werden, müssen Stereosignale in jeweils zwei Monosignale aufgeteilt werden. Im späteren Verlauf der Produktion ist es möglich beide Signale wieder zueinander zu positionieren.<sup>96</sup> Zusätzlich zu den Kanälen der wiederzugebenden Audiosignale wird eine weitere Mono-Spur benötigt, welche für die Synchronisation aller Auralite 3D Instanzen einen Timecode übermittelt.<sup>97</sup> Dieser kann von der DAW selbst generiert sein oder extern in die Signalkette eingebunden werden.

Die bearbeiteten Signale gilt es anschließend von der Player Unit zum Renderer weiterzuleiten. Dafür muss die Player Unit zunächst mit einem MADI Interface verbunden sein. Daraufhin müssen die einzelnen Monosignale innerhalb der DAW in Output-Kanäle geroutet werden, welche wiederum den jeweiligen MADI-Kanälen zugeordnet werden. Falls mehrere Signale sich bezüglich ihres Verhaltens im virtuellen dreidimensionalen Raum gleich verhalten sollen, können diese innerhalb der DAW in denselben Mono-Output geroutet werden.<sup>98</sup> Den Timecode sollte man nach Herstellerempfehlung stets dem letzten Übertragungskanal zuordnen. Dieser befindet sich demnach auf Kanal 64 des MADI-Interfaces.

---

<sup>95</sup> Vgl. 1.A.1.12.

<sup>96</sup> Vgl. Kirner, 12.

<sup>97</sup> Vgl. ebd., 9.

<sup>98</sup> Vgl. ebd., 11.

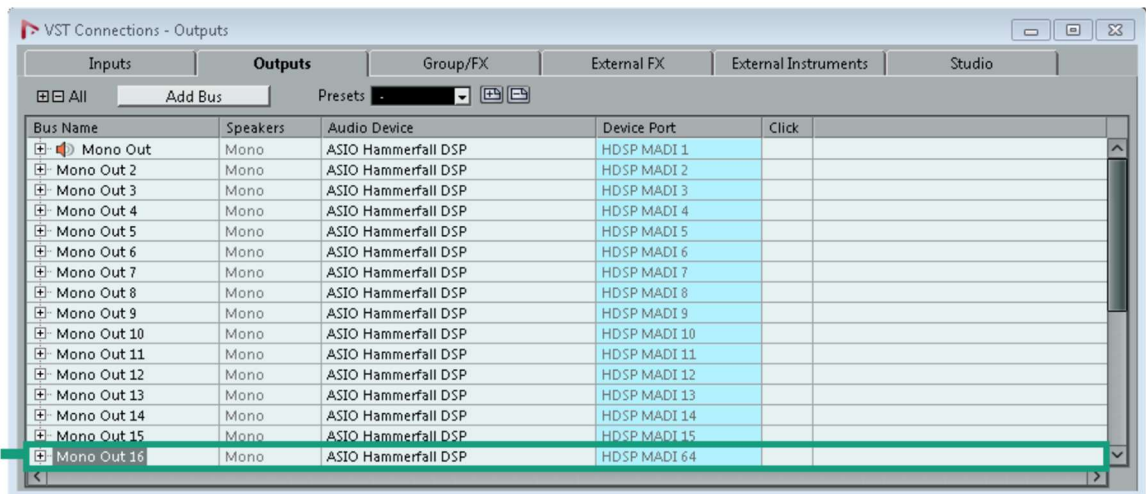


Abbildung 17 Output Routing in der DAW Nuendo<sup>99</sup>

Sobald alle Signale an den Renderer übermittelt werden, beginnt die immersive Mischung der Audioinhalte. Durch Eingabe der IP Adresse des Audioservers in Google Chrome gelangt man zu den grafischen Bedienelementen von Auralite 3D.<sup>100</sup> Verschiedene Apps mit jeweiligen Unterkategorien ermöglichen gezielten Zugriff auf einzelne Funktionen.

In der Configurator App werden grundlegende Systemeinstellungen vorgenommen. Die Dimensionierung des Beschallungsraumes spielt dabei eine fundamentale Rolle. Darin werden die Raumabmessungen in x, y sowie z Richtung, welche Breite, Länge und Höhe entsprechen, in das System eingegeben. Zusätzlich wird im dadurch entstandenen Koordinatensystem die Größe der Hörfläche definiert.<sup>101</sup> Von ebenso zentraler Bedeutung wie die Raumdimensionierung ist die Lautsprecherkonfiguration. Diese muss gemäß ihrem physischen Aufbau in das System übertragen werden. Dafür werden erneut Raumkoordinaten genutzt. Folglich erhält jeder Lautsprecher jeweils eine x, y und z Koordinate, welche seine Position definieren.<sup>102</sup> Selbstverständlich sollten sämtliche Lautsprecher von ihrer Position aus auf die Hörfläche gerichtet sein.

*„For a good sounding System it’s necessary that all loudspeakers have a similar frequency, time response and equal level. This depends on your room and your speakersetup.“<sup>103</sup>*

<sup>99</sup> Kirner, 10.

<sup>100</sup> Vgl. ebd., 13.

<sup>101</sup> Vgl. ebd., 15.

<sup>102</sup> Vgl. ebd., 17. f.

<sup>103</sup> Ebd., 22.



Das System unterscheidet zwischen drei Lautsprechertypen.<sup>104</sup> Reguläre Lautsprecher dienen der breitbandigen Wiedergabe von Audiosignalen und sind lediglich im Bassbereich limitiert. Subwoofer beschränken sich bei der Wiedergabe auf tieffrequente Signale. Reinforcement-Lautsprecher erweitern die Outputs eines Auralite 3D Systems. Sämtliche aktiven System-Inputs können allein oder in Kombination als Input für Reinforcement-Lautsprecher dienen. Dadurch ist es möglich gleiche Schallinformationen beispielsweise auf unterschiedlichen Theaterrängen wiederzugeben.<sup>105</sup>

Das Bewegungsverhalten virtueller Schallquellen in z-Richtung wird mithilfe eines sogenannten Corpus definiert. Dieser ist frei nach geometrischen Formen konfigurierbar. Dementsprechend bewegen sich die virtuellen Schallquellen auf der Oberfläche einer Kuppel, eines Ellipsoiden oder eines Pyramidenstumpfes, welche in ihren Dimensionen variabel gewählt werden können. Strebt man eine ebene Wiedergabe von Schallquellen an gibt es die Option keinen Corpus zu verwenden. Dabei verbleiben die virtuellen Schallquellen am Boden oder bei einem festgelegten z-Wert.<sup>106</sup> Eine Anpassung der Raumdimensionen kann sowohl spezifisch im jeweiligen Menüpunkt oder global vorgenommen werden.

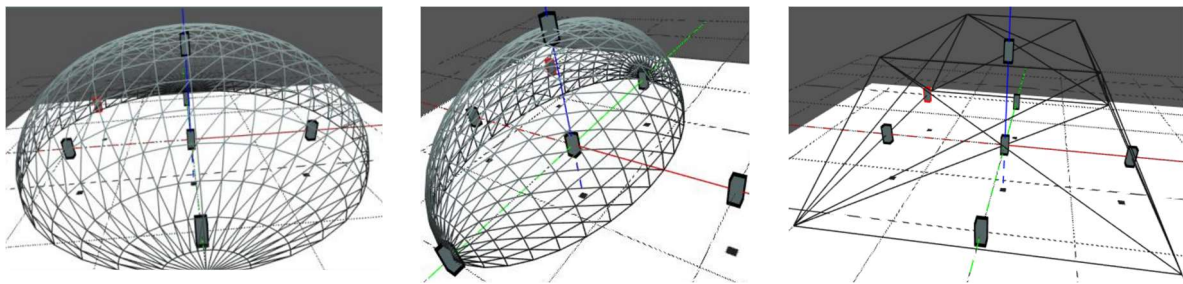


Abbildung 18 verschiedene Corpus-Konfigurationen<sup>107</sup>

Sind sämtliche Parameter im dreidimensionalen Raum in das System eingegeben worden folgt die eigentliche immersive Audiomischung im dreidimensionalen Raum. Dafür stehen dem Mischenden in der Sound Control App verschiedene Bearbeitungsmöglichkeiten zur Verfügung. Jeder Lautsprecherkanal kann in seinem Frequenzspektrum mithilfe eines eigenen 16-bandigen Equalizers angepasst werden.<sup>108</sup> Darüber hinaus dient ein globaler Equalizer der klanglichen Veränderung der gesamten Mischung und beeinflusst alle Outputs gleichermaßen. Ebenso wichtig, wie die klangliche Anpassung der Wiedergabe ist das Bass Management.

---

<sup>104</sup> Vgl. Kirner, 23.

<sup>105</sup> Vgl. A.1.26.

<sup>106</sup> Vgl. Kirner, 19. f.

<sup>107</sup> Ebd. f.

<sup>108</sup> Vgl. ebd., 25.

*„The fundamental principle of “bass management” is that “bass” content in the incoming signals should be directed only to loudspeakers capable of handling it. [...] The fullrange signals have to be divided into subwoofer signals and spatial loudspeaker signals, that the appropriate loudspeakers play correctly in their frequency ranges.”<sup>109</sup>*

Diese Signalteilung erfolgt mithilfe eines eigenen Bass Management Equalizers. Durch mehrere Tief- und Hochpassfilter erhält man einerseits breitbandige Wiedergabesignale für sämtliche Lautsprecher und andererseits gefilterte tieffrequente Signale für Subwoofer und LFE Kanäle, welche sowohl von zu hohen als auch von zu tiefen Signalfrequenzen befreit sind.<sup>110</sup> Die Sound Control App bietet darüber hinaus einen Überblick über sämtliche Pegelwerte der virtuellen Schallquellen und der Systemoutputs.

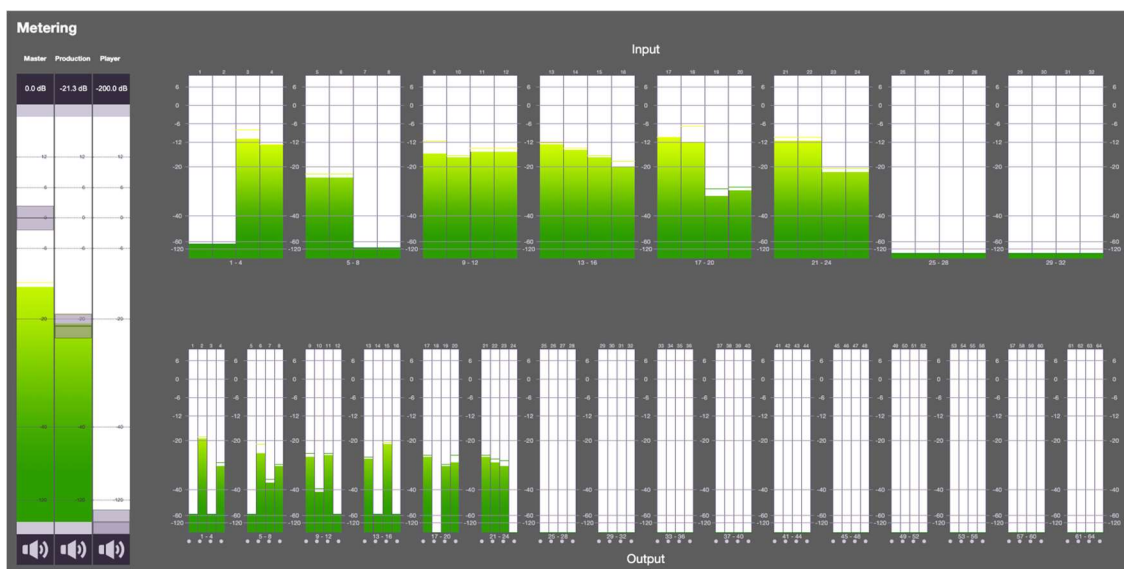


Abbildung 19 Übersicht aller aktiven Signalpegel während der Wiedergabe

Die Steuerung der komplexen Simulation von dreidimensionalen Raumakustiken erfolgt im Room Simulation Module – Configurator. Dafür werden insgesamt 22 Audioobjekte, welche einzelnen Reflexionsphasen eines Raumklangs entsprechen, im virtuellen Raum verteilt. Je nach eingestelltem Raumklang variiert die Nachhallzeit und somit der Anwendungsbereich. Neben einer 2D und 3D Ansicht aller Room Simulation Module Objekte gewährt das Bedienelement auch Zugriff auf die Lautstärke der einzelnen Reflexionsstufen, die Dimensionen des Hallraumes sowie die Gesamtlautstärke der Hallsimulation in Relation zur Lautstärke der restlichen Mischung.<sup>111</sup>

<sup>109</sup> Kirner, 27.

<sup>110</sup> Vgl. ebd.

<sup>111</sup> Vgl. ebd., 32. f.

Sind Mikrofone als Signalquellen in das System eingebunden können diese als Quellen für die Raumhallsimulation genutzt werden. Geräusche können demnach live aufgenommen und in den Raumklang eingebunden werden. Somit besteht die Möglichkeit Rezipienten interaktiv in das simulierte Schallfeld zu integrieren.

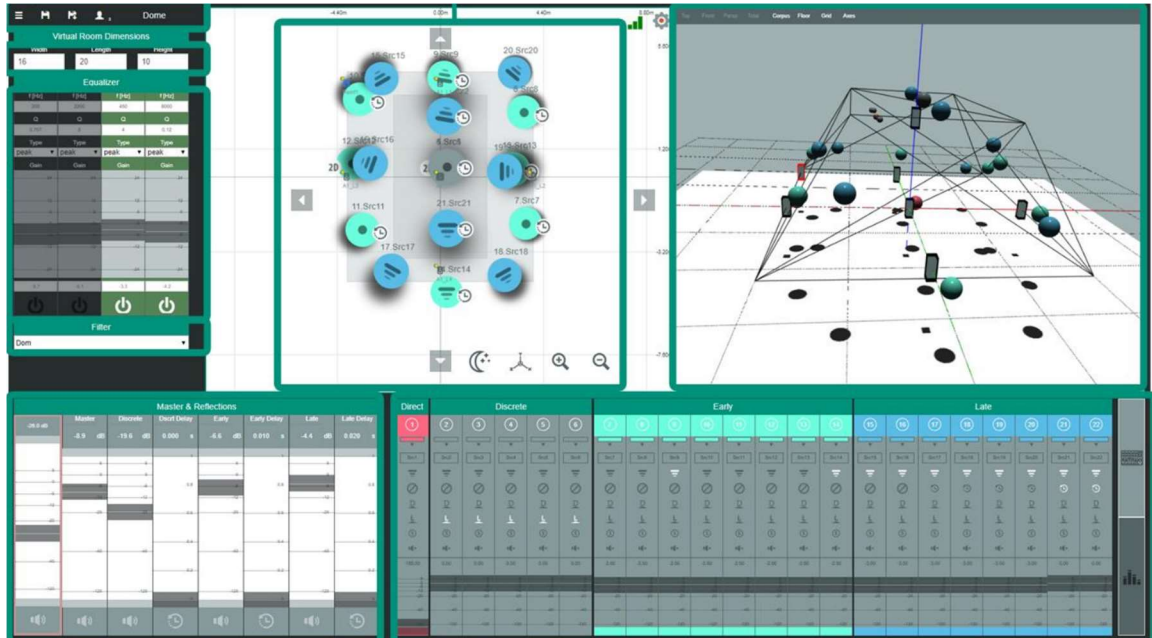


Abbildung 20 Bedienoberfläche des Room Simulation Module - Configurator<sup>112</sup>

Die Production App ist die größte Anwendung von Auralite 3D. Sie bietet einen Überblick über den virtuellen Raum, sämtliche Audioobjektkanäle, deren Position im Raum sowie essenzielle Parameter, wie den Timecode, die verwendete Einstellung der Raumhallsimulation oder den Fader der Gesamtlautstärke der Produktion.

<sup>112</sup> Kirner, 32.

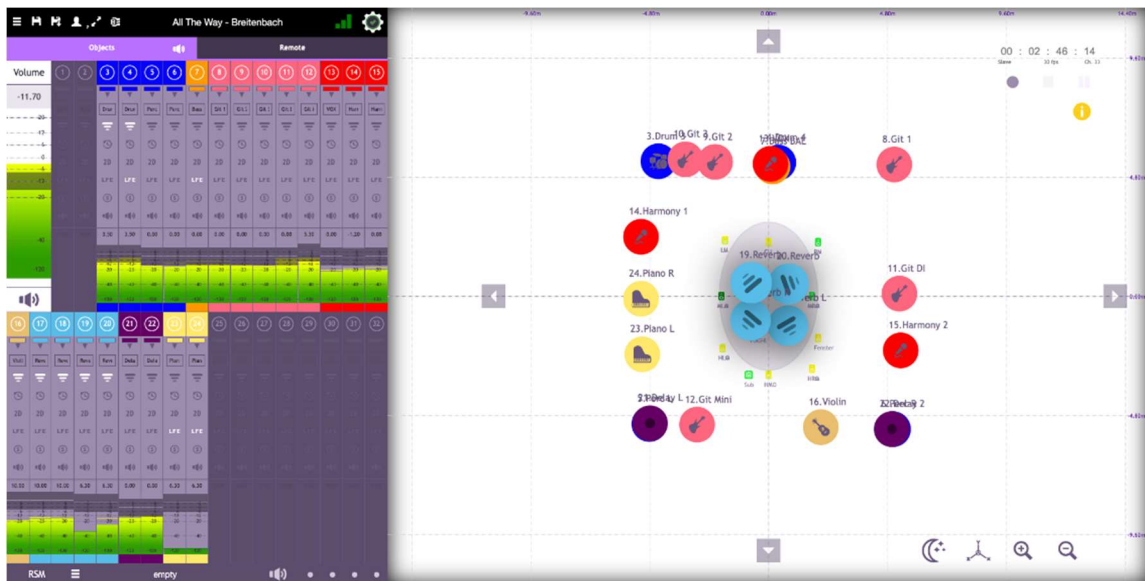


Abbildung 21 Bedienoberfläche der Production App

Alle Audioobjekte werden im Koordinatensystem des virtuellen Raums positioniert. Sie können entweder statisch an ihrer Position bleiben oder mithilfe von Automationsdaten ihre Koordinaten während der Wiedergabe verändern.<sup>113</sup> Jedes Audioobjekt wird durch einen Sourcstrip repräsentiert. Neben einer individualisierbaren Beschriftung und Färbung gewähren verschiedene Icons Zugriff auf verschiedene Bearbeitungsoptionen. Beispielsweise können Sourcstrips auf Mute oder Solo geschaltet werden, wie es in der Audiowelt üblich ist. Für eine diffusere Wiedergabe von Hintergrundgeräuschen kann das jeweilige Audioobjekt als ebene Welle wiedergegeben werden, anstelle der regulären Wiedergabe als Punktschallquelle. Die dadurch erzielte Wiedergabe über eine höhere Anzahl an Lautsprechern begünstigt die Übertragung der Schallwelle in weiter entfernte Bereiche des Beschallungsraumes aber senkt die Lokalisationsschärfe.<sup>114</sup> Dem Auftreten eines Doppler-Effekts bei der Bewegung einzelner Audioobjekte beugt die Minimum Delay Funktion vor.<sup>115</sup> Das 2D Icon limitiert die Bewegung des ursprünglich dreidimensionalen Audioobjektes auf eine Ebene. Um tieffrequente Signalanteile im Sinne räumlicher Effekte besonders hervorzuheben kann das LFE Symbol aktiviert werden.

Ein Zusammenspiel aller genannten Optionen ermöglicht die immersive Mischung von Audioinhalten in Auralite 3D. Die anschließende Wiedergabe der angefertigten Mischung ist dank der Anpassung der Soundinformationen durch einen Renderer auf einer beliebigen Lautsprecherkonfiguration möglich.

<sup>113</sup> Vgl. Kirner, 68.

<sup>114</sup> Vgl. ebd., 41.

<sup>115</sup> Vgl. ebd., 42.

### 3.1.4 Einsatzmöglichkeiten

Die Einsatzmöglichkeiten eines Auralite 3D Systems und der dahinterstehenden Technik der Wellenfeldsynthese sind kaum limitiert. Zunächst ist es an keinen konkreten Lautsprecheraufbau für die Wiedergabe gebunden. Selbst Surroundmischungen können mit wenig Aufwand über das System wiedergegeben werden. Im Falle einer herkömmlichen 5.1 Mischung müsste man für die Wiedergabe auf einem bestehenden Auralite 3D System lediglich die sechs Signalkanäle als individuelle Audioobjekte anlegen und anschließend gemäß einer 5.1 Lautsprecheranordnung im virtuellen Raum positionieren.<sup>116</sup> Für eine perfekte Nachbildung könnte man zusätzlich den physischen Lautsprecheraufbau gemäß einer 5.1-Anordnung vornehmen und diesen virtuell in das System übertragen. Zweifelsfrei würde eine explizit angefertigte dreidimensionale Mischung ein wesentlich immersiveres Hörerlebnis bieten. Die Möglichkeit der Wiedergabe kanalbasierter Mischungen unterstreicht dennoch die Flexibilität des Systems.

Darüber hinaus ist der Anwendungsbereich nicht auf die Vorproduktion limitiert. In entsprechender Ausbaustufe können mit live aufgenommenen Signalen völlig neue Klangwelten um das Publikum geschaffen werden. Immersive Konzerte würden jedem Rezipienten innerhalb der Hörzone den gleichen richtungsgetreuen Klangeindruck vermitteln können. „Es [gäbe] dann keine schlechten Plätze mehr“.<sup>117</sup>

In Verbindung mit der integrierten Raumhallsimulation können interaktive Installationen vorgenommen werden, bei denen Rezipienten das Schallfeld live beeinflussen können.<sup>118</sup> Sind im Beschallungsraum beispielsweise Mikrofone verbaut und jemand würde klatschen, würde dieses Klatschen in die simulierte Raumakustik integriert werden und einen entsprechenden Nachhall erzeugen.

Eine gekoppelte Nutzung von auditiven und visuellen Inhalten ist ebenfalls möglich. Einsatzgebiete wie Theater oder Eventhäuser können mithilfe immersiver Audioinhalte das Publikum mehr in das dargestellte Geschehen einbinden. In Verbindung mit Virtual Reality Brillen können Rezipienten sowohl mit den Augen als auch mit den Ohren in dreidimensionale Welten eintauchen. Zukünftig soll ein Projektor, welcher anhand der Raumdimensionen ein dreidimensionales Bild projiziert für Virtual Reality ohne Brille sorgen. Selbst touristische Einsatzzwecke stehen im Raum. „So könnte man heilige Orte, die sonst zu überlaufen sind, touristisch zur Verfügung stellen“<sup>119</sup> und dadurch den Flugverkehr entlasten.

---

<sup>116</sup> Vgl. Kirner, 49.

<sup>117</sup> Vgl. Elena Bruckner, „Ein Raum zum Mitnehmen.“ *Die Rheinpfalz*, 25. Mai 2021.

<sup>118</sup> Vgl. 1.A.1.15.

<sup>119</sup> Vgl. Bruckner, 25. Mai 2021.

## 3.2 Dolby Atmos

*„Dolby Atmos adds the flexibility and power of dynamic audio objects into traditional channel-based workflows. These audio objects allow moviemakers to control discrete sound elements irrespective of specific playback speaker configurations, including overhead speakers.“<sup>120</sup>*

Dolby Atmos ist ein immersives Audioformat, welches 2012 von den Dolby Laboratories für den Weltmarkt veröffentlicht wurde.<sup>121</sup> Ursprünglich wurde es für Kinos entwickelt, um Zuschauer durch eine differenziertere akustische Untermalung noch mehr in das Geschehen auf der Leinwand einzubinden, als es bis dato möglich war. Der zweidimensionale Filminhalt sollte sich akustisch auf den gesamten Kinosaal ausweiten und mithilfe gezielt platzierter Geräusche von bisherigen Surroundformaten abgrenzen. Mithilfe von klar nachvollziehbaren Bewegungsabläufen einzelner Schallquellen soll ein realistischeres Erlebnis geschaffen werden als bei einer kanalbasierten Signalaufteilung.

Der Schlüssel dazu liegt in der hybriden Form von Dolby Atmos. Hybrid bedeutet in diesem Fall, dass herkömmliches kanalbasiertes Audio, welches Dolby mit dem Begriff Bed betitelt, mit objektbasiertem Audio vereint wird. So werden die insgesamt 128 Signalkanäle zwischen kanalbasiert abgemischten Beds und frei positionierbaren Audioobjekten aufgeteilt.<sup>122</sup> Dadurch entstehen überaus vielseitige Möglichkeiten der Inhaltsvermittlung. Durch die freie Platzierung von Audioobjekten können visuelle Inhalte wesentlich detaillierter unterstützt werden als bei einer Unterbringung der Signale in diskreten Audiokanälen.<sup>123</sup> Zusätzlich werden die Audioinformationen mithilfe eines Renderers mit Metadaten verknüpft, welche beispielsweise deren Raumposition bestimmen. Diese Herangehensweise ermöglicht eine Wiedergabe auf verschiedensten Lautsprechersystemen, solange ein Renderer die übertragenen Informationen auf die Gegebenheiten anpasst. Eine weitere Innovation liegt in der Nutzung von Höhenlautsprechern, welche es ermöglichen Audioobjekte dreidimensional im Wiedergaberaum zu platzieren und zu bewegen.<sup>124</sup> Dadurch erreicht Dolby Atmos ein Klangerlebnis, welches aufgrund seiner Dreidimensionalität mehr Realismus und Immersion

---

<sup>120</sup> Dolby Laboratories Inc. 1, „Dolby Atmos: Next-Generation Audio for Cinema.“ <https://www.hollandfilmnieuws.nl/files/whitepaper-dolbyatmos.pdf> (letzter Zugriff: 3. Juni 2022), 1.

<sup>121</sup> Vgl. Felix Kraus, „3D Sound: Kino-Ton in drei Dimensionen.“ (Bachelorarbeit, University of Applied Sciences, 2015). <https://monami.hs-mittweida.de/frontdoor/index/index/start/2/rows/10/sortfield/score/sortorder/desc/searchtype/simple/query/felix+Kraus/docId/5782> (letzter Zugriff: 4. April 2022), 6.

<sup>122</sup> Vgl. Dolby Laboratories Inc. 1. 2014, 9.

<sup>123</sup> Vgl. ebd., 7.

<sup>124</sup> Vgl. ebd., 6.

transportiert als herkömmliche Audioformate. Gleichzeitig erfordert die Nutzung dieser Technik neben einer höheren Anzahl an Lautsprechern auch die individuelle Ansteuerbarkeit jedes einzelnen Lautsprechers im Wiedergaberaum. Bis zu 64 individuell nutzbare Lautsprecherkanäle dienen der immersiven Wiedergabe von Dolby Atmos Mischungen, wodurch nicht zuletzt hohe Anforderungen an die genutzte Hardware entstehen.<sup>125</sup> Demnach gehen die neuen Möglichkeiten in der inhaltlichen Gestaltung mit einem wachsenden Produktionsaufwand einher.

### 3.2.1 Technischer Aufbau

Ein Systemaufbau von Dolby Atmos existiert in verschiedenen Konfigurationen, welche individuellen Einsatzzwecken dienen. Sowohl der Umfang an verwendeter Hardware als auch die notwendige Software kann von Produktion zu Produktion variieren. Dennoch existieren zwischen allen individuellen Workflows Parallelen und Tools, welche essenziell für die Funktionalität von Dolby Atmos sind. Im Sinne der Übersichtlichkeit wird im Folgenden vordergründig auf Kinoinhalte eingegangen, da diese ursprünglich den primären Anwendungsbereich von Dolby Atmos darstellten. Ebenso wird sich auf eine Produktion mit der DAW Pro Tools festgelegt, um Verknüpfungen einzelner Instanzen klar aufzeigen zu können.

Somit fungiert Pro Tools als Verwaltungsinstanz für sämtliches Audio-Quellmaterial. Als Arbeitsgrundlage dienen Mono beziehungsweise Stereospuren, welche individuell voneinander nach herkömmlichen Herangehensweisen bearbeitet und abgemischt werden. Gleichzeitig erfolgt die Festlegung, welche Signale einem kanalbasiert abgemischten Bed zugeteilt werden und welche Schallquellen als frei positionierbare Audioobjekte agieren sollen. Die dabei hervorgegangenen Audiospuren werden softwareintern auf separate Outputs geroutet.<sup>126</sup> Je nachdem, wie viele von den insgesamt zur Verfügung stehenden 128 Kanälen genutzt werden, werden die bearbeiteten Audiodaten mithilfe von einem oder mehreren High Definition MADI Interfaces zum Herzstück von Dolby Atmos weitergeleitet, der sogenannten Rendering und Mastering Unit, kurz RMU.<sup>127</sup> Positions- und bewegungsbestimmende Metadaten von Audioobjekten werden mithilfe spezieller Dolby Atmos Software Plug-Ins ebenfalls bereits in der DAW erzeugt. Diese Metadaten werden separiert vom Audiomaterial über Ethernet an die RMU weitergeleitet.<sup>128</sup>

---

<sup>125</sup> Vgl. Dolby Laboratories Inc. 1. 2014, 9.

<sup>126</sup> Vgl. Dolby Laboratories Inc. 2, „Authoring for Dolby Atmos Cinema Sound Manual.“  
[https://www.dolby.com/uploadedFiles/Assets/US/Doc/Professional/Authoring\\_for\\_Dolby\\_Atmos\\_Cinema\\_Sound\\_Manual\(1\).pdf](https://www.dolby.com/uploadedFiles/Assets/US/Doc/Professional/Authoring_for_Dolby_Atmos_Cinema_Sound_Manual(1).pdf) (letzter Zugriff: 3. Juni 2022), 15. f.

<sup>127</sup> Vgl. ebd., 1.

<sup>128</sup> Vgl. ebd., 2.

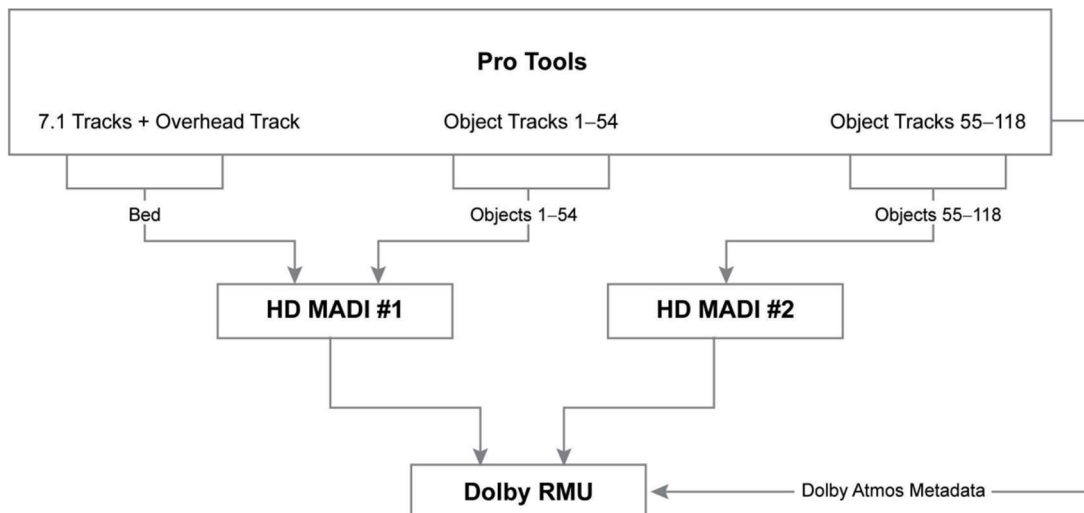


Abbildung 22 Signalwege zwischen Pro Tools und der Dolby RMU<sup>129</sup>

Das Herzstück von Dolby Atmos bildet die Renderer-Software der RMU. Obwohl bei einigen Anwendungen die Renderer-Software auf demselben Computer betrieben werden kann, wie die DAW, ist die RMU in der Regel ein dezidiertes Mac oder Windows Computer.<sup>130</sup> Innerhalb der RMU erfolgt die Verrechnung der einzelnen Audiosignale mit ihren zugehörigen Metadaten. Daraus resultierend erfolgt die Anpassung an eine dem System bekannte Wiedergabekonfiguration. Der Renderer benötigt dafür Informationen, wie die Art und Anzahl der Lautsprecher sowie die Dimensionen des Wiedergaberaumes. Diese Parameter müssen gemäß ihrem physischen Aufbau in das System eingegeben werden.<sup>131</sup>

Aufgrund der Kombination von kanal- und objektbasierten Audioelementen ist der wiedergabeseitige Lautsprecheraufbau nur begrenzt variabel. Würde man von den grundlegenden Abstands- und Winkelverhältnissen der Lautsprecher, wie sie sich bei Surroundwiedergabesystemen etabliert haben, abweichen, könnten die kanalbasiert abgemischten Beds nicht korrekt wiedergegeben werden und die Lokalisation einzelner Elemente innerhalb der Mischung wäre verfälscht. Deshalb wurden in zahlreichen Tests Lautsprecheranordnungen entwickelt, welche auf die bestehenden Surroundanordnungen aufbauen und eine zweckgemäße Dolby Atmos Wiedergabe ermöglichen. Die optimale Anzahl und Position der Lautsprecher hängt von der Beschaffenheit des Wiedergaberaumes ab.<sup>132</sup> Bestehende Kinosäle sollten für die Maximierung des Hörerlebnisses mit zusätzlichen Surround- und Deckenlautsprechern ausgestattet werden (s. blaue Elemente in Abbildung 23).

<sup>129</sup> Dolby Laboratories Inc. 2. 2013, 2.

<sup>130</sup> Vgl. Dolby Laboratories Inc. 3, „Dolby Atmos Renderer Guide.“  
[https://professional.dolby.com/siteassets/content-creation/dolby-atmos/dolby\\_atmos\\_renderer\\_guide.pdf](https://professional.dolby.com/siteassets/content-creation/dolby-atmos/dolby_atmos_renderer_guide.pdf) (letzter Zugriff: 14. Juni 2022), 12.

<sup>131</sup> Vgl. ebd., 53.

<sup>132</sup> Vgl. Dolby Laboratories Inc. 1. 2014, 17.



Dabei besteht die Anforderung, dass alle Lautsprecher sich klanglich möglichst identisch verhalten sollten.<sup>133</sup> Ein Audioobjekt soll sich stets gleich anhören, ungeachtet seiner Wiedergabe über einen Decken-, Surround- oder Frontal-Lautsprecher. Lediglich die Positionierung im Raum soll dem Rezipienten Richtungsinformationen vermitteln. Darüber hinaus ist es für eine klare Positionierung von Audioobjekten essenziell, jeden der 64 möglichen Lautsprecher über einen eigenen Audiokanal betreiben zu können.

Für eine optimale Wiedergabe hat Dolby eine Vielzahl an Empfehlungen erarbeitet, wie die jeweiligen Lautsprecher zueinander positioniert werden sollten. Beispielsweise sollten die Abstände zueinander zwischen zwei und drei Meter betragen, die Winkeldifferenzen benachbarter Lautsprecher darf 30 Grad nicht überschreiten und die Raummitte sollte eine Symmetrieachse für sämtliche Lautsprecher darstellen, mit Ausnahme der Subwoofer.<sup>134</sup>

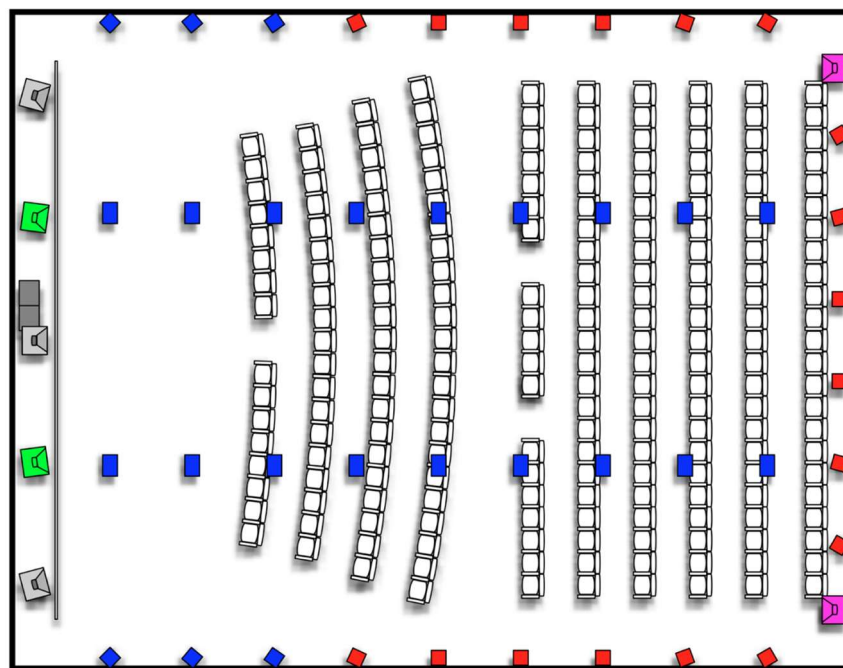


Abbildung 23 Lautsprecheranordnung von Dolby Atmos im Kinosaal<sup>135</sup>

Die Skalierung für den Heimkinobereich folgt für eine optimale Wiedergabe ebenfalls konkreten Vorgaben. Die notwendige Renderer Software ist in Dolby Atmos kompatiblen Audio/Visual Receivern enthalten. Diese können bis zu 24 Lautsprecher auf Ohrhöhe in Verbindung mit 10 Überkopflautsprechern betreiben.<sup>136</sup> Das Prinzip der individuellen Anpassung der Mischung an das Wiedergabemedium bleibt bestehen. Solange der Renderer

<sup>133</sup> Vgl. Dolby Laboratories Inc. 1. 2014, 8.

<sup>134</sup> Vgl. ebd., 18–22.

<sup>135</sup> Ebd., 18.

<sup>136</sup> Vgl. Dolby Laboratories Inc. 4, „Dolby Atmos for the Home Theater.“  
<https://professional.dolby.com/siteassets/tv/home/dolby-atmos/dolby-atmos-for-home-theater.pdf> (letzter Zugriff: 3. Juni 2022), 4.

das Wiedergabemedium kennt, werden sämtliche Audioinformationen darauf bestmöglich wiedergegeben.

Für Konsumenten, die nicht die Möglichkeit besitzen Lautsprecher an oder in der Decke zu montieren, hat Dolby sogenannte Dolby Atmos Enabled Speaker entwickelt. Diese speziellen Lautsprecher simulieren mithilfe von Schallreflexion an der Decke über dem Zuhörer die fehlenden Überkopflautsprecher.<sup>137</sup>

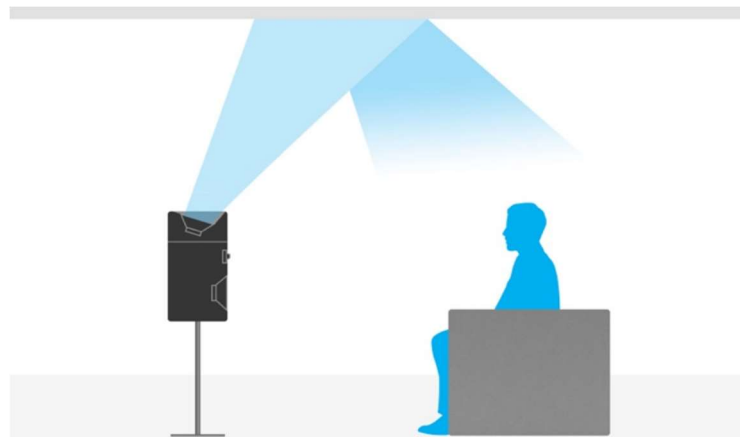


Abbildung 24 Funktionsprinzip eines Dolby Atmos Enabled Speaker<sup>138</sup>

Zusätzlich zur physisch von oben kommenden Einfallsrichtung des reflektierten Schalls wird das Signal in seiner Klangfärbung so bearbeitet, dass es die psychoakustische Wahrnehmung `von oben` begünstigt, wodurch der über Kopf Effekt verstärkt werden soll.<sup>139</sup> Voraussetzung dafür sind eine plane und parallel zum Boden verlaufende Reflexionsfläche, wie sie beispielsweise bei Holz oder Beton vorliegt.<sup>140</sup>

### 3.2.2 Funktionsweise

Dolby Atmos wurde vordergründig dafür entwickelt Filminhalte immersiv zu untermalen und die gesehene Szene akustisch wiederzugeben. "Every Sound is a scene".<sup>141</sup> Diese Szenen gilt es mit einem Dolby Atmos System realistisch zu reproduzieren und somit dem Rezipienten das Gefühl zu geben, in das dargestellte Geschehen involviert zu sein. Jedes individuelle Schallfeld hat seine eigenen Details mit unterschiedlichen Eigenschaften. Zwischen einhüllenden und großflächig vorhandenen Geräuschen, wie dem dreidimensionalen Rauschen des Windes auf einer Wiese, existieren in der Natur klar lokalisierbare Schallquellen, wie beispielsweise ein singender Vogel, welcher über den Beobachter

<sup>137</sup> Vgl. Dolby Laboratories Inc. 4. 2016, 5–8.

<sup>138</sup> Ebd., 5.

<sup>139</sup> Vgl. AVS Forum, „Revealing the Science Behind Dolby Atmos Enabled Speakers—Sponsored.“. <https://www.youtube.com/watch?v=qyqF07KLOVQ> (letzter Zugriff: 10. Juni 2022).

<sup>140</sup> Vgl. Dolby Laboratories Inc. 4. 2016, 6. f.

<sup>141</sup> Ebd., 3.

hinwegfliegt. Die akustische Reproduktion dieser Szene erfordert demnach Möglichkeiten beides wiederzugeben.

Dafür kombiniert Dolby Atmos kanal- und objektbasierte Komponenten in einem Audioformat. Atmosphärische Soundelemente, wie das Rauschen des Windes, bieten sich für eine herkömmliche kanalbasierte Unterbringung in einem sogenannten Bed an. Ihre einhüllenden Eigenschaften profitieren von einer diffusen Wiedergabe über eine Vielzahl von Lautsprechern, sogenannten Lautsprecherarrays.<sup>142</sup> Die richtungsgetreue Lokalisation einer solchen Bed-Mischung kann jedoch nur in einem bestimmten Sweetspot erfolgen, da sich die Richtungsinformationen aus den sich überlagernden Schallfeldern mehrerer Lautsprecher ergeben. Befindet sich der Zuhörer nicht im Zentrum der Lautsprecheranordnung, so verschiebt sich die Richtungswahrnehmung zu den nächstgelegenen Lautsprechern. Somit eignet sich eine herkömmliche kanalbasierte Mischung nur bedingt für die gezielte Platzierung einzelner Geräusche. Ebendiese gezielte Positionierung einzelner Schallquellen, wie den sich bewegenden Vogel, können diese als Audioobjekte im Mix untergebracht werden.

*“Audio objects can be considered as groups of sound elements that share the same physical location in the auditorium. Objects can be static or they can move. They are controlled by metadata that, among other things, details the position of the sound at a given point in time. When objects are monitored or played back in a theater, they are rendered according to the positional metadata using the speakers that are present, rather than necessarily being output to a physical channel.”<sup>143</sup>*

Die Nutzung von Audioobjekten bietet einen weiteren Vorteil. Dadurch, dass Geräusche physisch aus bestimmten Raumrichtungen und den ausschließlich dort befindlichen Lautsprechern wiedergegeben werden können, ist deren Lokalisation für jede Abhörposition innerhalb des Wiedergaberaums identisch.<sup>144</sup> Demnach ist die Richtungswahrnehmung der Audioobjekte nicht an einen Sweetspot gebunden. Ebenso entfallen ungewollte Artefakte, welche bei der Schallfeldüberlagerung mehrerer Lautsprecher entstehen würden.<sup>145</sup>

---

<sup>142</sup> Vgl. Dolby Laboratories Inc. 1. 2014, 9.

<sup>143</sup> Ebd., 8.

<sup>144</sup> Vgl. ebd., 7.

<sup>145</sup> Vgl. ebd., 8.

Die konkrete Berechnung der Wiedergabesignale wird in der für diese Arbeit genutzten Fachliteratur nicht thematisiert, weshalb sich basierend auf der Beschreibung der Wirkweise der Audioobjekte deren Wiedergabeprinzip lediglich vermuten lässt. Beschreibungen wie „content placed in the surrounds can be rendered to an individual speaker if desired“<sup>146</sup> eröffnen die Möglichkeit, dass Audioobjekte gemäß ihrer Raumposition ausschließlich im nächstgelegenen Lautsprecher abgespielt werden. Demnach würde das Wiedergabeprinzip von der ortsstabilen Lokalisation einer Mono-Schallquelle Gebrauch machen. Da einem Audioobjekt ausschließlich Signale zugeordnet werden, die dieselbe Raumposition innehaben sollen, wäre eine akustische Differenzierung innerhalb dieses Audioobjektes ohnehin irrelevant. Auch Berechnungsaspekte aus Bereichen des Amplitude Panning sind basierend auf Beschreibungen der fließenden Bewegung einzelner Schallquellen von Lautsprecher zu Lautsprecher denkbar.

### **3.2.3 Ablauf einer Mischung**

Ebenso wie bei herkömmlichen Audioformaten variiert die Produktion immersiver Tonmischungen mit Dolby Atmos je nach Produktionsinhalt und der Präferenz des jeweiligen Toningenieurs. Besonders die Variabilität des Wiedergabemediums bedingt ein neues Verständnis während einer Mischung. Im Folgenden wird sich deshalb auf die Beschreibung einer Produktionsmöglichkeit einer Kinotonmischung mit der DAW Pro Tools fokussiert.

Zu Beginn des Mischprozesses steht die Signalbearbeitung aller Audioquellen in der DAW. Die Aufnahme der Signale erfolgt in Mono oder Stereo.<sup>147</sup> Auf herkömmliche Weise werden die einzelnen Signale effektbearbeitet und klanglich finalisiert.

Der Unterschied zu regulären Audioformaten liegt in der Zuordnung zu Bed-Audio und Audioobjekten. Die Entscheidung, welches Geräusch wie behandelt wird, obliegt selbstverständlich dem jeweiligen Toningenieur, welcher die Mischung anfertigt. Dennoch lässt sich anhand einiger Kriterien eine generelle Zuordnung voraussagen. Für eine klassische Filmtone Mischung werden sämtliche Geräusche unter drei Gesichtspunkten betrachtet: Dialog, Soundeffekte und Musik.

Dialog kann sich situativ sowohl als frei bewegliches Objekt, als auch als einhüllendes Umgebungsgeräusch anbieten. Die Stimme eines Hauptcharakters könnte beispielsweise von einer freien Positionierung als Audioobjekt profitieren, wohingegen Dialog einer das Geschehen umgebenden Menschenmenge für eine diffuse Wiedergabe geeignet wäre.

---

<sup>146</sup> Dolby Laboratories Inc. 1. 2014, 9.

<sup>147</sup> Vgl. ebd., 13.

Bei Effekten verhält sich die Einteilung ähnlich. Sämtliche Geräusche, die frei positioniert und bewegt werden müssen, um das Filmgeschehen akustisch abzubilden, werden zu Audioobjekten. Alles Weitere wird im Bed-Audio untergebracht. Musikalische Inhalte werden im Normalfall ebenfalls als kanalbasiertes Bed abgemischt, wobei auch hier situativ Abweichungen möglich sind.<sup>148</sup>

Die Mischung des kanalbasierten Beds erfolgt direkt in der genutzten DAW ohne Einsatz zusätzlicher Software. Dolby Atmos sieht von seinen insgesamt 128 Kanälen die ersten 10 für das Bed vor. Dabei entsprechen acht der Kanäle einer herkömmlichen 7.1 Surroundmischung, also einem standardisierten Aufbau aus sieben Lautsprechern mit einem Subwoofer für tieffrequente Signalanteile. Die restlichen zwei Kanäle ergeben ein Stereo-Signalpaar für Überkopflautsprecher.<sup>149</sup> In diesen 10 Kanälen werden sämtliche Geräusche untergebracht, welche im Vorfeld als Bed-Elemente selektiert wurden. Anschließend werden die einzelnen Bed-Kanäle innerhalb der DAW auf individuelle Ausgangskanäle geroutet, welche mithilfe von HD MADI Interfaces an die Dolby Atmos RMU übertragen werden.

Die verbleibenden 118 Kanäle können für Audioobjekte genutzt werden. Ein Audioobjekt agiert als Audioinformation mit zugehörigen Metadaten, welche unter anderem seine Position und Bewegung bestimmen. Die Audioinformation wird über Monokanäle ebenfalls mithilfe von HD MADI Interfaces an die RMU übertragen. Stereosignale nehmen demnach zwei der 118 verbleibenden Kanäle ein. Geräusche, die im Mix dieselbe Position haben oder dieselbe Bewegung ausführen sollen, können zu einem Audioobjekt zusammengefasst werden. Dadurch wird die Anzahl an benötigten Kanälen minimiert. Die Erzeugung der Metadaten geschieht durch das sogenannte Dolby Atmos Panner Plug-In.

---

<sup>148</sup> Vgl. Dolby Laboratories Inc. 1. 2014, 12. f.

<sup>149</sup> Vgl. Dolby Laboratories Inc. 2. 2013, 15.

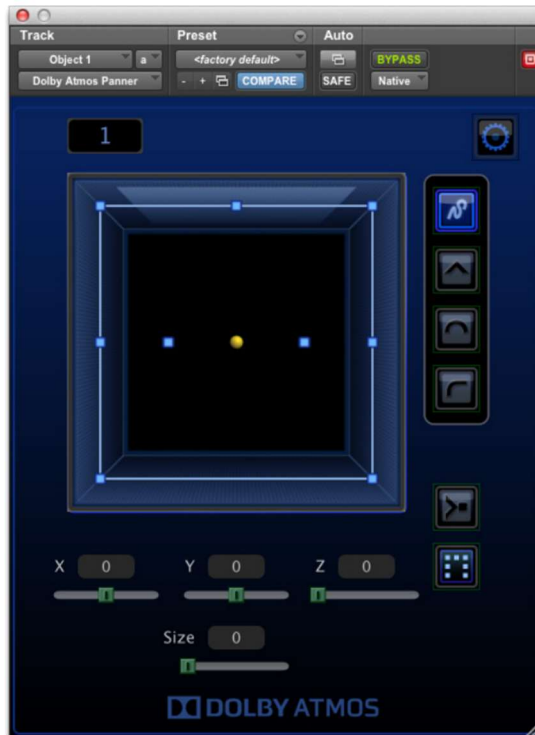


Abbildung 25 Dolby Atmos Panner Plug-In<sup>150</sup>

Dieses Plug-In muss in jede Audioobjektspur eingefügt werden, um die Positionsdaten für das zugehörige Audioobjekt zu generieren.<sup>151</sup> Die Bedienung erfolgt ähnlich wie bei herkömmlichen Surround-Pannern. Das Audioobjekt, welches grafisch als Kugel dargestellt ist, wird mithilfe einer x-, y- und z-Koordinate virtuell im Raum platziert. Dynamische Bewegungsabläufe können in Form von Automationsdaten definiert werden. Die daraus hervorgehenden Metadaten werden über Ethernet ebenfalls zur RMU übertragen.

Innerhalb der RMU werden die Audiodaten mit ihren Metadaten verrechnet. Dafür werden die ankommenden Audiokanäle innerhalb von Dolby Atmos entweder dem 7.1.2 Bed oder einzelnen Audioobjekten zugeordnet. Ebenso werden die per Ethernet übermittelten Metadaten ihren Audioobjekten zugeordnet.

---

<sup>150</sup> Dolby Laboratories Inc. 2. 2013, 5.

<sup>151</sup> Vgl. ebd.

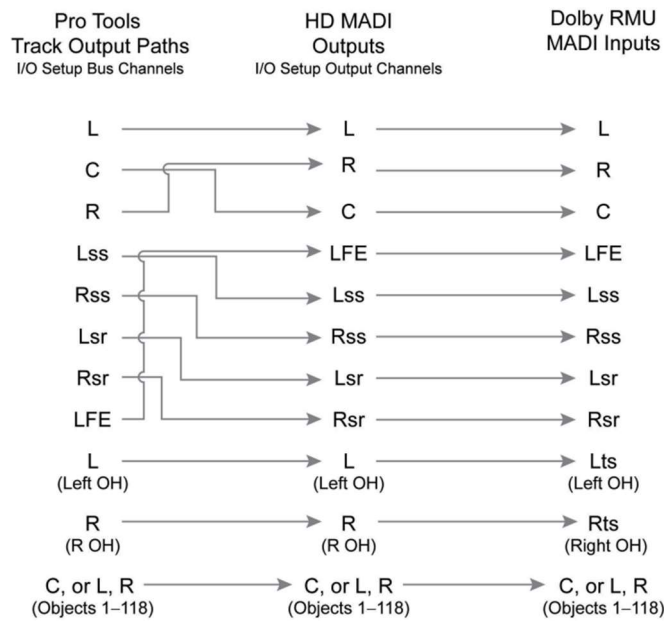


Abbildung 26 Routing der Audiokanäle von Pro Tools zur Dolby RMU<sup>152</sup>

Der in der RMU enthaltene Dolby Atmos Renderer ist für das Abhören während der Mischung erforderlich. In dieser Software werden die Dimensionen des Wiedergaberaumes sowie die Art und Position vorhandener Lautsprecher hinterlegt. Darauf angepasst wird der Mix aus Bed und Audioobjekten akustisch wiedergegeben. Ein enormer Unterschied zu herkömmlichen kanalbasierten Mischverfahren liegt darin, dass während des Mischprozesses die konkrete Wiedergabekonfiguration unbekannt ist, da der Mix wiedergabeseitig stets auf ein individuelles Lautsprechersystem gerendert wird. Für bestmögliche Ergebnisse empfiehlt Dolby bei der Mischung in Dolby Atmos den Mix mindestens auf einem 7.1.4 Lautsprechersystem wiederzugeben. Wahlweise kann jedoch ein Mix softwareintern auch in anderen Formaten abgehört werden, beispielsweise in 5.1 Surround, Stereo oder Binaural. Im Sinne der Kompatibilität und somit zur Sicherung der Wiedergabequalität auf möglichst vielen Wiedergabemedien empfiehlt es sich deshalb den Mix mehrfach in anderen Formaten abzuhören und gegebenenfalls anzupassen.

<sup>152</sup> Dolby Laboratories Inc. 2. 2013, 16.

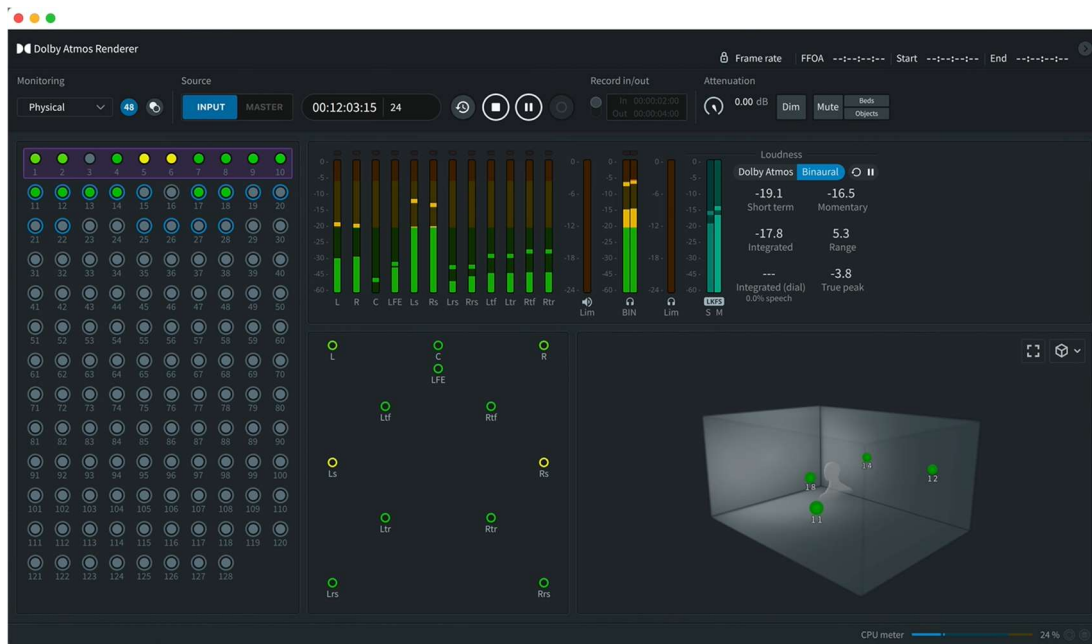


Abbildung 27 Dolby Atmos Renderer Hauptbildschirm<sup>153</sup>

Ist der Mix final abgemischt wird er innerhalb des Dolby Atmos Rederers zu einem Dolby Atmos Print Master File aufgenommen. Dieses fasst sämtliche Audio- und Metadaten in einem Dolby Atmos Mix zusammen. Parallel wird neben dem Atmos Mix auch ein regulärer kanalbasierter Main Audio Mix abgespeichert. Dieser umfasst die Dolby Atmos Mischung in einem 7.1 oder 5.1 Surround-Sound Format.<sup>154</sup>

*„Any installations that are not suitably equipped will simply ignore the additional track file containing the Dolby Atmos soundtrack and will use the existing main audio track file for standard playback. Those installations equipped with a Dolby Atmos cinema processor will be able to ingest and replay the Dolby Atmos soundtrack where applicable, reverting to the standard audio track if necessary.“<sup>155</sup>*

<sup>153</sup> Music Tech Explained, „Mixing in Dolby Atmos: How it Works.“  
<https://www.youtube.com/watch?v=W54VcLaSMjM> (letzter Zugriff: 28. Mai 2022).

<sup>154</sup> Vgl. Dolby Laboratories Inc. 1. 2014, 14.

<sup>155</sup> Ebd., 16.



### 3.2.4 Einsatzmöglichkeiten

Dolby Atmos hat sich im Laufe der Zeit über seinen ursprünglichen Anwendungsbereich des Kinotons hinausentwickelt. Nichtsdestotrotz stellt dieser noch immer einen enormen Anwendungsbereich der Technik dar. Seit der erstmaligen Verwendung 2012 im Disney Film "Brave" (im Deutschen: "Merida – Legende der Highlands") etablierte sich das neue Format in zahlreichen Filmen und Kinosälen. Mittlerweile sind allein in Deutschland über 50 Kinos mit Dolby Atmos fähiger Technik ausgestattet.<sup>156</sup>

Im Home Entertainment etablierte sich Dolby Atmos ebenfalls. Die Entwicklung von applizierbaren Heimkinosystemen in Verbindung mit Dolby Atmos Audio/Visual Receivern ermöglichen das Rendering und dadurch die Wiedergabe von Atmos-Mischungen in den eigenen vier Wänden. Dabei entscheidet das Wiedergabemedium über das Level an Immersion. Durch die Bereitstellung von Dolby Atmos fähigen Wiedergabemedien wie Soundbars, welche Audio lediglich frontal wiedergeben können, ist das Hören einer Dolby Atmos Mischung nicht zwangsläufig immersiv, gegebenenfalls nicht einmal Surround-Sound. Doch dadurch wird Dolby Atmos auch den Haushalten zugänglich gemacht, welche über kein 10-Kanal-Lautsprechersystem verfügen. Darüber hinaus können Atmos-Inhalte nicht nur über physische Blu-ray Discs, sondern auch über verschiedene Streaming Angebote übertragen werden. Beispielsweise wird Dolby Atmos Audio von Netflix, Amazon Prime oder Maxdome unterstützt.<sup>157</sup>

Auch in die Welt der Musik wurden die neuen Möglichkeiten implementiert. Durch die native Integration von Dolby Atmos Software in verschiedene DAWs, wie Nuendo oder Logic, wird es Privatpersonen ermöglicht ohne immensen Kostenaufwand Atmos Mischungen anzufertigen, wodurch die Bandbreite an konsumierbaren Inhalten zunehmend wächst. Auch die Wiedergabemöglichkeiten folgen diesem Trend des einfachen Einstiegs. Dolby Atmos Inhalte, welche über Apple Music gestreamt und über Apple Air Pods gehört werden, profitieren von Apple Spatial Audio. Dadurch werden Atmos-Mischungen auf ein binaurales Stereosignal gerendert und mit Head-Tracking Daten verknüpft, wodurch sich der Rezipient im Atmos Mix bewegen kann.<sup>158</sup>

Die Gamingwelt profitiert ebenfalls von immersivem Dolby Atmos Audio. „Hyper-realistic spatial audio makes you feel like you’re actually inside the experience rather than just hearing

---

<sup>156</sup> Vgl. Sebastian Rittau, „Übersicht: Dolby-Atmos-Kinos in Deutschland.“  
<https://blog.teufel.de/uebersicht-dolby-atmos-kinos-in-deutschland/> (letzter Zugriff: 14. Juni 2022).

<sup>157</sup> Vgl. „Die Streaming-Sound-Offensive: Dolby Atmos bei Netflix, Amazon Prime und maxdome.“  
<https://blog.teufel.de/dolby-atmos-bei-netflix/> (letzter Zugriff: 14. Juni 2022).

<sup>158</sup> Vgl. Michael Wagner 2, „Getting Started with Dolby Atmos Music.“  
[https://www.youtube.com/watch?v=\\_MGIGD2QAgA](https://www.youtube.com/watch?v=_MGIGD2QAgA) (letzter Zugriff: 13. Juni 2022).

it.“<sup>159</sup> Namenhafte Titel wie Elden Ring, Halo Infinite oder Forza Horizon 5 sind nur einige Beispiele für den Erfolg von Dolby Atmos in Videospielen. Die XBOX Series X|S ist zudem die erste Spielekonsole, welche Dolby Atmos unterstützt.<sup>160</sup>

Mithilfe des Dolby Atmos Live Panner wird der Anwendungsbereich von Dolby Atmos zusätzlich erweitert. Mit dieser Software kann Atmos nach nahezu zehnjährigem Bestehen nun auch in Liveevents integriert werden. Die Steuerung erfolgt dabei über ein I-Pad, welches über eine Netzwerkverbindung als Client mit dem Dolby Atmos Renderer kommuniziert.<sup>161</sup>

### 3.3 Gegenüberstellung von Auralite 3D und Dolby Atmos

Auralite 3D und Dolby Atmos verfolgen als immersive und dreidimensionale Audiosysteme grundlegend dasselbe Ziel. Rezipienten sollen, egal ob es sich um eine Live- oder Vorproduktion handelt, in eine technisch erzeugte Klangwelt eingebettet werden. Beide Systeme nutzen Mono- oder Stereosignale, welche in herkömmlichen Aufnahmeverfahren gewonnen werden können, als Quellmaterial. Ebenso parallel sorgen dreidimensionale Lautsprecheranordnungen und die Betrachtung einzelner Schallquellen als Audioobjekte für neue Möglichkeiten in einer Mischung. Audioobjekte können frei in einem virtuellen Raum positioniert und fließend von Lautsprecher zu Lautsprecher bewegt werden. Die dabei virtuell erzeugten Metadaten werden bei der Wiedergabe durch spezielle Renderer-Software mit ihren zugehörigen Audiodaten verknüpft und über das vorhandene Lautsprechersystem wiedergegeben. Durch die Variabilität der Wiedergabeordnung sind beide Systeme auch zu herkömmlichen Audioformaten, wie Surround oder Stereo, kompatibel. Dabei lässt sich bereits ein Unterschied feststellen.

Während Dolby Atmos in seinem finalen Print Master sowohl eine Dolby Atmos Mischung als auch einen 7.1 und 5.1 Surroundmix beinhaltet, müsste in Auralite 3D für die Kompatibilität zu diesen Formaten ein höherer Aufwand betrieben werden. Dieser umfasst das Anlegen der notwendigen Kanäle als Audioobjekte und deren virtuelle Anordnung nach den jeweiligen Wiedergabestandards. Für eine perfekte Surround-Reproduktion könnte zusätzlich der physische Lautsprecheraufbau dem Surround-Standard angepasst werden.

Eine weitere Differenzierung ist dahingehend möglich, dass Auralite 3D ausschließlich objektbasiert arbeitet und je nach Anwendungsbereich 32 oder 64 Audioobjekte verwalten

---

<sup>159</sup> Dolby Laboratories Inc. 5, „A whole new world of gaming: Get more intensely into the story, the visuals, and the action in Dolby.“. <https://www.dolby.com/gaming/> (letzter Zugriff: 14. Juni 2022).

<sup>160</sup> Vgl. Dolby Laboratories Inc. 6, „Gaming in Dolby, now on XBOX.“. <https://www.dolby.com/experience/xbox/> (letzter Zugriff: 14. Juni 2022).

<sup>161</sup> Vgl. Dolby Laboratories Inc. 7, „Dolby Atmos Live Panner: Intuitive Tool to Control & Automate Dolby Atmos Objects for Live Events.“. <https://www.youtube.com/watch?v=A7oT6rQ3c88> (letzter Zugriff: 14. Juni 2022).

kann. In Dolby Atmos hingegen werden kanalbasierte Bed-Mischungen mit bis zu 118 Audioobjekten kombiniert. Darüber hinaus werden die positionsbestimmenden Metadaten der Audioobjekte bei beiden Systemen in unterschiedlichen Phasen des Workflows erstellt. In Auralite 3D erfolgt die Positionierung einzelner Schallquellen erst innerhalb des Renderers. In Dolby Atmos erfolgt das dreidimensionale Panning bereits in der DAW durch dafür ausgelegte Software-Plug-Ins und somit vor der Signalübertragung an den Renderer.

Obwohl beide Systeme grundlegend mit variablen Lautsprecherkonfigurationen arbeiten, welche lediglich virtuell in die Software übertragen werden müssen, orientiert sich der Lautsprecheraufbau bei Dolby Atmos stark an bestehenden Surroundanordnungen. Dadurch wird sowohl die Wiedergabe des kanalbasierten Beds gewährleistet als auch die Kompatibilität zum im Print Master enthaltenen Main Audio Mix gesichert. Zusätzliche Surround- und Überkopflautsprecher grenzen Dolby Atmos von herkömmlichen Wiedergabesystemen ab. Auch die Anzahl an individuellen Lautsprecherkanälen zeigt Unterschiede. Während Auralite 3D 128 Ausgangskanäle berechnen kann sind bei Dolby Atmos lediglich 64 Ausgangskanäle vorgesehen.

Ein fundamentaler Unterschied liegt in der Berechnung der Lautsprechersignale. Das Funktionsprinzip von Auralite 3D basiert auf der Wellenfeldsynthese. Es wird also die physikalische Synthese eines Schallfeldes angestrebt, welches im Idealfall den gesamten Wiedergaberaum dreidimensional abdecken soll. Bei Dolby Atmos geht es nicht um die Reproduktion eines Schallfeldes, sondern vielmehr um die Weiterentwicklung kanalbasierter Wiedergabesysteme. Der Zuhörer soll sich dabei auch in ein produziertes Schallfeld integriert fühlen, welches jedoch nicht aus synthetisierten Wellenfronten besteht, sondern aus einem kanalbasierten Grundmix mit einzelnen, gezielt platzierten Geräuschen. Demzufolge arbeiten beide Systeme aus unterschiedlichen Perspektiven. In Auralite 3D wird zunächst virtuell ein Schallfeld berechnet, welches anschließend wiedergegeben wird. Das Schallfeld wird bei dieser Betrachtungsweise als physikalische Größe angesehen, welche es zu replizieren gilt. Dolby Atmos hingegen ordnet Schallinformationen basierend auf Metadaten einzeln angesteuerten Lautsprechern zu, welche infolgedessen ein dreidimensionales Schallfeld erzeugen. Dabei wird das Schallfeld als Resultat einer Wiedergabe angesehen.<sup>162</sup>

Ein Alleinstellungsmerkmal für Auralite 3D liegt in der im System enthaltenen Raumhallsimulation. Die dreidimensionale Nachbildung komplexer Raumakustiken in Verbindung mit der zugrundeliegenden Signalberechnung durch die Wellenfeldsynthese bedingen ein erweitertes Anwendungsspektrum im Vergleich zu anderen immersiven Audiosystemen.

---

<sup>162</sup> Vgl. Michael Wagner 1. 2021.

## 4 Fazit

Die Beantwortung der Frage „Wodurch erreicht man auditive Immersion im System Auralite 3D und wo liegen die Unterschiede zu Dolby Atmos?“ ist überaus komplex. Der höchst emotionale und subjektive Anspruch der Immersion des Rezipienten kann sich mithilfe zahlreicher Mechanismen manifestieren.

Allgegenwärtig ist dabei die Orientierung an der natürlichen Hörweise des Menschen. Schall umgibt einen Menschen täglich. Schallwellen unterschiedlichster Geräusche können aus sämtlichen Raumrichtungen auf den Rezipienten treffen und werden durch beide Ohren aufgenommen. Je nach Einfallsrichtung können zwischen den dadurch aufgenommenen Ohrsignalen Laufzeit- und Pegelunterschiede sowie Klangfarbenunterschiede auftreten, welche durch den Vergleich der Signale im Gehirn analysiert werden. Aus den daraus gewonnenen Richtungsinformationen wird der Ursprung der Schallwellen und somit die Position der Schallquelle bestimmt. Darüber hinaus bleibt diese Position bei einer statischen Schallquelle konstant, selbst wenn sich der Rezipient im Schallfeld bewegt.

Um ein sich real und einhüllend anfühlendes Schallfeld zu generieren, müssen somit vergleichbare Bedingungen bei einer technischen Wiedergabe in immersiven Audiosystemen geschaffen werden. Dafür nutzen sowohl Auralite 3D als auch Dolby Atmos dreidimensionale Lautsprecheranordnungen, um Schall aus nahezu allen Raumrichtungen wiedergeben zu können. Die bei einer Wiedergabe entstehenden physischen Schallwellen rufen die gleichen Wahrnehmungerscheinungen beim Rezipienten hervor wie ein reales Schallereignis.

Die Einflussnahme der Bewegung des Rezipienten auf die exakte Lokalisation innerhalb des Schallfeldes erfordert einen Einblick in die Funktionsweisen der beiden Systeme. Durch eine objektbasierte Positionierung virtueller Schallquellen in einem virtuellen Raum stellen sowohl Auralite 3D als auch Dolby Atmos revolutionäre Strategien der Audiomischung bereit. Die Audiodaten werden im Laufe der Mischung mit Metadaten verknüpft und mithilfe spezieller Rendering-Software auf ein dem System bekanntes Lautsprechersystem angepasst. Der klassische Sweetspot, welcher bei kanalbasierten Mischverfahren eine exakte Lokalisation ausschließlich an einem Punkt innerhalb der Lautsprecheranordnung gewährleistet, gehört demnach der Vergangenheit an. Dennoch existieren zwischen beiden Systemen Unterschiede in der tatsächlichen Berechnung der Wiedergabesignale.

Auralite 3D strebt mithilfe der Wellenfeldsynthese eine physikalische Reproduktion eines realen Schallfeldes an. Es wird die erste Wellenfront der jeweiligen virtuellen Schallquelle berechnet und gemäß ihrer Ausbreitung mithilfe einer dem System bekannten Lautsprecheranordnung wiedergegeben. Trotz einiger Abweichungen vom theoretischen Ideal der Wellenfeldsynthese ist mit dieser Technik die freie Bewegung innerhalb einer Hörzone

möglich, ohne dass der Klangeindruck verfälscht wird. Dolby Atmos hingegen kombiniert als hybrides Audioformat objektbasiertes und kanalbasiertes Audio. Das kanalbasierte Bed jeder Mischung ist demnach auf einen bestimmten Sweetspot ausgerichtet. Die objektbasierten Audioelemente sind hingegen durch ihre exakte Platzierung im gesamten Wiedergaberaum ortsstabil wahrnehmbar. Bedenkt man den ursprünglichen Einsatzzweck der Kinotonmischung ist es jedoch naheliegend, dass aufgrund der festgelegten Sitzplätze während eines Films keine Bewegung der Rezipienten innerhalb des Schallfeldes vorgesehen ist. Im durch den Dolby Atmos Live Panner neu erschlossenen Gebiet der Liveanwendungen ist jedoch eine Bewegung der Rezipienten im Beschallungsraum keinesfalls ausgeschlossen.

Des Weiteren unterscheiden sich beide Systeme in ihrer Lautsprecheranordnung. Im Sinne der Kompatibilität zu bestehenden Audioformaten orientiert sich der Lautsprecheraufbau bei Dolby Atmos stark an den etablierten Standards für Surround-Wiedergabesysteme. Nicht zuletzt wird somit die korrekte Wiedergabe der enthaltenen Beds sichergestellt. Als Innovation werden zusätzliche Höhenlautsprecher und die separate Ansteuerung jedes einzelnen Lautsprechers genutzt. Auralite 3D hingegen verzichtet auf jegliche Vorgaben der Lautsprecherpositionierung. Logisches Verständnis bedingt gewisse Grundlagen, wie beispielsweise die Ausrichtung auf eine festgelegte Hörzone, jedoch können jegliche Audioinhalte auf einer absolut variablen Lautsprecherkonfiguration wiedergegeben werden. Es ist lediglich erforderlich den jeweiligen physischen Aufbau virtuell in die Systemsoftware zu übertragen. Die Qualität der Wiedergabe steigt bei beiden Systemen mit der Anzahl der verwendeten Lautsprecher.

Auf Grundlage aller Faktoren ergeben sich für beide Systeme individuelle Anwendungsbereiche, welche vereinzelt Schnittmengen aufweisen. Da Auralite 3D in seinem Aufbau enorm variabel eingesetzt werden kann, eignet es sich für Festinstallationen sowie mobile Produktionen, welche einen Lautsprecheraufbau nach etablierten Standards erschweren. Darüber hinaus kann es durch die Einbindung der Raumhallsimulation für interaktive Installationen und die akustische Nachbildung von Raumklängen genutzt werden. Die wesentlich höhere Anzahl an nutzbaren Audiokanälen verschafft Dolby Atmos einen Vorteil in der Produktion umfangreicherer Inhalte. Zudem bietet es eine unmittelbarere Kompatibilität zu herkömmlichen Surround-Formaten.

Beide Systeme sind in der Lage immersive Audiowelten zu produzieren und zeigen durch die Nutzung objektbasierter Audioteknik, dass ein neues Verständnis von Schallwiedergabe enorme Möglichkeiten bietet. Individuelle Workflows und Funktionsprinzipien grenzen Auralite 3D und Dolby Atmos voneinander ab, wodurch beide Systeme ihre Legitimation besitzen.

## Literaturverzeichnis

- AVS Forum, „Revealing the Science Behind Dolby Atmos Enabled Speakers—Sponsored.“  
<https://www.youtube.com/watch?v=qyqF07KL0VQ> (letzter Zugriff: 10. Juni 2022).
- Backhaus, H., J. Friese, E. M. Hornbostel, A. Kalähne, H. Lichte, E. Lübcke und E. Meyer et al., Hrsg.,  
*Akustik*. Handbuch der Physik 8. Berlin, Heidelberg: Springer, 1927.
- Bruckner, Elena, „Ein Raum zum Mitnehmen.“ *Die Rheinpfalz*, 25. Mai 2021.
- Dickreiter, Michael, Volker Dittel, Wolfgang Hoeg und Martin Wöhr, Hrsg., *Handbuch der  
Tonstudioteknik*, 8. überarbeitete und erweiterte Auflage. München: Saur, 2014.
- „Die Streaming-Sound-Offensive: Dolby Atmos bei Netflix, Amazon Prime und maxdome.“  
<https://blog.teufel.de/dolby-atmos-bei-netflix/> (letzter Zugriff: 14. Juni 2022).
- Dolby Laboratories Inc. 1, „Dolby Atmos: Next-Generation Audio for Cinema.“  
<https://www.hollandfilmnieuws.nl/files/whitepaper-dolbyatmos.pdf> (letzter Zugriff: 3. Juni 2022).
- Dolby Laboratories Inc. 2, „Authoring for Dolby Atmos Cinema Sound Manual.“  
[https://www.dolby.com/uploadedFiles/Assets/US/Doc/Professional/Authoring\\_for\\_Dolby\\_Atmos\\_Cinema\\_Sound\\_Manual\(1\).pdf](https://www.dolby.com/uploadedFiles/Assets/US/Doc/Professional/Authoring_for_Dolby_Atmos_Cinema_Sound_Manual(1).pdf) (letzter Zugriff: 3. Juni 2022).
- Dolby Laboratories Inc. 3, „Dolby Atmos Renderer Guide.“  
[https://professional.dolby.com/siteassets/content-creation/dolby-atmos/dolby\\_atmos\\_renderer\\_guide.pdf](https://professional.dolby.com/siteassets/content-creation/dolby-atmos/dolby_atmos_renderer_guide.pdf) (letzter Zugriff: 14. Juni 2022).
- Dolby Laboratories Inc. 4, „Dolby Atmos for the Home Theater.“  
<https://professional.dolby.com/siteassets/tv/home/dolby-atmos/dolby-atmos-for-home-theater.pdf>  
(letzter Zugriff: 3. Juni 2022).
- Dolby Laboratories Inc. 5, „A whole new world of gaming: Get more intensely into the story, the  
visuals, and the action in Dolby.“. <https://www.dolby.com/gaming/> (letzter Zugriff: 14. Juni 2022).
- Dolby Laboratories Inc. 6, „Gaming in Dolby, now on XBOX.“. <https://www.dolby.com/experience/xbox/>  
(letzter Zugriff: 14. Juni 2022).
- Dolby Laboratories Inc. 7, „Dolby Atmos Live Panner: Intuitive Tool to Control & Automate Dolby  
Atmos Objects for Live Events.“. <https://www.youtube.com/watch?v=A7oT6rQ3c88> (letzter  
Zugriff: 14. Juni 2022).
- Fraunhofer - Institut für Digitale Medientechnologie 1, „Wellenfeldsynthese.“  
[https://www.idmt.fraunhofer.de/content/dam/idmt/documents/IL/wave\\_field\\_synthesis\\_de.pdf](https://www.idmt.fraunhofer.de/content/dam/idmt/documents/IL/wave_field_synthesis_de.pdf)  
(letzter Zugriff: 6. April 2022).
- Fraunhofer - Institut für Digitale Medientechnologie 2, „Object-based Spatial Audio Mastering.“  
<https://www.youtube.com/watch?v=cqC1rWSxT4c> (letzter Zugriff: 6. April 2022).

- Fraunhofer IIS, „MPEG-H AUDIO: Das Audiosystem der nächsten Generation für UHDTV, Streaming und VR.“. <https://www.youtube.com/watch?v=9fU90uhHBm8> (letzter Zugriff: 9. Juni 2022).
- Friesecke, Andreas, *Die Audio-Enzyklopädie: Ein Nachschlagewerk für Tontechniker*, 2. Auflage. Berlin: De Gruyter Saur, 2014.
- Henle, Hubert, *Das Tonstudio-Handbuch: Praktische Einführung in die professionelle Aufnahmetechnik*, 5., komplett überarbeitete Auflage. München: Carstensen, 2001.
- Hochschule der Medien. „Die Wellenfeldsynthese.“. <https://curdt.home.hdm-stuttgart.de/PDF/WFS%202.pdf> (letzter Zugriff: 28. August 2021).
- Kirner, Jakob. „Auralite3D Manual: Version 4.4.“.
- Kraus, Felix. „3D Sound: Kino-Ton in drei Dimensionen.“ Bachelorarbeit, University of Applied Sciences, 2015. <https://monami.hs-mittweida.de/frontdoor/index/index/start/2/rows/10/sortfield/score/sortorder/desc/searchtype/simple/query/felix+Kraus/docId/5782> (letzter Zugriff: 4. April 2022).
- Lossius, Trond, Pascal Baltazar und Théo de La Hogue, „DBAP - DISTANCE-BASED AMPLITUDE PANNING.“. [https://d1wqtxts1xzle7.cloudfront.net/50048560/DBAP\\_-\\_DISTANCE-BASED\\_AMPLITUDE\\_PANNING20161101-1709-6hyhe4-with-cover-page-v2.pdf?Expires=1655483932&Signature=BlhYIrl08olprainxA50dZsZvUiBokC84sDI7Cc4BfpTFMSNtDNOdLzSSSVRSZlGcNj2wxgW5sS390~rTZvPlr9iUwbV0VkYnYcRc1n2Ygyjc7FrROQtwVocNVluz243TzZxl7SA8s~dF2esUXthvHd51NcLqScVWhGi6nsTPoQz7G-5J~atRGFUiBKko19aD-5CkFa8AJBWP-H8gfp84uDURrAohlFzBXO4EuoprAlAUJfU84w~w3ocSlD14OrYD5kg1YZy3XBnr3ujQL16lrNHhxK2iyUI4C2Zxl5ntYimHEBld~f1eD814O48S8nsK0FyCOKMNCjh7ns5QxwHA\\_\\_&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA](https://d1wqtxts1xzle7.cloudfront.net/50048560/DBAP_-_DISTANCE-BASED_AMPLITUDE_PANNING20161101-1709-6hyhe4-with-cover-page-v2.pdf?Expires=1655483932&Signature=BlhYIrl08olprainxA50dZsZvUiBokC84sDI7Cc4BfpTFMSNtDNOdLzSSSVRSZlGcNj2wxgW5sS390~rTZvPlr9iUwbV0VkYnYcRc1n2Ygyjc7FrROQtwVocNVluz243TzZxl7SA8s~dF2esUXthvHd51NcLqScVWhGi6nsTPoQz7G-5J~atRGFUiBKko19aD-5CkFa8AJBWP-H8gfp84uDURrAohlFzBXO4EuoprAlAUJfU84w~w3ocSlD14OrYD5kg1YZy3XBnr3ujQL16lrNHhxK2iyUI4C2Zxl5ntYimHEBld~f1eD814O48S8nsK0FyCOKMNCjh7ns5QxwHA__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA) (letzter Zugriff: 17. Juni 2022).
- Lübcke, E., *Schallausbreitung*. [https://link.springer.com/chapter/10.1007/978-3-642-47352-4\\_15](https://link.springer.com/chapter/10.1007/978-3-642-47352-4_15) (letzter Zugriff: 28. Mai 2022).
- Maier, Markus. „Distance based amplitude Panning.“ Bachelorarbeit, Universität für Musik und darstellende Kunst, 2012. <https://phaidra.kug.ac.at/download/o:11015> (letzter Zugriff: 9. Juni 2022).
- Maunder, Paul, „An introduction to surround mixing in Pro Tools.“. <https://www.youtube.com/watch?v=tZeKaHyZ3ig> (letzter Zugriff: 12. Juni 2022).
- Music Tech Explained, „Mixing in Dolby Atmos: How it Works.“. <https://www.youtube.com/watch?v=W54VcLaSMjM> (letzter Zugriff: 28. Mai 2022).
- Pulkki, Ville, „Virtual Sound Source Positioning Using Vector Base Amplitude Panning.“. <http://lib.tkk.fi/Diss/2001/isbn9512255324/article1.pdf> (letzter Zugriff: 17. Juni 2022).
- Rittau, Sebastian, „Übersicht: Dolby-Atmos-Kinos in Deutschland.“. <https://blog.teufel.de/uebersicht-dolby-atmos-kinos-in-deutschland/> (letzter Zugriff: 14. Juni 2022).

Stahlmann, Rainer, *Die verschiedenen Dolby Surround-Sound-Verfahren: Überblick, Analyse und Funktionsweise*. München: GRIN Verlag GmbH, 2002. <https://www.grin.com/document/5732> (letzter Zugriff: 3. Mai 2022).

Stowasser, Josef M., Michael Petschenig und Franz Skutsch, *Stowasser: Lateinisch-deutsches Schulwörterbuch*, Auflage 1994. Wien, München: Verlag Holder-Pichler-Tempsky R. Oldenbourg Verlag, 1994.

Utke, Max. „Objekt- und kanalbasierte Audioformate: Vor- und Nachteile.“ Tonseminar, Hochschule der Medien, 2018. [https://curdt.home.hdm-stuttgart.de/PDF/Objekt-\\_u\\_kanalbasierte\\_Audioformate.pdf](https://curdt.home.hdm-stuttgart.de/PDF/Objekt-_u_kanalbasierte_Audioformate.pdf) (letzter Zugriff: 13. Mai 2022).

Wagner 1, Michael, „Q&A: Ambisonics vs. Atmos, Stereo vs. Binaural, Waves NX in Fairlight.“ <https://www.youtube.com/watch?v=R2QON1y5BCo> (letzter Zugriff: 24. Mai 2022).

Wagner 2, Michael, „Getting Started with Dolby Atmos Music.“ [https://www.youtube.com/watch?v=\\_MGIGD2QAgA](https://www.youtube.com/watch?v=_MGIGD2QAgA) (letzter Zugriff: 13. Juni 2022).

Weinzierl, Stefan, Hrsg., *Handbuch der Audiotechnik*. Berlin, Heidelberg: Springer, 2008.



# **Anhang**

## **A.1 Interview mit Nûjîn Kartal von der MediaApes GmbH**

### **A.1.1 Frage 1**

Seit wann gibt es Auralite 3D?

Auralite 3D gibt es jetzt seit 2018. Seit 2 – 3 Jahren ist es jetzt am Markt.

### **A.1.2 Frage 2**

Wie kam es zur Idee das System zu entwickeln?

Es gab schon diverse Vorstufen in anderer Form, von anderen Firmen und Herstellern, auch andere 3D Systeme. Wir haben uns seit Jahren schon verschiedenste Systeme angeguckt und vor allem angehört und kamen dann auf diese Technologie, die ja eine Frauenhofer Entwicklung ist, wo wir dann diese Entwicklung, wie sie bis dato war, gehört hatten und waren vom Klang sehr überzeugt. Das war der Schritt, wo wir überlegt haben. Wir waren bis dahin eher produzierendes Gewerbe und nicht mit eigenem Produkt, wie so einem System. Wir fanden das aber sehr reizvoll, weil wir auch ein bisschen in die Zukunft geguckt haben und die auch mitgestalten wollen als Player am Markt für immersive Systeme und immersiven Sound. Da hat sich dann die Idee manifestiert, dass wir ein eigenes Produkt ausbauen. Wir haben gesehen, viele können es am Markt gar nicht wirklich etablieren, weil es erklärungsbedürftig ist. Es ist eine Frage, wie schulst du Leute und wie schaffst du Produktionswelten dafür und dann noch am wichtigsten: wie gibst du Leuten Ideen an die Hand fürs Marketing und für die weitergehenden Business-Cases.

### **A.1.3 Frage 3**

Wofür wurde es konzipiert?

Live, als Player und für Installationen allgemein. Es ist ein System, was sowohl Live als auch nicht Livewelten vereint. Es ist eins der wenigen, was beides gut kann und was in seiner Installation superflexibel ist im Vergleich zu anderen Systemen. Es gibt ja auch andere Systeme die dann wirklich nur als Playback System funktionieren, nicht live einsetzbar wegen zu viel Delay, und dann gibt's dezidierte Livesysteme. Das System vereint im Grunde beides und hat dann auch noch Funktionen, die praktisch sind für bestimmte Eventhäuser, wenn es jetzt um Theaterhäuser geht oder wenn es um andere Venues (Veranstaltungsorte) geht, wo man auch szenenbasiert arbeitet und und und. Das hat einfach so vom Gesamtumfang Status Quo eine breite Fläche im Audiobereich, die man immersiv damit bedienen kann.

#### **A.1.4 Frage 4**

Was sind denn die ausschlaggebenden Punkte, die es von der Konkurrenz abgrenzt? Beispielsweise Dolby Atmos im Vergleich zu Auralite3D.

Andere Beispiele im Livesektor wären noch L-Isa oder Soundscape. Dann gibt's auch noch Sound of the Mountain oder Ambisonics-Systeme, die man auch als Lautsprecher System installieren kann. Ausschlaggebend, was es wirklich unterscheidet, ist zum einen meiner Meinung nach die Praktikabilität, wenn man es mit manchen Systemen vergleicht – von produzieren sowie auch einrichten, weil es im Verhältnis relativ schnell geht es einzurichten. Wenn man das auch noch weiter unterscheiden möchte: es ist ein Wellenfeldsynthese basiertes System. Das heißt die grundlegenden Algorithmen sind basierend darauf, wie sich Schall natürlich verbreitet und ausbreitet und das wird immer auf das ganze System in Echtzeit berechnet. Deswegen ist es auch ein dezidierter Renderer, der allein läuft, und nicht nur eine Software auf irgendeinem Rechner, wie es manch andere Systeme auch machen, weil diese Echtzeitberechnungen und dann auch noch die Raumsimulation, die man ja dazunehmen kann, entsprechend Power kostet. Einer der Anwendungsfälle ist auch noch die Raumhall-Simulation an sich, weil die in der Form und in der Qualität bisher ihresgleichen sucht, zumindest von dem, was ich kenne. Wir sind immer auf der Suche: was gibt es für Systeme am Markt, was gibt es für Sound am Markt und was ist wie gut, wie gut klingt es eigentlich - nicht nur wegen Business, nicht nur wegen unserem Produkt, sondern auch aus Interesse. Das vergleichen wir viel und sind dabei auch möglichst objektiv und ich nenne es mal leidenschaftslos. Nur weil wir ein Produkt haben, heißt das nicht, dass es für uns nur das Produkt gibt. Alle Produkte haben ihre Berechtigung. Die Frage ist nur manchmal, wie sehr ist es durch Marktmacht oder einfach nur Geld an den Markt gekommen. Wir setzen bei allem was wir tun immer gerne auf die Qualität des Produkts.

#### **A.1.5 Frage 5**

Wie sieht die interne Kanalbelegung aus?

Grundsätzlich ist es rein objektbasiert und hat keinen kanalbasierten Ansatz. Die Raumsimulation ist, basierend auf der Echtzeitberechnung, dass es stabil läuft und keine Aussetzer gibt, erstmal auf 16 Kanäle begrenzt. Trotzdem kann man diese Inputs auch nehmen und rumschieben. Was dazu kommt: es sind nicht nur die 16 Inputs, die mit der Raumsimulation berechnet werden, sondern gleichzeitig all diese virtuellen Inputs die dann direkte erste Reflektion und zweite Reflektion etc. wiedergeben. Plus das ganze Equalizing, was du betreiben kannst. Sprich wenn du die Raumsimulation einschaltest mit nur einem Mikro, ist der Renderer schon dabei eine multiple virtuelle Anzahl an Objekten zu berechnen. Deswegen gibt es diese Begrenzung, weil es einfach viel Leistung zieht.

### **A.1.6 Frage 6**

Was kostet so ein Setup von Auralite3D? Gibt es eine grobe Preisspanne von jemandem, der sich als Einsteiger dafür interessiert, bis zur Endausbaustufe?

Also das ist eine sehr sehr weite Range, weil das System in Modulen aufgebaut ist. Die einzelnen Module sind zu Anfang der Player. Das ist die kleinste Ausbaustufe. Damit kannst du nur Produktionen wiedergeben, die dafür gemacht und als Player File verpackt wurden. Dann gibt es den Player nochmal erweitert, damit du einzelne Inputs live reingeben kannst, also Sprechen zum Beispiel. Dann gibt es die Production-Suite, dann gibt es die Live-Suite und dann kommen wir zur voll ausgebauten Stufe mit Raumsimulation. Das heißt, ich habe eine sehr weite Range und es kommt immer auch drauf an: muss meine Hardware redundant sein, was für einen Rechner kann ich da konfigurieren, soll der auch zukünftig nur als Player agieren oder sollte er eventuell ausbaubar sein. Das macht allein schon bei der Hardware, die man grundlegend braucht, wo man drauf installiert, einen Unterschied. Hinzu kommt dann jedes Modul, was heißt, die Range ist locker von 20.000 bis 80.000€ je nachdem, welche Module man braucht. Das heißt, einen festen Preis kannst du so gar nicht sagen, weil es kein Produkt ist, was im Regal steht, sondern immer individuell auf denjenigen, der es einkaufen will, angepasst werden will. Das waren jetzt die Preise ohne, dass ich jetzt eine gute Hardware brauche. Wenn ich eine Live-Hardware haben will, die redundant laufen kann und alles, dann bist du damit auch bei 90.000 bis 100.000 €. Je nachdem ist es superschwer zu sagen, wie viel es kostet.

### **A.1.7 Frage 7**

Gibt es einen Preisvergleich zu Dolby Atmos beispielsweise?

Dolby Atmos ist etwas anders. Du kannst mittlerweile die Plugins der Production-Suite einfach kaufen. Dann brauchst du dementsprechend genug In/Out Kanäle in deiner DAW, um alles geben zu können oder du benutzt halt andere DAWs. Da funktioniert jedes Tool unterschiedlich. Eine andere Frage ist noch, ob du es mit Hardware nativ betreibst und und und. Das macht grundlegend schon mal einen Unterschied. Da das hier eine eigene Engine ist, die komplett für 3D ausgelegt und gebaut ist, von Software über Hardware, das ist nochmal sehr speziell. Andere Anbieter haben da ganz andere Preismodelle, ganz andere Lizenzmodelle von Systemen. Das variiert extrem. Von Ambisonics gibt's zum Beispiel ganz viele Tools als Freeware. Da musst du dir im Grunde nur noch Lautsprecher und den Weg zu den Lautsprechern als Hardware suchen und kannst theoretisch loslegen. Aber aus Erfahrung: ich finde das nicht so gut handhabbar von der Installationsseite her. Wenn ich Installation fahre, ist das Ganze klanglich nicht so schnell überzeugend. Da musst du viel mehr anpassen,

bis es zu so einem klanglichen Ergebnis kommt wie mit richtigen dezidierten Engines, die für Installationen gedacht sind.

#### **A.1.8 Frage 8**

Wie schnell kann man Auralite3D an andere Orte befördern? Wie lange dauert die Installation?

Auch das ist gar nicht so pauschal zu sagen, weil es immer drauf ankommt, was ein Kunde damit machen möchte. Was für Räumlichkeiten sind erstmal wichtig. Muss es eine mobile- oder eine Festinstallation sein? Das ist wieder total variabel. Deswegen haben wir uns darauf fokussiert, mit einer kleinen Installation mobil zu sein, mit der wir dann innerhalb von 2-3 Stunden einen Aufbau fahren können, was sehr schnell ist für ein 20 Lautsprecher Setup, was dann immersiv funktioniert. Mit anderen Systemen beziehungsweise, wenn ich größer werden und eine Truss aufbauen muss, dauert es dementsprechend länger. Es kommt immer komplett drauf an, was ein Kunde möchte und wie die Räumlichkeiten beschaffen sind. Es ist auch wichtig, ob ich sogar was an Wänden oder Decken installieren kann. Man sagt normalerweise, wenn man live unterwegs ist mit Truss-System etc.: am besten ist ein Arbeitstag Vorlauf. Wirklich ein Tag zum Aufbauen und Installieren gefolgt von einem oder mehreren Tagen Durchführung. Der Abbau geht dann bekanntlich immer etwas schneller.

#### **A.1.9 Frage 9**

Wie steht es um die Kompatibilität zu anderen Systemen? Können beispielsweise Auralite3D Mischungen auf anderen 3D-Systemen wiedergegeben werden und kann man die Mischungen als Stereo oder Surround ausgeben?

Eigentlich sind dezidierte Mischung immer die beste Lösung. Was aber funktioniert: dadurch, dass wir mit diesem System sehr flexibel sind und uns die Lautsprecherkonfiguration bauen können, wie wir es brauchen, sind wir damit auch sehr anpassbar. Wir können uns im Grunde dann virtuell einfach Lautsprecherkonfigurationen bauen und das System die Mischung dann darauf ausgeben lassen. Sagen wir mal, wir haben eine Auralite3D Mischung auf einem großen System mit 20 oder 30 Lautsprechern gemacht und dann will ich es aber auf einem Surround System haben, dann kann ich verschiedene Wege gehen. Ein Weg wäre zum Beispiel, dass ich mir virtuell Lautsprecher baue, die in einer bestimmten Entfernung zueinander stehen. Sagen wir mal das Ziel ist ein Surround-Setup mit 7.1. Das heißt, ich habe Left, Center, Right, die 90° von mir positionierten Lautsprecher rechts und links und dann noch ein Rearright und Rearleft plus Subwoofer. Die Lautsprecher baue ich mir so virtuell und lasse dann die gleiche Produktion darauf laufen. Das Ganze kann ich dann aufnehmen und habe dadurch eine kanalbasierte Mischung für Surround gebaut. So kann man da rangehen und sich verschiedenste Setups auch aus so einem System ausgeben lassen. Da muss man aber

auf jeden Fall für jedes Setup gucken: hat das so gepasst für meinen Sound, den ich da fahren will, aber in vielen Fällen funktioniert das sehr gut.

#### **A.1.10 Frage 10**

Könnte man eine binaurale Mischung parallel mitproduzieren?

Mit dem System allein nicht, weil es kein Binauralsystem ist oder noch angedockt hat. Das kann man dann parallel natürlich machen mit binauralen Mischtools. Für Liveveranstaltungen haben wir das auch schon so gemacht, dass wir eine binaurale Mischung gemacht haben, aber das ist dann auch ein dezidiertes Mix, den man extra anfertigt. Was sowas zukünftig vereinfachen wird, wenn es standardisiert wird, ist das Format MPEG-H. Zum Teil wird das schon in bestimmte Tools implementiert und in Südkorea ist es auch schon Rundfunkstandard. Im Grunde genommen ist es Mehrspurton mit Metadaten versehen, das heißt mit Koordinaten etc., plus, dass du auch Bilddaten übertragen kannst. Das Ganze in einem Container Format. Dann wird es egal bei welchem Receiver auf deinem Setup wieder decodiert. Sei es, du hast eine Soundbar, sei es, du hast ein Surround Setup oder du hast nur Kopfhörer und sowas. Das wird durch MPEG-H möglich, dass du das auf allen möglichen Systemen dann vice versa ausgeben kannst bzw. einspeisen kannst, weil du dann ein übergreifendes Format oder Formatierungstool hast. So weit ist es aber noch nicht, weil das noch nicht Standard ist. Ich bin auch noch nicht überzeugt davon, dass eine Lautsprechermischung, die dann durch MPEG-H irgendwo ausgespielt und auf Kopfhörern wiedergegeben wird, wirklich geil klingt. Das muss ich erst hören, damit ich überzeugt davon bin. Vielleicht muss man dann eine dezidierte Kopfhörermischung machen, weil, wie gesagt, das sind zwei Welten.

#### **A.1.11 Frage 11**

Könnte man zum Beispiel im Kinobereich von einem Dolby Atmos Setup auf ein Auralite3D Setup wechseln?

Klar, absolut. Lautsprecher anschließen, Positionen ins System eingeben und los geht's. Du kannst im Grunde sofort, wenn du ein fertiges Lautsprechersystem irgendwo stehen hast, mit dem Auralite-System hin, dem System sagen, wo die Lautsprecher sind und loslegen. Gucken, ob alles passt vom Raum, Lautsprecherposition nochmal abstimmen und durchgehen, aber dann brauchst du ja das Setup nicht mehr neu installieren. Es ist wirklich easy straight forward, wenn du die Position weißt, dann bist du richtig fix. Dann kannst du jedes bestehende System damit nutzen.

### **A.1.12 Frage 12**

Die Signalbearbeitung erfolgt vor der Weitergabe an Auralite3D in einer separaten DAW?

Genau, die Effektbearbeitung ist davor. Beispielsweise EQing, Kompression, wenn du Distortion haben willst, wenn es bestimmte Effekte gibt, die du einbinden möchtest durch irgendwelche Tools, Plugins oder auch wirklich durch Hardware - das machst du alles vorher bzw. in der DAW oder im Mischpult. Dann gibst du diese Signale effektbearbeitet an das 3D System und machst da dann alles, was EQing im 3D Bereich angeht, was Position und Bewegung angeht sowie auch 3D Effekte sage ich mal, wie die Raumsimulation.

### **A.1.13 Frage 13**

Die einzelnen Aufnahmen sind dann Mono bzw. Stereo oder braucht man da ein spezielles Aufnahmeverfahren?

Nein, die Signalquellen sind alle Mono oder Stereo.

### **A.1.14 Frage 14**

Gibt's es einen groben Überschlag, wie viele Anwendungen von Auralite3D bereits existieren?

Das ist eine gute Frage. Es gibt so keine direkte offizielle Zahl, weil es ja auch ein Forschungstool beziehungsweise ein Entwicklungstool von Fraunhofer ist. Das ist nicht immer das Produkt Auralite3D aber die gleiche Technologie als Grundlage. Das heißt das System mit diesen Funktionen oder mit diesen Grundlagen gibt es weltweit bestimmt locker 60-70 mal. Bei manchen ist es auch so, wenn es schon vor Jahren irgendwo installiert wurde: nicht alle Einrichtungen, die sowas haben, wollen, dass es bekannt ist. Aus unterschiedlichsten Gründen. Ich weiß auch nicht genau warum, aber das ist auch immer ein Thema. Was darf man nach außen geben, was nicht? Ich habe auf jeden Fall eine Liste, wo ich weiß, die darf ich veräußern, wo die installiert sind, dass man einen groben Überblick bekommt. Das beinhaltet dann auch Systeme, die schon vor dem Produkt Auralite3D installiert wurden, weltweit.

### **A.1.15 Frage 15**

Was wären potentiell zukünftige Einsatzmöglichkeiten von Auralite 3D?

Überall, wo Sound stattfindet oder wo man Sound wiedergeben möchte, bietet sich das System an. Das heißt auch diese Fläche ist überall, wo du Sound nutzt. Jetzt ist die Frage: wofür ist es dann aber noch dezidiert geeignet. Das ist also das Speziellere und das ist überall, wo man Leute irgendwo mehr hineinholen will. Ganz praktisch erklärt nutzen wir das sehr viel für Installation, wo wir verschiedenste Formen zeigen können. Von Produktionen im

Musikbereich oder auch in Kombination mit VR, wo man Filme oder Trailer zeigt und durch den Sound drumrum Leute nochmal mehr eintauchen lässt. Das ist dann auch manchmal eine Kombination aus Kopfhörer- und Lautsprecher-3D, aber auf jeden Fall brauchst du da immersiven Sound, wo du und dann auch wirklich 360 Grad beziehungsweise 3D hörst. Sonst macht es keinen Sinn, wenn du ein Bild überall, hast aber Ton nur von einer Seite irgendwo kommt. Das ist nicht Realismus nachbildend. Wenn du das irgendwie überhören kannst, dass der Sound nur aus einer Richtung kommt und dich trotzdem auf alles fokussieren kannst, dann funktioniert das irgendwie, aber du musst dich halt mehr drauf konzentrieren. Dann für interaktive Installationen. Allein, dass die Hörer ein Teil davon werden können durch die Raumsimulation oder wenn man jetzt noch andere Tools anbindet und dann wirklich interaktive Installationen schafft. Leute können irgendwo einen Button drücken oder sich darin bewegen und es passiert was anderes - solche Möglichkeiten bestehen ja auch mit solchen Systemen bloß, dass du eine Fläche hast, die Leute dazu befähigt, räumlich viel mehr zu agieren und nicht nur mit einem Sound, der irgendwie vor ihnen stattfindet oder von einer Quelle kommt. Das heißt, genauso vielfältig wie die Sound-Variationen weltweit sind, genau so kann man das Tool Auralite3D überall einsetzen. Was es jetzt auch als praktische Anwendungen gibt: einer will einen Club bauen, wo dann wirklich immersiv Clubbing betrieben werden kann. Das gab es auch schon in München, das P1. Da war auch ein System installiert. Das wird aber aktuell nicht mehr aktiv genutzt, aber da gibt es einige Installationsbereiche. Wo es ganz viel Anwendung findet, ist im Livebereich beziehungsweise in Theaterhäusern, Eventhäusern und und und. Da wird es super viel eingesetzt in multifunktionalen Räumen, wo man verschiedenste musikalische Acts hat, wo man aber auch teilweise Theater hat und das zusammenbringt. Die Seebühne in Bregenz ist ein großes Beispiel und es gibt noch etliche andere Häuser, die das installiert haben.

#### **A.1.16 Frage 16**

Die Technik hinter Auralite 3D ist ja vom Fraunhofer IDMT, der TU Delft etc. entwickelt worden. Wurde von euch technisch etwas ergänzt oder ist Auralite quasi das Bedienelement um die bestehende Technik herum?

Fraunhofer IDMT Forschung und Patente stecken im System. IOSONO war eine Auskopplung, die eine ältere Version am Markt anbringen wollte. Die TU Delft bzw. andere Forschungsinstitute waren eventuell in angrenzenden Forschungsbereichen beteiligt, aber die Forschung und Entwicklungen an sich gehen vom IDMT aus und waren wenn meist im Verbund vom Fraunhofer. Aber es gibt auch andere Unis und Institute, die zu Wellenfeldsynthese forschen. Wir haben die existierenden Grundlagen genommen und die Software mit dem Fraunhofer in unsere CI gebracht, so wie einige kleinere Fixes getätigt.

Weitere Anpassungen sind in der Planung. Vor allem die Kombination von live-tauglicher Hard- und Software ist unser Wirken. Die Hardware kommt von spezialisierten Partnern und ist ein wesentlicher Bestandteil von Auralite 3D.

#### **A.1.17 Frage 17**

Bildlich gesprochen berechnet das System ja die erste Wellenfront der jeweiligen virtuellen Schallquelle und gibt deren Ausbreitung mit den verfügbaren Lautsprechern wieder. Wie genau erfolgt die Wiedergabe? Hat das noch irgendwas mit stereophoner Wiedergabe zu tun? Werden die Richtungsinformationen beispielsweise über Laufzeit- und Pegeldifferenzen erzeugt oder gibt es Phantomschallquellen?

Ja genau, es wird die erste Wellenfront berechnet. Hier sind Grundlagen von Wellenfeldsynthese (bei Stereo nicht vorhanden) und weitere Kombinationen von VBAP (Vector Based Amplitude Panning), DBAP (Distance Based Amplitude Panning) ergänzend, damit es mit wenigen Lautsprechern und dynamisch funktioniert. Es bildet auch Phantomschallquellen und man findet bestimmt etwas, was auch bei Stereo angewandt wird, aber nicht in dieser komplexen Form und Zusammenstellung.

#### **A.1.18 Frage 18**

Kann ein Nutzer, der Auralite nur als Playback nutzt, Einfluss auf eine fertige Mischung nehmen? Bei Kanalbasierten Formaten hat man ja keinerlei Zugriff auf die ursprünglichen Audiosignale aber für objektbasierte Anwendungen müssen ja die einzelnen Quellen mit Metadaten übermittelt werden.

Wenn eine fertige Mischung nur als Player File vorhanden ist, kann man nur die Dimensionen beeinflussen, also die Skalierung von der relativen Entfernung auf x-, y- und z-Achse. Wenn man jedoch die Zuspiel-Session bzw. Einzelspuren hat und die Produktions-Session, dann kann man alles variieren und anpassen. Es ist dann auch eine Form Playback, aber ich denke nicht so wie du meinst. Es ist aber nicht wie bei MPEG-H, wo das alles dynamisch eingestellt werden kann. Eine Übersetzung auf MPEG-H ist aber auch eine Überlegung in naher Zukunft. Noch ist es aber nicht der Standard.

#### **A.1.19 Frage 19**

Normalerweise bewegen sich die virtuellen Schallquellen ja auf dem gewählten Corpus oder weiter entfernt. Laut Fraunhofer kann die Technik auch fokussierte Schallquellen erzeugen (Schallquellen, welche scheinbar zwischen Lautsprecher und Zuhörer auftreten). Kann Auralite das und wenn ja, wie funktioniert das?



Damit ist volle Wellenfeldsynthese gemeint, wo du viel mehr Lautsprecher benötigst. Auralite 3D berechnet es auch dementsprechend, aber durch die wesentlich geringere Anzahl an Lautsprechern ist das physikalisch nicht möglich.

#### **A.1.20 Frage 20**

Gibt es trotz variablem Lautsprecheraufbau konkrete Richtlinien (z.B. Symmetrie im Raum, gleichmäßige Abstände zueinander bzw. zum Rezipienten, mindestens 5 Höhenlautsprecher, Abstrahlungswinkel, etc.), die das Erlebnis begünstigen? 20 Lautsprecher sollen ja als immersive Untergrenze gelten. Wie viele Lautsprecher hat euer „Standard Setup“ im Studio?

Nein, nur wenn man etwas nachempfinden möchte, ist es leichter dahin zu kommen (z.B. wenn man ein Surround, Atmos oder Auto3D Setup simulieren möchte). Es kommt hier ansonsten wirklich auf den Raum und vor allem den Content an. Wenn ich beispielsweise einen Helikopter von oben hören möchte, benötige ich dementsprechend Deckenlautsprecher. Grundsätzlich gilt, je mehr Lautsprecher man hat, desto höher ist die Auflösung. Im Studio haben wir um die 22 Lautsprecher verbaut.

#### **A.1.21 Frage 21**

Im Ausschreibungstext steht, dass wahlweise 32 oder 64 Audioobjekte berechnet werden können. Wonach entscheidet sich das?

Das entscheidet sich aktuell noch nach dem Modul. Bis zu einer Produktions Engine sind 64 Quellen möglich, bei Live und RSM aktuell noch 32. Das wollen wir aber auch anpassen, sodass immer 64 Objekte/Inputs möglich sind.

#### **A.1.22 Frage 22**

Die einzelnen Mono-Busse, welche per MADI etc. übertragen werden, sind letztendlich die Audioobjekte. Kann man deshalb theoretisch mehr Ursprungssignale nutzen, solange die in die 32 bzw. 64 Objektkanäle eingeteilt werden?

Man kann in die 32/64 mono Kanäle im Auralite 3D System so viel leiten wie man möchte. Also ja, summieren ist möglich und macht auch bei vielen Produktionen Sinn. Man kann sie, wie du sagst, als Busse/immersive Busse sehen.

#### **A.1.23 Frage 23**

Was bedeutet LTC auf Seite 7 im Manual?

LTC steht für Longitudinal Timecode. Dieser dient zum Synchronisieren vom System mit anderen Systemen, oder auch zum Automatisieren von Objekten.

#### **A.1.24 Frage 24**

Liegt der Koordinatenursprung des Koordinatensystems zur Raumdimensionierung und Lautsprecherkonfiguration in der Raummitte oder an einem Eckpunkt?

Ja genau, der Ursprung liegt in der Raummitte.

#### **A.1.25 Frage 25**

Warum wird eine konkrete Hörfläche definiert? Wie groß kann diese gewählt werden? (Laut WFS Theorie soll ja annähernd der gesamte Beschallungsraum als Hörfläche dienen)

Die Hörfläche bzw. "Listening Area" definiert den Bereich, wo sich Zuhörer befinden. Dieser bestimmt kleine Veränderungen in der Sound Berechnung, die sich wirklich in sehr detaillierten Feinheiten widerspiegeln. Bei WFS ist es der gesamte Raum innerhalb des Lautsprecher-Arrays, aber bei Systemen, wo man wie bei Auralite 3D mit weniger Lautsprechern arbeitet, ist das so physikalisch nicht möglich.

#### **A.1.26 Frage 26**

Wofür werden Reinforcement Lautsprecher genutzt?

Reinforcement Speaker sind im Grunde eine Möglichkeit der Bildung von Delay-Lines. So kann also die Front-Abbildung vom System in einem Theatersaal auch in den ersten oder auch zweiten Stock auf entsprechenden Lautsprechern wiedergegeben werden.

## Eigenständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe. Stellen, die wörtlich oder sinngemäß aus Quellen entnommen wurden, sind als solche kenntlich gemacht. Diese Arbeit wurde in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegt.

Grimma, 24.06.2022  
Ort, Datum

